



**QUEEN'S
UNIVERSITY
BELFAST**

Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis

Murray, J. M., Brennan, S. F., French, D. P., Patterson, C. C., Kee, F., & Hunter, R. F. (2017). Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis. *Social Science & Medicine*, 192, 125-133.
<https://doi.org/10.1016/j.socscimed.2017.09.021>

Published in:
Social Science & Medicine

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

Copyright 2017 Elsevier.

This manuscript is distributed under a Creative Commons Attribution-NonCommercial-NoDerivs License

(<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits distribution and reproduction for non-commercial purposes, provided the author and source are cited.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access

This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: <http://go.qub.ac.uk/oa-feedback>

Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: A systematic review and meta-analysis.

Jennifer M. Murray^{a*}, Dr. Sarah F. Brennan^b, Professor David P. French^c, Prof Christopher C. Patterson^d, Prof Frank Kee^e, Dr. Ruth F. Hunter^{f*}

^aUKCRC Centre of Excellence for Public Health/Centre for Public Health Queen's University Belfast, School of Medicine, Dentistry and Biomedical Sciences, Institute of Clinical Science B, Royal Victoria Hospital, Grosvenor Road, Belfast, UK, BT 12 6BJ, Tel: 02890978944, Email: jmurray39@qub.ac.uk

^bUKCRC Centre of Excellence for Public Health, Queen's University Belfast, UK, sbrennan19@qub.ac.uk

^cManchester Centre of Health Psychology, School of Health Sciences, University of Manchester, UK, david.french@manchester.ac.uk

^dUKCRC Centre of Excellence for Public Health, Queen's University Belfast, UK, C.Patterson@qub.ac.uk

^eUKCRC Centre of Excellence for Public Health, Queen's University Belfast, UK, F.Kee@qub.ac.uk

^fUKCRC Centre of Excellence for Public Health/Centre for Public Health Queen's University Belfast, School of Medicine, Dentistry and Biomedical Sciences, Institute of Clinical Science B, Royal Victoria Hospital, Grosvenor Road, Belfast, UK, BT 12 6BJ; Tel: 02890978944; Email: ruth.hunter@qub.ac.uk

*Corresponding authors

*JM is supported by a PhD studentship funded by the Department for the Economy (DfE). RH is supported by a Career Development Fellowship from the National Institute of Health Research (NIHR) and acknowledges funding support from the HSC Research and Development Division.

Abstract

Background: Physical activity (PA) interventions are generally effective in supporting short-term behaviour change, but increases are not always maintained. This review examined the effectiveness of PA interventions for behaviour change *maintenance* in young and middle aged adults, and investigated which Behaviour Change Techniques (BCTs) and other intervention features were associated with maintenance.

Methods: Six databases (Medline, EMBASE, PsycINFO, Cochrane Database of Systematic Reviews, CINAHL, Web of Science) were systematically searched. Eligibility criteria were controlled trials investigating the effectiveness of PA interventions with adult (mean age 18-64 years) non-clinical populations using validated measures of PA behaviour at baseline and \geq six months post-baseline. Results were pooled in meta-analyses using standardised mean differences (SMD) at five time intervals (6-9, 9-15, 15-21, 21-24, >24 months). Moderator analyses investigated the influence of sample and intervention characteristics on PA maintenance at 6-9 months.

Results: Sixty-two studies were included. PA interventions had a significant effect on behaviour maintenance 6-15 months post-baseline relative to controls. Interventions had a larger effect on maintenance at 6-9 months (SMD=0.28; 95% CI: 0.20, 0.35; I-squared=73%) compared to 9-15 months (SMD=0.20; 95% CI: 0.13, 0.26; I-squared=70%). Beyond 15 months PA measurements were infrequent with little evidence supporting maintenance. Moderator analyses showed some BCTs and intervention settings moderated PA outcomes at 6-9 months. A multivariable meta-regression model showed interventions using the BCTs 'Prompt self-monitoring of behavioural outcome' ($b=1.46$, $P<0.01$) and 'Use of follow-up prompts' ($b=0.38$, $P<0.01$) demonstrated greater effectiveness at promoting PA maintenance at 6-9 months. Interventions implemented in Primary Care (versus community or workplace/university) settings ($b=-0.13$, $P=0.10$) tended to demonstrate less effectiveness.

Conclusions: This review provides evidence of some effective BCTs for maintaining behaviour to 15 months. Greater consideration must be given to how future interventions encourage and measure maintenance of changes, and investigate broader psychological, social and environmental influences of PA behaviour.

PROSPERO registration: PROSPERO 2015:CRD42015025462

Keywords:

Physical activity; Maintenance; Systematic review; Meta-analysis; Behaviour change; Public health

Introduction

Current World Health Organization (WHO) guidelines recommend that adults aged 18-64 years should perform at least 150 minutes of moderate-intensity aerobic physical activity (PA), or at least 75 minutes of vigorous-intensity aerobic PA, or an equivalent combination of moderate-vigorous intensity physical activity (MVPA) weekly for maintained health (World Health Organization, 2010). Approximately 6-10% of global mortality associated with non-communicable diseases (NCDs) is due to physical inactivity (Lee et al., 2012), making it the fourth principal risk factor for death (Kohl et al., 2012). However, encouraging regular PA initiation (i.e. taking up regular PA within six months) and maintenance (i.e. continuing to undertake regular PA longer than six months) is a challenge.

Previous PA interventions have shown modest effects on initiation of PA (Conn et al., 2011; Foster et al., 2005), and reviews have highlighted a lack of reporting of maintenance outcomes (Foster et al., 2013, 2005). A Cochrane review investigating the effectiveness of interventions for promoting PA in adults aged ≥ 16 years found that only six out of 19 included studies reported PA outcomes after six months, none of which found improvement in PA for intervention groups relative to controls (Foster et al., 2005). A more recent review examining the effectiveness of web-based PA interventions on long-term PA behavior among adults reported a small effect (standardised mean difference (SMD) and 95% confidence interval (CI)) at 12 (9 studies; SMD=0.20, 95% CI: 0.11, 0.28) and 24 (1 study; SMD=0.19, 95% CI: 0.08, 0.32) months compared to controls (Foster et al., 2013). Another review of PA interventions delivered face-to-face showed similar effectiveness on PA behaviors at 12 months (8 studies; SMD=0.19, 95% CI: 0.06, 0.31) compared to controls, which was not maintained at 24 months (Richards et al., 2013). Unfortunately, the majority of studies included in this last review failed to measure the long-term (i.e. >12 months) effects on PA (only 3 studies reported PA outcomes >12 months).

Investigation of PA maintenance is hindered by a number of unresolved conceptual issues. One of the most problematic is that researchers use different definitions of 'maintenance'. Some define it in

terms of a timeframe over which the new behaviour is carried out (typically 3-6 months after intervention completion) (Fjeldsoe et al., 2011; Prochaska and DiClemente, 1982), and others conceptualise it in terms of achieving behavioural automaticity (i.e. when it is efficiently and effortlessly carried out) (Rothman, 2000). Kwasnicka et al. (2016) suggest maintenance is achieved when the new behaviour becomes the 'dominant response' (i.e. has the highest probability of being enacted across times and contexts). In terms of behaviour change theories, the Transtheoretical Model (TTM) defines the individual as being in the 'maintenance' stage after being sufficiently active for six months (Prochaska and DiClemente, 1982).

In an attempt to address these issues, Fjeldsoe et al. (2011) adopted a two-fold definition of maintenance in their review of the effectiveness of PA and dietary interventions for behavior change maintenance among adults. With this definition, behaviour change maintenance was achieved when a significant intervention effect was reported (i.e. any increase in PA for the intervention compared to control group) at end of-intervention *as well as* at follow-up (i.e. three months post-intervention). However, the authors noted that this may have been restrictive with respect to longer-term community-based interventions with ongoing contact. This definition also restricts inclusion of studies with no defined end period, (e.g. altering the built environment). Furthermore, applying a threshold-based maintenance criterion is complicated by heterogeneity of outcomes and measurement tools across studies, and infrequent reporting of the magnitude of behaviour change between post-intervention and follow-up (Fjeldsoe et al., 2011). Evidently, there are many complexities inherent in conceptualising behaviour change maintenance and researchers have diverse opinions as to how it should be defined.

It is clear that this lack of consensus has impeded our understanding of what interventions are most effective for behaviour change maintenance (Ory et al., 2010). The present review adopts the operational definition of 'maintenance' as any positive and significant intervention effect at least six months post-baseline (with no restriction on intervention duration) in line with the TTM. This is also

consistent with the argument that individuals who maintain an increase in PA for six months are usually viewed as successful maintainers (Marcus et al., 2000). Further, research shows that the highest likelihood for relapse occurs within six months of starting an exercise program (Dishman, 1994). Physical activity interventions involving inactive participants usually aim to achieve any increase in PA rather than solely focusing on increasing PA to a target level (Conn et al., 2011). Indeed, current PA guidelines acknowledge that getting some PA is better than none (Office of Disease Prevention and Health Promotion, 2017). Given the lack of consensus in defining PA maintenance, we also examined an alternative definition (i.e. requiring a significant post-intervention effect on PA behaviour with PA outcomes measured again at least six months after the intervention-end), which is consistent with how intervention studies 'typically' operationalise maintenance (Marcus et al., 2000).

There are also weaknesses in the theoretical developments in understanding maintenance. In particular, there is limited evidence regarding which theories best support behavior change maintenance and whether they differ from those supporting PA initiation (Kwasnicka et al., 2016; Nigg et al., 2008; Rothman, 2000). A comprehensive review of 83 behaviour change theories reported that the most popular theories (e.g. TTM, Social Cognitive Theory (SCT)) rely on a static structure, rarely accounting for behavioural changes over time (Michie et al., 2014). More recent studies have compared theoretically-derived predictors of maintenance (Kassavou et al., 2014) and there is increasing evidence to support Rothman's hypothesis that the determinants of initiation and maintenance differ for a range of health behaviours, including PA (McAuley and Blissmer, 2000; Rothman, 2000). For example, Nigg et al. (2008) highlight that the correlates of PA initiation versus maintenance differ in terms of the focus of fundamental psychological variables (and temporal associations), and the inclusion of factors that impede PA. The Health Action Process Approach defines a phase-specific model of causality, separating the concept of self-efficacy according to the individual's phase (i.e. motivation or action) of behaviour (Schwarzer, 1992). Physical Activity

Maintenance Theory considers the impact of life stress in terms of re-directing personal resources away from focusing on PA, and increasing negative affect (Nigg et al., 2008).

One way to understand differences in the effectiveness of PA interventions for initiation and maintenance of behaviour is to examine the Behaviour Change Techniques (BCTs) employed. BCTs have been defined as the irreducible components or 'active ingredients' of interventions, designed to redirect the process of behaviour (Michie et al., 2013). Recent reviews examining whether BCTs, and other intervention features, are associated with increased effectiveness (i.e., greater improvements in behavioural changes) have mainly focused on initiation of behaviour change. These reviews suggest that self-regulation techniques such as goal setting, self-monitoring, action planning, and prompts are associated with more effective interventions in terms of initiation of PA (Dombrowski et al., 2012; Michie et al., 2009; Williams and French, 2011). In comparison, Fjeldsoe et al. (2011) specifically examined effectiveness of interventions for behaviour change maintenance and found that studies with specific methodologies (e.g. higher attrition and pre-trial behavioural screening) and intervention characteristics (e.g. longer duration, in-person contact, more intervention strategies, use of follow-up prompts, and targeting sample characteristics such as gender) were more effective at improving PA maintenance. Unfortunately, the authors did not report intervention components based on a standardised BCT taxonomy.

Given the lack of clarity on key issues in the literature, the present review builds upon previous research by focusing on the extent of behavioural maintenance in PA interventions among adults. Using systematic review and meta-analysis methods, objectives include: (1) To examine how effective PA interventions have been in achieving behaviour change maintenance at least six months post-baseline; (2) To investigate which sample (e.g. age, gender) and intervention characteristics (e.g. duration, setting, BCTs) are associated with PA behaviour maintenance.

Methods

Search strategy and data abstraction

Six electronic databases were systematically searched including Medline, EMBASE, PsycINFO, Cochrane Database of Systematic Reviews, CINAHL and Web of Science. Searches were carried out from inception to 16th January 2016. The search strategy included terms relating to PA, behaviour change, maintenance and study design (supplement 1). Reference lists of relevant systematic reviews, meta-analyses, and included studies were searched manually.

Eligibility criteria included:

- I. Primary aim of study was to promote PA behaviour. Interventions targeting multiple health behaviours were excluded (e.g. studies targeting PA and diet)
- II. Non-clinical, adult population aged 18-64 years. Populations with a diagnosed clinical condition (e.g. Diabetes, Cancer, Cardiovascular Disease, etc.) were excluded
- III. Measure of PA at baseline and at least six months post-baseline (i.e. in line with the study definition of maintenance)
- IV. Objective or validated self-report PA measure (i.e. a measure with at least one appropriate validation study for the target population)
- V. Prospective intervention design: Randomised controlled trial (RCT), controlled trial or cluster RCT. Interventions of any duration were considered. Due to the ongoing nature of some interventions (e.g. interventions altering the built environment), PA measurements may have taken place during an active intervention, directly post-intervention, or after a period of no-intervention
- VI. Control group (e.g. no intervention; wait-list) or comparison group (e.g. usual care; minimal intervention; intervention unrelated to PA)
- VII. Published in English language.

Titles and abstracts were screened by JM; 25% were independently screened by a second reviewer (RH). All full texts were examined by JM and independently by RH or SB. In cases of disagreement or uncertainty, consensus was reached via discussion with a third reviewer (DF). Percentage agreement and Cohen's kappa (Cohen, 1960) were calculated for all screening stages and BCT coding.

Relevant data were extracted by JM and verified by RH, SB and CP. Authors were contacted in an attempt to retrieve unreported or missing data. The results of studies which could not be included in meta-analyses due to missing data were summarised narratively.

Intervention components were independently extracted by three trained coders (JM, RH and SB) according to the 40-item 'Coventry, Aberdeen & London – Refined' (CALO-RE) BCT taxonomy (Michie et al., 2011). This taxonomy was specifically designed to be used with interventions targeting PA and dietary behaviours. Disagreements were resolved by discussion with another coder (DF). The CALO-RE taxonomy was supplemented with 'Credible source' from the 93-item 'BCT taxonomy v1' (Michie et al., 2013) to enable coding of BCTs regarding exercise advice, information or demonstration provided by a trained professional. Coded BCTs were categorised into domains (Michie et al., 2013).

Risk of bias was assessed using the Cochrane Risk of Bias tool (Higgins and Green, 2011), and categorised as 'high', 'low' or 'unclear' for each domain.

Physical activity outcomes and effect size calculation

Meta-analyses included objective measures (e.g. pedometer or accelerometer) or self-report estimates of time spent in PA, MVPA time, leisure-time PA, energy expenditure from PA and percentage of participants meeting recommendations for MVPA. Effect sizes (i.e. SMDs) represented PA behaviour maintenance at 6-months post-baseline. SMDs were calculated using between-group changes from baseline values for continuous outcomes (e.g. steps/day, MVPA minutes/week) and interpreted according to Cohen's classification as small (0.2-0.5), medium (0.5-0.8) and large (0.8) (Cohen, 1988; Higgins and Green, 2011). When studies reported post-intervention PA outcomes

adjusted for baseline differences, effect sizes were calculated from the adjusted estimates. When studies reported only mean and standard deviation values at baseline and follow-up ($k=26$), an assumption was made regarding the correlation between baseline and post-intervention PA levels (0.7 at 6-9 months, 0.5 at 9-21 months, 0.3 at >21 months). This captured the tendency for correlation between baseline and follow-up PA to decline at increasing time from baseline (Malina, 2001). For dichotomous outcomes (e.g. percentage of participants meeting PA recommendations), log odds ratios were calculated and converted to SMDs using the Chinn (2000) equation (Higgins and Green, 2011). When interventions employed more than one PA intervention arm ($k=13$), the control group sample size was split to provide separate comparisons (Higgins and Green, 2011). For studies which did not adjust results for clustering ($k=4$), sample sizes were adjusted according to recommendations outlined in the Cochrane handbook (Higgins and Green, 2011; Rao and Scott, 1992). When studies did not report intra-class correlations ($k=3$), an assumption was made based on a previously published report (Ukoumunne et al., 1999). Reported eta-squared values were converted to Cohen's d effect sizes using established procedures (Cohen, 1988) ($k=2$). All effect sizes were calculated using aggregate-level data (i.e., sample means and standard deviations for each study). A significant intervention effect ($P<0.05$) was determined when the 95% CI excluded 0 for the SMD.

Weighted SMD effect sizes were combined using the inverse variance method and a random-effects model (DerSimonian and Laird, 1986). The main meta-analysis employed measures taken 6-9 months post-baseline, and was repeated for measures at 9-15 months, 15-21 months, 21-24 months and >24 months. Examination of outcomes reported at different time intervals permitted the assessment of (1) the frequency with which studies measured PA at increasing time intervals from baseline; (2) how effective PA interventions have been in promoting maintained behaviour at increasing time intervals from baseline. The selected time intervals reflect the tendency of PA studies to measure effectiveness at six month time intervals (i.e. 6 months, 12 months and 24 months) (Seymour et al., 2010). Meta-analyses were conducted using Review Manager (The

Cochrane Collaboration, 2014). To facilitate interpretation of SMD results, the pooled SMD for each meta-analysis (i.e. 6-9, 9-15, 15-21, 21-24, >24 months) was converted to difference in minutes/week of PA between interventions and controls. SMDs were converted to minutes/week of PA by calculating a pooled standard deviation weighted by sample size (Cohen, 1988), for controls in studies reporting this outcome. The difference in minutes/week of PA between interventions and controls was calculated as the product of the pooled standard deviations and the pooled SMD effect sizes for each meta-analysis.

Cochran's Q was calculated and Chi-square tests determined whether variation observed between study results was consistent with sampling error only. Cochran's Q was transformed to I-squared statistics (i.e. the percentage of variability in SMDs due to heterogeneity between studies rather than chance (Higgins and Green, 2011)). Heterogeneity and reporting bias were assessed visually using forest and funnel plots created using Review Manager (The Cochrane Collaboration, 2014). The Egger et al. (1997) and precision-effect estimate with standard error (PEESE) (Stanley and Doucouliagos, 2014) tests for study size effects were used to check for publication bias in the 6-9 month meta-analysis. This was carried out by performing weighted least squares regressions with SMD effect sizes as the dependent variable. The independent variable for the Egger test was the SMD standard error (SE) whilst the squared standard error (SE^2) was used as the independent variable for the PEESE test. Regression weights were specified as $1/SE^2$. The intercept was interpreted as the mean SMD effect size adjusted for publication bias. Statistical significance was set at the 5% level ($P < 0.05$).

Sensitivity analyses on 6-9 month outcomes

Sensitivity analyses were conducted to determine whether the results of the 6-9 month analysis were robust to omission of outlying studies, studies not performing Intention-to-treat (ITT) analyses and studies with attrition rates higher than 20%. This determined whether pooled effect sizes differed for studies of higher versus lower risk of bias. The literature suggests that attrition rates

>20% should raise concern about the possibility of bias (Schulz and Grimes, 2002). Use of ITT analysis is suggested to minimise bias (Higgins and Green, 2011; Newell, 1992). The correlation assumption was tested repeating the 6-9 month meta-analysis with correlation coefficients of 0.3 and 0.5. To test the robustness of our maintenance definition, we performed a sensitivity analysis including studies reporting positive post-intervention effects, with PA behavioural outcomes measured at least six months after the intervention had actively ended. Interventions which did not actively end (e.g. ongoing contact/altering the built environment) were not included in this sensitivity analysis.

Moderator analyses for 6-9 month outcomes

Moderator analyses were carried out for studies reporting PA maintenance outcomes at 6-9 months to explore potential sources of heterogeneity. Study-level sample and intervention characteristics were investigated as potential moderators and included: age, percentage females, intervention setting (community, work/university, primary care), delivery method (individual, group, both), contact type (in-person/remote), intervention duration in months (i.e. period of time during which the intervention was active including a period of limited contact if applicable, or the length of the last follow-up from baseline for studies of interventions with no active end), ITT analysis, attrition rate (%), and individual BCTs. Continuous variables were mean-centred. Dummy variables were created for categorical moderators. Univariable and multivariable random-effects meta-regressions (i.e. regression models with each comparison weighted by the sum of the inverse of the within-trial variance, and residual between-trial variance (Thompson and Higgins, 2002)) with residual maximum likelihood estimation were performed. Variables significant in univariable analyses ($P \leq 0.10$) were retained and evaluated simultaneously in the multivariable model. The multivariable model was refined using a stepwise procedure with successive elimination of variables ($P > 0.10$), starting with the variable with the highest p-value. The model's I-squared statistic determined whether it was effective at explaining some of the between-study heterogeneity responsible for variability in SMDs

(indicated by a decrease compared to the I-squared statistic for the 6-9 month meta-analysis). The Stata 'metareg' command was used for meta regression analyses (StataCorp, 2013).

Weighted mean effect sizes (d_+) and 95% CIs were estimated at different levels of each moderator in the multivariable model using the moving constant technique (Johnson and Huedo-Medina, 2011). These estimates (\hat{d}_+) represent the change in PA levels from baseline (expressed in SMD units) for the intervention group relative to controls, while statistically controlling for each moderator in the model (held constant at 0) except for the moderator and level of interest. An optimal model was formed representing the greatest potential benefit to PA, in which all moderators were simultaneously evaluated at the level associated with the greatest increases in PA.

Subgroup analyses were conducted to complement meta-regressions and explore heterogeneity by showing how effect sizes differed between groups of studies. Study-level characteristics included age (above/below sample mean), gender (above/below sample mean), intervention setting (community, primary care, work/university), PA outcome (objective/self-report) and intervention duration (0-3 months, >3-6 months, >6 months). Chi-squared statistics and p-values were used to test for subgroup differences.

Results

5,029 titles and abstracts and 364 full texts were screened, resulting in 62 studies that met the inclusion criteria. Of these, 52 studies were included in meta-analyses (figure 1). Coders exhibited high reliability and agreement for the screening of titles and abstracts (Cohen's kappa=0.67, percentage agreement=90%), full-texts (Cohen's kappa=0.68, percentage agreement=84%), and coding of BCTs (Cohen's kappa=0.83, percentage agreement=96%) (Landis and Koch, 1977). Authors of 36 included studies were contacted via email with requests for data, or clarification of results reported in articles; 11 replied with the requested information and four replied stating that data was unavailable.

Study characteristics

A summary of study characteristics is provided in supplement 2. The majority of studies were conducted in the USA ($k=32$) and used a RCT design ($k=45$). Overall, 61,690 adults (mean age 52.5 years, 64.1% female) participated (i.e. were recruited/randomised). Sixty-one studies measured PA using self-report measures and 18 studies employed objective measures. Seventeen studies reported both self-report and objective PA measures.

Intervention characteristics

Interventions took place in the community ($k=39$), primary care ($k=13$) and work/university ($k=10$). Mean intervention duration (during which the intervention was active) was 8 months (median six months; range 24 months). The TTM and/or SCT were the most commonly employed theoretical frameworks ($k=42$). Forty studies reported ITT analyses and the average attrition rate at 6-9 months was 19.5% ($k=27$). Seven studies included group-based contact only, 18 had individual- and group-based contact, and 35 included in-person contact.

Thirty-four of 41 BCTs were coded in studies. BCTs such as 'Goal setting (behaviour)' ($k=44$), 'Barrier identification/problem solving' ($k=35$), 'Prompt self-monitoring of behaviour' ($k=33$), 'Provide information on consequences of behaviour in general' ($k=31$), 'Plan social support/social change' ($k=29$), and 'Credible source' ($k=29$) were most commonly employed (supplement 3).

Effectiveness of interventions for PA maintenance

Thirty-two studies reported PA behaviour maintenance between 6-9 months and could be incorporated into the meta-analysis. Eleven studies reported results separately for different intervention groups resulting in $k=43$ comparisons. The pooled SMD was 0.28 (95% CI: 0.20, 0.35), indicating that interventions had small, but significant effects on PA maintenance compared to controls. This SMD is the equivalent of 136 more minutes/week of PA for interventions compared to controls. The I-squared statistic showed substantial heterogeneity at 73% and the chi-square test

showed evidence to reject the null hypothesis that the studies were estimating the same underlying intervention effect ($P < 0.001$).

Thirty-one studies ($k=44$) reported PA outcomes at 9-15 months, showing that interventions had a small effect (SMD=0.20; 95% CI: 0.13, 0.26; I-squared=70%) on PA behavior that was maintained up to 15 months post-baseline. This SMD is the equivalent of 105 more minutes/week of PA for interventions compared to controls. Fewer studies reported PA outcomes at 15-21 months (four studies, $k=5$: SMD=0.04, 95% CI: -0.02, 0.10; I-squared=0%), 21-24 months (nine studies, $k=10$: SMD=0.08, 95% CI: 0.01, 0.14; I-squared=57%), and beyond 24 months (three studies, $k=5$: SMD=0.04, 95% CI: -0.05, 0.13; I-squared=0%), showing small, inconsistent effects on sustained PA behavior (supplement 4). These SMDs are the equivalent of 15, 50 and 15 more minutes/week of PA for interventions compared to controls, respectively. Thus, there was little evidence of significant intervention effects on PA maintenance beyond 15 months.

Ten studies were not included in meta-analyses as the required data were not available ($k=6$), or PA outcomes were not comparable with the other included studies ($k=4$). At 6-9 months, five out of seven studies reported significant effects on PA maintenance, while two reported non-significant effects. Four studies reported effects beyond nine months and found that PA behavior was not successfully maintained (supplement 2).

Publication bias

From visual examination of the funnel plot there was some suggestion of asymmetry possibly indicative of publication bias in the 6-9 month meta-analysis, although the plot did not conform particularly well to the expected funnel shape (supplement 4). Both the Egger and PEESE methods showed evidence of small-study effects ($P=0.002$). The intercept, representing the best estimate of intervention effect after adjustment for small-study effects, was non-significant (0.04; 95% CI: -0.07,

0.14) using Egger's method, but was significant (0.11; 95% *CI*: 0.03, 0.18) using the PEESE approach. The results of these tests should be interpreted cautiously due to the high degree of heterogeneity.

Risk of bias

Twenty-nine studies had low risk of bias for randomisation, 24 had unclear risk and nine had high risk. For allocation concealment, eleven studies had low risk of bias, 42 had unclear risk and nine had high risk. For blinding of outcome assessment, 19 studies had low risk of bias, 38 had unclear risk and five had high risk. For incomplete outcome data, 36 studies had low risk of bias and 26 had unclear risk. Results of the risk of bias assessment for all domains are included in supplement 5.

Sensitivity analyses

To test the robustness of our 6-9 month analyses and the assumed correlation coefficient of 0.7 between baseline and follow-up PA measures, we re-analysed these data using more conservative values; similar effect sizes estimates were observed assuming correlation coefficients of 0.3, 0.5 and 0.7 (SMDs 0.22, 0.24 and 0.28). Including only those studies carrying out an ITT analysis resulted in a SMD of 0.31 whilst pooling the effect sizes of studies with an attrition rate of less than or equal to 20% yielded a SMD of 0.29. Examination of the forest plot revealed two outlying studies (Castro et al., 2011; Sriramatr et al., 2014). Excluding the three comparisons contributed by these studies resulted in a slightly smaller, less heterogeneous but similar effect (SMD=0.22; 95% *CI*: 0.15, 0.28; I-squared=57%).

Sensitivity analysis using the alternative two-fold definition of 'maintenance' (i.e. reporting a significant post-intervention effect on PA behaviour and measuring PA again at least six months later) was conducted on a subset of 13 studies ($k=14$ comparisons). These studies revealed that PA interventions had a small but significant effect (SMD=0.18, 95% *CI*: 0.08, 0.27; I-squared=71%) on behavior maintenance, an effect that was substantially smaller than those reported in the primary analysis using a more inclusive definition (supplement 4).

Moderator analyses

Thirty-two studies ($k=43$) were included in moderator analyses. Univariable and multivariable meta-regression results are reported in supplement 6. Weighted mean SMD effect size estimates are presented in table 1. Variables significant in univariable analyses ($P \leq 0.10$) were all retained in the final multivariable model: 'Prompt self-monitoring of behavioural outcome' ($b=1.46$, $P < 0.01$), 'Use of follow-up prompts' ($b=0.38$, $P < 0.01$) and 'Primary care (versus community)' ($b=-0.13$, $P=0.10$). I^2 was reduced from 73% to 51%, indicating that the model was effective at explaining some of the between-study heterogeneity responsible for variability in SMDs. There was a non-significant mean SMD effect size estimate for interventions implemented in primary care ($\hat{d}_+ = 0.10$; 95% CI: -0.05, 0.24) controlling for the other moderators. The greatest potential PA benefit was achievable when interventions were implemented in non-primary care based settings, and used the BCTs 'Prompt self-monitoring of behavioural outcome' and 'Use of follow-up prompts'. The additive model produced a mean SMD effect size estimate of 1.94 (95% CI: 1.28, 2.59).

Subgroup analyses examining intervention setting (community, primary care or work/university) found a significant effect ($P < 0.001$). Although studies reporting objective PA measures had larger effect sizes at 6-9 months (SMD=0.50; 95% CI: 0.09, 0.91) than studies with self-report PA measures (SMD=0.24; 95% CI: 0.17, 0.32), this difference was not statistically significant. Subgroup analyses showed no significant differences for age, gender or intervention duration (supplement 4).

Discussion

The results of the present meta-analyses found that the effectiveness of PA interventions on PA behavior maintenance measured at 6-15 months was small (0.20-0.28) compared to control groups, although substantial between-study heterogeneity was indicated. There was little evidence of maintained change beyond 15 months due to the limited number of studies measuring PA and studies showing little effectiveness on PA maintenance at these longer time-points. Moderator

analyses revealed some potential sources of heterogeneity. A multivariable model with the variables: 'primary care (versus community)', 'prompt self-monitoring of behavioural outcome' (i.e. recording specific measures expected to be influenced by behaviour change) and 'use of follow-up prompts' (i.e. gradual reduction in intervention intensity or frequency of contact over time) reduced heterogeneity, from $I^2=73\%$, to moderate levels ($I^2=51\%$), indicating that differences in these variables amongst studies explained some of the variability in effect size.

These effect sizes are comparable with previous reviews finding small-moderate PA effect sizes to 12 months post-baseline (Foster et al., 2013; Williamson et al., 2015), and acknowledging that although PA maintenance outcomes are rarely reported, when they are reported maintenance is frequently achieved (Fjeldsoe et al., 2011; Gourlan et al., 2011). The effectiveness of PA interventions in the workplace (Malik et al., 2014; Schröder et al., 2014), university (Plotnikoff et al., 2015) and community (Bock et al., 2014) has been documented. In comparison, the present review showed null effect size estimates for interventions implemented in primary care. The limitations of primary care settings for promoting PA maintenance is acknowledged in the National Institute for Clinical Excellence (NICE) guidelines which currently make no reference to maintenance and have identified reporting of long-term effects as a key gap in the evidence (NICE, 2013). The present review focused solely on non-clinical populations, suggesting that primary care may not be the most appropriate setting for promoting PA maintenance in healthy populations but may have a significant role in promoting PA in clinical populations (i.e. disease specific populations not eligible for inclusion in our review) (Orrow et al., 2012; Sanchez et al., 2015).

The present review showed some self-regulation techniques (e.g. prompting participants to monitor outcomes of PA behaviour change) were effective strategies used in studies for PA maintenance. This is supported by the findings of a previous systematic review investigating effective techniques in changing PA and self-efficacy in obese adults (Olander et al., 2013). Samdal et al. (2017) also found that providing feedback on behavioural outcomes was a useful technique for promoting PA and

healthy eating in overweight/obese adults. The relevance of self-regulation techniques for PA behaviour has previously been evidenced (Michie et al., 2009; Samdal et al., 2017; Williams and French, 2011) and systematic reviews have found self-monitoring using pedometers to be effective for increasing PA and improving health outcomes (Bravata et al., 2007; Hobbs et al., 2013). Harkin et al. (2016) also found that monitoring goal progress can help promote behaviour change which is in line with a Control Theory framework (Carver and Scheier, 1982) and also Rothman's theory of behavioural maintenance which proposes that outcome satisfaction determines whether behaviour is maintained (Rothman, 2000). Evidence suggests that PA maintenance may be predicted by the realisation of positive social, mental-wellbeing and fitness goals (Brassington et al., 2002; Kassavou et al., 2014). Encouraging participants to regulate their behaviour by monitoring its outcomes and determining how satisfied they are may therefore be one potential method of inducing maintenance.

Use of follow-up prompts in studies, with gradual removal of intervention reinforcement, is associated with greater PA maintenance according to the present results and previous reviews (Avery et al., 2012; Fjeldsoe et al., 2011). Learning Theory states that positive reinforcement (e.g. rewards) is effective for maintenance only when an extinction schedule is implemented and sudden withdrawal causes behaviour to be extinguished (Johnston, 2016). Gradual withdrawal in the maintenance phase or intermittent reinforcement in the acquisition phase (with rewards given for a selection of correct behavioural responses) is recommended to increase behavioural resilience to extinction. Interventions promoting behavioural automaticity via gradual withdrawal of the support of interventions which effectively induce initiation should be more effective at promoting maintenance given the exchange between self-regulatory and automatic processes (Kwasnicka et al., 2016).

Methodological quality

Methods of randomisation and allocation concealment were poorly described but where reported, were appropriate. Studies classified as high risk of bias for randomisation and allocation concealment were mainly large community based interventions for which this was not feasible. Blinding of participants and personnel is also generally unfeasible in PA interventions and the quality assessment reflected this. Where attrition rates were high, reasons for drop-out and/or characteristics of drop-outs were moderately-well reported (e.g. 42% of studies had unclear risk of bias for incomplete outcome data). Some studies may have been subject to attrition biases. This is important given the focus on maintenance in the present review and increasing attrition rates as follow-up time increases. Where missing data was imputed, studies generally reported using conservative methods (e.g. baseline carried forward) decreasing the possibility that effect sizes were overstated. We conclude that our results are based on evidence that was generally reliable.

Strengths and Limitations

The present review is the first to investigate PA maintenance in adult non-clinical populations with a focus on potential effect modifiers, particularly BCTs, using meta-analysis and meta-regression. It expands upon the findings of previous reviews by exploring causes of variability in effectiveness between studies. There was a large degree of heterogeneity and moderator analyses did not explain all variability. Whilst several of the meta-analyses we conducted relied on results reported in a small number of studies (supplement 4), guidelines suggest that meta-analyses results with fewer than ten comparisons should be interpreted with caution due to reduced power (Higgins and Green, 2011). Egger's and PEESE tests indicated results may have been affected by small-study bias with the caveat that results of these tests should be interpreted with caution due to heterogeneity (Stanley and Doucouliagos, 2014). Furthermore, our search was limited to publications in the English language. Several reviews have defined PA maintenance as occurring when a period has elapsed after the intervention ends (Avery et al., 2012; Fjeldsoe et al., 2011; Gourlan et al., 2011) and our definition may be viewed as relatively lenient. Forty of the 62 (65%) studies in the present review reported PA

after the intervention had actively ended. Including an eligibility criterion for PA measurement after the intervention had actively ended would have penalised studies of longer intervention duration with no follow-up after intervention-end. Although our moderator analyses did not show an association between intervention duration and PA maintenance 6-9 months post-baseline, this may be an important consideration for maintenance at longer time-points. Previous reviews have found that interventions of longer duration more often promote maintenance (Fjeldsoe et al., 2011; Robertson et al., 2014). Previous research suggests that PA maintenance requires ongoing intervention support for some or all participants (Hall et al., 2010) and some recent PA interventions have been designed with specified maintenance components providing ongoing support (Yates et al., 2015). Interventions altering the physical environment, for which there is no active end, are also increasingly recommended (Butland et al., 2007) and have been the focus of some recent PA studies (Prins et al., 2016). In acknowledgement of the fact that researchers differ in opinion in terms of how PA maintenance should be defined, we conducted a sensitivity analysis requiring studies to report a significant post-intervention effect for intervention versus control, and additionally to measure PA behaviour six months after the intervention actively ended. This did not significantly change the results and we are therefore confident in the present review's conclusions regarding the literature on behavioural maintenance in PA interventions.

Recommendations

It is of critical importance that the research community establishes a standard definition of behaviour change maintenance to progress this research area as understanding is impaired by different authors describing different phenomena using the same label of 'maintenance'. Future PA interventions should be designed with maintenance of behaviour change in mind from the outset. For example, trial designs must consider attrition rates and associated sample size calculations, appropriate measurement time-points, and suitable techniques to reduce participant burden when

assessing behaviour change maintenance. Studies should investigate specific constructs hypothesised to induce initiation and maintenance to uncover how these processes differ.

Existing behaviour change theories generally tell us little about the dynamic nature of behaviour change over time (Michie et al., 2014) and behavioural measurement in the included PA studies reflected this (e.g. most studies measured PA at six month intervals with few using shorter, more continuous intervals or modelling maintenance trajectories (Seymour et al., 2010)). Included studies did not explicitly measure maintenance in terms of behaviour sustained at a target level and our operational definition of 'maintenance' reflected this. Future research must (1) collect measures that vary over the course of a study and influence behaviour change, (2) report maintenance outcomes in terms of target PA levels, and (3) focus on the specific individuals and times where behaviour changed to develop theoretical mechanisms for further investigation in intervention studies.

Future interventions should incorporate evidence-based BCTs for maintenance (i.e. self-monitoring of behavioural outcomes, use of follow-up prompts) and research should uncover which other BCTs foster behaviour change maintenance. Subsequently, research should outline how BCTs designed for maintenance are linked to theoretical components of behaviour change maintenance theories (e.g. self-efficacy, satisfaction). Consideration should be given to the setting in which PA maintenance might best be achieved for adult non-clinical populations. For example, interventions in workplace and community settings can provide habitual, supportive environments for PA. The Public Health White Paper (i.e. UK Government report informing the public on the government's stance on complex issues like PA) criticises the 'one-size-fits-all' approach and calls for interventions acknowledging the wide range of factors affecting health and well-being (Department of Health, 2010). By putting local communities 'at the heart of public health', neighbourhood-specific public health challenges can be more effectively addressed and the potential for community-based interventions to achieve PA behaviour change has been recently evidenced (Sallis et al., 2016). Local level interventions with ongoing support addressing the particular needs of the community are

potentially a more effective means of inducing lifelong PA behaviour change. Future research should examine population level interventions acknowledging the broader context (psychological, social and environmental) in which PA occurs.

Conclusion

This review and meta-analysis has shown that PA interventions have supported maintenance of behaviour change to 15 months post-baseline. Beyond 15 months, PA behaviour was measured infrequently with little evidence to support maintenance. Moderator analyses identified several intervention techniques associated with greater PA maintenance at 6-9 months and these previously received support. Interventions encouraging self-monitoring of behavioural outcomes, or using follow-up prompts were more effective. Interventions implemented in primary care were less effective than community- or work/university-based settings. Future interventions and trial design should consider how to induce and measure maintenance of changes, and investigate the broader psychological, social and environmental context within which PA occurs.

References

- Avery, L., Flynn, D., van Wersch, A., Sniehotta, F.F., Trenell, M.I., 2012. Changing physical activity behavior in type 2 diabetes: a systematic review and meta-analysis of behavioral interventions. *Diabetes Care* 35, 2681–2689.
- Bock, C., Jarczok, M., Litaker, D., 2014. Community-based efforts to promote physical activity: a systematic review of interventions considering mode of delivery, study quality and population subgroups. *J. Sci. Med. Sport* 17, 276–282.
- Brassington, G.S., Atienza, A.A., Perczek, R.E., DiLorenzo, T.M., King, A.C., 2002. Intervention-related cognitive versus social mediators of exercise adherence in the elderly. *Am. J. Prev. Med.* 23, 80–86.
- Bravata, D.M., Smith-Spangler, C., Sundaram, V., Gienger, A.L., Lin, N., Lewis, R., Stave, C.D., Olkin, I., Sirard, J.R., 2007. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA* 298, 2296–2304.
- Butland, B., Jebb, S., Kopelman, P., Mcpherson, K., 2007. Tackling obesities: future choices - project report (2nd edition).
- Carver, C.S., Scheier, M.F., 1982. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. *Psychol. Bull.* 92, 111–135.
- Castro, C.M., Pruitt, L.A., Buman, M.P., King, A.C., 2011. Physical activity program delivery by professionals versus volunteers: the TEAM randomized trial. *Health Psychol.* 30, 285–294.
- Chinn, S., 2000. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat. Med.* 19, 3127–3131.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. ed. Hillsdale (NJ): Lawrence Erlbaum Associates, Inc.
- Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46.
- Conn, V.S., Hafdahl, A.R., Mehr, D.R., 2011. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. *Am. J. Public Health* 101, 751–758.
- Department of Health, 2010. *Healthy lives, healthy people: our strategy for public health in England*.
- DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. *Control. Clin. Trials* 7, 177–188.
- Dishman, R.K., 1994. *Advances in Exercise Adherence*. Human Kinetics, Leeds.
- Dombrowski, S.U., Sniehotta, F.F., Avenell, A., Johnston, M., MacLennan, G., Araújo-Soares, V., 2012. Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: a systematic review. *Health Psychol. Rev.* 6, 7–32.
- Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315, 629–634.
- Fjeldsoe, B., Neuhaus, M., Winkler, E., Eakin, E., 2011. Systematic review of maintenance of behavior change following physical activity and dietary interventions. *Heal. Psychol.* 30, 99–109.
- Foster, C., Hillsdon, M., Thorogood, M., Kaur, A., Wedatilake, T., 2005. Interventions for promoting physical activity. *Cochrane Database Syst. Rev.*
- Foster, C., Richards, J., Thorogood, M., Hillsdon, M., 2013. Remote and web 2.0 interventions for promoting physical activity. *Cochrane Database Syst. Rev.*
- Gourlan, M.J., Trouilloud, D.O., Sarrazin, P.G., 2011. Interventions promoting physical activity among obese populations: a meta-analysis considering global effect, long-term maintenance, physical

- activity indicators and dose characteristics. *Obes. Rev.* 12, e633-645.
- Hall, K.S., Sloane, R., Pieper, C.F., Peterson, M.J., Crowley, G.M., Cowper, P.A., McConnell, E.S., Bosworth, H.B., Ekelund, C.C., Morey, M.C., 2010. Long-term changes in physical activity following a one-year home-based physical activity counseling program in older adults with multiple morbidities. *J. Aging Res.* 2011.
- Harkin, B., Webb, T.L., Chang, B.P.I., Prestwich, A., Conner, M., Kellar, I., Benn, Y., Sheeran, P., 2016. Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychol. Bull.* 142, 198–229.
- Higgins, J.P., Green, S., 2011. *Cochrane Handbook for Systematic Reviews of Interventions* version 5.1.0 [updated March 2011] [WWW Document]. Cochrane Collab. URL <http://handbook.cochrane.org/>
- Hobbs, N., Godfrey, A., Lara, J., Errington, L., Meyer, T.D., Rochester, L., White, M., Mathers, J.C., Sniehotta, F.F., 2013. 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.* 11, 75.
- Johnson, B.T., Huedo-Medina, T.B., 2011. Depicting estimates using the intercept in meta-regression models: the moving constant technique. *Res. Synth. Methods* 2, 204–220.
- Johnston, M., 2016. What more can we learn from early learning theory? The contemporary relevance for behaviour change interventions. *Br. J. Health Psychol.* 21, 1–10.
- Kassavou, A., Turner, A., Hamborg, T., French, D.P., 2014. Predicting maintenance of attendance at walking groups: testing constructs from three leading maintenance theories. *Heal. Psychol.* 33, 752–756.
- Kohl, H.W., Craig, C.L., Lambert, E.V., Inoue, S., Alkandari, J.R., Leetongin, G., Kahlmeier, S., 2012. The pandemic of physical inactivity: global action for public health. *Lancet* 380, 294–305.
- Kwasnicka, D., Dombrowski, S.U., White, M., Sniehotta, F., 2016. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychol. Rev.* 1–39.
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.
- Lee, I.-M., Shiroma, E.J., Lobelo, F., Puska, P., Blair, S.N., Katzmarzyk, P.T., 2012. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 380, 219–229.
- Malik, S.H., Blake, H., Suggs, L.S., 2014. A systematic review of workplace health promotion interventions for increasing physical activity. *Br. J. Health Psychol.* 19, 149–180.
- Malina, R.M., 2001. Tracking of physical activity across the lifespan. *Pres. Counc. Phys. Fit. Sports Res. Dig.* 3, 10.
- Marcus, B.H., Dubbert, P.M., McKenzie, T.L., Dunn, A.L., Blair, S.N., 2000. Physical activity behavior change: issues in adoption and maintenance. *Heal. Psychol.* 19, 32–41.
- McAuley, E., Blissmer, B., 2000. Self-efficacy determinants and consequences of physical activity. *Exerc. Sport Sci. Rev.* 28, 85–88.
- Michie, S., Abraham, C., Whittington, C., McAteer, J., Gupta, S., 2009. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Heal. Psychol.* 28, 690–701.
- Michie, S., Ashford, S., Sniehotta, F.F., Dombrowski, S.U., Bishop, A., French, D.P., 2011. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. *Psychol. Health* 26, 1479–1498.

- Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M.P., Cane, J., Wood, C.E., 2013. 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.* 46, 81–95.
- Michie, S., West, R., Campbell, R., Brown, J., Gainforth, H., 2014. *ABC of Behaviour Change Theories Book - An Essential Resource for Researchers, Policy Makers and Practitioners*, Silverback Publishing.
- Newell, D.J., 1992. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int. J. Epidemiol.* 21, 837–841.
- NICE, 2013. Physical activity: brief advice for adults in primary care | Guidance and guidelines | NICE [WWW Document]. URL <https://www.nice.org.uk/guidance/ph44> (accessed 10.21.16).
- Nigg, C.R., Borrelli, B., Maddock, J., Dishman, R.K., 2008. A theory of physical activity maintenance. *Appl. Psychol., International Review of Applied Psychology* 57, 544–560.
- Office of Disease Prevention and Health Promotion, 2017. Chapter 4: Active adults [WWW Document]. *Phys. Act. Guidel.* URL <https://health.gov/paguidelines/guidelines/chapter4.aspx> (accessed 1.9.16).
- Olander, E.K., Fletcher, H., Williams, S., Atkinson, L., Turner, A., French, D.P., 2013. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. *Int. J. Behav. Nutr. Phys. Act.* 10, 29.
- Orrrow, G., Kinmonth, A.-L., Sanderson, S., Sutton, S., 2012. Effectiveness of physical activity promotion based in primary care: systematic review and meta-analysis of randomised controlled trials. *BMJ* 344, e1389.
- Ory, M.G., Smith, M.L., Mier, N., Wernicke, M.M., Lee Smith, M., 2010. The science of sustaining health behavior change: the health maintenance consortium. *Am. J. Health Behav.* 34, 647–659.
- Plotnikoff, R.C., Costigan, S.A., Williams, R.L., Hutchesson, M.J., Kennedy, S.G., Robards, S.L., Allen, J., Collins, C.E., Callister, R., Germov, J., 2015. Effectiveness of interventions targeting physical activity, nutrition and healthy weight for university and college students: a systematic review and meta-analysis. *Int. J. Behav. Nutr. Phys. Act.* 12, 45.
- Prins, R.G., Panter, J., Heinen, E., Griffin, S.J., Ogilvie, D.B., 2016. Causal pathways linking environmental change with health behaviour change: natural experimental study of new transport infrastructure and cycling to work. *Prev. Med. (Baltim).* 87, 175–182.
- Prochaska, J.O., DiClemente, C.C., 1982. Transtheoretical therapy: toward a more integrative model of change. *Psychother. Theory, Res. Pract.* 19, 276–288.
- Rao, J.N.K., Scott, A.J., 1992. A simple method for the analysis of clustered binary data. *Biometrics* 48, 577.
- Richards, J., Hillsdon, M., Thorogood, M., Foster, C., 2013. Face-to-face interventions for promoting physical activity. *Cochrane Database Syst. Rev.*
- Robertson, C., Archibald, D., Avenell, A., Douglas, F., Hoddinott, P., van Teijlingen, E., Boyers, D., Stewart, F., Boachie, C., Fioratou, E., Wilkins, D., Street, T., Carroll, P., Fowler, C., 2014. Systematic reviews of and integrated report on the quantitative, qualitative and economic evidence base for the management of obesity in men. *Health Technol. Assess. (Rockv).* 18, 1–424.
- Rothman, A.J., 2000. Toward a theory-based analysis of behavioral maintenance. *Heal. Psychol.* 19, 64–69.
- Sallis, J.F., Bull, F., Guthold, R., Heath, G.W., Inoue, S., Kelly, P., Oyeyemi, A.L., Perez, L.G., Richards,

- J., Hallal, P.C., Lancet Physical Activity Series 2 Executive Committee, 2016. Progress in physical activity over the Olympic quadrennium. *Lancet* 388.
- Samdal, G.B., Eide, G.E., Barth, T., Williams, G., Meland, E., 2017. 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.* 14, 42.
- Sanchez, A., Bully, P., Martinez, C., Grandes, G., 2015. Effectiveness of physical activity promotion interventions in primary care: a review of reviews. *Prev. Med. (Baltim)*. 76, S56–S67.
- Schröer, S., Haupt, J., Pieper, C., 2014. Evidence-based lifestyle interventions in the workplace--an overview. *Occup. Med. (Lond)*. 64, 8–12.
- Schulz, K.F., Grimes, D.A., 2002. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet (London, England)* 359, 781–785.
- Schwarzer, R., 1992. Self-efficacy in the Adoption and Maintenance of Health Behaviors: Theoretical Approaches and a New Model., *Self-efficacy: Thought control of action*. Hemisphere Publishing Corp, Washington DC.
- Seymour, R.B., Hughes, S.L., Ory, M.G., Elliot, D.L., Kirby, K.C., Migneault, J., Patrick, H., Roll, J.M., Williams, G., 2010. A lexicon for measuring maintenance of behavior change. *Am. J. Health Behav.* 34, 660–668.
- Sriramatr, S., Berry, T.R., Spence, J.C., 2014. An internet-based intervention for promoting and maintaining physical activity: a randomized controlled trial. *Am. J. Health Behav.* 38, 430–439.
- Stanley, T.D., Doucouliagos, H., 2014. Meta-regression approximations to reduce publication selection bias. *Res. Synth. Methods* 5, 60–78.
- StataCorp, 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
- The Cochrane Collaboration, 2014. *Review Manager (RevMan) [Computer program]*. Version 5.3.
- Thompson, S.G., Higgins, J.P.T., 2002. How should meta-regression analyses be undertaken and interpreted? *Stat. Med.* 21, 1559–1573.
- Ukoumunne, O.C., Gulliford, M.C., Chinn, S., Sterne, J.A., Burney, P.G., 1999. Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review. *Health Technol. Assess.* 3, iii-92.
- Williams, S.L., French, D.P., 2011. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? *Health Educ. Res.* 26, 308–322.
- Williamson, W., Kluzek, S., Roberts, N., Richards, J., Arden, N., Leeson, P., Newton, J., Foster, C., 2015. Behavioural physical activity interventions in participants with lower-limb osteoarthritis: a systematic review with meta-analysis. *BMJ Open* 5, e007642.
- World Health Organization, 2010. *Global recommendations on physical activity for health*. World Health Organization.
- Yates, T., Griffin, S., Bodicoat, D.H., Brierly, G., Dallosso, H., Davies, M.J., Eborall, H., Edwardson, C., Gillett, M., Gray, L., Hardeman, W., Hill, S., Morton, K., Sutton, S., Troughton, J., Khunti, K., 2015. PRomotion Of Physical activity through structured Education with differing Levels of ongoing Support for people at high risk of type 2 diabetes (PROPELS): study protocol for a randomized controlled trial. *Trials* 16, 289.

Table 1: Multivariable model with moderators evaluated simultaneously ($k=43$)^a

Moderator level	\hat{d}_+ (95% CI)	θ	P-value
Setting		-0.13	0.10
Primary care	0.10 (-0.05, 0.24)		
Community/Work/University ^b	0.23 (0.15, 0.31)		
Prompt self-monitoring of behavioural outcome		1.46	<0.01
BCT used ^b	1.69 (1.10, 2.28)		
BCT not used	0.23 (0.15, 0.31)		
Use of follow-up prompts		0.38	<0.01
BCT used ^b	0.61 (0.41, 0.82)		
BCT not used	0.23 (0.15, 0.31)		
Optimal model^b	1.94 (1.28, 2.59)		

BCT: Behaviour change technique; θ : unique variance explained by moderator.

^a1-squared (variance unexplained by model)=50.7%, Tau-squared=0.02.

^bIndicates the moderator level that conferred the largest PA increase and were used to generate the optimal model.

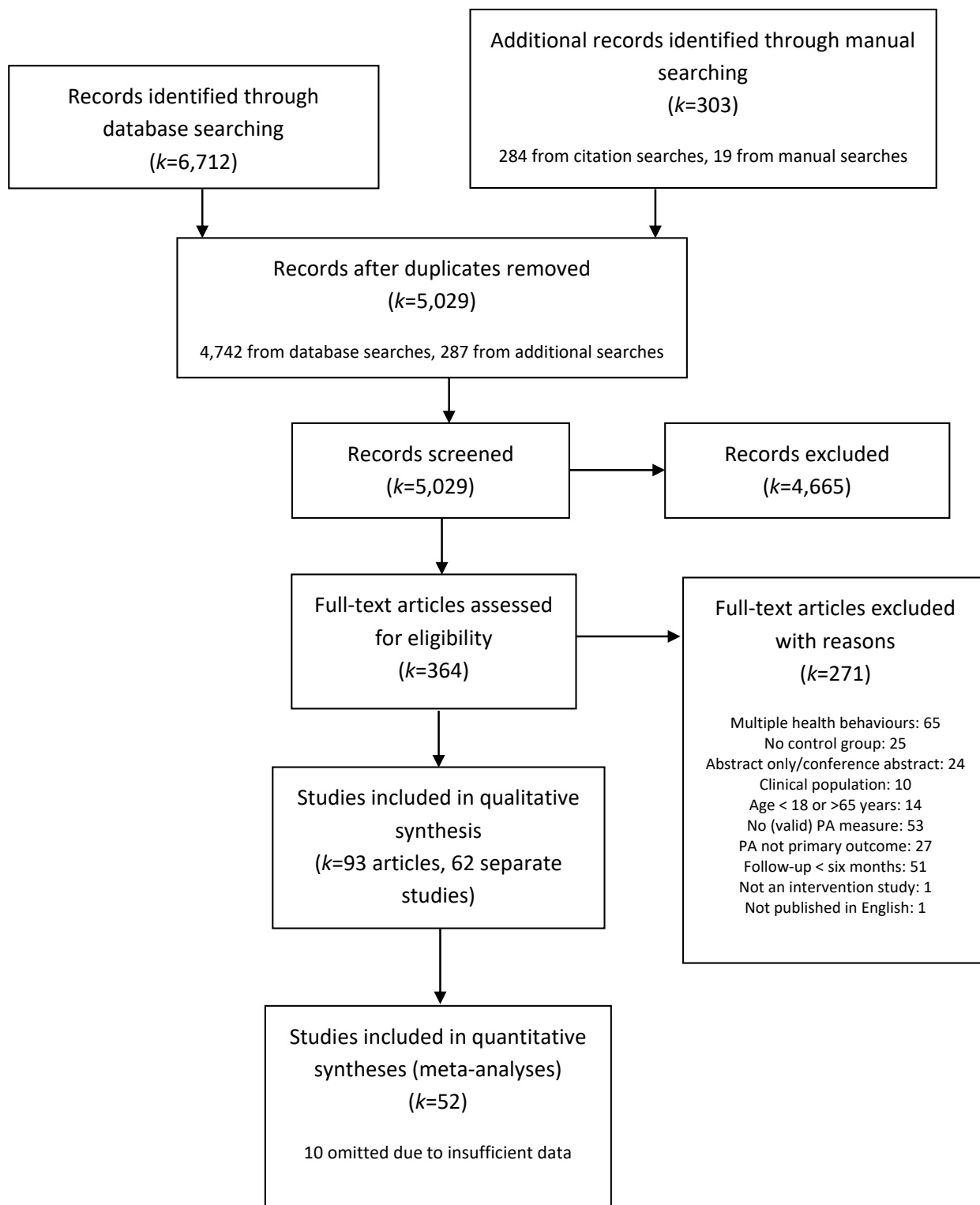


Figure 1: Study flow diagram