

REVIEW

Open Access



A systematic review of the use and reporting of evaluation frameworks within evaluations of physical activity interventions

Judith F. Fynn^{1*} , Wendy Hardeman², Karen Milton³, Joseph Murphy⁴ and Andy Jones¹

Abstract

Background: Evaluation of physical activity interventions is vital to inform, and justify, evidence-based policy and practice to support population-wide changes in physical activity. Several evaluation frameworks and guidance documents have been developed to facilitate the evaluation and reporting of evaluation studies in public health. However, there is a lack of evidence about whether frameworks are being used to guide evaluation. There continues to be claims of poor and inconsistent reporting in evaluation studies. The aim of this review was to assess the use of evaluation frameworks and the quality of reporting of how they were applied within evaluation studies of physical activity interventions.

Objectives:

1. To identify whether evaluation frameworks are reported to have been used within evaluation studies of physical activity interventions, and which frameworks have been used.
2. To appraise the quality of reporting with regards to how evaluation frameworks have been used.

Method: We developed a checklist of indicators to enable a critical appraisal of the use and reporting of different evaluation frameworks in evaluation studies. We conducted a systematic search and review of evaluation studies published between 2015 and the date of the search to appraise the use and reporting of evaluation frameworks. A narrative synthesis is provided.

Results: The review identified 292 evaluation studies of physical activity interventions, only 69 (23%) of these mentioned using an evaluation framework, and only 16 different frameworks were referred to. There was variation in the quality of reporting of framework use. 51 (74%) studies were identified as being explicitly based on the stated framework, however only 26 (38%) provided detailed descriptions consistently across all the checklist indicators. Details of adaptations and limitations in how frameworks were applied were less frequently reported. The review also highlighted variability in the reporting of intervention components. More consistent and precise reporting of framework and intervention components is needed.

(Continued on next page)

* Correspondence: jfynn@uea.ac.uk

¹UKCRC Centre for Diet and Activity Research (CEDAR) and Norwich Medical School, University of East Anglia, Norwich, UK

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Conclusion: Evaluation frameworks can facilitate a more systematic evaluation report and we argue their limited use suggests missed opportunities to apply frameworks to guide evaluation and reporting in evaluation studies. Variability in the quality of reporting of framework use limits the comparability and transferability of evidence. Where a framework has been used, the checklist of indicators can be employed to facilitate the reporting of an evaluation study and to review the quality of an evaluation report.

Keywords: Evaluation framework, Physical activity, Systematic review, Intervention

Introduction

Increasing physical activity levels among the population is a public health priority [1–3]. Yet the diversity of individual, environmental and societal influences on physical activity requires interventions that reflect that diversity [3]. This has led to various interventions targeting physical activity behaviour that are delivered to different populations and across many settings by a range of public, private and voluntary providers, many of which are multi-sectoral and multi-component. The complexity and heterogeneity in interventions poses challenges to understanding their effectiveness, and to generalising from one intervention to another [4, 5]. Given the high rates of inactivity [6, 7] and the importance of physical activity for health [8], it is vital that we learn from the interventions delivered about what works, for whom, and in what contexts [9].

Over the past 20 years, there has been a growing appreciation of the importance of evaluation to inform evidence-based interventions to support population-wide changes in physical activity and to justify policy and practice [9–11]. Evaluation can be defined as the “systematic examination and assessment of the features of an initiative and its effects, in order to produce information that can be used by those who have an interest in its improvement or effectiveness” [12], p3. Translation from one setting to another, and wider scale adoption of effective interventions, requires both rigorous evaluation and robust reporting of evaluations to build the evidence-base [11, 13].

Several frameworks and guidance documents have been developed to facilitate the evaluation and reporting of intervention studies in public health. In this review the term ‘evaluation framework’ is used to include any structured guidance which facilitates a systematic evaluation of the implementation or outcomes of an intervention. A recent scoping review that we conducted identified 68 evaluation frameworks that could be used to guide evaluation of physical activity interventions [14]. This included frameworks intended to support evaluation of physical activity interventions specifically (e.g. The Standard Evaluation Framework (SEF) for Physical Activity Interventions [15]), as well as frameworks intended to guide development and evaluation of

various public health interventions, such as RE-AIM [10], and the Medical Research Council (MRC) guidance on the development and evaluation of complex interventions [16]. We have included more general guidance, such as Logic Models [17], where these provide information or a structure to facilitate a systematic approach to identifying and reporting intervention objectives, activities and outcomes. Several checklists have also been developed to improve the completeness of reporting and quality of intervention descriptions; for example the STROBE Statement for Reporting Observational studies in Epidemiology [18, 19] and the Template for Intervention Description and Replication (TIDieR) [20]. Further, the Behaviour Change Wheel [21] and the Behaviour Change Technique (BCT) Taxonomy V1 [22] provide a framework to facilitate intervention development, that can also be applied to help standardise how the content of behaviour change interventions are specified. Despite the publication of these frameworks and guidance, there is a lack of evidence about whether frameworks are being used to guide evaluation.

There has been continued calls for better evaluation and reporting within public health [23, 24]. In particular, the need for more detailed descriptions of intervention components and contextual factors to help evaluate how, why and in what contexts interventions may be effective, and to allow implementation of good practice [21, 25]. Many of the frameworks and guidance have sought to address this and provide guidance on process evaluation and contextual factors. However, questions remain regarding if and how these frameworks are used within evaluation studies.

Two previous reviews have focused specifically on the use of RE-AIM [26] and the SEF for physical activity interventions [23]. These reviews concluded that the reporting of framework components was inconsistent, and that details related to participants, recruitment and broader effects were particularly poorly reported, despite these being components of the frameworks used. Both reviews also highlighted a need for greater clarity in the reporting of how frameworks have been used. Heterogeneity in the format and guidance provided by frameworks may lead to heterogeneity in the way they are applied. This creates difficulties for those interested in

further development of evaluation guidance, and those interested in understanding and comparing the effectiveness of interventions including reviewers of evaluation studies and practitioners or researchers wishing to implement or further develop interventions. This limits the contribution evaluation studies make to the evidence base. Given the extensive number of evaluation frameworks, a better understanding of current practices in the use and reporting of them is needed so that future recommendations related to the use of frameworks and evaluation can be developed appropriately.

The aim of this review was therefore to assess the use of evaluation frameworks and the quality of reporting of how they were used within evaluations of physical activity interventions. The primary objective was to explore whether evaluation frameworks are reported to have been used within evaluation studies of physical activity interventions, and which frameworks have been used. The second objective was to appraise the quality of reporting with regards to how evaluation framework use has been reported. Previous reviews [23, 26] have assessed use of a single evaluation framework against the criteria specified in that framework. To our knowledge,

no previous review has developed a set of generic indicators to facilitate the appraisal of the use of multiple evaluation frameworks in reported studies. We therefore developed and applied a set of indicators that would enable a critical appraisal of the use and reporting of different evaluation frameworks in evaluation studies.

Methods

Protocol & registration

Search methods and inclusion criteria were specified in advance and registered on PROSPERO (CRD42018089472). We applied the PRISMA statement for reporting items for systematic reviews [27].

Search strategy

We searched Scopus, CINAHL, and EMBASE for published evaluation studies of physical activity interventions. We used free search terms and MeSH terms relating to evaluation, e.g. program* evaluation, programme effectiveness, process evaluation and outcome evaluation. We also included names of specific evaluation frameworks that we had identified in our scoping review of evaluation frameworks [14], to

Table 1 Search strategy applied in CINAHL data base

	Search applied in CINAHL
1	TITLE-ABS-SUBJECT ("program* evaluation") Published Date: 20150101–20,191,231
2	TITLE-ABS-SUBJECT ("service evaluation") Published Date: 20150101–20,191,231
3	TITLE-ABS-SUBJECT ("process evaluation") Published Date: 20150101–20,191,231
4	TITLE-ABS-SUBJECT ("implementation evaluation") Published Date: 20150101–20,191,231
5	TITLE-ABS-SUBJECT ("program* effectiveness") Published Date: 20150101–20,191,231
6	TITLE-ABS-SUBJECT ("outcome evaluation") Published Date: 20150101–20,191,231
7	TITLE-ABS-SUBJECT ("re-aim") Published Date: 20150101–20,191,231
8	TITLE-ABS-SUBJECT ("standard evaluation framework") Published Date: 20150101–20,191,231
9	TITLE-ABS-SUBJECT ("intervention mapping") Published Date: 20150101–20,191,231
10	TITLE-ABS-SUBJECT ("program impact pathway") Published Date: 20150101–20,191,231
11	TITLE-ABS-SUBJECT ("process evaluation of complex interventions") Published Date: 20150101–20,191,231
12	TITLE-ABS-SUBJECT ("developing and evaluating complex interventions") Published Date: 20150101–20,191,231
13	TITLE-ABS-SUBJECT ("framework for program evaluation in public health") Published Date: 20150101–20,191,231
14	TITLE-ABS-SUBJECT ("logic model") Published Date: 20150101–20,191,231
15	1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14
16	TITLE-ABS-SUBJECT ("physical activity") Published Date: 20150101–20,191,231
17	TITLE (exercise) Published Date: 20150101–20,191,231
18	TITLE (MH "exercise") Published Date: 20150101–20,191,231
19	TITLE-ABS-SUBJECT (sedentary) Published Date: 20150101–20,191,231
20	TITLE-ABS-SUBJECT (sport*) Published Date: 20150101–20,191,231
21	TITLE-ABS-SUBJECT (inactiv*) Published Date: 20150101–20,191,231
22	TITLE-ABS-SUBJECT (fitness) Published Date: 20150101–20,191,231
23	16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22
24	15 AND 23

minimize the risk of missing frameworks that do not include the term evaluation in their title (e.g. RE-AIM). These terms were then combined with terms relating to physical activity behaviours (e.g. physical activity, sport, exercise, sedentary). Table 1 provides the full electronic search strategy for Scopus. The context of this review was to understand current practice and use of frameworks in evaluation studies of physical activity programmes. Therefore, the search was limited to studies published between 2015 and the date of the search (25th March 2019). Only studies published in the English language were included.

All studies identified from the searches were downloaded into the Endnote reference manager and duplicates were removed. Screening of all studies was completed by the lead author. At each stage of the screening process (title, abstract and full paper) a sample of 20 % of studies were checked and validated independently by a second author (JM). Disagreements were resolved through discussion.

Study selection

Inclusion and exclusion criteria were defined a priori and applied to all papers (see Table 2 for full details). Our interest was in evaluation studies, therefore other articles including conceptual papers, reviews, and research protocols were excluded. To assess the use, and any limitations in the use, of evaluation frameworks across the full range of physical activity interventions we screened the papers to identify studies where increasing physical activity was the stated primary goal, irrespective of whether they reported the use of specified frameworks. We included evaluation studies of any physical

activity intervention delivered in any individual, group or population setting (e.g. health care, schools, and geographical areas). We included studies of interventions delivered to the general population as well as to participants diagnosed with a disease (e.g. heart disease, diabetes) or as having one or more disease risk factors (e.g. inactive, obese). We then screened these to identify those studies that had referred to an evaluation framework, and to exclude those that had not mentioned one. We screened the reference lists of the included studies to identify any companion papers, for example, where process and outcome evaluations were reported separately.

Data extraction

To address the first objective, we extracted the names of any evaluation frameworks that had been reported as being used in any of the studies. For reporting purposes, we also noted the number of physical activity evaluation studies in which no framework was mentioned. To address the second objective, we extracted data from studies that reported the use of one or more evaluation frameworks. Criteria for data extraction were identified and agreed by all authors a priori. Data extraction was completed using a data extraction table.

To assess the context and circumstances in which evaluation frameworks had been used, we extracted data related to study characteristics. So that this review met PRISMA recommendations for the reporting of systematic reviews [27] we used PRISMA guidelines to inform the data we extracted from the studies. In addition we used STROBE for the reporting of observational studies and natural experiments [19], and the TIDieR checklist

Table 2 Inclusion and exclusion criteria

Included	Excluded
Published evaluation studies including real-world or service evaluations, randomised control trials, observational and natural experiments, feasibility and pilot studies, outcome and process evaluations, quasi-experimental, pre-post designs, effectiveness and impact studies. All types of evaluations using quantitative and/or qualitative methods will be included, whether they have used specified frameworks or not.	Commentaries or discussion papers, conceptual papers, published extracts, books, editorials, systematic reviews, clinical case-reports, research protocols and reported programme designs.
Reported evaluation studies of programmes that have increasing physical activity as the primary stated goal of the programme, including reduced sitting time or sedentary behaviour.	Reported evaluation studies of programmes that have other health behaviours as the primary stated goal of the programme, e.g. smoking, alcohol, substance abuse, eating disorder behaviours. Reported evaluation studies that state other behavioural outcomes or clinical measures as the primary goal of the programme, e.g. programmes aimed at weight loss, maintaining a healthy weight, prevention or management of diabetes, prevention of stroke or heart attack, improvement of aerobic or cognitive function, reduction of fall; improvement of physical performance/function through physical activity or exercise.
Evaluations of programmes that align with approaches to behaviour change, i.e. programmes that correspond to any of the nine intervention functions on the Behaviour Change Wheel (education, persuasion, incentivisation, coercion, training, enablement, modelling, environmental restructuring and restrictions) [21].	Evaluations of programmes that do not correspond to any of the nine intervention functions on the Behaviour Change Wheel (education, persuasion, incentivisation, coercion, training, enablement, modelling, environmental restructuring and restrictions).
Studies that referred to one or more evaluation frameworks.	Studies that did not refer to any evaluation framework.

[20] to guide our data extraction. We extracted data related to study population, intervention setting and components, study design, and process and outcome measures. To help us to characterise the intervention types we extracted data related to the nine intervention functions of the Behaviour Change Wheel, and the activities delivered, where these were explicitly reported. Intervention functions are broad categories to define the general means by which an intervention might change behaviour (e.g. Education, Enablement, and Incentivisation) [21, 28]. Their use in intervention development and reporting is intended to facilitate clearer descriptions of intervention components [21]. This is essential for evaluation and implementation [25]. We applied the nine intervention functions to guide a systematic approach to identify and report study characteristics.

To assess the quality of reporting of the use of the frameworks, we developed a set of data extraction criteria related to how the studies had described a framework and its application. To ensure that we identified a set of indicators that could be applied across any evaluation framework, rather than a specific framework, we used a similar approach to that described by Michie and Prestwich in their coding scheme for assessing the use and reporting of theory in intervention studies [29]. We developed a set of indicators that would allow a systematic examination of how the use of a framework had been reported within each study. Each indicator required a yes/no/not sure response and supporting evidence. We adapted their categories and indicators which aligned closely to our own objectives. For example, Category 1 "Reference to underpinning theory" aligned to our objective to identify any "Reference to an evaluation framework". Within this category we included four indicators that together assessed the extent to which the framework had been referred to and described to enable us to appraise whether or not the evaluation study was explicitly based on or informed by one or more frameworks. For other items, our indicators were more loosely based on those of Michie and Prestwich. Category 2 and 3 included three indicators to assess the extent to which the methods, data collection and outcomes reported were linked to the specified framework's components. Category 4 included two indicators to assess the extent to which additional information on how the framework had been used is reported. This last category is important, as there may be good justification for reporting on some rather than all of the components in a framework, or adapting how a framework is applied within a specific evaluation study, but without that information it is difficult to appraise its use and reporting. Any one indicator taken in isolation might seem deficient, so the indicators are best considered together within each category and across the full checklist to provide an overall assessment

of how use of a framework has been reported. The criteria were discussed and agreed by all authors. The checklist of categories and indicators is shown in Table 3.

Data extraction was completed by JF and validated by JM. For the data related to study characteristics, a sample of 20% of studies were checked and validated and any disagreements were resolved through discussion. For our checklist of indicators used to appraise the quality of reporting of framework use, we first tested the indicators by independently extracting data for a small sample of papers and discussed any differences to refine the process and reach a consensus in how to apply the indicators to extract data. We then independently validated a sample of 20% of studies and calculated the level of agreement as a percentage in order to validate the data extraction process. Any further disagreements were resolved through discussion. We used narrative synthesis to summarise the use and reporting of frameworks within the included studies.

Results

The search identified 1524 studies once duplicates had been removed. The PRISMA diagram for the screening is shown in Fig. 1. We identified a total of 292 evaluation studies of physical activity interventions. Only 69 (23%) of these mentioned using an evaluation framework. From the reference list of these 69 studies we identified an additional eight companion studies, however none mentioned using an evaluation framework so were not included. Three interventions were reported in more than one of the included studies; therefore the 69 included studies represent 64 different physical activity interventions.

Table 4 summarises the evaluation frameworks which were reported as being used and the number of studies using each framework. A total of 16 different evaluation frameworks were identified. These include frameworks that provide guidance on evaluation specifically, such as the Process Evaluation Plan [30], and frameworks that provide guidance on intervention planning and development but that facilitate evaluation and reporting, such as Precede-Proceed [34], Intervention Mapping [35] and Logic Models [17]. The frameworks most frequently reported were RE-AIM [10], Saunders and Joshi's process evaluation plan [30] and Steckler and Linnans' process evaluation guidance for public health [31]. RE-AIM [10] and the MRC guidance for development and evaluation of complex interventions [16] were the frameworks most frequently reported as being used as a single framework to inform the evaluation study. Realist evaluation [33] was only reported in four studies but was in all cases used as a standalone framework rather than in combination with other frameworks. Fourteen studies reported

Table 3 Categories and indicators for assessing the quality of reporting of the use of evaluation frameworks

Category	Data Extraction Indicators (options for responses)
1. Reference to Framework.	1. Is the framework mentioned even if the study is not explicitly based on it? <i>Yes/No/Not sure</i> 2. Does the study refer to 1 or more frameworks? <i>State number</i> 3. Is the framework mentioned in the introduction? <i>Yes/No/Not sure (Plus evidence)</i> 4. Is a description of the framework components provided? <i>Yes/No/Not sure (Plus evidence)</i>
2. How the framework has been used to develop the evaluation methods and data collection. Are relevant components applied?	5. Is the evaluation stated as explicitly based on the framework components? <i>Yes/No/Not sure (Plus evidence from the method of how the framework components have been applied to inform evaluation methods & data sources)</i>
3. How the framework has been applied to the reporting of outcomes.	6. Are the outcome measures discussed in the result/discussion sections linked to the relevant framework components? <i>Yes/No/Not sure (Plus evidence)</i> 7. How many of the framework components are linked to data sources/ measures? <i>All the main framework components / At least one, but not all /None of the components are linked to data (Plus evidence)</i>
4. Reporting use of framework fully.	8. Are any details of adaptations in how the framework has been applied provided? <i>Yes/No/Not sure (Plus evidence)</i> 9. Are any details of limitations and strengths in how the framework has been applied or suggestions for how it could be optimised provided? <i>Yes/No/Not sure (Plus evidence)</i>

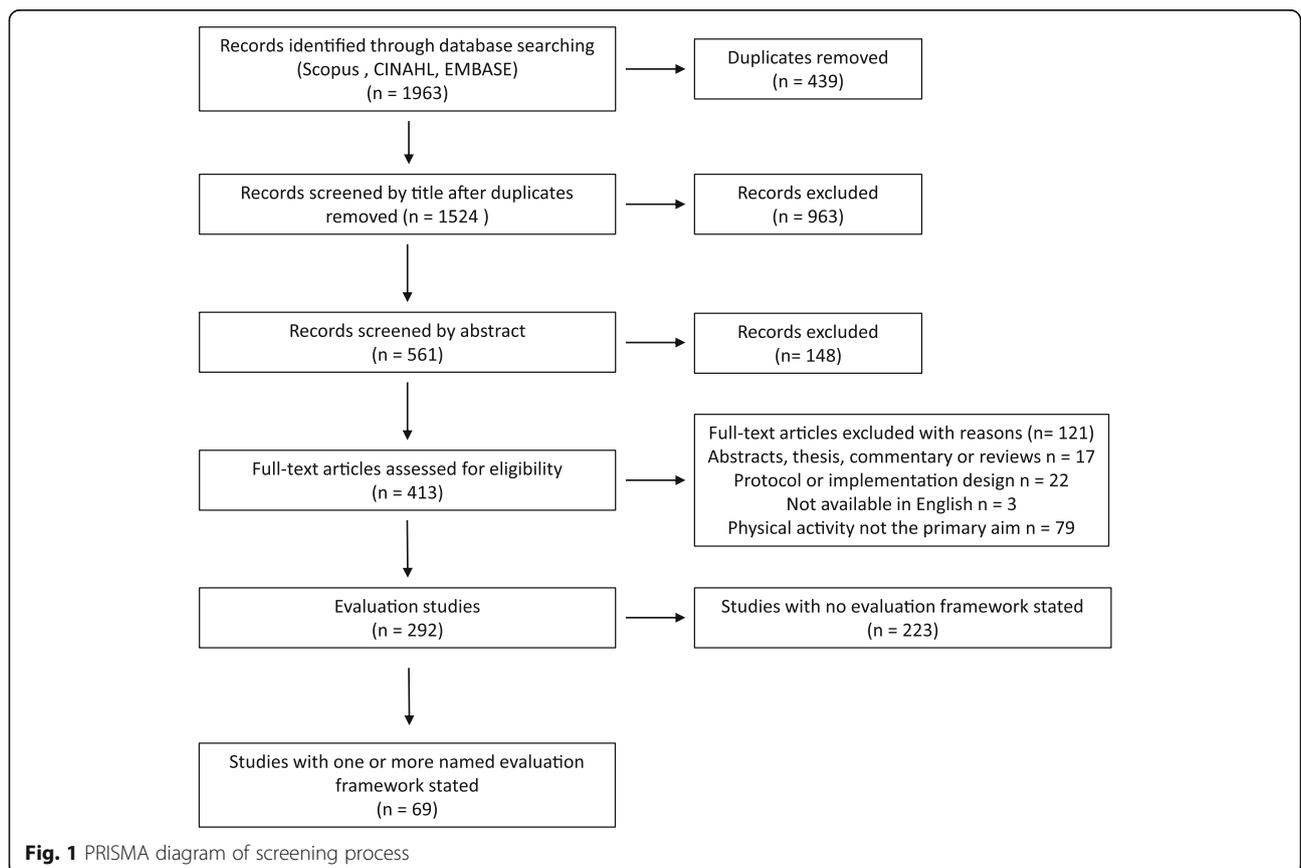


Fig. 1 PRISMA diagram of screening process

Table 4 Evaluation frameworks reported within the 69 studies

Named Framework	Number of studies reporting
RE-AIM [10]	27
Developing a process evaluation plan [30]	12
Process evaluation for public health [31]	10
MRC Guidance on evaluation of complex interventions [16]	8
MRC Guidance on process evaluation [32]	8
Logic Model [17]	7
Realist Evaluation [33]	4
Precede-Proceed [34]	3
Intervention Mapping [35]	2
Outcome Model [36]	2
CDC Framework [37]	1
Evaluation: a Systematic Approach [38]	1
Model of Implementation [39]	1
WHO Process Evaluation Workbook [40]	1
Swiss Model for Outcome Classification [41]	1
Concepts in process evaluation [42]	1

Note: 14 papers referred to more than one of these frameworks informing the evaluation

applying more than one framework (Table 6). The frameworks most frequently reported as being used in combination with others were Saunders and Joshi's [30] and Steckler and Linnan's [31] process evaluation frameworks. Both these frameworks provide a similar step-wise approach to process evaluation. The MRC guidance on process evaluation [32] and logic models [17] were also reported in several studies, both as a standalone framework and in combination with other frameworks.

Study characteristics

Study characteristics are shown in the supplementary material (Additional File 1). The frameworks have been used in a wide range of contexts and circumstances. Most of the criteria used to describe the interventions were clearly specified, and there was good agreement in the sample validated independently. The study population was reported in all studies; 37 studies (54%) reported interventions targeting children or young adults, 24 (35%) targeted adults, and five (7%) targeted older people. The remaining three (4%) studies did not specify an age group but implied the intervention was targeted at multiple population groups or the general public. Relevant details of demographic and/or health status of target populations were also described fully in studies where this was relevant: interventions targeting populations with or at increased risk of diabetes, the metabolic syndrome or heart disease; low socio-economic groups; and women or men only. Details of the included

population were reported variously as sample size, participants recruited, or the number of intervention sites. Intervention setting was described in all studies; 28 (40%) were implemented in schools (including pre-schools), 13 (20%) in health care settings, four (6%) in the workplace, and 24 (35%) in other community settings (e.g. youth groups, churches). All studies provided some description of the intervention components (i.e. activities delivered), although the level of detail was variable. For example, most studies described specific activities delivered (e.g. walking, dance, counselling, staff training, online tools), whilst fewer studies provided details of who delivered the intervention, the mode of delivery, the dose, or modifications to the delivery of the intervention. Most studies were multi-component and described several activities delivered together. Training ($n = 50$, 72%), education ($n = 47$, 68%) and enablement ($n = 42$, 61%) were the most frequently reported intervention functions stated in the studies. Studies less frequently reported modelling ($n = 12$, 17%), incentivisation ($n = 9$, 13%), environmental restructuring ($n = 9$, 13%) and persuasion ($n = 4$, 6%).

Additional File 1 shows the data we extracted related to the study objectives, study design and outcomes reported. Study designs included quantitative, qualitative and mixed-methods studies, controlled trials, quasi-experimental, case studies and hybrid designs. Thirty-five (51%) studies were described as a process evaluation and 15 (22%) as an outcome evaluation. In addition to physical activity outcomes, a range of secondary outcomes were reported: 52 (75%) reported on various implementation measures e.g. reach, dose, fidelity and maintenance; 14 (20%) reported outcomes related to anthropometric measures; and 15 (22%) reported details of participant demographics. Only nine (13%) studies reported outcome measures related to quality of life and only five (7%) reported on economic or cost analysis.

Appraisal of the quality of reporting on the use of evaluation frameworks

Table 5 shows the data extracted on the use and reporting of an evaluation framework for studies referring to a single framework, and Table 6 shows the data for studies referring to more than one framework. The level of agreement for the validation of data extracted for these items was 80%. Six studies mentioned a framework but did not state that the evaluation was informed by it. These included one study that provided a logic model but made no reference to this other than in the figure caption [64], and four studies that mentioned the MRC guidance on evaluating complex interventions and one that mentioned the MRC guidance on process evaluation of complex interventions but did not explicitly state that the study was informed by these guidance documents

[69–71, 75, 112] (four of these were companion studies relating to the same intervention). In three (4%) further studies the description lacked sufficient clarity to determine whether the study was intended to be based on the reported framework or not; for example these referred to the formulation of a logic model but did not describe the evaluation and outcomes as being based on the logic model [48, 49, 65]. The remaining 60 (87%) studies all stated that the evaluation was informed by one or more specified framework. However, based on the extracted data on how studies had reported framework components, how these had been applied and how the results linked to the framework components, we identified only 51 (74%) of the studies as being explicitly based on the reported framework.

Forty-four studies (64%) referred to the framework(s) in the introduction, while thirty-six (52%) provided a description of the framework components. Fifty-three (77%) reported outcomes linked to relevant framework components, the remaining sixteen (23%) studies provided no evidence of how the outcomes reported were linked to the framework components. Only 26 (38%) studies provided detailed descriptions consistently across all of the indicators; this included 13 that used RE-AIM, three that used Realist evaluation, two that used the MRC guidance on process evaluation, and two that used Saunders and Joshi's process evaluation framework. Four studies [46, 50, 55, 113] that had applied frameworks in combination also consistently reported details of the frameworks and their use across all indicators. Twenty-nine studies (42%) described strengths or limitations, whilst only 17 (25%) described adaptations in how the framework had been used.

Discussion

The extent to which evaluation frameworks have been used and reported

This is the first systematic review that has attempted to comprehensively assess the use of evaluation frameworks within evaluations of physical activity interventions. We identified 292 evaluation studies of interventions in which physical activity was the primary goal, published between 2015 and the date of our search. Only 69 (23%) of these studies reported using an evaluation framework; within these 16 different frameworks were mentioned. Given that we previously identified 68 published evaluation frameworks that could be used to facilitate evaluation of physical activity interventions [14], our findings highlight that evaluation frameworks are under-used and/or under-reported. Their limited use suggests missed opportunities to apply frameworks to guide evaluation and reporting in intervention studies. For example, despite recommendations in several guidance documents to use logic models to support intervention

development and evaluation [15, 32, 114], logic models were only referred to in seven of the studies, and their application was poorly reported. None of the studies reported using any frameworks that have been developed specifically for use in physical activity programme evaluation such as the SEF for physical activity interventions [15]. This may be explained by its more limited guidance on process evaluation, given that 51% of the studies were a process evaluation and 75% reported implementation measures. The SEF was developed for use in a UK practice context and may therefore be less likely to be used in a research led intervention than a real-world programme evaluation. Its absence from any of the studies in this review suggests not just a limited use made of it but also highlights the gap between research and practice and the challenges of reporting real world evaluations in the scientific literature. The more frequent use and reporting of RE-AIM may be because it provides guidance on both outcome and process evaluation components. However, its use may also be influenced by its greater exposure within the literature.

Framework use, choice of framework and the quality of reporting is likely to be influenced by the intervention's context and circumstances in which they are used. Many of the studies ($n = 35$, 51%) were process evaluations and it therefore follows that the most frequently reported frameworks were process evaluation frameworks. However, we found that a range of frameworks were used across different intervention types, contexts and study designs. This suggests that many evaluation frameworks are widely applicable and the decision to use and report a framework is more critical than the choice of which framework to use.

The quality of reporting with regards to how frameworks were used

Our checklist of indicators (Table 3) enabled us to appraise the quality of use and reporting of evaluation frameworks. There was considerable variation in the quality of reporting of framework use (Tables 5 and 6). Whilst some studies did report the framework and how it had been used consistently across all indicators in our checklist, others were less consistent in the quality of reporting and some only mentioned a framework without specifying the details of its use. In some studies, the evaluation was reported as being informed by a framework even where there was little evidence of the evaluation being based on it.

Studies tended to be poorer at describing framework components and adaptations or limitations in how these had been used, whilst links between outcome measures and framework components were more clearly described. For example, those which applied just one or more framework's components, rather than all the

Table 5 Appraisal of use and reporting of an evaluation framework in studies using a single evaluation framework

Framework(s)	Intervention name	First author & publication	Framework stated in introduction	Framework components described	Study stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
CDC Framework [37] Developing a Process-Evaluation Plan [30]	WAVE	Meng [57]	No	No	Yes	Yes	✓	✓	Not sure	No	No
	APAN	Blackford [58]	No	Yes	Yes	Yes		✓	All	Yes	Yes
Exercise Counselling NECaSP PACES ToyBox-study	Exercise Counselling	McCarthy [59]	Yes	Yes	Yes	Yes		✓	At least one	No	No
	NECaSP	Curry [60]	No	No	Yes	Not sure		✓	Not sure	Yes	Yes
	PACES	Webster [61]	Yes	Yes	Yes	No	✓		Not sure	No	No
	ToyBox-study	De Craemer [62]	Yes	Yes	Yes	Yes		✓	At least one	Yes	Yes
Evaluation: a Systematic Approach [38] Logic Model	FLEX	Wright [63]	No	Yes	Yes	Yes		✓	Not sure	Yes	No
	Girls Active	Harrington [64]	No	No	No	No	✓		Not sure	No	No
MAGNET MRC Guidance for Development & Evaluation of Complex Interventions [16]	GOTR	Ullrich-French [65]	No	No	Not sure	Yes	✓		Not sure	Yes	No
	Healthy Start	Chow [66]	No	No	Yes	Not sure	✓		At least one	No	Yes
	School-Community Linked PA	Griffiths [67]	No	No	Yes	Yes		✓	At least one	No	Yes
	MAGNET	Burkart [68]	Yes	Yes	Yes	Yes		✓	Not sure	No	No
MRC Process Evaluation of Complex Interventions [32]	Action 3.30	Jago [69]	No	No	No	No	✓		Not sure	No	No
	BGDP	Jago [70]	No	No	No	Yes	✓		At least one	No	No
	BGDP	Sebire [71]	Yes	No	No	No	✓		At least one	No	No
	GoActive	Corder [72]	Yes	No	Yes	No	✓		Not sure	No	No
Outcome Model [36] PRECEDE PROCEED [81]	Movement as Medicine	Avery [73]	Yes	Yes	Yes	Yes		✓	At least one	No	Yes
	STAND	Biddle [74]	Yes	No	No	No	✓		None	No	No
	BGDP	Sebire [75]	Yes	No	Yes	Not sure	✓		At least one	Not sure	Yes
	BGDP	Sebire [76]	No	No	No	No	✓		None	No	No
Healingo Fit	LPAW	Lefler [77]	Yes	No	Yes	No	✓		None	No	No
	PACE-UP	Furness [78]	Yes	Yes	Yes	Yes		✓	At least one	No	No
	We Act	Bonde [79]	Yes	Yes	Yes	Yes		✓	All	No	No
	Healingo Fit	Dadaczynski [80]	No	No	Yes	Yes		✓	At least one	No	No
SPACE	Tucker [82]	No	No	Yes	No	✓		Not sure	No	No	

Table 5 Appraisal of use and reporting of an evaluation framework in studies using a single evaluation framework (Continued)

Framework(s)	Intervention name	First author & publication	Framework stated in introduction	Framework components described	Study stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
Process Evaluation for Public Health [31]	Group fitness	Sofija [83]	No	Yes	Yes	Yes	✓	✓	All	No	No
	PAC	Matthews [84]	Yes	No	Yes	Yes	✓		At least one	No	No
RE-AIM [10]	5-As	Galaviz [85]	No	Yes	Yes	Yes	✓	✓	All	No	No
	ACTIVE	Christian [86]	No	No	Yes	Yes	✓	✓	All	No	No
	CHAM JAM	Reznik [87]	No	Yes	Yes	Yes	✓	✓	All	No	No
	COMMUNICATE	Kamada [4]	No	No	Yes	Yes	✓	✓	At least one	No	No
	Enhance®Fitness	Kohn [88]	Yes	No	Yes	Yes	✓	✓	At least one	No	No
	Enhance®Fitness	Petrescu-Prahova [89]	Yes	Yes	Yes	Yes	✓	✓	At least one	Yes	Yes
	FAN	Wilcox [90]	Yes	No	Yes	Not sure	✓	✓	At least one	No	No
	FitEx & ALED	Harden [91]	Yes	Yes	Yes	Yes	✓	✓	At least one	No	Yes
	Guided Walking	Baba [92]	Yes	Yes	Yes	Yes	✓	✓	All	No	Yes
	HKOS	Economos [93]	Yes	Yes	Yes	Yes	✓	✓	All	Yes	No
	Healthy Start-Départ Santé	Ward [94]	Yes	Yes	Yes	Yes	✓	✓	All	No	No
	Healthy Together	Jung [95]	Yes	Yes	Yes	Yes	✓	✓	All	No	No
	IMIL	Allar [96]	Yes	Yes	Yes	Yes	✓	✓	At least one	Yes	Yes
	ManUp	Caperchione [97]	Yes	Yes	Yes	Yes	✓	✓	All	No	Yes
	PAFES	Gonzalez-Viana [98]	Yes	Yes	Yes	Yes	✓	✓	All	No	Yes
Promotora Community Health Program	Schwingel [99]	Yes	Yes	Yes	Yes	✓	✓	All	Yes	Yes	
RCP & ACP	Paez [100]	Yes	Yes	Yes	Yes	✓	✓	All	Yes	Yes	
Sport England funded project	Koorts [101]	Yes	No	Yes	Yes	✓	✓	All	Yes	Yes	
Stair Climbing	Bellicha [102]	No	Yes	Yes	Yes	✓	✓	At least one	Yes	Yes	
STEPS & LET US Play	Beets [103]	No	No	Yes	Yes	✓	✓	At least one	Not sure	Yes	
STEPS & LET US Play	Beets [104]	No	No	Yes	Yes	✓	✓	At least one	Yes	Yes	
SAGE	Lee [105]	Yes	Yes	Yes	Yes	✓	✓	All	No	Yes	

Table 5 Appraisal of use and reporting of an evaluation framework in studies using a single evaluation framework (Continued)

Framework(s)	Intervention name	First author & publication	Framework stated in introduction	Framework components described	Study stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
Realist Evaluation [33]	TAME health	Lewis [106]	Yes	No	Yes	Yes		✓	All	No	Yes
	Walking Works	Adams [107]	Yes	Yes	Yes	Yes		✓	All	No	Yes
	CBHEPA	Herens [108]	Yes	Yes	Yes	Yes		✓	All	Not sure	Yes
	Local Authority Sport & PA	Daniels [109]	Yes	Yes	Yes	Yes		✓	All	Yes	Yes
	Local Environment Model	Willis [110]	Yes	Yes	Yes	Yes		✓	All	No	No
	Project SoL	Mikkelsen [111]	No	Yes	Yes	No	✓		Not sure	No	No

Table 6 Appraisal of use and reporting of the use of evaluation frameworks in studies using multiple frameworks

No. of Frameworks & references	Intervention Name	First author & publication	Framework stated in introduction	Framework components described	Stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
2 [17, 30]	WWPP	Fournier [43]	Yes	No	Yes	Yes		✓	At least one	No	No
2 [30, 34]	SPACE	Driediger [44]	Yes	No	Yes	Yes		✓	At least one	No	No
2 [30, 31]	IDEFICS	Verloigne [45]	Yes	No	Yes	Not sure	✓		Not Sure	No	Yes
2 [30, 31]	PA for grandparents	Young [46]	Yes	Yes	Yes	Yes		✓	At least one	No	No
2 [30, 31]	It's LiFe!	Verwey [47]	No	Yes	Yes	Yes		✓	At least one	No	Yes
2 [17, 41]	Classes in Motion	Grillich [48]	No	No	Not sure	Yes	✓		At least one	No	No
2 [16, 17]	ENGAGE-HD	Quinn [49]	Yes	No	Not sure	Not sure	✓		Not Sure	No	No
2 [10, 16]	Move for Well-being in School	Smedegaard [50]	Yes	Yes	Yes	Yes		✓	All	Yes	No
2 [31, 32]	WAVES	Griffin [51]	Yes	No	Yes	Yes		✓	At least one	Yes	Yes
2 [31, 40]	Walk Well	Matthews [52]	No	Yes	Yes	Yes		✓	At least one	No	No
3 [10, 30, 31]	'Beweegkuur'	Berendsen [53]	Yes	Yes	Yes	Yes		✓	All	Yes	Yes
3 [10, 32, 35]	Workplace intervention for Nurses	Torquati [54]	Yes	No	Yes	Yes		✓	All	Yes	Yes
3 [30, 31, 36]	SLIMMER	van Dongen [55]	Yes	Yes	Yes	Yes		✓	All	No	Not sure
5 [17, 31, 34, 35, 42]	SHAPES	Saunders [56]	Yes	No	Yes	Yes		✓	At least one	No	Yes

components, provided very little explanation or rationale for these adaptations. Publishing constraints can mean that reporting an evaluation study fully requires companion papers or supplementary files [16]. However, where this was done, we found that there was often inconsistency in reporting the use of frameworks across the different reported elements e.g. [61, 69–71, 75, 76, 79, 115]. More detailed and consistent reporting of the framework components and how these have been applied would help those trying to understand the intervention effectiveness fully.

It is inevitable that some frameworks lend themselves to better quality reporting. For example, studies using RE-AIM and Realist evaluation provided a more consistent report of their use across all indicators. RE-AIM is a structured framework; whilst Realist evaluation is a methodological approach, it too provides a guiding framework to facilitate a systematic evaluation and as such has been referred to as a framework within this paper. Both RE-AIM and Realist evaluation have a clear set of components that are relevant to both process and outcomes; they are therefore applicable to a range of evaluation objectives and can be used to identify appropriate data sources. Many of the studies using RE-AIM provided a full description of the components, an explanation of how these linked to data sources, and used the framework components to structure the reporting of findings. In this way the framework facilitated both a systematic evaluation and consistent reporting. RE-AIM was the most frequently used framework. There is a body of literature on how RE-AIM has been developed and used over time [116], and examples of its application. This may have helped to build a better understanding of how its components are defined and how they can be linked to data sources. Some of the less structured guidance documents, for example the MRC guidance on the development and evaluation of complex interventions [16], were used more loosely as a framework, particularly in studies that used more than one framework in combination. This does not necessarily equate to a poorer quality evaluation. However, we suggest those studies drawing on several frameworks and general guidance documents would benefit from a more detailed reporting of how these have been used to assist the reader in understanding which intervention components are reported on, and why. Whilst there is variability in the quality of reporting of how frameworks have been used, this review does highlight that evaluation frameworks can, when used appropriately, facilitate a systematic evaluation, and that studies that use a framework can facilitate systematic reporting of the evaluation process and outcomes.

Despite recommendations on the importance of fully reporting contextual factors and intervention components,

and guidance within the frameworks to facilitate this [16, 32], our review supports previous review findings [23, 26] that the reporting of intervention components is variable, with wider effects (e.g. quality of life and costs) and wider contextual factors (e.g. dose, intervention modifications) being particularly poorly reported. The Behaviour Change Wheel was developed to characterise intervention types and identify behaviour change techniques as “active ingredients” to improve the reporting and synthesis of evidence of what works in different populations and settings [21]. Yet we found ambiguity in the way in which studies reported intervention functions. It is noteworthy that intervention function was the item where we initially had most disagreement in the data extraction validation process and we would argue that clearer specification, or mapping of intervention functions against behaviour change techniques, would make them more useful in characterising interventions. Poor reporting of intervention components and types limits their comparability and transferability.

If evaluation studies are to contribute to an evidence base on which policymakers, practitioners and researchers can draw to inform the development and implementation of interventions, both the framework and intervention components need to be more clearly defined and documented. Clear, consistent and full reporting of interventions and their evaluation is essential to ensure that critical evidence gets shared and used to develop understanding of causal mechanisms, contextual factors and good practice [25]. This is vital to allow resources and efforts to address public health issues, such as increasing physical activity, to be focused on effective and efficient intervention components.

Where frameworks are used, their application to guide the full evaluation process from planning to reporting can improve the quality of reporting of their use. A focus on evaluation at the design and development stages of interventions and a clear understanding of the purpose of the evaluation can help to ensure outcome measures are linked to framework components. However, there is a need to improve understanding of how framework and intervention components are defined. Training and documentation can play a role, but more consistent and precise reporting within the scientific literature is needed. Our set of indicators (see Table 3) can be used to guide the reporting of framework use. Those reporting an evaluation study can apply the indicators as a checklist to provide a clear and consistent description of how framework components have been applied across all stages of the evaluation. Reviewers and journal editors can also play a role in using the checklists available to appraise evaluation reports.

Strengths and limitations

The strengths of this study are that we developed a comprehensive checklist of indicators to appraise the use and reporting of evaluation frameworks, based on a widely accepted coding scheme designed to assess the use and reporting of theory [29]. Our checklist and its use as a guide to data extraction was piloted and developed iteratively, and agreed by all authors. This enabled us to review the use and reporting of different frameworks.

Limitations of our study include the fact that some studies may use frameworks or framework components in a way that is implied but not explicitly stated, and we acknowledge that this may have led to underrepresentation of the full use made of evaluation frameworks. A more detailed assessment of evaluation studies against each specific framework's components may have provided greater insight into the limitations or fidelity of use and reporting of frameworks. This was not practical to do within a single review of multiple evaluation frameworks. Extracting details of outcome measures (findings) and intervention characteristics for all physical activity evaluation studies may have enabled a fuller appraisal of the quality of the studies and a comparison between those using and those not using an evaluation framework. This may have provided further insights on the impact of using evaluation frameworks on the quality of the evaluation study, however this was beyond the scope of this review.

Conclusion

Despite the use of evaluation frameworks being advocated to improve the rigour of evaluation studies, frameworks are underused and reported inconsistently in many studies. Applying an evaluation framework to inform both the evaluation and reporting of physical activity intervention studies facilitates a more systematic evaluation study. However, intervention and framework components need to be more precisely and consistently defined and documented to help improve the quality of reporting. Variability in the quality of reporting limits the comparability and transferability of evidence. This means that critical evidence that could be used to inform interventions to support the health of the population is not making it into the public domain. The indicators we developed enabled us to appraise the use and reporting of a range of different evaluation frameworks within evaluations of physical activity interventions. These indicators can be used by those reporting an evaluation to guide them in developing a systematic evaluation report, and by reviewers and journal editors to appraise evaluation studies that have reported the use of an evaluation framework.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12966-020-01013-7>.

Additional file 1. Study Characteristics. A summary of data extracted related to study characteristics to assess the context and circumstances in which evaluation frameworks have been used.

Abbreviations

ACTIVE: Active Children through Incentive Vouchers – Evaluation; ALED: Active Living Every Day; APAN: Albany Physical Activity and Nutrition; 5-As: Assess, Advise, Agree, Assist, Arrange for physical activity counselling; BCT: Behaviour Change Technique; BGDG: Bristol Girls Dance Project; CBHEPA: Dutch Community-Based Health-Enhancing Activity Programs; CDC: Center for Disease Control and Prevention; CHAM JAM: Children's Hospital at Montefiore Joining Academics and Movement; CoM: Communities on the Move; COMMUNICATE: COMMUNity-wide Campaign To promote Exercise; FAN: Faith, Activity, and Nutrition; FitEx: Fit Extension; FLEX: Fuelling Learning through Exercise; GOTR: Girls On The Run; Healingo Fit: Health Integrated Gaming Online; HKOS: Healthy Kids Out of School; HT: Healthy Together; IDEFICS: Identification & prevention of Dietary & lifestyle-induced health Effects In Children and infants; IMIL: I am Moving, I am Learning; LPAW: Lifestyle Physical Activity for Older Women; MAGN ET: Mothers And dauGhters daNcing togEther Trial; MRC: Medical Research Council; NECaSP: Newham's Every Child a Sports Person; PA: Physical Activity; PAC: Physical Activity Consultation; PACES: Partnerships for Active Children in Elementary Schools; PACE-UP: Primary care pedometer based walking trial; PAFES: Physical Activity, Sports, and Health Plan; RCP & ACP: Recreovia program & Academia da Cidade program; RE-AIM: Reach, Efficacy, Adoption, Implementation, Maintenance; SAGE: Sustainability via Active Garden Education; SEF: Standard Evaluation Framework; SHAPES : Study of Health and Activity in Preschool Environments; SLIMMER (SLIM): iMplementation Experience Region Noord- en Oost-Gelderland; SPACE: Supporting Physical Activity in the Childcare Environment; STAND : Sedentary Time ANd Diabetes; STEPs: Strategies To Enhance Practice in YMCA-operated After School Programmes; STROBE: Statement for Reporting Observational studies in Epidemiology; TAME health: Testing Activity Monitors' Effect on health; TIDieR: Template for Intervention Description and Replication; WAVE : WAVE~Ripples for Change; WWPP: Walking with Poles Program; WAVE S: West Midlands ActiVe lifestyle and healthy Eating in School children; WHO: World Health Organization

Acknowledgements

Not applicable.

Authors' contributions

JF, AJ, WH and KM conceptualised the research questions, developed and agreed the search strategy, data extraction and synthesis. JF conducted the search, screening, data extraction and analysis. JM checked and validated the screening and data extraction. All authors contributed to the manuscript, critically reviewed and approved the final manuscript.

Funding

The work was undertaken by the Centre for Diet and Activity Research (CEDAR), a UKCRC Public Health Research Centre of Excellence. Funding from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the National Institute for Health Research, and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. The funders had no role in any element of this research.

Availability of data and materials

The dataset(s) supporting the conclusions of this article are included within the article (and its additional files).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹UKCRC Centre for Diet and Activity Research (CEDAR) and Norwich Medical School, University of East Anglia, Norwich, UK. ²School of Health Sciences, University of East Anglia, Norwich, UK. ³Norwich Medical School, University of East Anglia, Norwich, UK. ⁴Physical Activity for Health Research Cluster, Physical Education and Sport Sciences Department, University of Limerick, Limerick, Ireland.

Received: 3 April 2020 Accepted: 12 August 2020

Published online: 24 August 2020

References

- World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. Geneva: World Health Organization; 2018.
- Public Health England. Everybody Active, Every Day: An evidence-based approach to physical activity. London: Public Health England; 2014.
- Rutter H, Cavill N, Bauman A, Bull F. Systems approaches to global and national physical activity plans. *Bull World Health Organ.* 2019;97(2):162.
- Kamada M, Kitayuguchi J, Abe T, Taguri M, Inoue S, Ishikawa Y, et al. Community-wide intervention and population level physical activity: A 5-year cluster randomized trial. *Int J Epidemiol.* 2018;47(2):642–53.
- Van Hoye A, Vandoorne C, Absil G, Lecomte F, Fallon C, Lombrail P, et al. Health enhancing physical activity in all policies? Comparison of national public actors between France and Belgium. *Health Policy.* 2019; 123(3):327–32.
- World Health Organization. Global health observatory data repository. Geneva: World Health Organization; 2018. Available from: <http://apps.who.int/gho/data/view>.
- Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1–6 million participants. *Lancet Child Adolescent Health.* 2020;4(1):23–35.
- Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet.* 2012;380(9838): 219–29.
- Pettman TL, Armstrong R, Doyle J, Burford B, Anderson LM, Hillgrove T, et al. Strengthening evaluation to capture the breadth of public health practice: ideal vs. real. *J Public Health.* 2012;34(1):151–5.
- Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health.* 1999;89(9):1322–7.
- Francis LJ, Smith BJ. Toward best practice in evaluation: A study of Australian health promotion agencies. *Health Promot Pract.* 2015;16(5):715–23.
- World Health Organization. Health promotion evaluation: recommendations to policy-makers. Geneva: World Health Organization; 1998.
- Milat AJ, King L, Bauman AE, Redman S. The concept of scalability: increasing the scale and potential adoption of health promotion interventions into policy and practice. *Health Promot Int.* 2013;28(3): 285–98.
- Fynn JF, Hardeman W, Milton K, Jones AP. A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. *BMC Public Health.* 2020;20(1):1–16.
- Cavill N, Roberts K, Rutter H. Standard evaluation framework for physical activity interventions. Oxford: National Obesity Observatory; 2012.
- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: The new Medical Research Council guidance. *Br Med J.* 2008;337.
- W K Kellogg Foundation. Logic Model Development Guide. Battle Creek Michigan: WK Kellogg Foundation; 2004.
- Knottnerus A, Tugwell P. STROBE—a checklist to strengthen the reporting of observational studies in epidemiology. *J Clin Epidemiol.* 2008;61(4):323.
- Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol.* 2008;61(4):344–9.
- Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj.* 2014;348:g1687.
- Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implement Sci.* 2011;6(1):42.
- Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, 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(1):81–95.
- Hanson S, Jones A. Missed opportunities in the evaluation of public health interventions: a case study of physical activity programmes. *BMC Public Health.* 2017;17(1):674.
- Smith BJ, Rissel C, Shilton T, Bauman A. Advancing evaluation practice in health promotion. *Health Promotion J Aust.* 2016;27(3):184–6.
- Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: the need for a scientific method. *Implement Sci.* 2009;4(1):40.
- Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: a systematic review of use over time. *Am J Public Health.* 2013;103(6):e38–46.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg.* 2010;8(5):336–41.
- Steinmo S, Fuller C, Stone SP, Michie S. Characterising an implementation intervention in terms of behaviour change techniques and theory: the ‘Sepsis Six’ clinical care bundle. *Implement Sci.* 2015;10(1):111.
- Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. *Health Psychol.* 2010;29(1):1.
- Saunders RP, Evans MH, Joshi P. Developing a process-evaluation plan for assessing health promotion program implementation: A how-to guide. *Health Promot Pract.* 2005;6(2):134–47.
- Steckler AB, Linnan L. Process evaluation for public health interventions and research. San Francisco, Calif: Jossey-Bass; 2002.
- Moore G, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process Evaluation of Complex Interventions: UK Medical Research Council (MRC) Guidance 2015.
- Pawson R, Tilley N. Realistic Evaluation. London: Sage Publications Ltd; 1997.
- Green LW, Kreuter MW. Health Promotion Planning : An Educational and Environmental Approach 2nd ed. Mountain View, Calif: Mayfield Publishing Co; 1991.
- Bartholomew LK, Parcel GS, Kok G, Gottlieb NH. Planning health promotion programs : an intervention mapping approach: John Wiley & Sons, Incorporated; 2006.
- Nutbeam D. Evaluating health promotion—Progress, Problems and solutions. *Health Promotion Int.* 1998;13(1):27–44.
- Center for Disease Control and Prevention. Introduction to Program Evaluation for Public Health Programs: A Self-Study Guide. 2012. <https://www.cdc.gov/eval/guide/step3/index.htm>.
- Rossi PH, Lipsey MW, Henry GT. Evaluation: A systematic approach: Sage publications; 2018.
- Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol.* 2008;41(3–4):327–50.
- World Health Organization (WHO). Process Evaluation Workbook 4. Geneva: WHO; 2001.
- Spencer B, Broesskamp-Stone U, Ruckstuhl B, Ackermann G, Spoerri A, Cloetta B. Modelling the results of health promotion activities in Switzerland: development of the Swiss model for outcome classification in health promotion and prevention. *Health Promot Int.* 2008;23(1):86–97.
- Baranowski T, Stables G. Process evaluations of the 5-a-day projects. *Health Educ Behav.* 2000;27(2):157–66.
- Fournier B, Manon P, Johanne F, Nathalie B, Lorthios-Guillement A, Marie-Ève M. Development and implementation of a community-based pole walking program for older adults. *Activities Adapt Aging.* 2019;43(1):1–22.
- Driediger M, Vanderloo LM, Burke SM, Irwin JD, Gaston A, Timmons BW, et al. The implementation and feasibility of the supporting physical activity in the childcare environment (SPACE) intervention: A process evaluation. *Health Educ Behav.* 2018;45(6):935–44.
- Verloigne M, Ahrens W, De Henauw S, Verbestel V, Marild S, Pigeot I, et al. Process evaluation of the IDEFICS school intervention: Putting the evaluation of the effect on children’s objectively measured physical activity

- and sedentary time in context. *Obesity Reviews*. 2015;16(Supplement 2):89–102.
46. Young T, Sharpe C. Process evaluation results from an intergenerational physical activity intervention for grandparents raising grandchildren. *J Phys Act Health*. 2016;13(5):525–33.
 47. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. Process evaluation of physical activity counselling with and without the use of mobile technology: A mixed methods study. *Int J Nurs Stud*. 2016;53:3–16.
 48. Grillich L, Kien C, Takuya Y, Weber M, Gartlehner G. Effectiveness evaluation of a health promotion programme in primary schools: a cluster randomised controlled trial. *BMC Public Health*. 2016;16:679.
 49. Quinn L, Trubey R, Gobat N, Dawes H, Edwards RT, Jones C, et al. Development and delivery of a physical activity intervention for people with Huntington disease: facilitating translation to clinical practice. *J Neurol Phys Ther*. 2016;40(2):71–80.
 50. Smedegaard S, Brondeel R, Christiansen LB, Skovgaard T. What happened in the 'Move for Well-being in School': A process evaluation of a cluster randomized physical activity intervention using the RE-AIM framework. *Int J Behav Nutri Phys Activity*. 2017;14(1):159.
 51. Griffin TL, Clarke JL, Lancashire ER, Pallan MJ, Adab P. Process evaluation results of a cluster randomised controlled childhood obesity prevention trial: the WAVES study. *BMC Public Health*. 2017;17:1–13.
 52. Matthews L, Mitchell F, Stalker K, McConnachie A, Murray H, Melling C, et al. Process evaluation of the walk well study: a cluster-randomised controlled trial of a community based walking programme for adults with intellectual disabilities. *BMC Public Health*. 2016;16(1):1–11.
 53. Berendsen BAJ, Kremers SPJ, Savelberg HHCM, Schaper NC, Hendriks MRC. The implementation and sustainability of a combined lifestyle intervention in primary care: mixed method process evaluation. *BMC Fam Pract*. 2015; 16(1):1–12.
 54. Torquati L, Kolbe-Alexander T, Pavey T, Leveritt M. Changing diet and physical activity in nurses: A pilot study and process evaluation highlighting challenges in workplace health promotion. *J Nutr Educ Behav*. 2018;50(10):1015–25.
 55. van Dongen EJ, Geerke D, Jansen SC, Ter Beek J, Huijg JM, Leerlooijer JN, Hiddink GJ, Feskens EJ, Haveman-Nies A. Process evaluation of a randomised controlled trial of a diabetes prevention intervention in Dutch primary health care: the SLIMMER study. *Public Health Nutr*. 2016;19(16): 3027–38.
 56. Saunders RP, Pfeiffer K, Brown WH, Howie EK, Dowda M, O'Neill JR, et al. Evaluating and refining the conceptual model used in the study of health and activity in preschool environments (SHAPES) intervention. *Health Educ Behav*. 2017;44(6):876–84.
 57. Meng Y, Wong SS, Manore MM, Patton-López M. WAVE-ripples for change obesity two-year intervention in high school soccer players: process evaluation, best practices, and youth engagement. *Nutrients*. 2018;10(6):711.
 58. Blackford K, Lee A, James AP, Waddell T, Hills AP, Anderson AS, et al. Process evaluation of the Albany physical activity and nutrition (APAN) program, a home-based intervention for metabolic syndrome and associated chronic disease risk in rural Australian adults. *Health Promotion J Australia*. 2017; 28(1):8–14.
 59. McCarthy MM, Dickson VV, Katz SD, Sciacca K, Chyun DA. Process evaluation of an exercise counseling intervention using motivational interviewing. *Appl Nurs Res*. 2015;28(2):156–62.
 60. Curry WB, Dagkas S, Wilson M. Newham's every child a sports person (NECaSP): A summative process evaluation of a school- and community-based intervention in East London, United Kingdom. *J Phys Act Health*. 2016;13(10):1124–31.
 61. Webster CA, Weaver RG, Egan CA, Brian A, Vazou S. Two-year process evaluation of a pilot program to increase elementary children's physical activity during school. *Evaluation and Program Planning*. 2018;67:200–6.
 62. De Craemer M, Verloigne M, De Bourdeaudhuij I, Andrououts O, Iotova V, Moreno L, et al. Effect and process evaluation of a kindergarten-based, family-involved cluster randomised controlled trial in six European countries on four- to six-year-old children's steps per day: the ToyBox-study. *Int J Behav Nutri Phys Activity*. 2017;14:1–16.
 63. Wright CM, Chomitz VR, Duquesnay PJ, Amin SA, Economos CD, Sackeck JM. The FLEX study school-based physical activity programs - Measurement and evaluation of implementation. *BMC Public Health*. 2019;19(1):73.
 64. Harrington DM, Davies MJ, Bodicoat DH, Charles JM, Chudasama YV, Gorely T, et al. Effectiveness of the 'Girls Active' school-based physical activity programme: A cluster randomised controlled trial. *Int J Behav Nutri Phys Activity*. 2018;15(1):40.
 65. Ullrich-French S, Cole AN, Montgomery AK. Evaluation development for a physical activity positive youth development program for girls. *Evaluation Program Planning*. 2016;55:67–76.
 66. Chow AF, Leis A, Humbert L, Muhajarine N, Engler-Stringer R. Healthy start – Départ santé: A pilot study of a multilevel intervention to increase physical activity, fundamental movement skills and healthy eating in rural childcare centres. *Can J Public Health*. 2016;107(3):e312–e8.
 67. Griffiths LA, Griffiths MA. Evaluation of a school-community linked physical activity intervention targeting 7- to 12-year-olds: a sociocultural perspective. *Am J Health Educ*. 2019;50(2):112–26.
 68. Burkart S, St. Laurent CW, Alhassan S. Process evaluation of a culturally-tailored physical activity intervention in African-American mother-daughter dyads. *Preventive Med Rep*. 2017;8:88–92.
 69. Jago R, Sebire SJ, Davies B, Wood L, Banfield K, Edwards MJ, et al. Increasing children's physical activity through a teaching-assistant led extracurricular intervention: Process evaluation of the action 3:30 randomised feasibility trial. *BMC Public Health*. 2015;15(1):156.
 70. Jago R, Edwards MJ, Sebire SJ, Tomkinson K, Bird EL, Banfield K, et al. Effect and cost of an after-school dance programme on the physical activity of 11–12 year old girls: The Bristol Girls Dance Project, a school-based cluster randomised controlled trial. *Int J Behav Nutri Phys Activity*. 2015;12(1):128.
 71. Sebire SJ, Jago R, Banfield K, Edwards MJ, Campbell R, Kipping R, et al. Results of a feasibility cluster randomised controlled trial of a peer-led school-based intervention to increase the physical activity of adolescent girls (PLAN-A). *Int J Behav Nutri Phys Activity*. 2018;15(1):50.
 72. Corder K, Brown HE, Schiff A, van Sluijs EM. Feasibility study and pilot cluster-randomised controlled trial of the GoActive intervention aiming to promote physical activity among adolescents: outcomes and lessons learnt. *BMJ Open*. 2016;6(11):e012335.
 73. Avery L, Charman SJ, Taylor L, Flynn D, Mosely K, Speight J, et al. Systematic development of a theory-informed multifaceted behavioural intervention to increase physical activity of adults with type 2 diabetes in routine primary care: Movement as Medicine for Type 2 Diabetes. *Implementation Sci*. 2016;11(1):99.
 74. Biddle SJ, Edwardson CL, Wilmot EG, Yates T, Gorely T, Bodicoat DH, et al. A randomised controlled trial to reduce sedentary time in young adults at risk of type 2 diabetes mellitus: project STAND (sedentary time AND diabetes). *PLoS One*. 2015;10(12):e0143398.
 75. Sebire SJ, Kesten JM, Edwards MJ, May T, Banfield K, Tomkinson K, et al. Using self-determination theory to promote adolescent girls' physical activity: exploring the theoretical fidelity of the Bristol girls dance project. *Psychol Sport Exerc*. 2016;24:100–10.
 76. Sebire SJ, Edwards MJ, Kesten JM, May T, Banfield KJ, Bird EL, et al. Process evaluation of the Bristol girls dance project. *BMC Public Health*. 2016;16(1):1–13.
 77. Lefler LL, Jones S, Harris B. Key strategies for physical activity interventions among older women: process evaluation of a clinical trial. *Am J Health Promot*. 2018;32(3):561–570.
 78. Furness C, Howard E, Limb E, Cook DG, Kerry S, Wahlich C, et al. Relating process evaluation measures to complex intervention outcomes: findings from the PACE-UP primary care pedometer-based walking trial. *Trials*. 2018; 19:1 N.PAG.
 79. Bonde AH, Stjernqvist NW, Sabinsky MS, Maindal HT. Process evaluation of implementation fidelity in a Danish health-promoting school intervention. *BMC Public Health*. 2018;18(1):1407.
 80. Dadaczynski K, Schieman S, Backhaus O. Promoting physical activity in worksite settings: Results of a German pilot study of the online intervention Healingo fit. *BMC Public Health*. 2017;17(1):696.
 81. Green J, Tones K. *Health promotion: planning and strategies*. 3rd edition ed. London: SAGE; 2010.
 82. Tucker P, Vanderloo LM, Johnson AM, Burke SM, Irwin JD, Gaston A, et al. Impact of the Supporting Physical Activity in the Childcare Environment (SPACE) intervention on preschoolers' physical activity levels and sedentary time: A single-blind cluster randomized controlled trial. *Int J Behav Nutri Phys Activity*. 2017;14(1):120.
 83. Sofija E, Plugge M, Wiseman N, Harris N. 'This is the beginning of the new me': Process evaluation of a group fitness intervention to promote wellbeing in formerly homeless individuals. *BMC Public Health*. 2018;18(1):290.
 84. Matthews L, Kirk A, McCallum M, Mutrie N, Gold A, Keen A. The feasibility of a physical activity intervention for adults within routine diabetes care: a process evaluation. *Pract Diabetes*. 2017;34(1):7–12a.

85. Galaviz KI, Estabrooks PA, Ulloa EJ, Lee RE, Janssen I, Lopez y Taylor J, et al. Evaluating the effectiveness of physician counseling to promote physical activity in Mexico: an effectiveness-implementation hybrid study. *Transl Behav Med*. 2017;7(4):731–40.
86. Christian D, Todd C, Hill R, Rance J, Mackintosh K, Stratton G, et al. Active children through incentive vouchers - evaluation (ACTIVE): a mixed-method feasibility study. *BMC Public Health*. 2016;16(1):1–10.
87. Reznik M, Wylie-Rosett J, Kim M, Ozuah PO. A classroom-based physical activity intervention for urban kindergarten and first-grade students: A feasibility study. *Child Obes*. 2015;11(3):314–24.
88. Kohn M, Belza B, Petrescu-Prahova M, Miyawaki CE. Beyond strength: participant perspectives on the benefits of an older adult exercise program. *Health Education and Behavior*. 2015;43(3):305–12.
89. Petrescu-Prahova M, Belza B, Kohn M, Miyawaki C. Implementation and maintenance of a community-based older adult physical activity program. *Gerontologist*. 2016;56(4):677–86.
90. Wilcox S, Saunders RP, Kaczynski AT, Forthofer M, Sharpe PA, Goodwin C, et al. Faith, activity, and nutrition randomized dissemination and implementation study: countywide adoption, reach, and effectiveness. *Am J Prev Med*. 2018;54(6):776–85.
91. Harden SM, Johnson SB, Almeida FA, Estabrooks PA. Improving physical activity program adoption using integrated research-practice partnerships: an effectiveness-implementation trial. *Transl Behav Med*. 2017;7(1):28–38.
92. Baba CT, Martins Oliveira I, Ferreira Silva AE, Vieira LM, Cerri NC, Florindo AA, et al. Evaluating the impact of a walking program in a disadvantaged area: using the RE-AIM framework by mixed methods. *BMC Public Health*. 2017;17:1–11.
93. Economos CD, Anzman-Frasca S, Koomas AH, Bakun PJ, Brown CM, Brown D, et al. Dissemination of healthy kids out of school principles for obesity prevention: A RE-AIM analysis. *Prev Med*. 2019;119:37–43.
94. Ward S, Chow AF, Humbert ML, Belanger M, Muhajarine N, Vatanparast H, et al. Promoting physical activity, healthy eating and gross motor skills development among preschoolers attending childcare centers: process evaluation of the healthy start-depart Sante intervention using the RE-AIM framework. *Evaluation and Program Planning*. 2018;68:90–8.
95. Jung ME, Bourne JE, Gainforth HL. Evaluation of a community-based, family focused healthy weights initiative using the RE-AIM framework. *Int J Behav Nutr Phys Activity*. 2018;15(1):13.
96. Allar I, Jones E, Elliott E, Kristjansson A, Taliadro A, Mann M, et al. The perceived impact of i am moving, i am learning on physical activity and family involvement: A preliminary investigation. *Am J Health Behav* 2017; 41(6):683–692.
97. Caperchione CM, Duncan M, Kolt GS, Vandelanotte C, Rosenkranz RR, Maeder A, et al. Examining an Australian physical activity and nutrition intervention using RE-AIM. *Health Promot Int*. 2016;31(2):450–8.
98. Gonzalez-Viana A, Violan Fors M, Castell Abat C, Rubinat Masot M, Oliveras L, Garcia-Gil J, et al. Promoting physical activity through primary health care: the case of Catalonia. *BMC Public Health*. 2018;18(1):968.
99. Schwingel A, Gálvez P, Linares D, Sebastião E. Using a mixed-methods RE-AIM framework to evaluate community health programs for older Latinas. *J Aging Health*. 2017;29(4):551–93.
100. Paez DC, Reis RS, Parra DC, Hoehner CM, Sarmiento OL, Barros M, et al. Bridging the gap between research and practice: an assessment of external validity of community-based physical activity programs in Bogotá, Colombia, and Recife, Brazil. *Transl Behav Med*. 2015;5(1):1–1.
101. Koorts HG. Mixed method evaluation of a community-based physical activity program using the RE-AIM framework: Practical application in a real-world setting Energy balance-related behaviors. *BMC Public Health*. 2015; 15(1):1102.
102. Bellicha A, Kieusseian A, Fontvieille A-M, Tataranni A, Copin N, Charreire H, et al. A multistage controlled intervention to increase stair climbing at work: effectiveness and process evaluation. *Int J Behav Nutr Phys Act*. 2016;13:1–9.
103. Beets MW, Glenn Weaver R, Brazendale K, Turner-McGrievy G, Saunders RP, Moore JB, et al. Statewide dissemination and implementation of physical activity standards in afterschool programs: Two-year results. *BMC Public Health*. 2018;18(1):819.
104. Beets MW, Glenn Weaver R, Turner-McGrievy G, Saunders RP, Webster CA, Moore JB, et al. Evaluation of a statewide dissemination and implementation of physical activity intervention in afterschool programs: a nonrandomized trial. *Transl Behav Med*. 2017;7(4):690–701.
105. Lee RE, Parker NH, Soltero EG, Ledoux TA, Mama SK, McNeill L. Sustainability via active garden education (SAGE): results from two feasibility pilot studies. *BMC Public Health*. 2017;17:1–11.
106. Lewis ZH, Ottenbacher KJ, Fisher SR, Jennings K, Brown AF, Swartz MC, et al. The feasibility and RE-AIM evaluation of the TAME health pilot study. *Int J Behav Nutr Phys Act*. 2017;14:1–15.
107. Adams EJ, Chalkley AE, Eslinger DW, Sherar LB. Evaluation of the implementation of a whole-workplace walking programme using the RE-AIM framework. *BMC Public Health*. 2017;17:1–18.
108. Herens M, Wagemakers A, Vaandrager L, van Ophem J, Koelen M. Contexts, mechanisms, and outcomes that matter in Dutch community-based physical activity programs targeting socially vulnerable groups. *Evaluation Health Professions*. 2017;40(3):294–331.
109. Daniels JE, Bell B, Horrocks C. Capturing the realities of sports programmes: systematic ‘messiness’? *Int J Sport Pol Politics*. 2018;10(4):779–94.
110. Willis CE, Reid S, Elliott C, Rosenberg M, Nyquist A, Jahnsen R, et al. A realist evaluation of a physical activity participation intervention for children and youth with disabilities: What works, for whom, in what circumstances, and how? *BMC Pediatrics*. 2018;18(1):113.
111. Mikkelsen BE, Bloch P, Reinbach HC, Buch-Andersen T, Winkler LL, Toft U, et al. Project SoL—A community-based, multi-component health promotion intervention to improve healthy eating and physical activity practices among danish families with young children part 2: Evaluation. *Int J Environ Res Public Health*. 2018;15(7):1513.
112. Biddle SJ, Edwardson CL, Gorely T, Wilmot EG, Yates T, Nimmo MA, et al. Reducing sedentary time in adults at risk of type 2 diabetes: process evaluation of the STAND (sedentary time AND diabetes) RCT. *BMC Public Health*. 2017;17(1):80.
113. Berendsen BA, Hendriks MR, Verhagen EA, Schaper NC, Kremers SP, Savelberg HH. Effectiveness and cost-effectiveness of ‘BeweegKuur’, a combined lifestyle intervention in the Netherlands: rationale, design and methods of a randomized controlled trial. *BMC Public Health*. 2011;11(1): 815.
114. Bauman A, Nutbeam D. *Evaluation in a nutshell*: McGraw-Hill education Australia; 2013.
115. Stjernqvist NW, Sabinsky M, Morgan A, Trolle E, Thyregod C, Maingal HT, et al. Building school-based social capital through ‘we act - together for health’ – a quasi-experimental study. *BMC Public Health*. 2018;18(1):1141.
116. Harden SM, Strayer TE 3rd, Smith ML, Gaglio B, Ory MG, Rabin B, et al. National Working Group on the RE-AIM planning and evaluation framework: goals, resources, and future directions. *Front Public Health*. 2019;7:390.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

