



HAL
open science

A scoping review of outcome selection and accuracy of conclusions in complex digital health interventions for young people (2017–2023): methodological proposals for population health intervention research

Claire Collin, Clara Eyraud, Philippe Martin, Morgane Michel, Enora Le Roux,
Corinne Alberti

► To cite this version:

Claire Collin, Clara Eyraud, Philippe Martin, Morgane Michel, Enora Le Roux, et al.. A scoping review of outcome selection and accuracy of conclusions in complex digital health interventions for young people (2017–2023): methodological proposals for population health intervention research. *BMC Medicine*, 2025, 23 (1), pp.400. <10.1186/s12916-025-04245-1>. <hal-05340597>

HAL Id: hal-05340597

<https://cnrs.hal.science/hal-05340597v1>

Submitted on 31 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

RESEARCH

Open Access



A scoping review of outcome selection and accuracy of conclusions in complex digital health interventions for young people (2017–2023): methodological proposals for population health intervention research

Claire Collin^{1*} , Clara Eyraud¹ , Philippe Martin^{1,2,3} , Morgane Michel^{1,2} , Enora Le Roux^{1,2,4}  and Corinne Alberti^{1,2} 

Abstract

Background Determining the success of population health interventions often involves assessing multiple, multidimensional outcomes rather than a single one, which presents significant methodological challenges under the evidence-based medicine paradigm. This scoping review examines outcome selection, analysis, and interpretation, and the accuracy of conclusions in complex digital health interventions promoting health among adolescents and young adults (DHI-AYA).

Methods A comprehensive search of PubMed, EMBASE, ClinicalTrials.gov, PsycINFO, and CINAHL identified DHI-AYA implemented between 2017 and 2023. Studies were categorised by methodological choice regarding outcome hierarchical position: unique primary, multiple primary, or non-hierarchised outcomes. Outcomes were further classified into effectiveness, process, or economic categories. The authors' conclusions on intervention success were compared with conclusions drawn by the research team based on the reported outcome analysis strategy. Secondly, four analytical strategies were applied to a subset of selected interventions to illustrate the impact of outcome hierarchical position and number on conclusions about intervention success.

Results Analysis of 100 studies linked to 26 DHI-AYA identified 251 distinct outcomes: 164 effectiveness, 78 process, and 9 economic outcomes. Seven interventions were evaluated using a unique primary outcome, 10 using multiple primary outcomes, and 9 using multiple non-hierarchised outcomes. Primary and secondary outcomes were predominantly effectiveness endpoints. The research team reclassified nine interventions (35%) deemed successful by authors as non-conclusive due to statistically conflicting results across outcomes. Most interventions deemed non-conclusive by the research team were evaluated using non-hierarchised outcomes (7/10, 70%). The choice of outcome analysis strategy substantially affected conclusions on intervention success.

Conclusions Discrepancies in intervention success assessments highlight the need for enhanced transparency, robustness, and trustworthiness in conclusion-drawing processes. In response, five methodological proposals are

*Correspondence:

Claire Collin
ccollin.contact@gmail.com

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



© The Author(s) 2025. **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/>.

formulated: (1) developing core outcome sets specific to population health intervention research (PHIR), (2) collaboratively selecting multidimensional outcomes through a steering committee that accounts for stakeholder preferences and existing theoretical models, (3) exploring multi-criteria decision analysis and consensus-driven methods to transparently combine outcomes, (4) enhancing methodological reporting through intervention development and evaluation to improve scientific integrity and reproducibility, and (5) increasing PHIR expert involvement in ethics, funding, and evaluation committees to improve recognition of evidence produced in this field.

PROSPERO Registration number CRD42023401979.

Keywords Methodological research, Outcomes, Interventions, Population health intervention research, Scoping review

Background

Health research spans a continuum of fields from (pre) clinical sciences to population health intervention research (PHIR), each characterised by distinct focuses and methodologies [1]. While clinical research has long adhered to standardised practices, notably randomised controlled trials often centred on a unique primary outcome under the evidence-based medicine (EBM) paradigm [2, 3], PHIR is still developing and refining its methodological framework [4].

The UK Medical Research Council (MRC) has significantly advanced PHIR methodology [5] through guidance for developing and evaluating complex interventions [6], which constitute the majority of PHIR studies [7]. This guidance outlines four research phases—development, feasibility, evaluation (encompassing efficacy or effectiveness, process and efficiency evaluations), and implementation—alongside six core elements to consider throughout these phases (i.e. context, programme theory, stakeholder engagement, key uncertainties, intervention refinement, economic considerations). While providing a valuable framework for evaluating PHIR, the guidance offers limited direction regarding the practical application of evaluation methodologies [6].

One key methodological challenge in PHIR is the selection, analysis, and interpretation of multiple outcomes. The use of multiple outcomes is recommended to assess the full range of intervention effects across several dimensions, including effectiveness, process, efficiency, and implementation contexts [6]. While this approach aims to support conclusions about interventions' success and deployment [8], it raises several methodological issues: increased risk of type I errors [9], complex sample size calculations [10], and potential interpretation difficulties when measured outcomes yield conflicting results [11].

These methodological issues are particularly reported in the evaluation of digital health interventions, which are emerging as a growing focus in PHIR [12]. Digital health interventions are typically considered complex interventions, composed of multiple interactive components,

often targeting several health behaviours simultaneously and implemented across diverse contextual settings [13]. They offer broad reach at a relatively low cost [14] and are particularly well-suited to promote behaviour change among adolescents and young adults, a group that extensively uses digital platforms for socialising and accessing health information [15]. The limited availability of robust evidence hinders the widespread implementation of evidence-based digital health interventions, highlighting the need for further methodological research [16, 17].

This study aims to examine the methodological decisions made by researchers in selecting, analysing, and interpreting outcomes to determine the success of complex digital interventions promoting health among adolescents and young adults. While these methodological issues extend beyond digital interventions, they serve as an illustrative example for broader applications within PHIR.

Methods

A scoping review was conducted following PRISMA guidelines [18] (see Additional file 1: Tables S1–S2 [18] for the completed PRISMA 2020 Checklists) and registered with PROSPERO (CRD42023401979). The review examined digital health interventions (DHIs) promoting health among adolescents and young adults (AYA), focusing on the methods and outcomes used for evaluation and the impact of different methodological approaches on intervention conclusions.

Eligibility of digital health interventions

Eligibility was assessed at the intervention level, identifying complex primary prevention or health promotion interventions for AYA aged 10 to 24, delivered exclusively via digital technologies (e.g. computer programmes, mobile applications, websites, social media, text messaging services, wearable devices) [19]. Interventions were classified as complex when comprising at least two distinct components (e.g. videos, quizzes, information sheets) and engaging users through multiple modalities, ranging from passive information delivery to interactive

engagement with peers or healthcare professionals. To ensure the contemporary relevance of included technologies in a rapidly evolving digital landscape, interventions had to be implemented between 2017 and 2023, with 2017 specifically selected as the search start date following the publication of several seminal papers that established fundamental frameworks for digital health intervention design and evaluation [12, 13, 20–24]. Eligible interventions were required to report outcome results from an efficacy or effectiveness evaluation (evaluation phase) and from at least one additional MRC-defined research phase among development, feasibility, or implementation [6]. This approach aimed to identify interventions already adhering to the best methodological recommendations to explore gaps in their practical application rather than highlight the need for improved guideline adherence. Publications not in English or French were excluded.

Search strategy

PubMed, EMBASE, ClinicalTrials.gov, PsycINFO, and CINAHL were searched on February 9th, 2023, with a last update on January 2nd, 2024. Five groups of relevant keywords were combined: “eHealth”, “intervention research”, “evaluation”, “health promotion and prevention”, and “adolescents and young adults”. The search strategy was iteratively developed with input from the research team and support from the university library (see Additional file 2 for detailed search strategies). For each eligible intervention, all related protocols, published articles, and preprints were retrieved through citation and reference searching, regardless of publication date. Ancillary studies were excluded [25].

Selection process, data extraction, and risk-of-bias assessment

Two researchers (CC, CE) independently screened titles and abstracts, reviewed full texts, and extracted data using Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia). Disagreements were resolved through discussion with a third researcher (ELR or PM).

Data were extracted using a standardised form (see Additional file 3), collecting the following: (1) intervention characteristics, including health topics, population demographics, types and number of digital technologies employed, and intervention components; (2) reported outcomes, including variable domain, specific measurement, metric and time points, method of aggregation [26], hierarchical position (primary, secondary, non-hierarchical), number of outcomes per position, and statistical significance of the result for each outcome

(statistically significant or not statistically significant at level 5% with bilateral testing potentially corrected for multiple testing, or not tested); (3) authors’ conclusions regarding their interventions, categorised into three groups: (i) success (S) in achieving intervention objectives, (ii) failure (F) in achieving objectives, or (iii) non-conclusive (NC) due to statistically conflicting results; and (4) risk of bias for individual studies, assessed using Cochrane’s tools for randomised controlled trials [27] and non-randomised studies [28].

Outcome hierarchical positions were extracted from study protocols or, when unavailable, from the *Methods* section of primary effectiveness papers.

Data synthesis

Outcomes categorisation

Outcomes were categorised according to their nature and role in conclusions.

Nature of outcomes Outcomes were classified into three categories: effectiveness (encompassing efficacy and effectiveness outcomes), process, and economic outcomes.

Effectiveness outcomes were grouped into clinical outcomes (e.g. pregnancy, body mass index), healthcare utilisation (e.g. HIV or sexually transmitted infections (STIs) testing), behavioural outcomes (e.g. condom use, substance use), psychosocial outcomes (e.g. stress, coping), and behavioural determinants (i.e. knowledge, attitudes, perceived barriers, outcomes expectancy, norms, self-efficacy, skills, and intentions) as identified in behavioural change theories [29].

Process outcomes were grouped according to the MRC guidance for process evaluation (i.e. dose, fidelity, reach, contextual factors), supplemented by the Proctor’s outcomes for implementation research (i.e. acceptability, adoption, appropriateness, feasibility, retention, perceived effectiveness, usability) [30, 31]. Additionally, safety, defined as participants feeling safe using the intervention, was included as a process outcome, consistent with its classification in the original publication [32].

Economic outcomes, as defined in this study, encompass both cost estimates (i.e. absolute costs) and cost-effectiveness metrics (i.e. incremental cost-effectiveness ratios (ICER), incremental-cost utility ratios (ICUR), and costs per disability-adjusted life year (DALY)).

Role in conclusions Outcomes were classified as ‘used to conclude’ if mentioned in the justification of intervention success or failure, or ‘not used’ if only reported in *Results*.

Interventions categorisation

At the intervention level, studies were classified into three groups based on the outcomes defined in the effectiveness evaluation phase, distinguishing those using (A) a unique primary outcome (with or without secondary or additional non-hierarchised outcomes), (B) multiple primary outcomes (with or without secondary or additional non-hierarchised outcomes), or (C) multiple non-hierarchised outcomes.

Data analysis

The analysis comprised the following: (1) a comparison between the authors' conclusions about intervention success, as stated on the articles' discussion or conclusion sections, and conclusions drawn by the research team based on the authors' reported outcome analysis strategy for each intervention, and (2) a sub-analysis examining how different outcome analysis strategies (i.e. outcome hierarchical positions and numbers) may influence overall intervention conclusions.

First, the research team assigned their conclusions to each intervention, based on outcomes' hierarchical position specified in the study protocols or *Methods* sections and results (effect size and statistical significance) reported in the *Results* sections. Interventions deemed successful by authors but unsuccessful by researchers were classified as spin. A spin is defined as "the intentional or unintentional specific reporting that unfaithfully reflects the nature and range of findings" [33]. Agreement percentages between authors' and researchers' conclusions were calculated.

Second, three sexual and reproductive health interventions were selected from the analysed articles to explore the influence of outcome analysis strategy on conclusions. The sub-analysis was restricted to effectiveness outcomes. The three interventions were purposefully selected to represent each of the three approaches to outcome analysis (i.e. use of a unique primary outcome (category A), multiple primary outcomes (category B), or multiple non-hierarchised outcomes (category C)), while being as comparable as possible on other characteristics, including type of digital technology, target population age, theoretical framework, and at least one shared effectiveness outcome (i.e. number of protected sex acts).

Four strategies reflecting varying outcome hierarchical positions and numbers, identified in the review and the existing literature [34], were applied to the three interventions:

1) *Unique primary outcome, with conclusions based on this outcome.*

2) *Multiple non-hierarchised outcomes, with conclusions based on all measured outcomes, without prioritisation; this strategy required all outcomes to show significant improvement to deem the intervention successful.*

For the following strategies 3 and 4, only a limited number of outcomes, comparable across the chosen interventions or considered important by the original authors, were included in determining success. Each selected outcome was assigned a score and a weight. The score was binary: 1 for improvement (i.e. statistically significant) or 0 for no improvement (i.e. not statistically significant). The weight was chosen between 0 and 1 (totalling 1). Each outcome's weight was then multiplied by its score, and the total intervention score was calculated by summing all obtained products. Interventions scoring above 0.5 were classified as successful.

3) *Multiple non-hierarchised outcomes equally weighted, with conclusions based on the equally weighted total score of the selected outcomes;*

4) *Multiple non-hierarchised outcomes differentially weighted, with conclusions based on the differentially weighted total score of the selected outcomes. One researcher (CC) assigned the weights according to the relevance of the outcomes to the health condition (Additional file 4: Table S3 [35]).*

Results are presented using descriptive statistics. Categorical variables are expressed as counts and percentages, and numerical variables as medians with 1st and 3rd quartiles (med, (Q1-Q3)).

Results

Study selection

From the 6691 records initially identified, 26 interventions described across 100 individual reports were analysed (Fig. 1, Additional file 5 for the full list of included reports [32, 36–134]). A median of four articles per intervention was included (Q1–Q3: 3–5).

Intervention characteristics

Included interventions predominantly targeted sexual and reproductive health (13/26, 50%), were delivered through a dedicated mobile app or website (20/26, 77%), using a single technology (17/26, 65%), over a median duration of 3 months (Q1–Q3: 1–7). Nearly three-quarters of interventions were implemented and evaluated in daily-life settings (19/26, 73%) (Table 1). Most interventions were evaluated using individual or cluster randomised controlled trials (23/26, 88%), with the remainder using quasi-experimental pre-post designs (3/26, 12%).

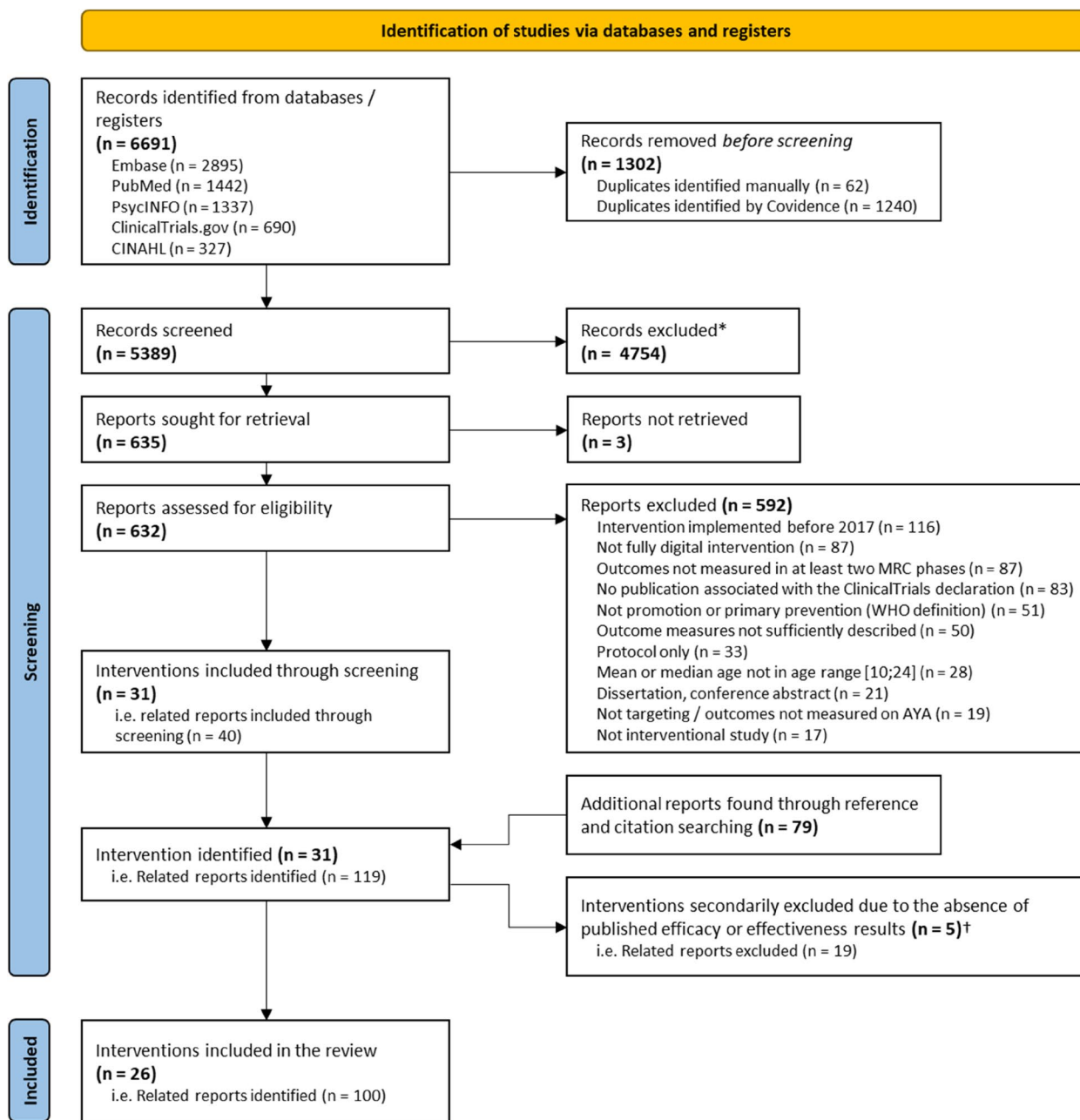


Fig. 1 PRISMA flow diagram of literature search

*Exclusion reasons are similar to those for full-text assessment; †These interventions were excluded because only efficacy or effectiveness evaluation protocols were available, which would have resulted in insufficient data for comprehensive analysis

Seven interventions (7/26) specified a unique primary outcome, with secondary outcomes and with or without additional non-hierarchised outcomes (Fig. 2, category A); 10 interventions (10/26) were assessed on multiple primary outcomes (category B), with or without secondary or non-hierarchised outcomes, and 9 interventions (9/26) were evaluated using solely multiple non-hierarchised outcomes (category C). None of the interventions

using multiple outcomes applied a correction for multiplicity.

Outcomes

Figure 2 presents the distribution of outcomes by their nature (effectiveness, process, economic), hierarchical position (primary, secondary, non-hierarchised), and statistical results (significant, not significant, not tested)

Table 1 Distribution of characteristics for included interventions ($n = 26$)

General characteristics	All interventions n (%) [*]
Health topics	
Sexual and reproductive health	13 (50)
Substance use (alcohol, tobacco, drugs)	6 (23)
Mental health	5 (19)
Physical activity and/or nutrition	2 (8)
Main type of digital technology used in interventions	
Mobile app or website	20 (77)
Social media	3 (12)
Text-messaging services (e.g. SMS ^x , emails)	3 (12)
Number of digital technologies used in interventions	
Single technology	17 (65)
Multiple technologies	9 (35)
Recruitment setting	
Educational setting (school, university)	10 (38)
Online / Internet	6 (23)
Multi-setting [†]	6 (23)
Healthcare facilities	2 (8)
Other (community-based organisations, phone)	2 (8)
Implementation setting	
Daily life (online, home, etc.)	19 (73)
Educational setting (school, university)	7 (27)
Intervention categorisation	
Intervention evaluated using a unique primary outcome	7 (27)
Intervention evaluated using multiple primary outcomes	10 (38)
Intervention evaluated using non-hierarchised outcomes	9 (35)
Outcomes	med (Q1–Q3)
Number of reported outcomes per intervention [‡]	8 (5–11)
Effectiveness outcomes ($n = 26$)	5 (4–7)
Process outcomes ($n = 17$)	3 (2–6)
Economic outcomes ($n = 3$)	3 (3–4)

^{*}Percentages may not add up to 100 due to rounding; ^xSMS: Short Message Service; [†]: multi-setting refers to recruitment that occurred in at least two different settings, including educational settings, online or internet-based platforms, healthcare facilities, or other settings (e.g. community-based organisations, phone); [‡]: number of outcomes regardless of outcome hierarchical position; excludes outcomes mentioned in protocols only

across intervention categories (A, B, C). It also compares the conclusions drawn by authors and researchers regarding intervention success.

A total of 251 reported outcomes were identified across the interventions, including 164 effectiveness outcomes from all 26 interventions, 78 process outcomes from 17 interventions, and 9 economic outcomes from 3 interventions. Of the effectiveness outcomes, 132 measured behaviours or behavioural determinants. An additional

42 outcomes, primarily process ones, were reported only in protocols (shown in grey in Fig. 2) and were not included in the total outcome count. Additional file 6: Tables S4–S5 provide detailed characteristics of outcome hierarchical position and measurement.

Interventions evaluated with a unique primary outcome (category A, $n = 7$)

All primary and secondary outcomes were effectiveness-related and statistically tested. Five interventions additionally reported non-hierarchised process or economic outcomes. These outcomes were generally not statistically tested, notably because many were exclusively measured in the intervention group.

Interventions evaluated with multiple primary outcomes (category B, $n = 10$)

All primary outcomes were effectiveness-related, except for one usability outcome, and statistically tested. Process outcomes were evaluated as secondary outcomes in only two interventions. Half of the interventions reported additional non-hierarchised outcomes, all process and economic endpoints.

Interventions evaluated with multiple non-hierarchised outcomes (category C, $n = 9$)

These interventions predominantly assessed effectiveness outcomes, all statistically tested. Six interventions also reported process outcomes, while none evaluated economic endpoints.

Rationales for outcome hierarchical positions were generally not stated for all interventions (categories A, B, and C).

Role of outcomes in conclusions

In 8/26 interventions, conclusions were based on all reported outcomes, primarily effectiveness, with one also considering a process outcome (see Additional file 7: Figure S1 for details on outcomes used to determine intervention success). In the remaining 18/26 interventions, conclusions were drawn from a subset of reported outcomes, though the rationale for their selection was not stated.

Interventions evaluated with a unique primary outcome (category A, $n = 7$)

Two interventions used all reported outcomes to draw conclusions. All interventions used all reported primary and secondary effectiveness outcomes to conclude, with four incorporating non-hierarchised process outcomes. None included economic outcomes.

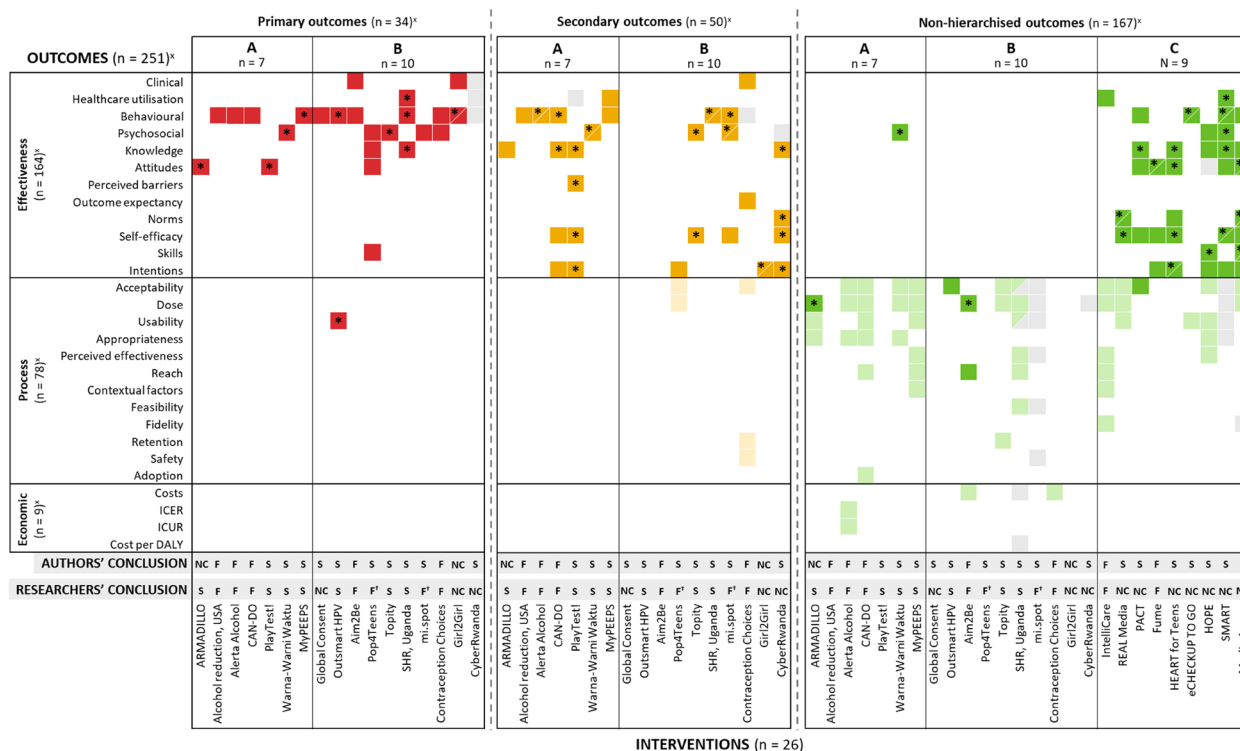


Fig. 2 Description of outcome nature, hierarchical position, and statistical results by intervention category, with authors’ and researchers’ conclusions on success

*Number of reported outcomes excludes the 42 outcomes mentioned in protocols only. Intervention names are listed at the bottom of the figure. **Intervention categorisation:** A: intervention studies defining a unique primary outcome, with or without additional outcomes. B: intervention studies defining multiple primary outcomes, with or without additional outcomes. C: intervention studies defining multiple non-hierarchical outcomes. Interventions in categories A and B are displayed across the three outcome hierarchical positions—primary, secondary, and non-hierarchical—and therefore appear three times in the figure. **Outcome hierarchical position:** Red: primary outcomes. Yellow: secondary outcomes. Green: non-hierarchical outcomes. Grey: outcomes in protocols only ($n=42$), not included in the total outcome count. **Outcome statistical result:** Dark shade with an asterisk (*): statistically significant outcomes ($p < 0.05$). Dark shade without an asterisk: not statistically significant outcomes ($p \geq 0.05$). Light shade: outcomes not subjected to statistical testing. **Conclusions on intervention success:** S, success; F, failure; NC, non-conclusive; F+, spin (when authors reported success despite all primary outcomes being not statistically significant). **Abbreviations:** ICER, incremental cost-effectiveness ratio; ICUR, incremental cost-utility ratio; DALY, disability-adjusted life years; SRH, sexual and reproductive health. **How to read the figure:** for example, in the ARMADILLO study (category A intervention), the unique effectiveness primary outcome (attitudes) was statistically significant, while the effectiveness secondary outcome (knowledge) was not. Three exploratory (non-hierarchical) process outcomes were measured: intervention dose (statistically significant), usability and appropriateness (both not statistically tested). Both the authors and the researchers concluded that the intervention results were non-conclusive (NC)

Interventions evaluated with multiple primary outcomes (category B, n = 10)

Four interventions used all reported outcomes to draw conclusions. All interventions included all primary effectiveness outcomes, with four supplementing these with secondary effectiveness outcomes. One intervention also included a secondary process outcome, and three used non-hierarchical process outcomes in their conclusions. None included economic outcomes.

Interventions evaluated with multiple non-hierarchical outcomes (category C, n = 9)

Two interventions used all reported outcomes to draw conclusions. Six used only effectiveness outcomes, while two included process outcomes. One used all but one effectiveness outcome. None used economic outcomes.

Authors’ and researchers’ conclusions on intervention success

Interventions evaluated with a unique primary outcome (category A, n = 7)

The research team agreed with the authors’ conclusions for six interventions (three successful, three unsuccessful;

Fig. 2, Additional file 8: Tables S6-S7 for detailed agreement percentages). The remaining intervention was deemed successful by the team based on improvement in the unique primary outcome, while the authors considered it non-conclusive due to secondary outcome results.

Interventions evaluated with multiple primary outcomes (category B, n = 10)

Agreement was reached for six interventions (three successful, two unsuccessful, one non-conclusive). Two interventions were spin, with authors reporting success despite no statistically significant primary outcomes. The remaining two were classified as non-conclusive by the research team, as only some primary outcomes were statistically significant and no correction for multiplicity was applied.

Interventions evaluated with multiple non-hierarchicalised outcomes (category C, n = 9)

Agreement was reached for two unsuccessful interventions. The remaining seven, deemed successful by their authors, were considered non-conclusive by the research team due to statistically conflicting results and lack of multiplicity correction.

Impact of outcome hierarchical position and number on intervention conclusions

Table 2 illustrates the impact of outcome hierarchical position and number on intervention conclusions for three interventions originally evaluated using (1) a single primary outcome, (2) multiple primary outcomes, and (3) multiple non-hierarchicalised outcomes. Additional file 9: Table S8 provides additional details on the hierarchical position and result of each outcome for these interventions.

Original authors' conclusions

Intervention 1 was considered successful based on improvement in its unique primary outcome. Intervention 2 was deemed non-conclusive, as only two of its four primary outcomes showed improvement. Intervention 3 was judged successful, although only 10 of the 18 non-hierarchicalised outcomes showed improvement.

Conclusions under alternative strategies

When applying the four predefined strategies, the conclusions on intervention success varied. When success was determined based on the “number of protected sex acts” as the unique primary outcome (Strategy 1), interventions 1 and 2 were successful, while intervention 3 was a failure. Notably, this outcome was measured using three different instruments across the interventions. Under Strategy 2, all interventions were deemed

non-conclusive, as none showed improvement across all outcomes. Weighted strategies (Strategies 3 and 4) were applied to four outcomes similarly measured in interventions 1 and 3, and to the four primary outcomes measured in intervention 2. Both weighted strategies classified intervention 2 as successful and intervention 1 as a failure. Intervention 3 was successful under Strategy 3 (equal weighting) but failed under Strategy 4 (differential weighting).

Risk-of-bias assessment

Most outcomes were self-reported by participants aware of their intervention allocation, leading to some concerns or high risk of bias in outcome measurement for 87% (20/23) of interventions evaluated within a randomised controlled design. Quasi-experimental studies also showed a high risk of outcome measurement bias (3/3). Selective reporting was at low risk of bias in 65% (17/26) of all interventions (see Additional file 10: Figures S2–S3 for the full risk assessment).

Discussion

Summary of main findings

This study demonstrates that population health intervention research (PHIR) on adolescents and young adults (AYA) using digital health interventions (DHIs) is evaluated using a wide range of outcomes (effectiveness, process and economic outcomes), as recommended by the MRC guidance [6]. However, intervention success remains predominantly determined by effectiveness outcomes, particularly behaviours and behavioural determinants. While defining primary outcomes simplifies success determination, it requires researchers to prioritise. However, the scientific rationale and processes underpinning this prioritisation are rarely documented, corroborating previous findings [135]. Similarly, authors rarely report their interpretative process for drawing conclusions from multiple outcomes, particularly when results are statistically conflicting and outcomes non-hierarchicalised. Even in studies with clearly defined primary outcomes, secondary measures frequently influence conclusions, suggesting a more complex interpretation process than initially reported. Finally, this study highlights how methodological decisions on outcome selection and prioritisation substantially shape intervention conclusions.

Interpretation of results

The reliance on effectiveness outcomes for determining intervention success supports existing evidence of their overemphasis in DHI evaluations [136] and PHIR more broadly [137], despite methodological guidelines advocating for multidimensional assessments [6, 17, 138]. This

Table 2 Impact of outcome hierarchical position and number on intervention conclusions ($n = 3$)

	Intervention 1	Intervention 2	Intervention 3
General characteristics			
Objective	HIV prevention	Unwanted pregnancy prevention	HIV prevention
Intervention theories	Social-Personal framework, Social Learning Theory	IMB skills model for HIV prevention	IMB skills model for HIV prevention
Target population	MSM	Cis-gender LGB+ girls	MSM
Target age	13-18	14-18	13-18
Technology	App	Text messages	App + Online live sessions
Intervention duration	3 months	4,5 months	9 months
Implementation setting	Daily life	Daily life	Daily life
Methods			
Study design	2-arm parallel RCT	2-arm parallel RCT	Pretest-posttest design
N participants: intervention/control (M3)	380 / 381	389 / 410	983 / 983
Original outcome hierarchisation	Unique primary outcome	Multiple primary outcomes	Non-hierarchised outcomes
N outcomes assessed	8	7	18
Authors' conclusion at M3			
<i>N outcomes improved</i>	1 / 1	2 / 4	10 / 18
<i>/ N outcomes used to conclude</i>			
<i>Original conclusion</i>	Success	Non-conclusive	Success
Strategy 1 – Unique primary outcome at M3			
<i>Outcome (result)</i>	N protected sex acts (1)	N protected sex acts (1)	N protected sex acts (0)
<i>Conclusion</i>	Success	Success	Failure
Strategy 2 – Multiple non-hierarchised outcomes at M3, considering all assessed outcomes			
<i>N outcomes improved</i>	1 / 8	3 / 7	10 / 18
<i>/ N outcomes used to conclude</i>			
<i>Conclusion</i>	Non-conclusive	Non-conclusive	Non-conclusive
Strategy 3 – Multiple non-hierarchised equally weighted outcomes at M3, considering a subset of outcomes			
<i>Outcome (result ; weight)</i>	N protected sex acts (1 ; 0.25)	N protected sex acts (1 ; 0.25)	N protected sex acts (0 ; 0.25)
	N condomless anal sex partners (0 ; 0.25)	Non-condom birth control use (1 ; 0.25)	N condomless anal sex partners (0 ; 0.25)
	Testing for STIs (0 ; 0.25)	Abstinence (penetrative sex) (0 ; 0.25)	Testing for STIs (1 ; 0.25)
	Testing for HIV (0 ; 0.25)	Pregnancy (0 ; 0.25)	Testing for HIV (1 ; 0.25)
<i>Total[†]</i>	0.25 [†]	0.5 [†]	0.5 [†]
<i>Conclusion</i>	Failure	Success	Success
Strategy 4 – Multiple non-hierarchised differentially weighted outcomes at M3, considering a subset of outcomes			
<i>Outcome (result ; weight)</i>	N protected sex acts (1 ; 0.35)	N protected sex acts (1 ; 0.35)	N protected sex acts (0 ; 0.35)
	N condomless anal sex partners (0 ; 0.35)	Non-condom birth control use (1 ; 0.35)	N condomless anal sex partners (0 ; 0.35)
	Testing for STIs (0 ; 0.15)	Abstinence (penetrative sex) (0 ; 0.15)	Testing for STIs (1 ; 0.15)
	Testing for HIV (0 ; 0.15)	Pregnancy (0 ; 0.15)	Testing for HIV (1 ; 0.15)
<i>Total[†]</i>	0.35 [†]	0.70 [†]	0.3 [†]
<i>Conclusion</i>	Failure	Success	Failure

Table 2 (continued)

[†]Total score calculated by summing the products of each outcome result by its weight; [‡]arbitrary threshold for intervention success was set at 0.5. **Abbreviations:** *M*, months; *N*, number of; *HIV*, human immunodeficiency virus; *IMB*, Information-Motivation-Behavioural; *MSM*, men who have sex with men; *LGB+*, lesbian, gay, bisexual, and other sexual minority; *App*, mobile application; *RCT*, randomised controlled trial; *STIs*, sexually transmitted infections. **How to read the table:** for example, in intervention 1, under Strategy 1, the unique primary outcome is improved, deeming the intervention successful. Under Strategy 2, eight non-hierarchised outcomes are evaluated, with one improvement, leading to a non-conclusive intervention. Under Strategy 3, four outcomes are equally weighted (0.25 each) with one improved outcome, yielding an intervention score of 0.25, below the success threshold of 0.5, indicating failure. Under Strategy 4, four outcomes are differentially weighted (0.35, 0.35, 0.15, 0.15) with one improved outcome, yielding an intervention score of 0.35, below the success threshold of 0.5, indicating failure

tendency may arise from the relative ease of interpreting effectiveness outcomes, which are typically assessed through statistical testing with a clear significance threshold, whereas process outcomes are often qualitative and measured exclusively within intervention arms. However, this reflects a key methodological gap: while comprehensive evaluations are needed for informed decision-making, practical guidance on integrating heterogeneous findings across different evaluation phases remains limited (e.g. efficacy, effectiveness, implementation, efficiency, theory-based, and system-based outcomes) [6, 137]. Evidence synthesis is further complicated by inconsistencies in outcome selection and measurement across studies and, while diverse stakeholder perspectives on outcome selection enrich the analysis, they add to this challenge [139]. Thus, the lack of methodological guidance may explain why researchers struggle to articulate their interpretative processes, especially when incorporating non-primary outcomes into conclusions. However, reported processes suggest that authors distinguish intervention effectiveness from intervention success, with success encompassing a broader interpretation of findings across multiple dimensions. Yet, reproducing conclusions is challenging without transparent documentation of interpretative procedures, potentially explaining discrepancies between authors' and researchers' conclusions in this study. As in this study, researchers must rely on published *Methods* sections, which typically limit interpretation to the statistical results of primary or secondary outcomes, most often effectiveness outcomes, whereas authors may draw on a broader range of information, including additional secondary or non-hierarchised outcomes, unpublished data or contextual knowledge when determining success.

The overreliance on effectiveness outcomes may also reflect the prevailing research environment shaped by the dominance of the evidence-based medicine paradigm since the 1990s [140], which emerged as a means to prevent the implementation of untested and sometimes harmful interventions. However, longstanding investment in clinical research has reinforced an evaluation model prioritising randomised controlled trials and hierarchised outcomes, even in fields such as health promotion and population-health intervention research where such designs may not always be suitable. Notably, the

requirement to pre-specify a hierarchy of outcomes may overshadow valuable insights from secondary and non-hierarchised outcomes [6, 141–143]. Furthermore, unlike many pharmacological interventions, complex interventions showing positive results in some contexts may not 'work' in others, highlighting the importance of complementary studies, such as implementation and transferability research.

Additionally, while effectiveness outcomes may appear fundamental for intervention assessment, this overlooks critical implementation and resource constraints. An effective intervention that cannot be feasibly scaled or is cost-prohibitive ultimately fails to serve public health objectives. This underscores the necessity of integrating implementation considerations into intervention design from inception to reduce research waste [144].

The lack of clarity in outcome selection, measurement, and prioritisation for determining intervention success, particularly when multiple outcomes are involved, presents a significant challenge. Key risks include heterogeneity in outcome definitions, inconsistent measures for assessing similar interventions (e.g. sexual health), and difficulties ensuring result reproducibility. These issues hinder meta-analyses, cross-study comparisons, and interpretation of findings, especially when similar interventions are evaluated in different contexts. In clinical research, the Core Outcome Measures in Effectiveness Trials (COMET) Initiative was established to mitigate outcome heterogeneity within the same clinical context [145]. Given its unique complexities, PHIR could benefit from a similar framework tailored to its needs. Developing such a framework would require engaging diverse stakeholders, including public health researchers, citizens, policymakers, health technology assessment (HTA) bodies, developers, and information technology (IT) specialists (notably for DHIs), and economists. Additionally, outcome selection and prioritisation should align with established theoretical frameworks and the intervention's theoretical model or be informed by existing evidence from systematic reviews, similar interventions, and contexts [146].

Beyond outcome selection, given that multiple outcomes are necessary to assess the complexity of interventions from different stakeholders' perspectives, new methodological approaches to outcome combination

are needed. Multiple criteria decision analysis (MCDA), widely used in other research fields and increasingly in HTA [34], employs multi-stakeholders qualitative discussions, quantitative weighting, or decision rules to prioritise, rank, or select interventions based on multiple quantitative and qualitative criteria [34, 147]. Similarly, consensus-based approaches could be adapted to synthesise multidimensional information to determine intervention success [148]. Seldom used in PHIR, these methods warrant further exploration and adaptation to integrate the wide range of outcomes measured within complex interventions.

Finally, failure to document outcome selection and prioritisation processes undermines research integrity, as conclusions may be drawn post hoc rather than based on pre-specified measures [149]. To ensure transparency, success criteria should be pre-defined, alongside prioritisation and combination strategies, acknowledging stakeholders' diverse objectives. While not all measured outcomes will inform final decisions, they may provide valuable insights into unintended effects, mechanisms of actions, contextual influences, and scalability potential [6, 150].

Future directions

These findings underscore the need for greater transparency in communicating prioritisation criteria and decision-making processes to enhance methodological rigour in assessing the success of population health interventions based on multiple, multidimensional outcomes.

To support these improvements, we propose structuring the governance of research projects through a multidisciplinary steering committee established at the project's outset, comprising key stakeholders such as academic researchers, funders, implementers, and end-users.

Prior to the study, this committee would define a set of theoretically grounded, valid, reproducible, measurable, and responsive outcomes across at least the effectiveness, process, and economic dimensions. Each stakeholder would prioritise these outcomes according to existing evidence and their preferences and develop a predefined analytical framework that integrates quantitative and qualitative outcomes, accounting for contextual influences to determine intervention success [150]. Where possible, the committee would develop and document a consensus-based analytical plan for combining outcomes and formulating conclusions.

During the study, the committee would monitor project progress, document contextual influences, and identify challenges related to recruitment, data collection, and intervention fidelity.

Upon study completion, the committee would ensure the rigorous application of predefined analytical and interpretative procedures, considering any challenges encountered during implementation. The committee would critically appraise the results—particularly when both beneficial and adverse effects are observed—assess their generalisability in light of contextual factors, and deliver a well-reasoned, context-specific, and consensus-based conclusion regarding intervention success. Additionally, it would outline the risks and benefits of scaling up the intervention and transparently report any deviations from the initial analytical framework.

Even with improved outcome prioritisation and interpretation reporting, statistical challenges related to multiple outcome testing remain, particularly the risk of type I error inflation. While various statistical adjustments exist, their conservatism or underlying assumptions of hypothesis independence may limit their applicability in PHIR [151, 152]. Future research should explore the statistical implications of multiple outcome assessments and the inclusion of outcomes not typically associated with *p*-values to complement theoretical reflections in PHIR.

Strengths and limitations

This study represents the first systematic methodological examination of how multiple outcomes are interpreted to determine DHI success. By comprehensively reviewing all published materials related to eligible interventions and reporting on each measured outcome, the research offers novel insights into evaluation practices and methodological challenges often overlooked when considering articles individually.

Several limitations must be acknowledged. The research team's assessment of intervention success was based solely on outcomes reported in peer-reviewed publications. This may be considered a conservative analytical approach that does not account for effect sizes, clinical significance, or partial improvements of overarching behaviours. In practice, authors would typically consider these elements alongside unpublished data, contextual knowledge, and professional expertise when evaluating success. However, the chosen analytical approach was deliberately adopted to underscore the gap that may arise in the conclusion-drawing processes between authors and readers in the absence of transparent interpretation procedures. This conservative methodology was also chosen for our sub-analysis to demonstrate how the analytical approach influences interpretation, rather than to provide definitive assessments of intervention success. The limited sample size and exclusive focus on individual-level DHIs may restrict generalisability, particularly as PHIR often evaluates community-level

interventions. However, as the study focuses on methodological decisions rather than the health implications of results, the findings likely reflect key challenges in evaluating multiple outcomes. The proposed methodological improvements are expected to be relevant to other health promotion interventions, especially given that identifying a unique, widely accepted primary outcome in population-wide studies is likely to be even more challenging than in individual-level interventions.

The impact of outcome hierarchical position on conclusions was explored using only effectiveness outcomes, excluding process, economic, qualitative, and non-statistically tested outcomes. While it simplified the demonstration of how methodological choices influence conclusions, this approach did not capture the complexity of integrating quantitative and qualitative outcomes, warranting further investigation.

Finally, this study focused solely on outcomes, while other methodological aspects such as research programme design, study designs, and comparators also influence intervention success determination. Future analysis should explore these to complement existing PHIR literature.

Conclusions

This study critically examines the methodological processes underpinning outcome selection, analysis, and interpretation in determining the success of complex digital health interventions. Findings reveal that limited transparency in outcome prioritisation and inadequate reporting of interpretative processes may undermine the trustworthiness and reproducibility of conclusions regarding success.

To improve evaluation practices, we propose (1) developing core outcome sets specific to PHIR, (2) collaboratively selecting multidimensional outcomes through a steering committee that accounts for stakeholder preferences and existing theoretical models, (3) exploring MCDA and consensus-driven methods to transparently combine outcomes, with biostatisticians tailoring these methods to PHIR, (4) enhancing methodological reporting at each stage of intervention development and evaluation to improve scientific integrity and reproducibility, and (5) increasing the involvement of PHIR experts in ethics, funding, and evaluation committees to improve recognition of evidence produced in this field.

These methodological proposals aim to ensure robust, transparent evaluations that can inform scalable interventions, ultimately strengthening evidence-based population health.

Abbreviations

AYA	Adolescents and young adults
COMET	Core Outcome Measures in Effectiveness Trials
CONSORT	Consolidated Standards of Reporting Trials
DALY	Disability-adjusted life year

DHI	Digital health intervention
EBM	Evidence-based medicine
HIV	Human immunodeficiency virus
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
ICUR	Incremental cost-utility ratio
IMB	Information-motivation-behavioural
IT	Information technology
LGB+	Lesbian, gay, bisexual, and other sexual minority
MCDA	Multiple criteria decision analysis
MSM	Men who have sex with men
PHIR	Population health intervention research
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomised controlled trial
SMS	Short message service
SRH	Sexual and reproductive health
STIs	Sexually transmitted infections
UK MRC	United Kingdom Medical Research Council
WHO	World Health Organisation

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-025-04245-1>.

Additional file 1: Table S1. PRISMA checklist for report. Table S2. PRISMA checklist for abstract.

Additional file 2: Search strategies.

Additional file 3: Data extraction form.

Additional file 4: Table S3. Rationale for weights' assignment to outcomes.

Additional file 5: List of references included in the review, by intervention.

Additional file 6: Table S4. Outcomes' characteristics. Table S5. Outcomes' measurement methods.

Additional file 7: Figure S1. Role of outcomes in conclusions.

Additional file 8: Table S6. Inter-rater agreement on conclusions. Table S7. Inter-rater agreement on conclusions by intervention categorisation.

Additional file 9: Table S8. Outcomes assessment in the sub-analysis.

Additional file 10: Fig. S2. Risk-of-bias assessment (ROB-2) for randomised controlled trials. Fig. S3. Risk-of-bias assessment (ROBINS-I) for non-randomised studies.

Acknowledgements

The authors thank Bruno Giraudeau for his valuable and insightful comments on this paper.

Authors' contributions

CC, ELR and CA conceptualised and designed the study. CC wrote the study protocol. CC, ELR and PM created the search strategy. CC and CE screened abstracts and full texts and extracted data. ELR resolved disagreements. CC conducted the formal analysis with contributions from PM, ELR and CA for the visualisation of results and contributions from MM, ELR, and CA for data interpretation. CC drafted the original manuscript, which was reviewed and edited by CE, PM, MM, ELR and CA. ELR and CA supervised the research. All authors read and approved the final manuscript. CC and CA are the guarantors for the study.

Funding

The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors. CC is supported by a doctoral contract funded by the École des Hautes Études en Santé Publique (EHESP).

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Université Paris Cité, Inserm, ECEVE, Paris F-75010, France. ²AP-HP, Nord-Université Paris Cité, Inserm, Hôpital Universitaire Robert Debré, Clinical Epidemiology Unit, CIC 1426, Paris F-75019, France. ³Institut National d'Études Démographiques (INED), UR14 - Sexual and Reproductive Health and Rights, Aubervilliers F-93300, France. ⁴University Center for Adolescent and Young Adult Health, Fondation Santé des Étudiants de France, Paris F-75014, France.

Received: 5 March 2025 Accepted: 23 June 2025

Published online: 02 July 2025

References

- Fletcher RH, Fletcher SW, Wagner EH. *Clinical epidemiology*. Philadelphia: Lippincott Williams & Wilkins; 1996.
- Schulz KF, Altman DG, Moher D, the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med*. 2010;8(1):18.
- Sackett DL. Evidence-based medicine. *Semin Perinatol*. 1997;21(1):3–5.
- Hawe P, Potvin L. What is population health intervention research? *Can J Public Health Rev Can Sante Publique*. 2009;100(1):Suppl 18–14.
- Hawe P. Lessons from complex interventions to improve health. *Annu Rev Public Health*. 2015;36:307–23.
- Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*. 2021;374:2061.
- Alla F, Cambon L, Ridde V. *Population health intervention research. Concepts, methods, applications*. IRD Editions; 2023.
- Stufflebeam D. *Evaluation models: new directions for evaluation*. All Books and Monographs by WMU Authors. San Francisco: Jossey-Bass; 2001.
- Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: the perils of multiple testing. *Perspect Clin Res*. 2016;7(2):106–7.
- Vickerstaff V, Omar RZ, Ambler G. Methods to adjust for multiple comparisons in the analysis and sample size calculation of randomised controlled trials with multiple primary outcomes. *BMC Med Res Methodol*. 2019;19(1):129.
- Pocock SJ. Clinical trials with multiple outcomes: a statistical perspective on their design, analysis, and interpretation. *Control Clin Trials*. 1997;18(6):530–45.
- Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med*. 2016;51(5):843–51.
- Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and Evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *J Med Internet Res*. 2017;19(6):e7126.
- Barbosa W, Zhou K, Waddell E, Myers T, Dorsey ER. Improving access to care: telemedicine across medical domains. *Annu Rev Public Health*. 2021;42:463–81.
- Wong CA, Madanay F, Ozer EM, Harris SK, Moore M, Master SO, et al. Digital health technology to enhance adolescent and young adult clinical preventive services: affordances and challenges. *J Adolesc Health*. 2020;67(2, Supplement):S24–33.
- Rouleau G, Wu K, Ramamoorthi K, Boxall C, Liu RH, Maloney S, et al. Mapping theories, models, and frameworks to evaluate digital health interventions: scoping review. *J Med Internet Res*. 2024;26(1):e51098.
- Segur-Ferrer J, Moltó-Puigmartí C, Pastells-Peiró R, Vivanco-Hidalgo RM. Methodological frameworks and dimensions to be considered in digital health technology assessment: scoping review and thematic analysis. *J Med Internet Res*. 2024;26:e48694.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
- West PR, Michie PS. A guide to development and evaluation of digital behaviour change interventions in healthcare. Sutton: Silverback Publishing; 2016. 82 p.
- Yardley L, Choudhury T, Patrick K, Michie S. Current issues and future directions for research into digital behavior change interventions. *Am J Prev Med*. 2016;51(5):814–5.
- Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing models and theories for digital behavior change interventions. *Am J Prev Med*. 2016;51(5):825–32.
- Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med*. 2016;51(5):833–42.
- McNamee P, Murray E, Kelly MP, Bojke L, Chilcott J, Fischer A, et al. Designing and undertaking a health economics study of digital health interventions. *Am J Prev Med*. 2016;51(5):852–60.
- Patrick K, Hekler EB, Estrin D, Mohr DC, Riper H, Crane D, et al. The pace of technologic change: implications for digital health behavior intervention research. *Am J Prev Med*. 2016;51(5):816–24.
- Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ*. 2013;346:e7586.
- Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database — update and key issues. *N Engl J Med*. 2011;364(9):852–60.
- Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
- Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919.
- Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. Making psychological theory useful for implementing evidence based practice: a consensus approach. *BMJ Qual Saf*. 2005;14(1):26–33.
- Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015;350:1258.
- Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*. 2011;38(2):65–76.
- Reupert A, Maybery D, Bartholomew C, Cuff R, Foster K, Matar J, et al. The acceptability and effectiveness of an online intervention for youth with parents with a mental illness and/or substance use issue. *J Adolesc Health*. 2020;66(5):551–8.
- Boutron I, Ravaud P. Misrepresentation and distortion of research in biomedical literature. *Proc Natl Acad Sci*. 2018;115(11):2613–9.
- Baltussen R, Marsh K, Thokala P, Diaby V, Castro H, Cleemput I, et al. Multicriteria decision analysis to support health technology assessment agencies: benefits, limitations, and the way forward. *Value Health*. 2019;22(11):1283–8.
- Desrosiers A, Betancourt T, Kergoat Y, Servilli C, Say L, Kobeissi L. A systematic review of sexual and reproductive health interventions for young people in humanitarian and lower-and-middle-income country settings. *BMC Public Health*. 2020;20(1):666.
- Nuwamanya E, Nuwasiima A, Babigumira JU, Asiimwe FT, Lubinga SJ, Babigumira JB. Study protocol: using a mobile phone-based application to increase awareness and uptake of sexual and reproductive health services among the youth in Uganda. *A randomized controlled trial Reprod Health*. 2018;15(1):216.
- Nuwamanya E, Nalwanga R, Nuwasiima A, Babigumira JU, Asiimwe FT, Babigumira JB, et al. Effectiveness of a mobile phone application to increase access to sexual and reproductive health information, goods, and services among university students in Uganda: a randomized controlled trial. *Contracept Reprod Med*. 2020;5(1):31.

38. Nalwanga R, Nuwamanya E, Nuwasiima A, Babigumira JU, Asimwe FT, Babigumira JB. Utilization of a mobile phone application to increase access to sexual and reproductive health information, goods, and services among university students in Uganda. *Reprod Health*. 2021;18(1):95.
39. Riordan BC, Conner TS, Flett JA, Scarf D. A brief orientation week ecological momentary intervention to reduce university student alcohol consumption. *J Stud Alcohol Drugs*. 2015;76(4):525–9.
40. Riordan BC, Conner TS, Flett JA, Scarf D. A text message intervention to reduce first year university students' alcohol use: a pilot experimental study. *Digit Health*. 2017;3:2055207617707627.
41. Riordan BC, Moradi S, Carey KB, Conner TS, Jang K, Reid KE, et al. Effectiveness of a combined web-based and ecological momentary intervention for incoming first-year university students: protocol for a 3-arm randomized controlled trial. *JMIR Res Protoc*. 2018;7(5): e10164.
42. Riordan BC, Winter T, Carey KB, Conner TS, Moradi S, Jang K, et al. A combined web based intervention and ecological momentary intervention for reducing alcohol use among incoming first-year university students: results from a three-arm randomised controlled trial. *Addict Behav*. 2023;136: 107471.
43. Piatkowski C, Faulkner GE, Guhn M, Mässe LC. User characteristics and parenting practices associated with adolescents' initial use of a lifestyle behavior modification intervention. *Child Obes*. 2020;16(6):367–78.
44. Lin Y, Mässe LC. A look at engagement profiles and behavior change: a profile analysis examining engagement with the Aim2Be lifestyle behavior modification app for teens and their families. *Prev Med Rep*. 2021;24: 101565.
45. Mässe LC, Vlaar J, Macdonald J, Bradbury J, Warshawski T, Buckler EJ, et al. Aim2Be mHealth intervention for children with overweight and obesity: study protocol for a randomized controlled trial. *Trials*. 2020;21(1):132.
46. Tugault-Lafleur CN, De-Jongh González O, Macdonald J, Bradbury J, Warshawski T, Ball GDC, et al. Efficacy of the Aim2Be intervention in changing lifestyle behaviors among adolescents with overweight and obesity: randomized controlled trial. *J Med Internet Res*. 2023;25: e38545.
47. De-Jongh González O, Tugault-Lafleur CN, Buckler EJ, Hamilton J, Ho J, Buchholz A, et al. The Aim2Be mHealth intervention for children with overweight or obesity and their parents: person-centered analyses to uncover digital phenotypes. *J Med Internet Res*. 2022;24(6): e35285.
48. Deslippe AL, González ODJ, Buckler EJ, Ball GDC, Ho J, Buchholz A, et al. Do individual characteristics and social support increase children's use of an mhealth intervention? Findings from the evaluation of a behavior change MHealth App, Aim2Be. *Child Obes*. 2023;19(7):435–442.
49. Buckler EJ, González ODJ, Ball GDC, Hamilton J, Ho J, Morrison KM, et al. Recruiting families using social media versus pediatric obesity clinics: a secondary analysis of the Aim2Be RCT. *Contemp Clin Trials*. 2023;133: 107322.
50. Lima-Serrano M, Martínez-Montilla JM, Lima-Rodríguez JS, Mercken L, de Vries H. Design, implementation and evaluation of a web-based computer-tailored intervention to prevent binge drinking in adolescents: study protocol. *BMC Public Health*. 2018;18(1):449.
51. Vargas-Martínez AM, Trapero-Bertran M, Lima-Serrano M, Anokye N, Pokhrel S, Mora T. Measuring the effects on quality of life and alcohol consumption of a program to reduce binge drinking in Spanish adolescents. *Drug Alcohol Depend*. 2019;205: 107597.
52. Martínez-Montilla JM, Mercken L, de Vries H, Candel M, Lima-Rodríguez JS, Lima-Serrano M. A web-based, computer-tailored intervention to reduce alcohol consumption and binge drinking among Spanish adolescents: cluster randomized controlled trial. *J Med Internet Res*. 2020;22(1): e15438.
53. Vargas-Martínez AM, Lima-Serrano M, Trapero-Bertran M. Cost-effectiveness and cost-utility analyses of a web-based computer-tailored intervention for prevention of binge drinking among Spanish adolescents. *Alcohol Clin Exp Res*. 2023;47(2):319–35.
54. Gonsalves L, L'Engle KL, Tamrat T, Plourde KF, Mangone ER, Agarwal S, et al. Adolescent/Youth Reproductive Mobile Access and Delivery Initiative for Love and Life Outcomes (ARMADILLO) Study: formative protocol for mHealth platform development and piloting. *Reprod Health*. 2015;12:67.
55. Guerrero F, Lucar N, Garvich Claux M, Chiappe M, Perez-Lu J, Hindin MJ, et al. Developing an SMS text message intervention on sexual and reproductive health with adolescents and youth in Peru. *Reprod Health*. 2020;17(1):116.
56. Gonsalves L, Hindin MJ, Bayer A, Carcamo CP, Gichangi P, Habib N, et al. Protocol of an open, three-arm, individually randomized trial assessing the effect of delivering sexual and reproductive health information to young people (aged 13–24) in Kenya and Peru via mobile phones: adolescent/youth reproductive mobile access and delivery initiative for love and life outcomes (ARMADILLO) study stage 2. *Reprod Health*. 2018;15(1):126.
57. Perez-Lu JE, Guerrero F, Cárcamo CP, Alburqueque M, Chiappe M, Hindin MJ, et al. The ARMADILLO text message intervention to improve the sexual and reproductive health knowledge of adolescents in Peru: results of a randomized controlled trial. *PLoS ONE*. 2022;17(2): e0262986.
58. Gichangi P, Gonsalves L, Mwaisaka J, Thiongo M, Habib N, Waithaka M, et al. Busting contraception myths and misconceptions among youth in Kwale County, Kenya: results of a digital health randomised control trial. *BMJ Open*. 2022;12(1): e047426.
59. Gonsalves L, Njeri WW, Schroeder M, Mwaisaka J, Gichangi P. Research and implementation lessons learned from a youth-targeted digital health randomized controlled trial (the ARMADILLO Study). *JMIR MHealth UHealth*. 2019;7(8): e13005.
60. Mwaisaka J, Gonsalves L, Thiongo M, Waithaka M, Sidha H, Alfred O, et al. Young people's experiences using an on-demand mobile health sexual and reproductive health text message intervention in Kenya: qualitative study. *JMIR MHealth UHealth*. 2021;9(1): e19109.
61. Rouf A, Allman-Farinelli M. Messaging for interventions aiming to improve calcium intake in young adults—a mixed methods study. *Nutrients*. 2018;10(11):1673.
62. Rouf A, Nour M, Allman-Farinelli M. Improving calcium knowledge and intake in young adults via social media and text messages: randomized controlled trial. *JMIR MHealth UHealth*. 2020;8(2): e16499.
63. Stephenson J, Bailey JV, Blandford A, Brima N, Copas A, D'Souza P, et al. An interactive website to aid young women's choice of contraception: feasibility and efficacy RCT. *Health Technol Assess*. 2020;24(56):1–44.
64. Stephenson J, Bailey JV, Gubijev A, D'Souza P, Oliver S, Blandford A, et al. An interactive website for informed contraception choice: randomised evaluation of Contraception Choices. *Digit Health*. 2020;6:2055207620936435.
65. Bailey JV, Bennett KF, Gubijev A, Shawe J, Stephenson J. Participant views and experiences of sexual health research: the Contraception Choices online trial. *Digit Health*. 2021;7:20552076211033424.
66. Nolan C, Packel L, Hope R, Levine J, Baringer L, Gatare E, et al. Design and impact evaluation of a digital reproductive health program in Rwanda using a cluster randomized design: study protocol. *BMC Public Health*. 2020;20(1):1701.
67. Ippoliti N, Sekamana M, Baringer L, Hope R. Using human-centered design to develop, launch, and evaluate a national digital health platform to improve reproductive health for Rwandan youth. *Glob Health Sci Pract*. 2021;9(Suppl 2):S244–60.
68. Hémono R, Packel L, Gatare E, Baringer L, Ippoliti N, McCoy SI, et al. Digital self-care for improved access to family planning and reproductive health services among adolescents in Rwanda: preliminary findings from a pilot study of CyberRwanda. *Sex Reprod Health Matters*. 2022;29(3):2110671.
69. Hémono R, Gatare E, Kayitesi L, Packel L, Hunter LA, Kunesh J, et al. CyberRwanda's pathway to impact: results from a cluster-randomized trial of adolescent family planning knowledge, beliefs, self-efficacy, and behavior. *J Adolesc Health*. 2024;74(6):1239–48.
70. Dumas DM. Web-based personalized feedback: is this an appropriate approach for reducing drinking among high school students? *J Subst Abuse Treat*. 2015;50:76–80.
71. Dumas DM, Esp S, Johnson J, Trull R, Shearer K. The eCHECKUP TO GO for high school: impact on risk factors and protective behavioral strategies for alcohol use. *Addict Behav*. 2017;64:93–100.
72. Dumas DM, Esp S, Flay B, Bond L. A randomized controlled trial testing the efficacy of a brief online alcohol intervention for high school seniors. *J Stud Alcohol Drugs*. 2017;78(5):706–15.

73. Doumas DM, Esp S. Reducing alcohol-related consequences among high school seniors: efficacy of a brief, web-based intervention. *J Couns Dev*. 2019;97(1):53–61.
74. Doumas DM, Esp S, Turrissi R, Bond L. A randomized controlled trial of the eCHECKUP to GO for high school seniors across the academic year. *Subst Use Misuse*. 2021;56(13):1923–32.
75. Doumas DM, Esp S, Turrissi R, Bond L, Porchia S, Flay B. Sex differences in the acceptability and short-term outcomes of a web-based personalized feedback alcohol intervention for high school seniors. *Psychol Sch*. 2020;57(11):1724–40.
76. Doumas DM, Esp S, Turrissi R, Bond L, Flay B. Efficacy of the eCHECKUP TO GO for high school seniors: sex differences in risk factors, protective behavioral strategies, and alcohol use. *J Stud Alcohol Drugs*. 2020;81(2):135–43.
77. Parisod H, Pakarinen A, Axelin A, Danielsson-Ojala R, Smed J, Salanterä S. Designing a health-game intervention supporting health literacy and a tobacco-free life in early adolescence. *Games Health J*. 2017;6(4):187–99.
78. Parisod H, Pakarinen A, Axelin A, Löytyniemi E, Smed J, Salanterä S. Feasibility of mobile health game “Fume” in supporting tobacco-related health literacy among early adolescents: a three-armed cluster randomized design. *Int J Med Inf*. 2018;113:26–37.
79. Pinto DL, Parisod H, Nyman J, Barroso TMM de A. Effectiveness of the Portuguese version of Fume in adolescents’ health literacy about tobacco. *Rev Lat Am Enfermagem*. 2022;30:e3513.
80. Ybarra ML, Price-Feeney M, Prescott T, Goodenow C, Saewyc E, Rosario M. Girl2Girl: How to develop a salient pregnancy prevention program for cisgender sexual minority adolescent girls. *J Adolesc*. 2020;85(1):41–58.
81. Ybarra M, Goodenow C, Rosario M, Saewyc E, Prescott T. An mHealth intervention for pregnancy prevention for LGB teens: an RCT. *Pediatrics*. 2021;147(3):e2020013607.
82. Ybarra M, Rosario M, Saewyc E, Goodenow C, Dunsiger S. One-year follow-up after a pregnancy prevention intervention for LGB1 teens: an RCT. *Pediatrics*. 2023;151(4):e2022059172.
83. Ybarra ML, Saewyc E, Rosario M, Dunsiger S. Subgroup analyses of Girl2Girl, a text messaging-based teen pregnancy prevention program for sexual minority girls: results from a national RCT. *Prev Sci*. 2023;24(2):292–9.
84. Yount KM, Minh TH, Trang QT, Cheong YF, Bergenfeld I, Sales JM. Preventing sexual violence in college men: a randomized-controlled trial of GlobalConsent. *BMC Public Health*. 2020;20(1):1331.
85. Yount KM, Cheong YF, Bergenfeld I, Trang QT, Sales JM, Li Y, et al. Impacts of GlobalConsent, a web-based social norms edutainment program, on sexually violent behavior and bystander behavior among university men in Vietnam: randomized controlled trial. *JMIR Public Health Surveill*. 2023;9:e35116.
86. Yount KM, Bergenfeld I, Anderson KM, Trang QT, Sales JM, Cheong YF, et al. Theoretical mediators of GlobalConsent: an adapted web-based sexual violence prevention program for university men in Vietnam. *Soc Sci Med*. 2022;313:115402.
87. Yount KM, Anderson KM, Trang QT, Bergenfeld I. Preventing sexual violence in Vietnam: qualitative findings from high school, university, and civil society key informants across regions. *BMC Public Health*. 2023;23(1):1114.
88. Javidi H, Widman L, Lipsey N, Brasileiro J, Javidi F, Jhala A. Redeveloping a digital sexual health intervention for adolescents to allow for broader dissemination: implications for HIV and STD prevention. *AIDS Educ Prev*. 2021;33(2):89–102.
89. McCrimmon J, Widman L, Javidi H, Brasileiro J, Hurst J. Evaluation of a brief online sexual health program for adolescents: a randomized controlled trial. *Health Promot Pract*. 2024 Jul;25(4):689–97.
90. Tay JL, Goh YSS, Klainin-Yobas P. Online HOPE intervention on mental health literacy among youths in Singapore: an RCT protocol. *J Adv Nurs*. 2020;76(7):1831–9.
91. Tay JL, Goh YSS, Sim K, Klainin-Yobas P. Impact of the HOPE intervention on mental health literacy, psychological well-being and stress levels amongst university undergraduates: a randomised controlled trial. *Int J Environ Res Public Health*. 2022;19(15):9772.
92. Tay JL. Online HOPE intervention on help-seeking attitudes and intentions among young adults in Singapore: a randomized controlled trial and process evaluation. *Arch Psychiatr Nurs*. 2022;41:286–94.
93. Cohen KA, Graham AK, Lattie EG. Aligning students and counseling centers on student mental health needs and treatment resources. *J Am Coll Health*. 2022;70(3):724–32.
94. Lattie E, Cohen KA, Winquist N, Mohr DC. Examining an app-based mental health self-care program, intellicare for college students: single-arm pilot study. *JMIR Ment Health*. 2020;7(10):e21075.
95. Lattie EG, Cohen KA, Hersch E, Williams KDA, Krusan KP, MacIver C, et al. Uptake and effectiveness of a self-guided mobile app platform for college student mental health. *Internet Interv*. 2022;27:100493.
96. Scull T, Malik C, Morrison A, Keefe E. Promoting sexual health in high school: a feasibility study of a web-based media literacy education program. *J Health Commun*. 2021;26(3):147–60.
97. Scull TM, Malik CV, Morrison A, Keefe EM. Study protocol for a randomized controlled trial to evaluate a web-based comprehensive sexual health and media literacy education program for high school students. *Trials*. 2020;21(1):50.
98. Scull TM, Dodson CV, Geller JG, Reeder LC, Stump KN. A media literacy education approach to high school sexual health education: immediate effects of media aware on adolescents’ media, sexual health, and communication outcomes. *J Youth Adolesc*. 2022;51(4):708–23.
99. Matar JL, Maybery DJ, McLean LA, Reupert A. Web-based health intervention for young people who have a parent with a mental illness: Delphi study among potential future users. *J Med Internet Res*. 2018;20(10):e10158.
100. Reupert A, Bartholomew C, Cuff R, Foster K, Matar J, Maybery DJ, et al. An online intervention to promote mental health and wellbeing for young adults whose parents have mental illness and/or substance use problems: theoretical basis and intervention description. *Front Psychiatry*. 2019;10:59.
101. Maybery D, Reupert A, Bartholomew C, Cuff R, Duncan Z, Foster K, et al. A web-based intervention for young adults whose parents have a mental illness or substance use concern: protocol for a randomized controlled trial. *JMIR Res Protoc*. 2020;9(6):e15626.
102. Maybery D, Reupert A, Bartholomew C, Cuff R, Duncan Z, McAuliffe C, et al. An online intervention for 18–25-year-old youth whose parents have a mental illness and/or substance use disorder: A pilot randomized controlled trial. *Early Interv Psychiatry*. 2022;16(11):1249–58.
103. Schnell R, Kuhns LM, Hidalgo MA, Powell D, Thai J, Hirshfield S, et al. Adaptation of a group-based HIV RISK reduction intervention to a mobile app for young sexual minority men. *AIDS Educ Prev*. 2018;30(6):449–62.
104. Cho H, Powell D, Pichon A, Thai J, Bruce J, Kuhns LM, et al. A mobile health intervention for HIV prevention among racially and ethnically diverse young men: usability evaluation. *JMIR MHealth UHealth*. 2018;6(9):e11450.
105. Ignacio M, Garofalo R, Pearson C, Kuhns LM, Bruce J, Scott Batey D, et al. Pilot feasibility trial of the MyPEEPS mobile app to reduce sexual risk among young men in 4 cities. *JAMIA Open*. 2019;2(2):272–9.
106. Kuhns LM, Garofalo R, Hidalgo M, Hirshfield S, Pearson C, Bruce J, et al. A randomized controlled efficacy trial of an mHealth HIV prevention intervention for sexual minority young men: MyPEEPS mobile study protocol. *BMC Public Health*. 2020;20(1):65.
107. Schnell R, Kuhns LM, Pearson C, Batey DS, Bruce J, Hidalgo MA, et al. Efficacy of MyPEEPS mobile, an HIV prevention intervention using mobile technology, on reducing sexual risk among same-sex attracted adolescent males: a randomized clinical trial. *JAMA Netw Open*. 2022;5(9):e2231853.
108. Cordoba E, Idnay B, Garofalo R, Kuhns LM, Pearson C, Bruce J, et al. Examining the Information Systems Success (ISS) of a mobile sexual health app (MyPEEPS Mobile) from the perspective of very young men who have sex with men (YMSM). *Int J Med Inf*. 2021;153:104529.
109. Anderson A, Karczmar A, Kuhns LM, Garofalo R, Radix A, Bruce J, et al. A qualitative study to inform adaptation of MyPEEPS mobile for transmasculine youth. *J Health Care Poor Underserved*. 2022;33(1):301–16.
110. Reiter PL, Katz ML, Bauermeister JA, Shoben AB, Paskett ED, McRee AL. Increasing human papillomavirus vaccination among young gay and bisexual men: a randomized pilot trial of the outsmart HPV intervention. *LGBT Health*. 2018;5(5):325–9.

111. McRee AL, Shoben A, Bauermeister JA, Katz ML, Paskett ED, Reiter PL. Outsmart HPV: acceptability and short-term effects of a web-based HPV vaccination intervention for young adult gay and bisexual men. *Vaccine*. 2018;36(52):8158–64.
112. Reiter PL, Gower AL, Kiss DE, Malone MA, Katz ML, Bauermeister JA, et al. A web-based human papillomavirus vaccination intervention for young gay, bisexual, and other men who have sex with men: protocol for a randomized controlled trial. *JMIR Res Protoc*. 2020;9(2):e16294.
113. Reiter PL, Gower AL, Kiss DE, Shoben AB, Katz ML, Bauermeister JA, et al. Effects of a web-based HPV vaccination intervention on cognitive outcomes among young gay, bisexual, and other men who have sex with men. *Hum Vaccines Immunother*. 2022;18(6):2114261.
114. Reiter PL, Gower AL, Kiss DE, Shoben AB, Katz ML, Bauermeister JA, et al. Efficacy of the outsmart HPV intervention: a randomized controlled trial to increase HPV vaccination among young gay, bisexual, and other men who have sex with men. *Cancer Epidemiol Biomark Prev*. 2023;32(6):760–7.
115. Javidi H, Widman L, Maheux AJ, McCrimmon J, Evans-Paulson R, Becker W. PACT: developing and evaluating a digital sexual consent program for youth. *J Sex Res*. 2024;61(3):466–80.
116. Pendergrass T, Hieftje K, Duncan L, et al. Videogame intervention to encourage HIV testing and counseling among adolescents. *Mhealth*. 2020;6:26.
117. Boomer TP, Larkin K, Duncan LR, Fernandes CSF, Fiellin LE. A serious video game targeting HIV testing and counseling: a randomized controlled trial. *J Adolesc Health*. 2024;74(2):252–9.
118. Moore SK, Grabinski M, Bessen S, Borodovsky JT, Marsch LA. Web-based prescription opioid abuse prevention for adolescents: program development and formative evaluation. *JMIR Form Res*. 2019;3(3):e12389.
119. Marsch LA, Moore SK, Grabinski M, Bessen SY, Borodovsky J, Scherer E. Evaluating the effectiveness of a web-based program (POP4Teens) to prevent prescription opioid misuse among adolescents: randomized controlled trial. *JMIR Public Health Surveill*. 2021;7(2):e18487.
120. Ray AE, Greene K, Hecht ML, Barriage SC, Miller-Day M, Glenn SD, et al. An E-learning adaptation of an evidence-based media literacy curriculum to prevent youth substance use in community groups: development and feasibility of REAL media. *JMIR Form Res*. 2019;3(2):e12132.
121. Ray AE, Greene K, Pristavec T, Hecht ML, Miller-Day MA, Banerjee SC. Exploring indicators of engagement in online learning as applied to adolescent health prevention: a pilot study of REAL media. *Educ Technol Res Dev*. 2020;68(6):3143–63.
122. Greene K, Ray AE, Choi HJ, Glenn SD, Lyons RE, Hecht ML. Short term effects of the REAL media e-learning media literacy substance prevention curriculum: an RCT of adolescents disseminated through a community organization. *Drug Alcohol Depend*. 2020;214:108170.
123. Peña-Alves S, Greene K, Ray AE, Glenn SD, Hecht ML, Banerjee SC. 'Choose today, live tomorrow': a content analysis of anti-substance use messages produced by adolescents. *J Health Commun*. 2019;24(6):592–602.
124. Greene K, Choi HJ, Glenn SD, Ray AE, Hecht ML. The Role of engagement in effective, digital prevention interventions: the function of engagement in the REAL media substance use prevention curriculum. *Prev Sci Off J Soc Prev Res*. 2021;22(2):247–58.
125. Li DH, Moskowitz DA, Macapagal K, Saber R, Mustanski B. Using intervention mapping to developmentally adapt an online HIV risk reduction program for adolescent men who have sex with men. *Prev Sci Off J Soc Prev Res*. 2020;21(7):885–97.
126. Ventuneac A, Li DH, Mongrella MC, Moskowitz DA, Weingardt KR, Brown CH, et al. Exploring potential implementation barriers and facilitators of the SMART Program, a stepped-care package of eHealth HIV prevention interventions for adolescent men who have sex with men. *Sex Res Soc Policy*. 2020;17(3):378–88.
127. Mustanski B, Moskowitz DA, Moran KO, Newcomb ME, Macapagal K, Rodriguez-Diaz C, et al. Evaluation of a stepped-care eHealth HIV prevention program for diverse adolescent men who have sex with men: protocol for a hybrid type 1 effectiveness implementation trial of SMART. *JMIR Res Protoc*. 2020;9(8):e19701.
128. Saber R, Mongrella M, Matson M, Carty-Fickes E, Novotny J, Pirog S, et al. Use of a custom testing center locator tool to improve STI and HIV testing rates in adolescent men who have sex with men as part of an online sexual health program. *Procedia Comput Sci*. 2022;206:92–100.
129. Mustanski B, Saber R, Macapagal K, Matson M, Laber E, Rodriguez-Diaz C, et al. Effectiveness of the SMART Sex Ed program among 13–18 year old English and Spanish speaking adolescent men who have sex with men. *AIDS Behav*. 2023;27(2):733–44.
130. Sinno J, Macapagal K, Mustanski B. Social media and online dating safety practices by adolescent sexual and gender diverse men: mixed-methods findings from the SMART study. *J Adolesc Health*. 2024;74(1):113–22.
131. Matheson EL, Smith HG, Amaral ACS, Meireles JFF, Almeida MC, Mora G, et al. Improving body image at scale among Brazilian adolescents: study protocol for the co-creation and randomised trial evaluation of a chatbot intervention. *BMC Public Health*. 2021;21(1):2135.
132. Matheson EL, Smith HG, Amaral ACS, Meireles JFF, Almeida MC, Linardon J, et al. Using chatbot technology to improve Brazilian adolescents' body image and mental health at scale: randomized controlled trial. *JMIR MHealth UHealth*. 2023;11(1):e39934.
133. Garbett KM, Craddock N, Haywood S, Nasution K, White P, Saraswati LA, et al. A novel, scalable social media-based intervention ("Warna-Warni Waktu") to reduce body dissatisfaction among young Indonesian women: protocol for a parallel randomized controlled trial. *JMIR Res Protoc*. 2022;11(1):e33596.
134. Garbett KM, Haywood S, Craddock N, Gentili C, Nasution K, Saraswati LA, et al. Evaluating the Efficacy of a social media-based intervention (Warna-Warni Waktu) to improve body image among young Indonesian women: parallel randomized controlled trial. *J Med Internet Res*. 2023;25(1):e42499.
135. Pocock SJ, Rossello X, Owen R, Collier TJ, Stone GW, Rockhold FW. Primary and secondary outcome reporting in randomized trials: JACC state-of-the-art review. *J Am Coll Cardiol*. 2021;78(8):827–39.
136. Kracht CL, Hutchesson M, Ahmed M, Müller AM, Ashton LM, Brown HM, et al. E-health interventions targeting nutrition, physical activity, sedentary behavior and/or obesity amongst children: a scoping review of systematic reviews and meta-analyses. *Obes Rev Off J Int Assoc Study Obes*. 2021;22(12):e13331.
137. Cambon L, Alla F. Current challenges in population health intervention research. *J Epidemiol Community Health*. 2019;73(11):990–2.
138. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ*. 2016;352:i1174.
139. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. *Health Technol Assess*. 2021;25(57):1–132.
140. Wyer PC, Silva SA. Where is the wisdom? I – a conceptual history of evidence-based medicine. *J Eval Clin Pract*. 2009;15(6):891–8.
141. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health*. 2004;94(3):400–5.
142. Atkins D. Creating and synthesizing evidence with decision makers in mind: integrating evidence from clinical trials and other study designs. *Med Care*. 2007;45(10 Supl 2):S16–22.
143. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health*. 2012;66(12):1182–6.
144. Klaić M, Kapp S, Hudson P, Chapman W, Denehy L, Story D, et al. Implementability of healthcare interventions: an overview of reviews and development of a conceptual framework. *Implement Sci*. 2022;17(1):10.
145. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. *Trials*. 2017;18(3):280.
146. Moore G, Cambon L, Michie S, Arwidson P, Ninot G, Ferron C, et al. Population health intervention research: the place of theories. *Trials*. 2019;20:285.

147. Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, et al. Multiple Criteria decision analysis for health care decision making—an introduction: report 1 of the ISPOR MCDA emerging good practices task force. *Value Health*. 2016;19(1):1–13.
148. Jones J, Hunter D. Qualitative research: consensus methods for medical and health services research. *BMJ*. 1995;311(7001):376–80.
149. Hawe P. The truth, but not the whole truth? Call for an amnesty on unreported results of public health interventions. *J Epidemiol Community Health*. 2012;66(4):285–285.
150. Leroy JL, Frongillo EA, Kase BE, Alonso S, Chen M, Dohoo I, et al. Strengthening causal inference from randomised controlled trials of complex interventions. *BMJ Glob Health*. 2022;7(6):e008597.
151. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic Physiol Opt*. 2014;34(5):502–8.
152. Li A, Barber RF. Multiple testing with the structure-adaptive Benjamini-Hochberg algorithm. *J R Stat Soc Ser B Stat Methodol*. 2019;81(1):45–74.

Publisher's Note

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