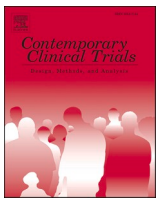
Contemporary Clinical Trials 142 (2024) 107575

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial



Developing an international consensus Reporting guideline for intervention Fidelity in Non-Drug, non-surgical trials: The ReFiND protocol

Luis Fernando Sousa Filho [a,*], Melanie K. Farlie [a], Terry Haines [a], Belinda Borrelli [b], Christopher Carroll [c], Catherine Mathews [d,e], Daniel C. Ribeiro [f], Julie M. Fritz [g], Martin Underwood [h,i], Nadine E. Foster [j,k], Sarah E. Lamb [l], Zila M. Sanchez [m], Peter Malliaras [a]

[a] School of Primary and Allied Health Care, Monash University, Melbourne, Australia
[b] Center for Behavioral Science Research, Department of Health Policy and Health Services Research, Boston University, Henry M. Goldman School of Dental Medicine, Boston, MA, USA
[c] Division of Population Health, University of Sheffield, Sheffield, UK
[d] Health Systems Research Unit, South African Medical Research Council, Cape Town, South Africa
[e] School of Public Health and Family Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa
[f] Centre for Health, Activity and Rehabilitation Research, School of Physiotherapy, University of Otago, Dunedin, New Zealand
[g] Department of Physical Therapy & Athletic Training, University of Utah, Salt Lake City, UT, USA
[h] Warwick Clinical Trials Unit, University of Warwick, Warwick, UK.
[i] University Hospitals of Coventry and Warwickshire, Coventry, UK
[j] STARS Education and Research Alliance, Surgical Treatment and Rehabilitation Service (STARS), The University of Queensland and Metro North Health, Brisbane, Australia
[k] School of Medicine, Keele University, Keele, UK
[l] University of Exeter, St Luke's Campus, Exeter, UK
[m] Department of Preventive Medicine, Universidade Federal de São Paulo (UNIFESP), São Paulo, SP, Brazil

ARTICLE INFO

Keywords:
Fidelity
Research report
Research design
Reproducibility of results
Clinical trials as topic
Program evaluation

ABSTRACT

Background: Inadequate reporting of fidelity to interventions in trials limits the transparency and interpretation of trial findings. Despite this, most trials of non-drug, non-surgical interventions lack comprehensive reporting of fidelity. If fidelity is poorly reported, it is unclear which intervention components were tested or implemented within the trial, which also hinders research reproducibility. This protocol describes the development process of a reporting guideline for fidelity of non-drug, non-surgical interventions (ReFiND) in the context of trials.
Methods: The ReFiND guideline will be developed in six stages. Stage one: a guideline development group has been formed to oversee the guideline methodology. Stage two: a scoping review will be conducted to identify and summarize existing guidance documents on the fidelity of non-drug, non-surgical interventions. Stage three: a Delphi study will be conducted to reach consensus on reporting items. Stage four: a consensus meeting will be held to consolidate the reporting items and discuss the wording and structure of the guideline. Stage five: a guidance statement, an elaboration and explanation document, and a reporting checklist will be developed. Stage six: different strategies will be used to disseminate and implement the ReFiND guideline.
Discussion: The ReFiND guideline will provide a set of items developed through international consensus to improve the reporting of intervention fidelity in trials of non-drug, non-surgical interventions. This reporting guideline will enhance transparency and reproducibility in future non-drug, non-surgical intervention research.

* Corresponding author.
E-mail addresses: fernando.sousa@monash.edu (L.F. Sousa Filho), melanie.farlie@monash.edu (M.K. Farlie), terry.haines@monash.edu (T. Haines), belindab@bu.edu (B. Borrelli), c.carroll@sheffield.ac.uk (C. Carroll), catherine.mathews@mrc.ac.za (C. Mathews), daniel.ribeiro@otago.ac.nz (D.C. Ribeiro), julie.fritz@utah.edu (J.M. Fritz), m.underwood@warwick.ac.uk (M. Underwood), n.foster@uq.edu.au (N.E. Foster), S.E.Lamb@exeter.ac.uk (S.E. Lamb), zila.sanchez@unifesp.br (Z.M. Sanchez), peter.malliaras@monash.edu (P. Malliaras).

https://doi.org/10.1016/j.cct.2024.107575
Received 7 February 2024; Received in revised form 23 April 2024; Accepted 11 May 2024
Available online 13 May 2024

1. Background

Fidelity refers to the extent to which an intervention is implemented in the trial as described and planned in the trial protocol [1,2]. There are multiple definitions of fidelity in the literature, but it commonly involves how the intervention is delivered by intervention providers and how it is received and enacted by trial participants [1,2]. In a trial context, fidelity evaluation is crucial to understanding any effects, positive or negative, or lack thereof, of an intervention [3]. For example, if an intervention appears to be ineffective, or harmful, but has poor intervention fidelity (e.g., core components of the intervention were not delivered as planned), the observed null, or negative, findings may result from inadequate intervention delivery or participant enactment in the trial rather than from the intervention itself. Inadequate or incomplete reporting of fidelity limits the interpretation of trial findings and can lead to potentially effective interventions being discarded or to the implementation of potentially ineffective interventions (i.e., the intervention was found to be effective, but the core components of the intervention were not delivered or were different from the ones planned).

Despite the importance of fidelity in intervention research, both drug and non-drug trials often do not consistently monitor or report intervention fidelity, with some indication that overall intervention reporting is poorer in trials of non-drug interventions [4–7]. In this study, we focus on the challenges of fidelity monitoring and reporting in non-drug, non-surgical interventions, which are likely to be less standardised (i.e., procedures are usually tailored) and more complex (e.g., group delivery, different settings and formats, etc), involving multiple components and requiring greater participant engagement levels.

The poor monitoring and reporting of fidelity components occur across different types of non-drug, non-surgical interventions and healthcare contexts. A 2016 systematic review ($n = 193$ studies) showed that only 7% of studies investigating non-drug, non-surgical interventions for obesity in the paediatric field reported the length of treatment sessions and only 4% reported methods for assessing participants' comprehension of the intervention [8]. For non-drug, non-surgical interventions targeting smoking cessation ($n = 755$ trials), the proportion of trials reporting fidelity of delivery, receipt, and enactment can be as low as 15%, 16%, and 25%, respectively [9]. For exercise interventions ($n = 757$ studies), reporting of fidelity is missing or

Box 1
Glossary of relevant terms used in this study protocol.

Intervention delivery
The degree to which an intervention is delivered to trial participants by intervention providers as planned in the trial protocol.
Intervention receipt
The degree to which trial participants understand and demonstrate the ability to use or replicate intervention components delivered by intervention providers.
Intervention enactment
The degree to which trial participants apply the intervention components in their daily life.
Guidance document
A document that provides actionable guidance on the monitoring of intervention fidelity, irrespective of the methods used to develop the guidance. For example, literature reviews focused on providing actionable recommendations on fidelity can be considered a guidance document.
Reporting guideline
A document that provides a structured set of recommendation items for reporting fidelity, developed through robust, reproducible, and international consensus-based methods.
Conventional clinical or public health trials
Trials aiming to assess whether a clinical or public health intervention works to change participant or population outcomes under ideal conditions [24]. These trials are commonly referred to as 'explanatory' or 'efficacy' trials.
Effectiveness-implementation hybrid trials type I
Trials aiming to assess the effectiveness of a clinical or public health on participant or population outcomes under usual conditions. These trials may also have a secondary aim of exploring the context for implementation [24]. They are commonly referred to as 'pragmatic' or 'effectiveness' trials.
Effectiveness-implementation hybrid trials type II
Trials aiming to assess the effectiveness of a clinical or public health on participant or population outcomes under usual conditions AND to assess the effects of an implementation strategy on implementation outcomes [24].
Effectiveness-implementation hybrid trials type III
Trials aiming to assess whether an implementation strategy works to change implementation outcomes of a clinical or public health intervention. These trials may also have a secondary aim of exploring the participant or population outcomes associated with the implementation of the intervention [24].
Implementation trial
Trials aiming to assess the effects of an implementation strategy on implementation outcomes of a clinical or public health intervention [24].
Drug
A substance, other than food, intended to affect the structure or any function of the body [25]. Not all drugs are medications.
Medication
A drug or medicine used to treat or cure illness.
Therapeutic biological agents or products
Agents or products isolated from natural sources and living materials (e.g., cells or tissues) or produced by biotechnology methods used to treat or cure disease. They are included within the definition of drugs [25].
Process evaluation
Evaluation of the implementation of evidence-based interventions in trials aiming to understand implementation processes (e.g., training, resources, fidelity), mechanisms of impact (e.g., mediators), and contextual factors that can affect intervention outcomes [26,27].

incomplete for 95% of interventions [10].

The reporting of fidelity in trials remains limited despite the existence of a range of 'frameworks', 'guides', and 'practical recommendations' that have been described in the last two decades to guide the monitoring of intervention fidelity [1–3,11–14]. In this protocol, we use the term 'guidance document' (see Glossary in Box 1) to refer to these publications. This term is distinct from 'reporting guideline', which denotes a structured set of recommendation items developed through robust, reproducible, and international consensus-based methods, as proposed in this study's methodology. Many existing guidance documents primarily focus on how to monitor fidelity in trials, limiting attention to fidelity reporting. Furthermore, fidelity definitions, terminology, and concepts vary significantly across these documents, challenging the establishment of standard practices [2,3,11,15,16]. In addition, several of these existing documents do not describe their development methods, are based on researchers' experience, or employed non-reproducible approaches when formulating fidelity recommendations [2,3,11,15,16]. These factors may contribute to explaining why the understanding of fidelity by researchers and the adoption of recommendations from existing fidelity guidance documents are limited [4,17]. Notably, we did not identify any reporting guideline specific to fidelity of non-drug, non-surgical interventions incorporating international consensus building approaches to develop recommendations, as endorsed by the EQUATOR Network [18].

Reporting guidelines for intervention research, such as the template for intervention description and replication (TIDieR) [19], the consensus on exercise reporting template (CERT) [20], and the recommendations for the development, implementation, and reporting of control interventions in efficacy and mechanistic trials of physical, psychological, and self-management therapies (CoPPS) [21], developed with accepted standards for reporting guidelines (e.g., Delphi technique for consensus) [22], include items related to fidelity. However, these guidelines do not cover all aspects of fidelity discussed in the literature [3]. While TIDieR has a few fidelity-related items, a systematic review has indicated that the reporting of fidelity (i.e., fidelity of intervention receipt) in RCTs for complex interventions has not significantly improved since its publication in 2014 [23]. Therefore, researchers have been calling for the development of reporting guidelines for intervention fidelity within non-drug, non-surgical trials [3,5,17]. In this protocol, we aim to detail the development of the reporting guideline for fidelity of non-drug, non-surgical interventions (ReFiND). This study seeks to fill the existing gap in fidelity reporting by developing international consensus on a standardised set of items for reporting fidelity within trials.

2. Methods

2.1. Scope and design

The protocol for developing this reporting guideline is based on the premise that intervention research functions as a continuum, encompassing the early stages of intervention design, specification, delivery, feasibility, acceptability and efficacy testing to the later stages of comparative effectiveness, evaluation, and implementation in practice [28]. Decisions made at one stage of the continuum can influence others. Therefore, this protocol outlines a methodology deliberately planned to achieve our goal of developing fidelity reporting guidelines considering all stages of intervention research and trials, including conventional clinical or public health trials (e.g., efficacy trials), effectiveness-implementation hybrid trials (Types I, II, and III), and implementation trials (i.e., trials assessing implementation outcomes rather than health outcomes), as previously described [24] (see Glossary in Box 1).

The fidelity reporting guideline will be applicable to the reporting of non-drug interventions, which is a general term used to refer to interventions that do not involve the use of any form of drug, medication, or biological agents. It includes behavioral interventions (e.g., lifestyle modifications), physical interventions (e.g., exercise), psychological interventions (e.g., psychotherapy), mind-body interventions (e.g., mindfulness), or some complementary and integrative health (CIH) strategies (e.g., acupuncture). In certain instances, multifaceted interventions may encompass both drug and non-drug components. Nevertheless, the guideline is primarily focused on the non-drug components of such programs. Approaches involving herbal remedies or dietary supplements containing active ingredients likely to promote physiological effects on the body, as well as surgical interventions and homeopathic preparations, are out of the scope of this guideline. Similar to other reporting guidelines for intervention research [19,21,29], ReFiND will be a high-level conceptual guideline. This means it will focus on complex and overarching concepts that can be applied or adapted to different disciplines, rather than concentrating on superficial or discipline-specific tasks.

We registered the ReFiND project with the Open Science Framework registry and EQUATOR Network. The project was approved by the Monash University Human Research Ethics Committee (ID 41579). The ReFiND will be developed and reported according to the EQUATOR Guidance for Developers of Health Research Reporting [18] and the ACCORD guideline, which is a reporting guideline for consensus methods in biomedicine research [30,31]. Accordingly, we will (1) form an international guideline development group, (2) review the literature to identify existing non-drug, non-surgical intervention fidelity guidance and systematically extract fidelity definitions, domains, and recommendations (scoping review is in progress), (3) conduct an online, 3-round international Delphi study to achieve consensus on fidelity reporting items, (4) hold a consensus meeting to discuss the Delphi results, fidelity definition, and the final structure and wording of the guideline, (5) elaborate a guideline statement, an explanation and elaboration document and a reporting checklist, and (6) formulate dissemination and implementation strategies (see Fig. 1).

2.2. Stage (1): Guideline development group (GDG)

After conducting a literature review to confirm the need for a reporting guideline for intervention fidelity in the context of non-drug, non-surgical interventions, the project leadership (LFSF, MKF, TH, PM) started to invite other researchers to join the GDG. We considered the following premises to select potential members: (1) most GDG members should have interest in and experience with fidelity/process evaluation (e.g., publications related to fidelity or process evaluation), (2) at least one GDG member should have participated in the development of reporting guidelines previously, (3) the GDG should be multidisciplinary, (4) the GDG should be composed of members with expertise in the different stages and processes of intervention research (e.g., intervention design, efficacy trials, effectiveness trials, hybrid design trials, and implementation trials), and (5) the GDG should be diverse in gender and geographical location. Based on these variables, LFSF and PM examined the research profiles of authors from relevant publications identified in our literature review. They checked for information on their respective institutional websites or on Google Scholar. Subsequently, LFSF and PM categorized researchers by geographical location and created a list of potential members. They then invited 10 researchers to be part of the GDG, ensuring diversity across geographical location, disciplines, and expertise. Nine researchers accepted to participate in this project.

The role of the ReFiND GDG is to provide informed advice on research standards and methodology to facilitate the development of the reporting guideline. The group is also expected to finalise the working definitions and scope of the guideline. The GDG members will determine a working definition of fidelity through a two-step iterative process. In the first step, the project leadership will invite GDG members to participate in an online survey on Qualtrics®, a management software including a survey platform, where they can select their preferred published definition of fidelity. The choices in the survey will be based on existing fidelity definitions identified in the studies included in the

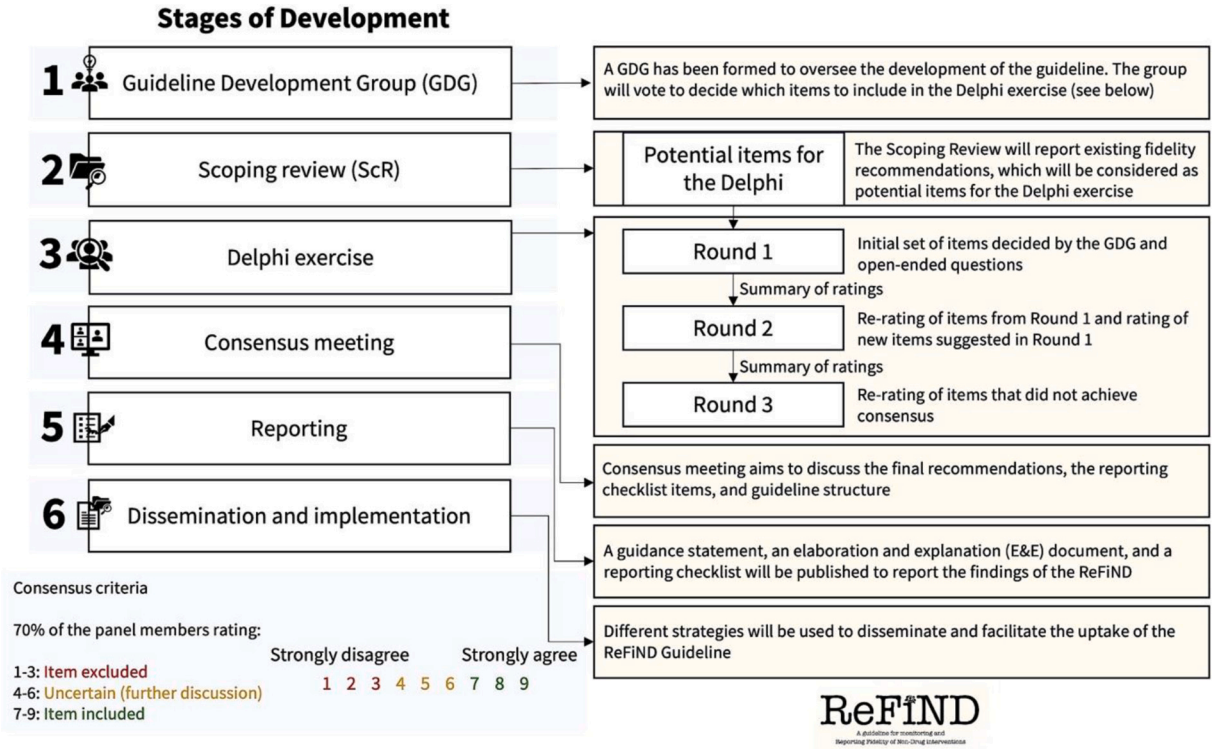


Fig. 1. Reporting guideline for Fidelity of Non-Drug interventions (ReFiND) development process.

scoping review (see subsection 2.3.). A free-text box will be provided on the survey for GDG members to provide comments or suggestions to enhance the existing definition they have selected. In the second step, the project leadership will discuss the definitions selected and the suggestions provided by the GDG members. Based on this discussion, they will reach a consensus on two possible working definitions of fidelity. GDG members will be then asked to accept one or the other definition and provide minor suggestions (e.g., wording), if needed. This accepted working definition of fidelity will be used in the Delphi study. We chose this iterative process to determine the working definition of fidelity because it considers both the existing definitions from the literature and the judgment of GDG members, rather than relying solely on stakeholder input. Additionally, the online surveys will be conducted anonymously, which may reduce social desirability bias in this process.

2.3. Stage (2): Scoping review

Given the broad scope of our study and the need for exploring and describing the literature rather than performing analytical approaches, we chose to conduct a scoping review to identify and summarize existing guidance documents on the fidelity of non-drug, non-surgical interventions. The protocol has been developed according to the JBI (formerly known as Joanna Briggs Institute) recommendations for developing scoping reviews [32].

2.3.1. Eligibility criteria

We will include any document providing guidance on the monitoring or reporting of fidelity of non-drug, non-surgical interventions, regardless of the study design. We will exclude studies that solely report fidelity assessment results from specific projects and lack general, explicit, fidelity recommendations applicable beyond the scope of those projects. Guidance documents targeting fidelity of drug or surgical interventions will be excluded. No date or language restrictions will be applied. Documents in languages other than English will be translated using the DeepL Translator [33], a web-based tool that uses artificial intelligence (AI) through neural networks to translate documents with more accuracy and nuances.

2.3.2. Evidence sources and search strategy

We will search the following databases: MEDLINE, Embase, PsycINFO, CINAHL, Cochrane Central Register of Controlled Trials (CENTRAL), and Allied and Complementary Medicine Database (AMED) from inception to the date of searching. The search strategy was developed in discussion with a medical librarian. We developed a set of eight known references from our literature review and conducted pilot searches to evaluate the search strategy retrieval performance for these references. Initially, we piloted the search across databases using free and mapped terms (Medical Subject Heading) related to two main concepts: (1) fidelity and (2) reporting guidelines. This broad search retrieved all the known references, but also retrieved a huge number of irrelevant references. Following, we conducted pilot searches using a third concept: (3) non-drug interventions. This search performed well by reducing the number of irrelevant references while successfully identifying the known references. Therefore, we decided to use the three concepts in the final search strategy. To increase the comprehensiveness of the search and to maximize the likelihood of identifying all relevant documents, we will perform backward citation searching (search the reference list of included studies and published relevant reviews) according to the TARCiS statement [34].

2.3.3. Selection process and data management

Two reviewers will independently screen the titles and abstracts based on the eligibility criteria and subsequently review the full text. Any disagreements will be resolved by consensus. A third reviewer will be consulted if the two reviewers do not achieve consensus. We will use Covidence [35], a web-based collaboration software platform, for data management including searching for duplicates, screening titles, abstracts and full text studies, extracting data and resolving disagreements.

2.3.4. Data extraction and synthesis

One reviewer will extract the data from included studies using an extraction form. A second reviewer will check the extracted data and

any disagreements will be resolved by consensus among them or in discussion with a third reviewer. The extraction form will be piloted with five references and will be refined as needed during this phase. We will extract the following data: publication characteristics (first author, year of publication, country, organizations, number of authors, funding, conflict of interest, document type, accessibility, and protocol registration), scope and purpose, concepts and domains (terminology, definition of fidelity, domains of fidelity), methodology (core methods, literature review, voting approach, consensus approach, steering committee, Delphi exercise, involvement of stakeholders), recommendations, dissemination and implementation, and development of reporting checklist/tools. Narrative synthesis of the extracted data will be presented in text and tables. The scoping review will be reported according to the PRISMA for Scoping Reviews [36].

2.4. Stage (3): Delphi exercise

For this study we will use a Delphi method to develop consensus on recommendations for reporting fidelity of non-drug, non-surgical interventions. Delphi is a structured but flexible approach that expands the knowledge around a determined topic by gathering inputs from stakeholders and building consensus on it [37]. The Delphi exercise involves a broad representation of key stakeholders, including a specialised panel from relevant disciplines (e.g., methodologists, clinical trialists) and different geographical locations, which is likely to increase the impact and uptake of recommendations. We will conduct the Delphi study in three online survey rounds using Qualtrics®. Each round will be open for 3–4 weeks and one email reminder will be sent for panellists each week. Information about the panellists, the Delphi survey rounds, items rating criteria, and consensus definition/threshold are provided in the following sections. The Delphi will, where appropriate, be consistent with relevant published guidance [30,37–40].

2.4.1. Panellists' recruitment

Researchers who have published fidelity or reporting methods studies and trials of non-drug, non-surgical interventions will be identified initially from the existing fidelity guidance documents included in the scoping review. Reporting methods studies refer to publications providing methods or guidance on the reporting of non-drug, non-surgical interventions. We will invite first, senior, and corresponding authors of these documents who satisfy any of the two criteria shown in Fig. 2. The criteria are broad, aiming to involve researchers with diverse expertise in trials (e.g., fidelity, process evaluation, reporting methods, etc.) and at different stages of intervention research (e.g., development, efficacy trials, effectiveness trials, implementation, etc.). This approach will contribute to a comprehensive understanding and enhance the breadth of perspectives and generalisability of the findings. Previous reporting guidelines have used similar approaches (e.g., broad criteria, more than one panel category, etc.) to ensure the scope of the guideline is covered and to reduce the risk of circularity (i.e., replication of existing practices) [19,21,41].

The invitation emails will ask for recommendations for additional potential panellists to participate in the study. We will also advertise this study through professional and research networks to maximize the likelihood of reaching panellists from diverse geographical locations, disciplines, and professional positions. We will monitor gender identity and geographical location of panellists (asked in the survey) accepting the invitations to identify any gender and/or geographic imbalance. If that happens, we will try to balance the email invitations accordingly and communicate with professional and research networks to reach underrepresented populations. No monetary incentives will be offered, but we will ask Delphi panellists their willingness to be named in the acknowledgements of the published work. Panellists participating in the consensus meeting and meeting all requirements for authorship will be offered co-authorship of the guidance statement.

The sample size for Delphi studies is typically determined by factors such as context, panel distribution (e.g., heterogeneity), and availability of resources [39]. In this Delphi study, our goal is to incorporate a broad range of perspectives from panellists with diverse backgrounds. Given the absence of clear standard guidelines for defining the sample size, we have chosen not to pre-specify one. Instead, our approach will be to recruit the largest number of panellists possible within the resources and timeline of this study, ensuring a wide variety of experiences and perspectives. Consistent with other studies [40], a minimal number of 25 complete responses per round is anticipated.

2.4.2. Delphi rounds

Before asking panellists to rate the actual Delphi items, we will provide them with a brief overview of the project goals and a working definition of fidelity elaborated by the GDG (see subsection 2.2.).

In the first Delphi round, we will ask panellists to rate a set of fidelity items and answer a set of open-ended questions. The selection of these initial items will be done by the GDG prior to the first round in two steps as suggested by the ACCORD guideline [30,31]. In the first step, we will extract existing recommendations on the reporting of fidelity from the scoping review, refine them, and put them as individual items into a survey. In the second step, GDG members will rate each item in a 9-point scale (1 = strongly disagree to 9 = strongly agree) for inclusion in the actual Delphi survey. Items rated 1–3 (disagree) by 70% of the GDG members will be excluded from the survey. The remaining items will be included in the Delphi survey to be rated by Delphi panellists in the first round.

Delphi panellists will rate each item on a 9-point scale (e.g., This item should be reported in trials: 1 'strongly disagree' to 9 'strongly agree'). We chose the 9-point scale because smaller scales have limited discriminating power, and larger scales present reduced reliability [42]. In addition, it has been showed that participants prefer larger scales such as the 9-point scale rather than smaller scales [42]. We will include a free-text box after each item to encourage further input, allowing panellists to provide additional comments. The open-ended questions in the

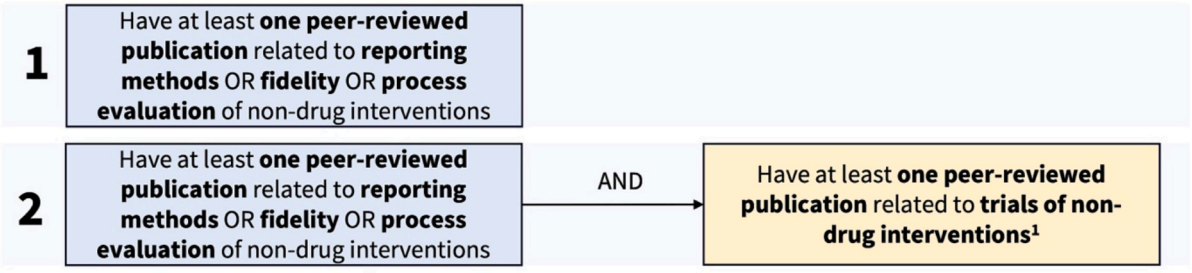


Fig. 2. Eligibility criteria for the Delphi panel. Researchers satisfying the criteria of any of these two strands will be eligible to be part of the Delphi panel. [1]Trials encompasses conventional clinical or public health trials (e.g., efficacy trials), effectiveness-implementation hybrid trials (Types I, II, and III), and implementation trials (i.e., trials assessing implementation outcomes rather than health outcomes), as described by Wolfenden et al. [24]. Further description of trials can be found in Box 1.

first round will cover aspects of monitoring fidelity that were identified as challenging in previous reviews and guidance documents.

In the second Delphi round, we will provide panellists with a summary of the ratings and justifications from the first round and new items developed from the analysis of open-ended responses. Panellists will be asked to rate each item again using the same scale. In the third round, we will ask panellists to rate items which did not achieve consensus for inclusion or exclusion in the previous round. Then, we will analyse the final data and group items that achieved consensus to present in the consensus meeting.

Consensus will be based on the percentage of agreement among panellists for each item. An item will be excluded when more than 70% of the panel members rate it as 3 or below. Conversely, it will be included if they rate it as 7 to 9. Items rated as 4 to 6 will be considered 'uncertain' and will be rated again by panellists in the next round. Items not reaching a consensus after the third round (rated as 'uncertain') will be discussed during the consensus meeting.

We will include a question about the applicability of each item in the Delphi survey to the different types of non-drug, non-surgical intervention trials. We will ask panellists to mark the option(s) they consider the Delphi item is applicable to: Conventional clinical or public health trial, Effectiveness-implementation hybrid trials type I, II, and III, and Implementation trial. A description of each trial design according to Wolfenden et al. 2021 [24] will be provided. We will report the total percentage of panellists selecting each type of trial for each Delphi item. These data will be presented in the consensus meeting and will inform the applicability of each item included in the reporting guideline to different types of trials.

We will use descriptive statistics to present demographic data (panellists' characteristics). In addition, we will present the retention rate as the percentage of panellists completing each round in relation to the previous round. We will also report the number of invitations sent to potential panellists in relation to the number of panellists completing the first round. We will conduct a content analysis of the responses to free text and open-ended questions in the first round. We will create codes for the responses and then group them into categories and subcategories using an inductive approach, as proposed by Elo and Kyngäs (2008) [43]. This analysis will inform the formulation of new items for the second round. We will present a summary of ratings for each item according to each round, including the median rating (total panel) and the percentage of panellists rating 1–3 (exclusion), 4–6 (uncertain), and 7–9 (inclusion).

2.5. Stage (4): On-line consensus meetings

The purpose of the consensus meeting is to develop and discuss the final recommendations, the reporting checklist items, and guideline structure. The working definition of fidelity will be assessed in the consensus meeting to check whether it reflects the items that achieved consensus in the Delphi study. Refinements will be made by the meeting members if needed. To ensure that all Delphi panellists willing to participate have the chance to contribute to the consensus meeting, we will adopt the following strategies: share Delphi results and meeting agenda with panellists at least two weeks before the meeting, schedule multiple online meetings considering different time-zones, set duration of the meeting beforehand, limit the number of panellists to eight per meeting, record the meetings and provide the recordings to panellists interested in contributing but not able to attend, and offer co-authorship for panellists attending the consensus meeting and commenting on the final paper prior to submission.

2.6. Stage (5) and (6): Reporting, dissemination, and implementation

A guidance statement, an elaboration and explanation (E&E) document, and a reporting checklist will be published to report the findings of the ReFiND study. The content, format, and structure of these documents will be discussed in the consensus meeting and reviewed by the ReFiND GDG. Then, we will pilot the documents with relevant stakeholders, including those who benefit from non-drug, non-surgical interventions, to ensure the documents are clear and readable for research consumers.

We will use different strategies to disseminate the ReFiND and facilitate its uptake: submit the guideline for publication in an appropriate healthcare or research methods journal, provide worked examples, collaborate with research networks and organizations from different geographical locations to promote the guideline, present the guideline at targeted conferences and events, conduct webinars and workshops with guideline authors at partner universities, and encourage the implementation of an easy-to-complete checklist on manuscript submissions with journal editors.

3. Discussion

The ReFiND protocol incorporates robust methods that will generate guidelines to enhance the transparency and reproducibility of non-drug, non-surgical interventions. While intervention fidelity has been more commonly discussed in the literature in the context of efficacy trials, ReFiND will consider the reporting of fidelity in all stages of intervention research [24], including effectiveness and implementation trials. The eligibility criteria for the Delphi panel are deliberately broad to capture a breadth of perspectives and to extend the capacity for generalizability [39]. These approaches, along with the dissemination and implementation strategies, are intended to facilitate the uptake of fidelity guidelines, achieving our goal of improving the quality and trustworthiness of non-drug, non-surgical intervention trials.

CRediT authorship contribution statement

Luis Fernando Sousa Filho: Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Conceptualization. **Melanie K. Farlie:** Writing – review & editing, Project administration, Methodology, Conceptualization. **Terry Haines:** Writing – review & editing, Project administration, Methodology, Conceptualization. **Belinda Borrelli:** Writing – review & editing, Methodology, Conceptualization. **Christopher Carroll:** Writing – review & editing, Methodology, Conceptualization. **Catherine Mathews:** Writing – review & editing, Methodology, Conceptualization. **Daniel C. Ribeiro:** Writing – review & editing, Methodology, Conceptualization. **Julie M. Fritz:** Writing – review & editing, Methodology, Conceptualization. **Martin Underwood:** Writing – review & editing, Methodology, Conceptualization. **Nadine E. Foster:** Writing – review & editing, Methodology, Conceptualization. **Sarah E. Lamb:** Writing – review & editing, Methodology, Conceptualization. **Zila M. Sanchez:** Writing – review & editing, Methodology, Conceptualization. **Peter Malliaras:** Writing – review & editing, Resources, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

NE Foster is supported by a National Health and Medical Research Council (NHMRC) Investigator Grant Leadership 2 (#2018182).

DC Ribeiro is supported by the Health Research Council New Zealand

(18/111).

LF Sousa Filho is supported by an Australian Government Research Training Program (RTP) Scholarship.

References

[1] B. Borrelli, The assessment, monitoring, and enhancement of treatment Fidelity in public health clinical trials, J. Public Health Dent. 71 (s1) (2011) S52–S63.

[2] L.R. Ginsburg, M. Hoben, A. Easterbrook, R.A. Anderson, C.A. Estabrooks, P. G. Norton, Fidelity is not easy! Challenges and guidelines for assessing fidelity in complex interventions, Trials 22 (1) (2021) 372.

[3] E. Toomey, W. Hardeman, N. Hankonen, M. Byrne, J. McSharry, K. Matvienko-Sikar, F. Lorencatto, Focusing on fidelity: narrative review and recommendations for improving intervention fidelity within trials of health behaviour change interventions, Health Psychol. Behav. Med. 8 (1) (2020) 132–151.

[4] H. Walton, A. Spector, I. Tombor, S. Michie, Measures of fidelity of delivery of, and engagement with, complex, face-to-face health behaviour change interventions: a systematic review of measure quality, Br. J. Health Psychol. 22 (4) (2017) 872–903.

[5] T.A. Rookes, A. Schrag, K. Walters, M. Armstrong, Measures of fidelity of delivery and engagement in self-management interventions: a systematic review of measures, Clin. Trials 19 (6) (2022) 665–672.

[6] S.C. Adams, J. McMillan, K. Salline, J. Lavery, C.S. Moskowitz, K. Matsoukas, M.M. Z. Chen, D. Santa Mina, J.M. Scott, L.W. Jones, Comparing the reporting and conduct quality of exercise and pharmacological randomised controlled trials: a systematic review, BMJ Open 11 (8) (2021) e048218.

[7] T.C. Hoffmann, C. Erueti, P.P. Glasziou, Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials, BMJ 347 (2013) f3755.

[8] M.M. JaKa, J.L. Haapala, E.S. Trapl, A.S. Kunin-Batson, B.A. Olson-Bullis, W. J. Heerman, J.M. Berge, S.M. Moore, D. Matheson, N.E. Sherwood, Reporting of treatment fidelity in behavioural paediatric obesity intervention trials: a systematic review, Obes. Rev. 17 (12) (2016) 1287–1300.

[9] R.G. Salloum, A.M. Rojewski, M.E. Piper, J.A. Blalock, B. Borrelli, L.M. Boyce, J. A. Minnix, O. Dogar, R.L. Tomko, D.E. Jorenby, C. Kotsen, J.S. Ostroff, Reporting treatment Fidelity in behavioral tobacco treatment clinical trials: scoping review and measurement recommendations, Nicotine Tob. Res. 24 (2) (2022) 150–159.

[10] H.J. Hansford, M.A. Wewege, A.G. Cashin, A.D. Hagstrom, B.K. Clifford, J. H. McAuley, M.D. Jones, If exercise is medicine, why don't we know the dose? An overview of systematic reviews assessing reporting quality of exercise interventions in health and disease, Br J Sports Med 56 (12) (2022) 692–700.

[11] A.J. Bellg, B. Borrelli, B. Resnick, J. Hecht, D.S. Minicucci, M. Ory, G. Ogedegbe, D. Orwig, D. Ernst, S. Czajkowski, N.I.H.B.C.C, Treatment Fidelity workgroup of the, enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH behavior change consortium, Health Psychol. 23 (5) (2004) 443–451.

[12] B. Borrelli, D. Sepinwall, D. Ernst, A.J. Bellg, S. Czajkowski, R. Breger, C. DeFrancesco, C. Levesque, D.L. Sharp, G. Ogedegbe, B. Resnick, D. Orwig, A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research, J. Consult. Clin. Psychol. 73 (5) (2005) 852–860.

[13] H. Walton, A. Spector, M. Williamson, I. Tombor, S. Michie, Developing quality fidelity and engagement measures for complex health interventions, Br. J. Health Psychol. 25 (1) (2020) 39–60.

[14] T. Mars, D. Ellard, D. Carnes, K. Homer, M. Underwood, S.J. Taylor, Fidelity in complex behaviour change interventions: a standardised approach to evaluate intervention integrity, BMJ Open 3 (11) (2013) e003555.

[15] C. Carroll, M. Patterson, S. Wood, A. Booth, J. Rick, S. Balain, A conceptual framework for implementation fidelity, Implement. Sci. 2 (2007) 40.

[16] F. Perepletchikova, On the topic of treatment integrity, Clin Psychol (New York) 18 (2) (2011) 148–153.

[17] D. McGee, F. Lorencatto, K. Matvienko-Sikar, E. Toomey, Surveying knowledge, practice and attitudes towards intervention fidelity within trials of complex healthcare interventions, Trials 19 (1) (2018) 504.

[18] D. Moher, K.F. Schulz, I. Simera, D.G. Altman, Guidance for developers of health research reporting guidelines, PLoS Med. 7 (2) (2010) e1000217.

[19] T.C. Hoffmann, P.P. Glasziou, I. Boutron, R. Milne, R. Perera, D. Moher, D. G. Altman, V. Barbour, H. Macdonald, M. Johnston, S.E. Lamb, M. Dixon-Woods, P. McCulloch, J.C. Wyatt, A.W. Chan, S. Michie, Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide, BMJ 348 (2014) g1687.

[20] S.C. Slade, C.E. Dionne, M. Underwood, R. Buchbinder, Consensus on exercise reporting template (CERT): explanation and elaboration statement, Br. J. Sports Med. 50 (23) (2016) 1428–1437.

[21] D. Hohenschurz-Schmidt, L. Vase, W. Scott, M. Annoni, O.K. Ajayi, J. Barth, K. Bennell, C. Berna, J. Bialosky, F. Braithwaite, N.B. Finnerup, A.C.C. Williams, E. Carlino, F. Cerritelli, A. Chaibi, D. Cherkin, L. Colloca, P. Cote, B.D. Darnall, R. Evans, L. Fabre, V. Faria, S. French, H. Gerger, W. Hauser, R.S. Hinman, D. Ho, T. Janssens, K. Jensen, C. Johnston, S. Juhl Lunde, F. Keefe, R.D. Kerns, H. Koechlin, A. Kongsted, L.A. Michener, D.E. Moerman, F. Musial, D. Newell, M. Nicholas, T.M. Palermo, S. Palermo, K.J. Peerdeman, E.M. Pogatzki-Zahn, A. A. Puhl, L. Roberts, G. Rossettini, S. Tomczak Matthiesen, M. Underwood, P. Vaucher, J. Vollert, K. Wartolowska, K. Weimer, C.P. Werner, A.S.C. Rice, J. Draper-Rodi, Recommendations for the development, implementation, and reporting of control interventions in efficacy and mechanistic trials of physical, psychological, and self-management therapies: the CoPPS statement, BMJ 381 (2023) e072108.

[22] M. Banno, Y. Tsujimoto, Y. Kataoka, The majority of reporting guidelines are not developed with the Delphi method: a systematic review of reporting guidelines, J. Clin. Epidemiol. 124 (2020) 50–57.

[23] T.A. Rookes, A. Barat, R. Turner, S. Taylor, Reporting dose in complex self-management support interventions for long-term conditions: is it defined by researchers and received by participants? A systematic review, BMJ Open 12 (8) (2022) e056532.

[24] L. Wolfenden, R. Foy, J. Presseau, J.M. Grimshaw, N.M. Ivers, B.J. Powell, M. Taljaard, J. Wiggers, R. Sutherland, N. Nathan, C.M. Williams, M. Kingsland, A. Milat, R.K. Hodder, S.L. Yoong, Designing and undertaking randomised implementation trials: guide for researchers, BMJ 372 (2021) m3721.

[25] F.U.F.a.D. Administration, FDA Glossary of Terms. https://www.fda.gov/drugs/drug-approvals-and-databases/drugsfda-glossary-terms#D, 2023.

[26] T. Quasdorf, L. Clack, F. Laporte Uribe, D. Holle, M. Berwig, D. Purwins, M. T. Schultes, M. Roes, Theoretical approaches to process evaluations of complex interventions in health care: a systematic scoping review protocol, Syst. Rev. 10 (1) (2021) 268.

[27] G.F. Moore, S. Audrey, M. Barker, L. Bond, C. Bonell, W. Hardeman, L. Moore, A. O'Cathain, T. Tinati, D. Wight, J. Baird, Process evaluation of complex interventions: Medical Research Council guidance, BMJ 350 (2015) h1258.

[28] H. Pinnock, M. Barwick, C.R. Carpenter, S. Eldridge, G. Grandes, C.J. Griffiths, J. Rycroft-Malone, P. Meissner, E. Murray, A. Patel, A. Sheikh, S.J. Taylor, R.I. G. Sta, Standards for reporting implementation studies (StaRI) statement, BMJ 356 (2017) i6795.

[29] K. Skivington, L. Matthews, S.A. Simpson, P. Craig, J. Baird, J.M. Blazeby, K. A. Boyd, N. Craig, D.P. French, E. McIntosh, M. Petticrew, J. Rycroft-Malone, M. White, L. Moore, A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance, BMJ 374 (2021) n2061.

[30] W.T. Gattrell, P. Logullo, van Zuuren, A. Price, E.L. Hughes, P. Blazey, C. C. Winchester, D. Tovey, K. Goldman, A.P. Hungin, N. Harrison, ACCORD (ACcurate COnsensus Reporting Document): A reporting guideline for consensus methods in biomedicine developed via a modified Delphi, PLoS medicine 21 (1) (2024) e1004326, https://doi.org/10.1371/journal.pmed.1004326.

[31] W.T. Gattrell, A.P. Hungin, A. Price, C.C. Winchester, D. Tovey, E.L. Hughes, E. J. van Zuuren, K. Goldman, P. Logullo, R. Matheis, N. Harrison, ACCORD guideline for reporting consensus-based methods in biomedical research and clinical practice: a study protocol, Res Integr Peer Rev 7 (1) (2022) 3.

[32] M.D.J. Peters, C. Godfrey, P. McInerney, H. Khalil, P. Larsen, C. Marnie, D. Pollock, A.C. Tricco, Z. Munn, Best practice guidance and reporting items for the development of scoping review protocols, JBI Evid Synth 20 (4) (2022) 953–968.

[33] DeepL, DeepL Translator. https://www.deepl.com/whydeepl, 2024.

[34] J. Hirt, T. Nordhausen, T. Fuerst, H. Ewald, C. Appenzeller-Herzog, TARCiS study group, Guidance on terminology, application, and reporting of citation searching: the TARCiS statement, BMJ (Clinical research ed.) 385 (2024) e078384. https://doi.org/10.1136/bmj-2023-078384.

[35] V.H. Innovation, Covidence Systematic Review Software. www.covidence.org, 2023.

[36] A.C. Tricco, E. Lillie, W. Zarin, K.K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M. D.J. Peters, T. Horsley, L. Weeks, S. Hempel, E.A. Akl, C. Chang, J. McGowan, L. Stewart, L. Hartling, A. Aldcroft, M.G. Wilson, C. Garritty, S. Lewin, C. M. Godfrey, M.T. Macdonald, E.V. Langlois, K. Soares-Weiser, J. Moriarty, T. Clifford, O. Tuncalp, S.E. Straus, PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation, Ann. Intern. Med. 169 (7) (2018) 467–473.

[37] S. Junger, S.A. Payne, J. Brine, L. Radbruch, S.G. Brearley, Guidance on conducting and REporting DElphi studies (CREDES) in palliative care: recommendations based on a methodological systematic review, Palliat. Med. 31 (8) (2017) 684–706.

[38] S.C. Slade, C.E. Dionne, M. Underwood, R. Buchbinder, Standardised method for reporting exercise programmes: protocol for a modified Delphi study, BMJ Open 4 (12) (2014) e006682.

[39] E.G. Trevelyan, P.N. Robinson, Delphi methodology in health research: how to do it? Eur. J. Integr. Med. 7 (4) (2015) 423–428.

[40] I.R. Diamond, R.C. Grant, B.M. Feldman, P.B. Pencharz, S.C. Ling, A.M. Moore, P. W. Wales, Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies, J. Clin. Epidemiol. 67 (4) (2014) 401–409.

[41] O. Solovyeva, M. Dimairo, C.J. Weir, S.W. Hee, A. Espinasse, M. Ursino, D. Patel, A. Kightley, S. Hughes, T. Jaki, A. Mander, T.R.J. Evans, S. Lee, S. Hopewell, K. R. Rantell, A.W. Chan, A. Bedding, R. Stephens, D. Richards, L. Roberts, J. Kirkpatrick, J. de Bono, C. Yap, Development of consensus-driven SPIRIT and CONSORT extensions for early phase dose-finding trials: the DEFINE study, BMC Med. 21 (1) (2023) 246.

[42] C.C. Preston, A.M. Colman, Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences, Acta Psychol. (Amst) 104 (1) (2000) 1–15.

[43] S. Elo, H. Kyngas, The qualitative content analysis process, J. Adv. Nurs. 62 (1) (2008) 107–115.