Designing and undertaking randomised implementation trials: guide for researchers

Luke Wolfenden,[1,2] Robbie Foy,[3] Justin Presseau,[4,5] Jeremy M Grimshaw,[4,6] Noah M Ivers,[7,8,9,10] Byron J Powell,[11] Monica Taljaard,[4,5] John Wiggers,[1,2] Rachel Sutherland,[1,2] Nicole Nathan,[2] Christopher M Williams,[1,2,12] Melanie Kingsland,[1,2] Andrew Milat,[12] Rebecca K Hodder,[1,2] Sze Lin Yoong[13]

Implementation science is the study of methods to promote the systematic uptake of evidence based interventions into practice and policy to improve health. Despite the need for high quality evidence from implementation research, randomised trials of implementation strategies often have serious limitations. These limitations include high risks of bias, limited use of theory, a lack of standard terminology to describe implementation strategies, narrowly focused implementation outcomes, and poor reporting. This paper aims to improve the evidence base in implementation science by providing guidance on the development, conduct, and reporting of randomised trials of implementation strategies. Established randomised trial methods from seminal texts and recent developments in implementation science were consolidated by an international group of researchers, health policy makers, and practitioners. This article provides guidance on the key components of randomised trials of implementation strategies, including articulation of trial aims, trial recruitment and retention strategies, randomised design selection, use of implementation science theory and frameworks, measures, sample size calculations, ethical review, and trial reporting. It also focuses on topics requiring special consideration or adaptation for implementation trials. We propose this guide as a resource for researchers, healthcare and public health policy makers or practitioners, research funders, and journal editors with the goal of advancing rigorous conduct and reporting of randomised trials of implementation strategies.

Investments in health research are not fully realised because of delayed and variable uptake of effective interventions by health systems and professionals.[1-3] Implementation science seeks to resolve this problem by generating evidence to facilitate the use and integration of evidence based interventions into health policy and practice.[4] Just as well conducted randomised clinical trials can provide robust estimates of the effects of medical and surgical treatments, well conducted randomised trials of implementation strategies (which we refer to as implementation trials) can provide robust assessments of the effects of implementation strategies. These strategies include audit and feedback, training, or reminders, on measures of the uptake and integration of evidence based interventions in healthcare and public health practice.[5]

Although randomised trials are central to evidence based medicine[6] and are a common evaluation design in the field of implementation science,[7] concerns have been raised about the quality of implementation trials. Criticisms include high risks of bias, limited use of theory, a lack of standardised terminology to describe implementation strategies, limited measures, and poor reporting.[7-11] Progress in the field, however, has been rapid with recent advances in implementation science theory, concepts, terminology, measures, and reporting standards to resolve many of these limitations.[12-14]

This article draws on recent developments in implementation science with established randomised

SUMMARY POINTS

Criticisms of current implementation trials include risks of bias, lack of theory use, lack of standardised terminology to describe implementation strategies, and limited measures and poor reporting

This article consolidates recent methodological developments in implementation science with established guidance from seminal texts of randomised trial methods to provide best practice guidance to improve the development and conduct of randomised implementation trials

Consideration of such guidance will improve the quality and use of randomised implementation trials for healthcare and public health improvement

trial methods to provide a best practice guide to improve the development, conduct, and reporting of randomised implementation trials. This guidance was authored by an international interdisciplinary group with expertise spanning implementation science, health services research, behavioural science, public health, trial methods, biostatistics, and health policy and practice. It discusses application of randomised trial methods in the context of large scale trials of implementation strategies, focusing on aspects that might be unique to implementation studies. Table 1 defines key implementation terms used in the guide.

Recommendations for the development, conduct, and reporting of randomised implementation trials

When is an implementation trial warranted?

Implementation trials generate scientific knowledge to improve the uptake of evidence based interventions in practice. Researchers should consider several factors when deciding whether a trial of implementation strategies is needed,[19] primarily the following:

- A healthcare or public health intervention that is supported by evidence as effective (ideally by a systematic review of trials);
- A known evidence-practice gap—that is, verification that the evidence based intervention is not routinely implemented in practice;[19 20] and
- Equipoise regarding the effects of an implementation strategy.

The need for a trial and the trial methods used should also be guided by the needs, values, and input of end users and other stakeholder groups. A range of guidance documents are available to identify appropriate groups to engage and undertake meaningful research co-design across all phases of trial design, conduct, and dissemination.[21-23] Key features of successful co-design include clearly articulated roles and responsibilities in the process, research training to end users, clear communication pathways, and frequent interactions between researchers and end users.[24]

Statement of the implementation trial aim

Randomised implementation trials should have precisely stated aims, defining the population, intervention, comparison, and outcome under investigation. They should also distinguish clearly between the aims of the implementation strategy and the therapeutic intent of the targeted evidence based intervention.[12] For example: "The study aimed to assess the effectiveness of audit and feedback (implementation strategy), relative to usual practice (implementation comparison) for improving clinician (implementation population) provision (implementation outcome, and target of the implementation strategy) of nicotine replacement therapy (clinical intervention) to inpatients of a cardiac ward to support smoking cessation (therapeutic intent of the clinical intervention)."

Randomised implementation trials can assess the effect of a given strategy on implementation outcomes alone, or assess both the effectiveness of the intervention on clinical or population health therapeutic outcomes as well as the effect of the implementation strategy on implementation outcomes.[25] Trials with a dual focus are known as effectiveness-implementation hybrid trials (table 2). Type I effectiveness-implementation hybrid designs aim to evaluate the effects of an evidence based intervention and describe or better understand the context for implementation, but do not test an implementation strategy.[25] Type II and III hybrid trials test implementation strategies on implementation outcomes.[25] Although hybrid designs are suggested to be an efficient means of accumulating evidence to inform implementation, the contribution of type I and II trials to this end could be limited. This limitation could be the case when research design considerations to preserve the robust assessment of clinical effectiveness questions are prioritised over those considerations to assess the effect of an implementation strategy (on implementation outcomes).

As an example, a type II hybrid trial could express dual aims as follows: "The primary aims of the study were to: i) assess the effectiveness of audit and feedback (implementation strategy), relative to usual practice (implementation comparison) for improving clinician (implementation population) provision (implementation outcome, and target of the implementation strategy) of nicotine replacement therapy (clinical intervention); and ii) to assess the effectiveness of nicotine replacement therapy (clinical intervention), relative to usual care, in improving smoking cessation (therapeutic outcome and therapeutic intent of the clinical intervention) among cardiac inpatients (therapeutic population)."

Table 1 | Definitions of key terms in implementation science

Term	Definition
Implementation science	Scientific study of methods to promote the systematic uptake of evidence based interventions into practice and policy to improve health[4]
Implementation strategy	Method or technique used to enhance the adoption, implementation, and sustainability of an evidence based intervention[15]
De-implementation	Process of identifying and removing non-evidence based interventions that are harmful, not cost effective, or ineffective[16]
Evidence based intervention	Evidence based practice, model of care, programme, policy, process, or guideline recommendation that is being implemented[17]
Implementation outcomes	Process-of-care or quality measures (or related measures for public health) to assess the effects of the implementation strategy[15 17]
Implementation trial	Research design testing the effects of implementation strategies on implementation outcomes[5]
Clinical (therapeutic) trial	Research that investigates the effect of a treatment or other intervention on patient health outcomes[18]
Adaptation	Degree to which an evidence based intervention is changed (eg, during intervention delivery) to suit the needs of the setting or the target population[17]

Table 2 | Typical characteristics of conventional clinical or public health trials, effectiveness-implementation hybrid trials, and implementation trials. Adapted from Curran et al, 2012, with permission[25]

Conventional clinical (therapeutic) or public health trial	Effectiveness-implementation hybrid trials (type)			Implementation trial
	I	II	III	
Research aim				
To assess the therapeutic effects of a clinical or public health intervention on individual patient or population health outcomes	Primary: to assess the therapeutic effectiveness of a clinical or public health intervention on individual patient or population health outcomes; secondary: to describe or better understand the context for implementation	Co-primary: to assess the therapeutic effectiveness of a clinical or public health intervention on individual patient or population health outcomes; and to assess the effects of a strategy to implement a clinical or public health intervention on implementation outcomes	Primary: to assess the effects of a strategy to implement a clinical or public health intervention on implementation outcomes; secondary: to describe individual or population therapeutic health outcomes associated with implementation of an intervention	To assess the effects of a strategy to implement a clinical or public health intervention on implementation outcomes
Target of experimental manipulation (intervention or implementation strategy)				
Individual patients, community members, or populations	Individual patients, community members, or populations	Both individual patient's community members, or populations; and clinicians, policy makers, service providers or medical or public health systems responsible for implementation	Primary: clinicians, policy makers, service providers or medical or public health systems responsible for implementation; secondary: individual patients, community members or populations	Clinicians, policy makers, service providers or medical or public health systems responsible for implementation
Effects of therapeutic intervention on patient or population health outcomes of interest				
Explicitly tested	Explicitly tested	Explicitly tested	Not tested; known to be effective	Not tested; known to be effective
Effects of implementation strategy on implementation outcomes				
Typically not considered or required as intervention delivery is usually at the control of, or administered by researchers	Not tested	Explicitly tested	Explicitly tested	Explicitly tested
Trial outcome measures				
Clinical conditions, patient symptoms, health behaviours, disease risk factors, or other patient or population health related outcomes	Clinical conditions, patient symptoms, health behaviours, disease risk factors, or other patient or population health related outcomes	Both: clinical conditions, patient symptoms, health behaviours, disease risk factors, or other patient or population health related outcomes; professional practice improvement, changes in processes of care, adherence to clinical standards, quality of intervention delivery or other implementation outcomes	Primary: professional practice improvement, changes in processes of care, adherence to clinical standards, quality of intervention delivery or other implementation outcomes; secondary: health service use, clinical conditions, patient symptoms, health behaviours, or other health related outcomes	Professional practice improvement, changes in processes of care, adherence to clinical standards, quality of intervention delivery or other implementation outcomes

Recruitment and retention

Implementation trials usually recruit and randomise staff or organisations rather than individual patients. Intervention effects on clinical practice are often assessed using routinely collected, anonymised data. Therefore, implementation trials can be conducted at relatively low cost, with potentially more complete trial data than those from clinical trials that require intensive recruitment and follow-up of patients.[26 27] Nonetheless, effective recruitment and retention approaches are needed to ensure that all participant groups (patients, clinicians, health services) are broadly representative of the populations for which the findings are intended to generalise. Minimising barriers to participation is therefore critical to maximise external validity. Consent procedures for participants to opt out could be appropriate in some circumstances and can result in high levels of participation,[28] recruitment of more typical participants groups, and more generalisable effects.[29-31] Opt out consent was recently used, for example, in a randomised trial of mail-outs and phone calls to improve adherence to secondary preventive treatment after myocardial infarction that used administrative data for outcome assessment.[32]

For research using active consent procedures, recruitment and retention strategies recommended for patients in clinical trials (such as dedicated recruitment coordinators) and reminders for non-responders also apply to the recruitment of patient groups in implementation trials. Researchers can also leverage the networks of relevant professional associations or governing health authorities,[33 34] engage potential trial sites in the design of the study and its recruitment and retention strategies to minimise the potential burden of participation, ensure acceptability, and facilitate the recruitment of health organisations and clinicians. Because implementation trials aim to promote evidence based practice, they could be more attractive to clinicians and organisations than other types of research, particularly when stepped wedge or delayed control group designs are used as all sites receive implementation support as part of, or immediately following, follow-up data collection.

Underlying trial philosophy: pragmatic and explanatory trials

Explanatory trials use methods that prioritise internal validity, and are undertaken in more ideal research conditions.[35] Pragmatic trials emphasise external validity using methods more closely aligned to real world contexts.[35] Explanatory trials focus on questions asking whether the intervention (or implementation strategy) "can" work. Implementation trials are inherently pragmatic because they usually focus on

whether an intervention (or implementation strategy) "does" work when delivered in routine clinical or public health contexts.[35] As such, the effect sizes of interventions tested in pragmatic trials are typically smaller than those reported in explanatory trials.[36 37]

The pragmatic explanatory continuum indicator summary tool (PRECIS-2) describes the methodological characteristics of explanatory and pragmatic trials and can help researchers undertaking implementation trials to make design decisions consistent with the intended purpose and pragmatic nature of implementation trials.[38] The tool requires users to consider trial eligibility criteria, recruitment methods, setting, the expertise and resources required for intervention implementation, the degree of flexibility in the implementation and adherence to the intervention, follow-up procedures, the selection of relevant primary outcome measures, and analysis. Furthermore, pragmatic trials might require departures from conventional safety and integrity monitoring processes, which have been largely designed for explanatory studies. Simon et al offer some guidance of adaptations that could be appropriate across each of the key participant safety and trial integrity obligations.[39]

Research trial design considerations

Non-randomised study designs are often used in implementation research on the basis that they might be more appropriate or feasible than a randomised controlled trial. However, these designs could report misleading estimates of effect even when experimental groups appear similar on important prognostic factors, and when such factors are considered in analyses.[40] Randomised trials have also been suggested to be unnecessary in instances when extreme effects are anticipated, for example, when relative risks are less than 0.25 or greater than 4.[41] However such effect sizes are rarely reported in implementation trials. Because the process of random assignment of an adequate number of units can effectively eliminate the risk of confounding, randomised trials provide the most robust evidence of the effects of implementation strategies. Further, with improving access and opportunity to use existing routinely collected data such as registries and electronic medical records, such designs are increasingly feasible.[41 42]

Nonetheless, randomised trials require interventions that can feasibly be assigned at random. Examination of the impact of national level legislative or regulatory changes on professional practice, for example, are unlikely to be amenable to evaluation using randomised designs. Complex, adaptive systems based strategies, and those developed using complexity theory, have been tested as part of randomised implementation trials,[43 44] but there are many challenges to doing so, particularly for interventions in open systems without clearly defined boundaries.[45] Randomised trials of such strategies may include mixed method research approaches, in-depth case studies, and ethnographic narratives

to better understand system interconnectedness, interactions, and impact.[45] The development of evaluation methods of these types of interventions has been identified as a priority, and are beginning to emerge.[46 47]

A variety of randomised trial designs can be used in implementation trials (table 3). Researchers undertaking implementation trials should be aware of the relative merits of different randomised designs to inform appropriate design selection.[55 56] A thorough description of randomised trial design limitations (and strengths) is provided elsewhere and summarised in supplementary file 1.[55 57] Here, we discuss the level of randomisation considerations, and describe randomised trial designs that can be applied to assess the effects of implementation strategies.

Level of randomisation

In an individually randomised trial, individual participants (that is, patients)[55] are randomised to one of two or more parallel groups, and outcomes (eg, clinical effectiveness) are measured at the same level as the unit of randomisation (patient). Such trials are relatively uncommon in implementation research given that interventions often operate at multiple levels and involve changes to health systems. Most implementation trials using random assignment, therefore, use cluster randomised designs (also called group randomised designs).[7] In these designs, clusters such as hospitals or clinicians are randomised to receive support to implement an evidence based intervention (an implementation strategy) or a comparison condition, but where implementation outcome data can be collected from multiple individuals (that is, patients) within each cluster.[55] Such outcome data are usually correlated, and this clustering must be accounted for in the design and analysis to obtain valid statistical inferences.[58]

Many levels of clustering are possible in implementation trials: for example, patients can be clustered within clinicians, who could themselves be clustered within a hospital, and hospitals could be clustered within a healthcare organisation. The unit of randomisation should be carefully chosen to reflect the trial aims, and should consider trade-offs between randomising at a higher level to prevent contamination versus randomising at a lower level to increase the number of units available for randomisation. Contamination likely occurs even in cluster randomised trial designs where individual clinicians within a hospital are allocated to implementation training and support, and then pass on such implementation resources or knowledge to clinicians in the same hospital allocated to a control condition. In such cases, randomising at the level of the hospital or organisation rather than the clinician can help mitigate this risk. On the other hand, if the contamination is not substantial, randomising at a lower level might be preferable, from a statistical efficiency perspective.[59] The higher the level of randomisation, the fewer groups (eg, clinics, hospital) may be available to be randomised.

Table 3 | Description and key considerations of randomised designs for assessing the effects of implementation interventions

Description	Considerations	Example
Two arm, parallel randomised trial[48]		
Individuals or groups (eg, clinics or schools) consisting of multiple individuals (eg, patients) are randomly assigned to receive a treatment (implementation strategy) or an alternative condition (eg, usual practice or control)	Most appropriate when sample size or trial resources are limited, and when there is an interest in assessing the effect of one implementation strategy compared with current practice or an alternative implementation strategy	To evaluate the effectiveness of an implementation intervention to improve six guideline recommended, health professional behaviours in managing type 2 diabetes in primary care, 44 general practices were randomised to implementation support or usual care control. Implementation support was provided to clinicians within general practices allocated to receive it, while the primary outcome included a patient survey of a random sample of patients per practice that reported receipt of updated diabetes education advice as well as routinely collected prescribing data for blood pressure, insulin initiation for glycaemic control, and foot examinations from practice records across practices[26]
Multi-arm randomised trial[49]		
Investigate the effects of two or more implementation strategies versus a comparison (or alternative strategy) at the same time. Such designs can involve individual or cluster randomisation	Most appropriate when sample sizes are large, when there is an interest in assessing the relative effects of different implementation strategies alone or in combination, and where there is good control over the implementation strategies provided to each group	To promote the uptake of evidence based guidance on blood transfusion in surgery, a 2×2 factorial, cross sectional, cluster-randomised controlled trial allocated NHS trusts* to receive one of the following: standard feedback reports (usual care), standard reports with follow-on support, enhanced reports, or enhanced reports with follow-on support. The primary outcome for each topic will be the proportion of patients receiving a transfusion coded as unnecessary using data from a national audit[50]
Stepped wedge randomised trials[51]		
Following a baseline period, an implementation strategy is sequentially provided to clusters. The order in which the different clusters are assigned to receive the implementation strategy is randomised. Over time, all units will have received implementation support	Most appropriate when a decision has been made to roll out an implementation strategy across a health system, when risks of bias are low, and when routinely collected data are available for outcome assessment	To improve the delivery of evidence based cardiovascular care in primary care, practices were randomly assigned by region to receive implementation support 12, 24, or 36 months after initiation of baseline data collection. The primary outcome was mean adherence to indicators of evidence based care as measured by chart review of a randomly selected cohort of 66 patients per practice (measured before, during, and after receipt of implementation support)[52]
Sequential trial design: sequential multiple assignment randomised trial[53]		
Intervention dose, type, or delivery of an implementation strategy (or intervention) is modified at several stages based on specified decision rules. At each stage, the participant is randomly (re)assigned to one of several implementation strategy (intervention) options	Can be used to help many practical decisions regarding how best to support improvements in implementation. Most appropriate for the development of adaptive implementation strategies when a sufficient sample is available, and where there is good control over the implementation strategies provided to each group	To evaluate the effectiveness of a sequential approach to sustainment of a postpartum depression prevention programme (EIAU) in outpatient clinics, clinics at risk of not sustaining programme implementation will be randomised to receive either no additional implementation support (that is, EIAU only), or low intensity coaching and feedback (LICF). If clinics receiving LICF are still at risk at subsequent assessments, they will be randomised to either LICF or high intensity coaching and feedback. The primary outcome includes percent sustainment of implementation of core programme elements[54]

*Trusts in the United Kingdom's health service.

Parallel, two arm, randomised trial

Parallel, two arm, randomised implementation trials compare the effects of an implementation strategy with those of a control or alternative implementation strategy. Conduct of two arm trials is useful when the effects of one implementation strategy are primarily of interest. These trials are more feasible than multi-arm trials and are the most common randomised design used to assess the effects of implementation strategies.[60 61]

Multi-arm randomised trials

Multi-arm randomised trials provide information about the comparative effects of multiple implementation approaches. They represent a more efficient method of testing the effects of implementation strategies than performing sequential two arm trials.[49] For example, including three arms in a randomised implementation trial could enable the comparison of two implementation strategies with each other as well as a comparison condition. In randomised factorial designs, participants (or clusters) are randomised into groups comprised of combinations of the experimental conditions. Researchers interested in testing the effects of implementation strategy A as well as those of implementation strategy B within the same trial, for example, might randomise participants into four groups: A alone, B alone, both A and B, and neither A nor B.[55] Such designs enable exploration of interactions between groups, and the effects of implementation strategies separately and in combination. Fractional factorial randomised trials include larger numbers of strategies, however, and allocate participants to selected (rather than all) strategy combinations, eliminating comparisons that are of no interest to reduce the potential sample size requirements of the trial.[62 63]

When an intervention must, for practical, logistical, or organisational reasons, be rolled out to all units in a health system, a stepped wedge design might be useful. In stepped wedge randomised trials,[57 64] all units such as hospitals (clusters) are first recruited, then randomised to receive the implementation intervention at regular intervals (or steps) sequentially over time, until all units have been exposed to the intervention.[65 66] Trial outcome data are collected at regular intervals throughout the trial, with each unit providing data for both experimental and control conditions (periods). Under some circumstances, the design might require fewer units to participate than parallel arm, cluster randomised trials, particularly when the intraclass correlation is high and cluster period sizes are large. Stepped wedge trials require repeated assessment of outcomes across the trial periods, making these designs most suited for outcomes that can be assessed using routinely collected data.

Such designs are increasingly being used in health services and implementation research, although they are vulnerable to increased risks of bias and other complexities that could make them less attractive than parallel arm designs.[64][65][67]

Sequential trial designs

Sequential multiple assignment randomised trials (SMART) are a type of adaptive design used to inform the development of adaptive implementation strategies (or interventions).[53][68] In an adaptive implementation strategy, the dose, type, or delivery of strategies is modified across several stages based on prespecified decision rules, providing individualised approaches to better meet the specific needs and evolving status of participants. With this design, participants are randomised to different implementation strategy options at each stage.[68] For example, clinicians who do not improve implementation of an intervention following the provision of an initial package of implementation strategies could receive different or more intensive implementation support subsequently than clinicians who do improve implementation. The design allows researchers to assess the effect of adaptive approaches and the isolation of the effects of specific strategy modifications. Such designs involve complex statistical considerations.

Hybrid trials

Hybrid trials can use any type of randomised trial design. However, because they focus on assessing the effects of implementation strategies on both clinical effectiveness and implementation outcomes, design modification might be needed (table 3).[25] Design modifications may often be required because clinical effectiveness outcomes are usually assessed at an individual level, while implementation outcomes could be assessed at a provider or organisational level. This duality of purpose of hybrid trials can result in research designs to assess outcomes at one level being nested within a design determined by an outcome at another level. For example, a randomised trial of the introduction of a school nutrition policy might require 100 schools to participate to detect meaningful change in school level policy implementation (implementation outcome), but need only to assess students in a nested random sample of 20 participating schools to identify meaningful improvements in child dietary intake (effectiveness outcome).

Reducing bias in randomised implementation trials

Researchers should be aware that randomised trials are prone to threats to internal validity and seek to avoid major risks of bias.[56] As implementation trials often include multiple outcomes assessed at different levels (organisation, clinician, patient), research design characteristics and risk of bias need consideration at each level. For cluster trials, baseline comparability of groups at both the cluster and individual levels can be difficult to achieve if only a small number of clusters such as hospitals are available for randomisation.[69][70]

In many cluster implementation trials, study sites (clusters) such as clinics, might be randomised and allocated before individual (that is, patient level) recruitment. If those identifying and recruiting participants (or the potential participants themselves) are not blinded to allocation, differential recruitment and study participation can occur (selection bias).[71] Selection bias is a common problem in clustered designs.[72] In the UK BEAM trial, for example, primary care practices were recruited and randomised.[73] Clinicians at primary care practices allocated to the experimental arm then received training in guideline based management of back pain after which patient recruitment commenced. In the study, practice nurses recruited twice as many patients among primary care practices allocated to receive training as those patients allocated to usual care, and the characteristics of patients differed between groups. Gatekeepers can also withdraw their health site (cluster) from a trial once informed of group allocation but before individual participant level recruitment.[71] Such circumstances can be particularly challenging for intention-to-treat approaches to analyses of trial outcomes, because little is known about the characteristics of those individuals who would have participated in that cluster.[74] Selection bias can best be avoided by allocating units after consent and baseline data collection.

In clinical trials, a lack of blinding of participants and personnel delivering an intervention in a clinical trial could increase the risk of bias,[55] because knowledge of assignment to an intervention might lead to contamination, protocol deviations, or co-intervention. However, the blinding of participants and personnel is often inappropriate (and not possible) in implementation trials because they seek to assess the effect of an implementation strategy in individuals or organisations aware of the care given. A range of other strategies could reduce the risks of such biases including the use of clustered designs,[75] simply asking clinicians or patients not to share information, trial intervention or implementation strategy sessions that are spatially or temporally separate, and systems to avoid transfer of patients between clinicians.[76] The effectiveness of these strategies, however, is unclear. If adequately assessed, statistical approaches can also be used to adjust for contamination in analyses.[77-79] The Cochrane risk-of-bias tool (version 2)[56] for randomised trials provides a comprehensive description of potential risks of bias for various randomised designs and strategies to help identify and reduce such risks.

Models, theories, and frameworks

The lack of explicit descriptions of the mechanism by which implementation strategies are hypothesised to exert their effects is suggested to reduce the ability to judge the generalisability of trial findings across settings and contexts, to limit understanding of implementation processes and to slow the cumulative progression of the field.[80-83] As such, implementation trials should include an explicit programme theory,[81] or a logic model that details the rationale and assumptions

Table 4 | Description of models, theories, and frameworks used in implementation strategy design. Adapted from Nilsen, 2015[85]

Theory or framework type	Description	Application
Classic theories (eg, theory of planned behaviour, social cognitive theory, situated change theory)[86-88]	Originate from related disciplines (eg, psychology) and help understand or explain individual, group, or organisational behaviour. They describe precise mechanisms of behaviour change	Classic and implementation theories describe precise mechanisms of behaviour and behaviour change. One or more of these theories can be used to developed targeted implementation strategies and describe how change in the behaviour of those involved in an implementation process is anticipated to occur
Implementation theories (eg, implementation climate, organisational readiness to change, normalisation process theory)[89-91]	Theories developed (or adapted classical theories) specifically to understand, explain, and inform implementation. They describe precise mechanisms of change for one or more aspect of implementation	
Determinants frameworks (eg, consolidated framework for implementation research, theoretical domains framework)[14 92]	Often developed through the consolidation of constructs from of a range of theories, they aim to understand and explain factors that could influence (facilitate or impede) implementation. They typically do not describe mechanisms for change	Determinants frameworks can help identify factors thought to be associated with implementation, and implementation strategies that can be used to address these, from which programme theory can be developed

about the mechanisms linking implementation strategy (and intervention),[84] processes, and inputs to trial outcomes. A programme theory can be developed using informal theory—that is, understanding of the problem and its determinants gained through experience or tacit knowledge by the developers of the intervention. However, we recommend that the use of informal theory is coupled with the formal behavioural or implementation theories or frameworks (table 4).[85] Although a range of theories and frameworks exist, few are supported empirically,[93] and some are known to be of little use in predicting or explaining behaviour.[94] Determinant frameworks can be particularly useful in implementation strategy development because they consolidate several behavioural theories and identify a comprehensive range of multilevel factors that are theoretically (or empirically) linked with implementation outcomes. In addition to the extent to which a theory or framework is empirically supported, criteria including usability, testability, familiarity, and applicability should be considered when comparing and selecting a model, theory, or framework.[95]

Several useful resources are available to support the application of formal theory in the development of broader programme models and specific implementation strategies.[96] French et al propose a four step process for such a development (table 5).[97] Other systematic methods for developing implementation strategies also exist,[99 100] which typically involve four common steps: barrier identification, linking barriers to implementation strategy component selection, use of theory, and user engagement.[99] Importantly, the development of programme theory and implementation strategies requires a thorough understanding of the problem, its determinants, and context in which implementation needs to occur and so should involve considerable end user engagement and formative evaluation.[100]

Measures

Trial outcome measures

The selection of outcome measures should be linked directly to trial primary and secondary aims and enable the robust quantification of an effect. Proctor and colleagues proposed a taxonomy of eight conceptually distinct implementation outcomes, namely acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration, and sustainability.[101] From a trial design perspective, the collective labelling of such measures as "outcomes," is a misnomer that has created some confusion,[102] because many of these measures do not lend themselves to the reporting of an effect size. For example, measures of the acceptability of an intervention (or implementation strategy) can only be reported in the trial group in receiving it, precluding between group comparisons. Many of these measures might be better aligned to the assessment of implementation processes and other factors influencing implementation.[42 102]

Most implementation trials primarily focus on measuring the extent to which an implementation strategy achieved implementation of the targeted evidence based intervention (eg, a guideline) such as measures of professional practice improvement, changes in processes of care, adherence to clinical standards, or the amount or quality of programme or intervention delivery.[7] As measures of such outcomes are often unique to the intervention being implemented and its context, generic standard measures are unlikely to be available. Instead, researchers might identify or develop measures that assesses their specific implementation outcome and context, for example, using data collected as part of environmental observations, routinely collected administrative records, or questionnaires. The limitations of each of these approaches need to be considered,[103] but as

Table 5 | Suggested steps for the development of a theory informed implementation strategy. Adapted from French et al, 2012[97]

Steps	Description
1	Identify who (eg, individuals or professional groups) needs to do what differently in order for implementation to be improved[98]
2	Using informal and formal theory and frameworks, identify barriers and enablers that need to be resolved, and articulate a pathway of change for the targeted behaviour change to occur. A variety of research methods, including literature reviews and local qualitative and quantitative data collection, should be used to support the development of the change pathway (programme theory)
3	Select implementation strategies (behaviour change techniques, modes of delivery) that might be effective, locally relevant, acceptable, and feasible to overcome identified barriers and enhance facilitators to change. Selection of strategies could be based on matrices recommended by determinant frameworks, empirical evidence, and engagement with end users
4	Decide how change in implementation can be robustly and feasibly measured, including factors on the hypothesised casual pathway (mediators) and appropriate implementation outcomes

Table 6 | Implementation measures used to establish and support evidence based interventions. Adapted from Proctor et al, 2011, with permission[101]

Measures	Description
Acceptability	Perception among implementation stakeholders that an evidence based intervention (or implementation strategy) is agreeable, palatable, or satisfactory
Adoption	Intention, initial decision, or action to try or use an evidence based intervention (or implementation strategy). Adoption also can be referred to as "uptake"
Appropriateness	Perceived fit, relevance, or compatibility of an evidence based intervention (or implementation strategy) for a given practice setting, provider, or consumer; or perceived fit of the innovation to resolve a particular issue or problem
Feasibility	Extent to which an evidence based intervention (or implementation strategy) can be successfully used or carried out
Fidelity	Degree to which an evidence based intervention (or implementation strategy) was delivered as it was intended
Cost (incremental or implementation cost)	Cost or relative cost of the implementation of an evidence based intervention
Penetration	Integration of an evidence based intervention within a service setting and its subsystems
Sustainability	Extent to which a newly implemented evidence based intervention is maintained or institutionalised within a service setting's ongoing, stable operations

trial outcomes, such measures should be robust, and sensitive to change. Multiple outcome measures should also be used in trials to provide a more comprehensive appraisal of the effects of an implementation strategy, acknowledging how these measures are related to each other and the inherent limitations of single measures of implementation.[42 103] For trials focused on assessment of individual patient level outcomes, clinical outcomes should be sufficiently proximal and arise exclusively (or mostly) from the improvements in clinical practice targeted by the implementation strategy.[104] For example, in a study to improve survival from heart attack, researchers noted that even if perfect compliance with care standards in a hospital could be achieved, the anticipated changes in cardiac mortality (or survival) would be insufficient to feasibly detect in a trial.[105]

Process evaluation

Process evaluation provides important depth to the interpretation of trial outcomes. Qualitative and mixed method approaches can elucidate insights to better understand how and why implementation might improve (or not) following the application of an implementation strategy, and key contextual factors that might influence it. Several publications, including a white paper by the Qualitative Research in Implementation Science (QualRIS) group (an expert group convened by the National Institute of Health), provide guidance for the use of qualitative methods in implementation science, including discussion of design, data collection, and analytical methods as well as recent developments in the field.[106 107] While several approaches have been suggested to undertake process evaluations,[108-111] here we offer guidance consistent with the United Kingdom's Medical Research Council, which suggests process evaluations include assessment of implementation processes, mechanism of impact, and contextual factors that shape outcomes.[112]

Implementation processes

Implementation processes are specific policies, practices, and strategies that are used to establish and support an intervention.[101] Table 6 provides a range of measures proposed by Proctor et al[101] that might be useful for exploring implementation processes. Such measures, for example, could be used to describe characteristics of the evidence based intervention, or the implementation strategy (table 6). The psychometric properties of a range of existing tools that assess these have recently been reported.[113 114] Additionally, because evidence based interventions are often adapted by end users (such as clinicians) in the process of their implementation, the documentation, recording, and reporting of adaptations has been suggested to be important to understanding the effects of efforts to implement evidence based interventions.[12] A framework by Stirman et al provides more detailed guidance of how to do so.[115] The use of qualitative inquiry has also been recommended by QualRIS to assess adaptation and other implementation processes while ethnography has been suggested to be well suited to assess implementation microprocesses at the level of individual interactions.[107]

Implementation mechanisms

The mechanism by which an implementation strategy exerts its effects is important to understand in order to identify how these effects might be replicated and improved.[112] To develop such an understanding, specific analytical methods can be applied to assess casual assumptions of the pathways specified by the programme theory.[116-119] Such mechanistic evaluations require clear specification of implementation strategies, links between strategy and mechanism, identification of outcomes, and (if relevant) articulation of effect modifiers.[119] Some classic theories, implementation theories, and determinants frameworks have existing measures of factors theoretically linked to implementation outcomes. Several reviews of such measures have been published,[120] of which the most comprehensive is the Instrument Review Project, funded by the National Institutes of Health.[13] Reviews, however, suggest that implementation mechanisms are rarely tested in trials of implementation strategies,[121 122] and where testing has occurred, often it is undertaken inappropriately. To best understand the multilevel and interdependence of factors that might influence implementation, sophisticated quantitative and qualitative methods are required.[123 124] Lewis and colleagues suggest that common quantitative approaches to mediation

testing in implementation trials are suboptimal, and that the product of coefficients approach might be preferable given its capacity to examine single level and multilevel mediation and maximise power.[122] Further, qualitative approaches have been suggested to be particularly useful in the absence of established quantitative measures, and structured qualitative inquiry can help deepen an understanding of mechanistic processes.[107 122] Contemporary guidance on mechanistic evaluation, including how it is applied in implementation science, is provided in more detail elsewhere.[122]

Implementation contexts

Context refers to external factors that might act as a barrier or facilitator to implementation, or influence the effects of an implementation strategy.[12 112] Descriptions of context, therefore, provide critical information regarding the external validity of trial findings and enable readers to assess the applicability of the findings to their own setting. Context measures can include measures of the social, political, or economic environment that might influence implementation.[12] These measures include leadership, workforce capacity, readiness to change, and other organisational or patient characteristics.[125] Some randomised implementation trials have also used systematic reviews of news archives, and of websites of relevant agencies to assess changes in government policy, guidelines, accreditation standards or funded programmes that might influence implementation or confound trial outcomes.[126 127] Quantitative or qualitative measures of context can also be assessed analytically to examine their potential role in shaping implementation processes or outcomes in the context of the broader programme theory.[42]

Sample size calculation

Sample size calculations estimate the number of participants required to detect the hypothesised effect of an implementation strategy with acceptable power.[128 129] While sample size calculations for clinical effectiveness trials are based on treatment effects identified as of sufficient magnitude to provide a clinical therapeutic benefit to a patient,[129] sample size calculations for implementation trials need to consider a meaningful or worthwhile effect size for an implementation outcome from a population or system level perspective. Because implementation strategies typically seek to improve the implementation of existing evidence based interventions of known therapeutic benefit, any improvement in implementation may increase the number of patients or the community exposed to (and benefiting from) evidence based healthcare. Strategies that lead to small improvements in implementation might be meaningful from a system perspective if they can be delivered, easily, at low cost, and at a population level. Sample size calculations need to use parameters required for the type of randomised design undertaken and researchers should follow design specific advice

to do so.[130] Because implementation trials can have participants at multiple levels, sample size calculations are usually more complicated than those for clinical effectiveness trials, and might need to consider the relative contributions to the power of increasing the numbers of participants at each level.

Research ethics review

As implementation trials meet the definition of research (a systematic investigation designed to produce generalisable knowledge) and involve human research participants (which could include health professionals),[131] ethical review by an institutional review board is required before trial commencement. Implementation trials can occur in the context of usual service improvement activities that can complicate the nature of consent for research participation.[132 133] Implementation trials often involve participants at multiple levels, so research ethics review is more complicated. Although no specific ethical statements exist pertaining to implementation trials,[133] the Ottawa Statement on the Ethical Design and Conduct of Cluster Randomised Trials covers such issues, and has recently been applied to trials of knowledge translation interventions.[134 135] The statement provides guidance to help identify research participants (patients, clinicians, and managers), and lists requirements for organisational governance, assessing benefits and harms, and protecting vulnerable participants (table 7). A key consideration when submitting a protocol to a research ethics committee is identifying the human research participants in the trial.[136] Research participants can be identified as any individual whose interests might be affected as a result of study interventions or data collection procedures.[136] In some implementation trials, patients might not be considered research participants (that is, they do not have any study interventions directed at them, or do not have their identifiable data collected for the purposes of research). When patients are not research participants, their informed consent is not required.[137] However, when employees such as clinicians are the recipients of an implementation strategy, and are involved in data collection or where identifiable data are collected about them, their consent is required. Approval might also be required from gatekeepers such as an organisational leader for such research to be undertaken in their facility.

Reporting

The Standards for Reporting Implementation Studies (StARI) guide has been designed specifically to facilitate the better reporting of implementation trials and should be used in conjunction with the CONSORT reporting guideline (and extension) specific to the type of randomised trial design used.[12] Efforts to test the effectiveness of implementation strategies have been hindered by a lack of conceptual clarity owing to inconsistent definitions and insufficient detail to enable replication.[9] To resolve this, StaRI recommend the use of the Template for Intervention Description

Table 7 | Selected ethical issues included in the Ottawa Statement on the Ethical Design and Conduct of Cluster Randomised Trials that are relevant to implementation trials. Adapted from Taljaard et al, 2013[134]

Ethical issue	Summary of recommendation
Identifying research participants	Research participants are any individuals who are the intended recipients of an implementation strategy (or control) or are the target of an experimental manipulation of their environment; investigators interact for the purpose of collecting data about that individual; or who provide personal data for research purposes. Participants could include clinicians, health service, or other staff, where implementation initiatives are occurring
Informed consent	Informed consent is required from all individuals who meet the criteria for research participants before data collection or intervention exposure unless a waiver is granted from an ethics review board. Waiver or alternate consent procedures may be granted when the research poses no more than minimal risk and the study is not feasible without the alteration of consent
Organisational governance approval	Where research might substantially affect organisations (or other cluster unit) interests, permission to undertake it should be sought from stakeholders who have legitimate authority to make decisions on behalf of the organisation. When research might substantially affect cluster interests, researchers should seek to protect cluster interests through cluster consultation with organisations (eg, gatekeepers) to inform study design conduct and reporting. Such organisational stakeholders may not provide consent on behalf of research participants
Assessing benefits and harms	Researchers must justify the intervention and data collection procedures, as well as the selection of the control condition. The research should not deny access to effective care or programmes to which would otherwise be accessible to patients or providers of care. Benefits and harms of participation must be considered, and stand in reasonable relation to anticipated knowledge gain
Protecting vulnerable participants	Additional protection might be needed for research, including vulnerable participant groups (eg, those unable to provide informed consent, at particular risk of harm, or in subordinate organisational or social positions)

and Replication (TIDieR) checklist when describing the evidence base intervention that is subject to implementation.[12 138] Similar recommendations have been proposed for standardising description of implementation strategies,[15] and implementation researchers should describe implementation strategies using an established taxonomy (eg, the Behaviour Change Technique or Expert Recommendations for Implementation Change taxonomies).[9 15 139 140] The identification of core and non-core components of the implementation strategy, based on the underlying programme theory, should also be articulated.

Conclusion

High quality randomised trials have a key role in advancing implementation science by providing robust evidence on the effects of approaches to improve the uptake and integration of evidence based practice. With the emergence of more accepted concepts, terminology, processes, and reporting standards in the field, the opportunity to improve the development, conduct, and reporting of such trials is considerable.[12-14] This article summarises the latest guidance on the best practice randomised trial and implementation science methods to fulfil this need for improvement. The development of guidance documents have proved a useful resource in improving the rigour of randomised controlled trials in healthcare and public health.[141] This guide is also aimed at journal editors, reviewers, and funders of implementation research as a resource to improve the quality of the implementation science evidence base.

Author affiliations
[1]School of Medicine and Public Health, Faculty of Health and Medicine, University of Newcastle, Callaghan, NSW, Australia

[2]Hunter New England Population Health, Locked Bag 10, Wallsend, NSW 2287, Australia

[3]Leeds Institute of Health Sciences, University of Leeds, Leeds, UK

[4]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

[5]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

[6]Department of Medicine, University of Ottawa, Ottawa, ON, Canada

[7]Women's College Research Institute, Women's College Hospital, Toronto, ON, Canada

[8]Institute for Health Systems Solutions and Virtual Care, Women's College Hospital, Toronto, ON, Canada

[9]Department of Family Medicine and Community Medicine, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[10]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[11]Brown School and School of Medicine, Washington University in St Louis, St Louis, MI, USA

[12]School of Public Health, Faculty of Medicine and Health, University of Sydney, Sydney, NSW, Australia

[13]Swinburne University of Technology, School of Health Sciences, Faculty Health, Arts and Design, Hawthorn, VIC, Australia

discrepancies from the study as planned (and, if relevant, registered) have been explained.

Patient and public involvement: Patients and the public were not involved during the process of this research.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Lenfant C. Shattuck lecture--clinical research to clinical practice--lost in translation? *N Engl J Med* 2003;349:868-74. doi:10.1056/NEJMsa035507

2 Wolfenden L, Ziersch A, Robinson P, Lowe J, Wiggers J. Reducing research waste and improving research impact. *Aust N Z J Public Health* 2015;39:303-4. doi:10.1111/1753-6405.12467

3 National Academies of Sciences, Engineering, and Medicine. *Crossing the global quality chasm: improving health care worldwide* . National Academies Press2018.

4 Eccles MP, Mittman BS. *Welcome to implementation science.* BioMed Central, 2006:1.

5 Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. *BMC Psychol* 2015;3:32. doi:10.1186/s40359-015-0089-9

6 Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312:71-2. doi:10.1136/bmj.312.7023.71

7 Wolfenden L, Reilly K, Kingsland M, et al. Identifying opportunities to develop the science of implementation for community-based non-communicable disease prevention: a review of implementation trials. *Prev Med* 2019;118:279-85. doi:10.1016/j.ypmed.2018.11.014

8 Powell BJ, McMillen JC, Proctor EK, et al. A compilation of strategies for implementing clinical innovations in health and mental health. *Med Care Res Rev* 2012;69:123-57. doi:10.1177/1077558711430690

9 Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci* 2015;10:21. doi:10.1186/s13012-015-0209-1

10 McKibbon KA, Lokker C, Wilczynski NL, et al. A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: a Tower of Babel? *Implement Sci* 2010;5:16. doi:10.1186/1748-5908-5-16

11 Hodder RK, Wolfenden L, Kamper SJ, et al. Developing implementation science to improve the translation of research to address low back pain: a critical review. *Best Pract Res Clin Rheumatol* 2016;30:1050-73. doi:10.1016/j.berh.2017.05.002

12 Pinnock H, Barwick M, Carpenter CR, et al, StaRI Group. Standards for reporting implementation studies (StaRI) statement. *BMJ* 2017;356:i6795. doi:10.1136/bmj.i6795

13 Lewis CC, Stanick CF, Martinez RG, et al. The Society For Implementation Research Collaboration Instrument Review Project: a methodology to promote rigorous evaluation [correction: *Implement Sci* 2020;15:3]. *Implement Sci* 2015;10:2. doi:10.1186/s13012-014-0193-x

14 Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009;4:50. doi:10.1186/1748-5908-4-50

15 Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. *Implement Sci* 2013;8:139. doi:10.1186/1748-5908-8-139

16 Upvall MJ, Bourgault AM. De-implementation: a concept analysis. *Nurs Forum* 2018. doi:10.1111/nuf.12256

17 US Department of Health & Human Services. *Implementation science at a glance: a guide for cancer control practitioners* . National Cancer Institute, 2019:43-4.

18 World Health Organisation. Clinical trials. 2018. https://www.who.int/health-topics/clinical-trials/#tab=tab_1

19 Wolfenden L, Kingsland M, Yoong SL, et al. Improving the impact of public health service delivery and research: a decision tree to aid evidence-based public health practice and research. *ANZJPH* 2020. doi:10.1111/1753-6405.13023.

20 Lane-Fall MB, Curran GM, Beidas RS. Scoping implementation science for the beginner: locating yourself on the "subway line" of translational research. *BMC Med Res Methodol* 2019;19:133. doi:10.1186/s12874-019-0783-z

21 Forsythe LP, Ellis LE, Edmundson L, et al. Patient and stakeholder engagement in the PCORI pilot projects: description and lessons learned. *J Gen Intern Med* 2016;31:13-21. doi:10.1007/s11606-015-3450-z

22 Forsythe L, Heckert A, Margolis MK, Schrandt S, Frank L. Methods and impact of engagement in research, from theory to practice and back again: early findings from the Patient-Centered Outcomes Research Institute. *Qual Life Res* 2018;27:17-31. doi:10.1007/s11136-017-1581-x

23 Gray-Burrows KA, Willis TA, Foy R, et al. Role of patient and public involvement in implementation research: a consensus study. *BMJ Qual Saf* 2018;27:858-64. doi:10.1136/bmjqs-2017-006954

24 Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. *Health Res Policy Syst* 2020;18:17. doi:10.1186/s12961-020-0528-9

25 Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care* 2012;50:217-26. doi:10.1097/MLR.0b013e3182408812

26 Presseau J, Mackintosh J, Hawthorne G, et al. Cluster randomised controlled trial of a theory-based multiple behaviour change intervention aimed at healthcare professionals to improve their management of type 2 diabetes in primary care. *Implement Sci* 2018;13:65. doi:10.1186/s13012-018-0754-5

27 Mc Cord KA, Al-Shahi Salman R, Treweek S, et al. Routinely collected data for randomized trials: promises, barriers, and implications. *Trials* 2018;19:29. doi:10.1186/s13063-017-2394-5

28 Willis TA, Collinson M, Glidewell L, et al, ASPIRE programme team. An adaptable implementation package targeting evidence-based indicators in primary care: a pragmatic cluster-randomised evaluation. *PLoS Med* 2020;17:e1003045. doi:10.1371/journal.pmed.1003045

29 Lord PA, Willis TA, Carder P, West RM, Foy R. Optimizing primary care research participation: a comparison of three recruitment methods in data-sharing studies. *Fam Pract* 2016;33:200-4. doi:10.1093/fampra/cmw003

30 Treweek S, Lockhart P, Pitkethly M, et al. Methods to improve recruitment to randomised controlled trials: Cochrane systematic review and meta-analysis. *BMJ Open* 2013;3:e002360. doi:10.1136/bmjopen-2012-002360

31 Treweek S, Pitkethly M, Cook J, et al. Strategies to improve recruitment to randomised trials. *Cochrane Database Syst Rev* 2018;2:MR000013. doi:10.1002/14651858.MR000013.pub6

32 Ivers NM, Schwalm J-D, Bouck Z, et al. Interventions supporting long term adherence and decreasing cardiovascular events after myocardial infarction (ISLAND): pragmatic randomised controlled trial. *BMJ* 2020;369:m1731. doi:10.1136/bmj.m1731

33 Foy R, Penney GC, Grimshaw JM, et al. A randomised controlled trial of a tailored multifaceted strategy to promote implementation of a clinical guideline on induced abortion care. *BJOG* 2004;111:726-33. doi:10.1111/j.1471-0528.2004.00168.x

34 Eccles M, Steen N, Grimshaw J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. *Lancet* 2001;357:1406-9. doi:10.1016/S0140-6736(00)04564-5

35 Sackett DL. Explanatory and pragmatic clinical trials: a primer and application to a recent asthma trial. *Pol Arch Med Wewn* 2011;121:259-63. doi:10.20452/pamw.1071

36 Yoong SL, Wolfenden L, Clinton-McHarg T, et al. Exploring the pragmatic and explanatory study design on outcomes of systematic reviews of public health interventions: a case study on obesity prevention trials. *J Public Health (Oxf)* 2014;36:170-6. doi:10.1093/pubmed/fdu006

37 Finch M, Jones J, Yoong S, Wiggers J, Wolfenden L. Effectiveness of centre-based childcare interventions in increasing child physical activity: a systematic review and meta-analysis for policymakers and practitioners. *Obes Rev* 2016;17:412-28. doi:10.1111/obr.12392

38 Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147. doi:10.1136/bmj.h2147

39 Simon GE, Shortreed SM, Rossom RC, Penfold RB, Sperl-Hillen JAM, O'Connor P. Principles and procedures for data and safety monitoring in pragmatic clinical trials. *Trials* 2019;20:690. doi:10.1186/s13063-019-3869-3

40 Deeks JJ, Dinnes J, D'Amico R, et al, International Stroke Trial Collaborative Group, European Carotid Surgery Trial Collaborative Group. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003;7:iii-x, 1-173. doi:10.3310/hta7270

41 Gerstein HC, McMurray J, Holman RR. Real-world studies no substitute for RCTs in establishing efficacy. *Lancet* 2019;393:210-1. doi:10.1016/S0140-6736(18)32840-X

42 Wensing M, Grol R, Grimshaw J, eds. *Improving patient care. the implementation of change in health care.* 3rd ed. Wiley Blackwell, 2020:352. doi:10.1002/9781119488620

43 Brainard J, Hunter PR. Do complexity-informed health interventions work? A scoping review. *Implement Sci* 2016;11:127. doi:10.1186/s13012-016-0492-5

44 Simpson KM, Porter K, McConnell ES, et al. Tool for evaluating research implementation challenges: a sense-making protocol for addressing implementation challenges in complex research settings. *Implement Sci* 2013;8:2. doi:10.1186/1748-5908-8-2

45 Greenhalgh T, Papoutsi C. Studying complexity in health services research: desperately seeking an overdue paradigm shift. *BMC Med* 2018;16:95. doi:10.1186/s12916-018-1089-4

46 Bonell C, Fletcher A, Morton M, Lorenc T, Moore L. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Soc Sci Med* 2012;75:2299-306. doi:10.1016/j.socscimed.2012.08.032

47 Braithwaite J, Churruca K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation science: a theoretical and empirical analysis of systems change. *BMC Med* 2018;16:63. doi:10.1186/s12916-018-1057-z

48 Kabisch M, Ruckes C, Seibert-Grafe M, Blettner M. Randomized controlled trials: part 17 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 2011;108:663-8.

49 Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2019;321:1610-20. doi:10.1001/jama.2019.3087

50 Hartley S, Foy R, Walwyn REA, et al, AFFINITIE programme. The evaluation of enhanced feedback interventions to reduce unnecessary blood transfusions (AFFINITIE): protocol for two linked cluster randomised factorial controlled trials. *Implement Sci* 2017;12:84. doi:10.1186/s13012-017-0614-8

51 Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol* 2006;6:54. doi:10.1186/1471-2288-6-54

52 Liddy C, Hogg W, Singh J, et al. A real-world stepped wedge cluster randomized trial of practice facilitation to improve cardiovascular care. *Implement Sci* 2015;10:150. doi:10.1186/s13012-015-0341-y

53 Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A "SMART" design for building individualized treatment sequences. *Annu Rev Clin Psychol* 2012;8:21-48. doi:10.1146/annurev-clinpsy-032511-143152

54 Johnson JE, Wiltsey-Stirman S, Sikorskii A, et al. Protocol for the ROSE sustainment (ROSES) study, a sequential multiple assignment randomized trial to determine the minimum necessary intervention to maintain a postpartum depression prevention program in prenatal clinics serving low-income women. *Implement Sci* 2018;13:115. doi:10.1186/s13012-018-0807-9

55 Cook TD, Campbell DT, Shadish W. *Experimental and quasi-experimental designs for generalized causal inference.* Houghton Mifflin, 2002.

56 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898. doi:10.1136/bmj.l4898

57 NSW Ministy of Health. *Study design for evaluating population health and health service interventions: a guide.* Evidence and Evaluation Guidance Series, Population and Public Health Division, 2019.

58 Campbell MK, Mollison J, Grimshaw JM. Cluster trials in implementation research: estimation of intracluster correlation coefficients and sample size. *Stat Med* 2001;20:391-9. doi:10.1002/1097-0258(20010215)20:3‹391::AID-SIM800›3.0.CO;2-Z

59 Hewitt CE, Torgerson DJ, Miles JN. Individual allocation had an advantage over cluster randomization in statistical efficiency in some circumstances. *J Clin Epidemiol* 2008;61:1004-8. doi:10.1016/j.jclinepi.2007.12.002

60 Wolfenden L, Nathan NK, Sutherland R, et al. Strategies for enhancing the implementation of school-based policies or practices targeting risk factors for chronic disease. *Cochrane Database Syst Rev* 2017;11:CD011677. doi:10.1002/14651858.CD011677.pub2

61 Wolfenden L, Barnes C, Jones J, et al. Strategies to improve the implementation of healthy eating, physical activity and obesity prevention policies, practices or programmes within childcare services. *Cochrane Database Syst Rev* 2020;2:CD011779. doi:10.1002/14651858.CD011779.pub3

62 Brown CH, Curran G, Palinkas LA, et al. An overview of research and evaluation designs for dissemination and implementation. *Annu Rev Public Health* 10.1146/annurev-publhealth-031816-044215

63 Chakraborty B, Collins LM, Strecher VJ, Murphy SA. Developing multicomponent interventions using fractional factorial designs. *Stat Med* 2009;28:2687-708. doi:10.1002/sim.3643

64 Hooper R, Eldridge SM. Cutting edge or blunt instrument: how to decide if a stepped wedge design is right for you. *BMJ Qual Saf* 2020. doi:10.1136/bmjqs-2020-011620

65 Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* 2015;350:h391. doi:10.1136/bmj.h391

66 Hemming K, Taljaard M, McKenzie JE, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. *BMJ* 2018;363:k1614. doi:10.1136/bmj.k1614

67 Hemming K, Taljaard M. Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice?*Int J Epidemiol* 2020;49:1043-52. doi:10.1093/ije/dyaa077

68 Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. *Transl Behav Med* 2014;4:260-74. doi:10.1007/s13142-014-0265-0

69 Taljaard M, Teerenstra S, Ivers NM, Fergusson DA. Substantial risks associated with few clusters in cluster randomized and stepped wedge designs. *Clin Trials* 2016;13:459-63. doi:10.1177/1740774516634316

70 Miller CJ, Smith SN, Pugatch M. Experimental and quasi-experimental designs in implementation research. *Psychiatry Res* 2020;283:112452. doi:10.1016/j.psychres.2019.06.027

71 Eldridge S, Kerry S, Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done?*BMJ* 2009;339:b4006. doi:10.1136/bmj.b4006

72 Eldridge S, Ashby D, Bennett C, Wakelin M, Feder G. Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ* 2008;336:876-80. doi:10.1136/bmj.39517.495764.25

73 Farrin A, Russell I, Torgerson D, Underwood M, UK BEAM Trial Team. Differential recruitment in a cluster randomized trial in primary care: the experience of the UK back pain, exercise, active management and manipulation (UK BEAM) feasibility study. *Clin Trials* 2005;2:119-24. doi:10.1191/1740774505cn073oa

74 Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med* 2009;6:e1000065. doi:10.1371/journal.pmed.1000065

75 Robinson K, Allen F, Darby J, et al. Contamination in complex healthcare trials: the falls in care homes (FinCH) study experience. *BMC Med Res Methodol* 2020;20:46. doi:10.1186/s12874-020-00925-z

76 Magill N, Knight R, McCrone P, Ismail K, Landau S. A scoping review of the problems and solutions associated with contamination in trials of complex interventions in mental health. *BMC Med Res Methodol* 2019;19:4. doi:10.1186/s12874-018-0646-z

77 Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med* 1997;16:1017-29. doi:10.1002/(SICI)1097-0258(19970515)16:9‹1017::AID-SIM508›3.0.CO;2-V

78 Sussman JB, Hayward RA. An IV for the RCT: using instrumental variables to adjust for treatment contamination in randomised controlled trials. *BMJ* 2010;340:c2073. doi:10.1136/bmj.c2073

79 Keogh-Brown MR, Bachmann MO, Shepstone L, et al. Contamination in trials of educational interventions. *Health Technol Assess* 2007;11:iii, ix-107. doi:10.3310/hta11430

80 Birken SA, Powell BJ, Shea CM, et al. Criteria for selecting implementation science theories and frameworks: results from an international survey. *Implement Sci* 2017;12:124. doi:10.1186/s13012-017-0656-y

81 Davidoff F, Dixon-Woods M, Leviton L, Michie S. Demystifying theory and its use in improvement. *BMJ Qual Saf* 2015;24:228-38. doi:10.1136/bmjqs-2014-003627

82 Renger R, Titcomb A. A three-step approach to teaching logic models. *Am J Eval* 2002;23:493-503. doi:10.1177/109821400202300409

83 Eccles M. The Improved Clinical Effectiveness through Behavioural Research Group (ICEBeRG). Designing theoretically-informed implementation interventions. *Implement Sci* 2006;1:4. doi:10.1186/1748-5908-1-4

84 NSW Ministry of Health. *Centre for Epidemiology and Evidence. Developing and Using Program Logic: A Guide. Evidence and Evaluation Guidance Series.* Ministry of Health, Population and Public Health Division, 2017: 4-16.

85 Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci* 2015;10:53. doi:10.1186/s13012-015-0242-0

86 Ajzen I. *Attitudes, personality, and behaviour* . McGraw-Hill Education (UK), 2005;1-144.

87 Orlikowski WJ. Improvising organizational transformation over time: A situated change perspective. *Inf Syst Res* 1996;7:63-92. doi:10.1287/isre.7.1.63

88 Bandura A. *Social foundations of thought and action* . Prentice Hall, 1986.

89 Klein KJ, Sorra JS. The challenge of innovation implementation. *Acad Manage Rev* 1996;21:1055-80. doi:10.5465/amr.1996.9704071863

90 Weiner BJ. A theory of organizational readiness for change. *Implement Sci* 2009;4:67. doi:10.1186/1748-5908-4-67

91 May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociol* 2009;43:535-54. doi:10.1177/0038038509103208

92 Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implement Sci* 2012;7:37. doi:10.1186/1748-5908-7-37

93 Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. *Am J Prev Med* 2012;43:337-50. doi:10.1016/j.amepre.2012.05.024

94 Sniehotta FF, Presseau J, Araújo-Soares V. *Time to retire the theory of planned behaviour.* Health Psych Rev, 2014:1-7.

95 Birken SA, Rohweder CL, Powell BJ, et al. T-CaST: an implementation theory comparison and selection tool. *Implement Sci* 2018;13:143. doi:10.1186/s13012-018-0836-4

96 Moullin JC, Dickson KS, Stadnick NA, et al. Ten recommendations for using implementation frameworks in research and practice. *Implement Sci Commun* 2020;1:42. doi:10.1186/s43058-020-00023-7

97 French SD, Green SE, O'Connor DA, et al. Developing theory-informed behaviour change interventions to implement evidence into practice: a systematic approach using the Theoretical Domains Framework. *Implement Sci* 2012;7:38. doi:10.1186/1748-5908-7-38

98 Presseau J, McCleary N, Lorencatto F, Patey AM, Grimshaw JM, Francis JJ. Action, actor, context, target, time (AACTT): a framework for specifying behaviour. *Implement Sci* 2019;14:102. doi:10.1186/s13012-019-0951-x

99 Colquhoun HL, Squires JE, Kolehmainen N, Fraser C, Grimshaw JM. Methods for designing interventions to change healthcare professionals' behaviour: a systematic review. *Implement Sci* 2017;12:30. doi:10.1186/s13012-017-0560-5

100 Powell BJ, Beidas RS, Lewis CC, et al. Methods to improve the selection and tailoring of implementation strategies. *J Behav Health Serv Res* 2017;44:177-94. doi:10.1007/s11414-015-9475-6

101 Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011;38:65-76. doi:10.1007/s10488-010-0319-7

102 McKay H, Naylor P-J, Lau E, et al. Implementation and scale-up of physical activity and behavioural nutrition interventions: an evaluation roadmap. *Int J Behav Nutr Phys Act* 2019;16:102. doi:10.1186/s12966-019-0868-4

103 Smith JD, Hasan M. Quantitative approaches for the evaluation of implementation research studies. *Psychiatry Res* 2020;283:112521. doi:10.1016/j.psychres.2019.112521

104 Brown C, Hofer T, Johal A, et al. An epistemology of patient safety research: a framework for study design and interpretation. Part 3. End points and measurement. *Qual Saf Health Care* 2008;17:170-7. doi:10.1136/qshc.2007.023655

105 Mant J, Hicks N. Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *BMJ* 1995;311:793-6. doi:10.1136/bmj.311.7008.793

106 Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. *Psychiatry Res* 2019;280:112516. doi:10.1016/j.psychres.2019.112516

107 US Department of Health and Human Services. *Qualitative methods in implementation science. National Institutes of Health Bethesda.* National Cancer Institute, 2018:1-20.

108 Grant A, Treweek S, Dreischulte T, Foy R, Guthrie B. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials* 2013;14:15. doi:10.1186/1745-6215-14-15

109 Hawe P, Shiell A, Riley T, Gold L. Methods for exploring implementation variation and local context within a cluster randomised community intervention trial. *J Epidemiol Community Health* 2004;58:788-93. doi:10.1136/jech.2003.014415

110 Murtagh MJ, Thomson RG, May CR, et al. Qualitative methods in a randomised controlled trial: the role of an integrated qualitative process evaluation in providing evidence to discontinue the intervention in one arm of a trial of a decision support tool. *Qual Saf Health Care* 2007;16:224-9. doi:10.1136/qshc.2006.018499

111 Byng R, Norman I, Redfern S, Jones R. Exposing the key functions of a complex intervention for shared care in mental health: case study of a process evaluation. *BMC Health Serv Res* 2008;8:274. doi:10.1186/1472-6963-8-274

112 Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;350:h1258. doi:10.1136/bmj.h1258

113 Lewis CC, Mettert KD, Dorsey CN, et al. An updated protocol for a systematic review of implementation-related measures. *Syst Rev* 2018;7:66. doi:10.1186/s13643-018-0728-3

114 Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. *Implement Sci* 2017;12:108. doi:10.1186/s13012-017-0635-3

115 Stirman SW, Miller CJ, Toder K, Calloway A. Development of a framework and coding system for modifications and adaptations of evidence-based interventions. *Implement Sci* 2013;8:65. doi:10.1186/1748-5908-8-65

116 Grimshaw JM, Zwarenstein M, Tetroe JM, et al. Looking inside the black box: a theory-based process evaluation alongside a randomised controlled trial of printed educational materials (the Ontario printed educational message, OPEM) to improve referral and prescribing practices in primary care in Ontario, Canada. *Implement Sci* 2007;2:38. doi:10.1186/1748-5908-2-38

117 Presseau J, Grimshaw JM, Tetroe JM, et al. A theory-based process evaluation alongside a randomised controlled trial of printed educational messages to increase primary care physicians' prescription of thiazide diuretics for hypertension [ISRCTN72772651]. *Implement Sci* 2016;11:121. doi:10.1186/s13012-016-0485-4

118 Grimshaw JM, Presseau J, Tetroe J, et al. Looking inside the black box: results of a theory-based process evaluation exploring the results of a randomized controlled trial of printed educational messages to increase primary care physicians' diabetic retinopathy referrals [Trial registration number ISRCTN72772651]. *Implement Sci* 2014;9:86. doi:10.1186/1748-5908-9-86

119 Lewis CC, Klasnja P, Powell BJ, et al. From classification to causality: Advancing Understanding of Mechanisms of change in implementation science. *Front Public Health* 2018;6:136. doi:10.3389/fpubh.2018.00136

120 Clinton-McHarg T, Yoong SL, Tzelepis F, et al. Psychometric properties of implementation measures for public health and community settings and mapping of constructs against the Consolidated Framework for Implementation Research: a systematic review. *Implement Sci* 2016;11:148. doi:10.1186/s13012-016-0512-5

121 McIntyre SA, Francis JJ, Gould NJ, Lorencatto F. The use of theory in process evaluations conducted alongside randomized trials of implementation interventions: a systematic review. *Transl Behav Med* 2020;10:168-78.

122 Lewis CC, Boyd MR, Walsh-Bailey C, et al. A systematic review of empirical studies examining mechanisms of implementation in health. *Implement Sci* 2020;15:21. doi:10.1186/s13012-020-00983-3

123 Lapointe-Shaw L, Bouck Z, Howell NA, et al. Mediation analysis with a time-to-event outcome: a review of use and reporting in healthcare research. *BMC Med Res Methodol* 2018;18:118. doi:10.1186/s12874-018-0578-7

124 Whitaker RG, Sperber N, Baumgartner M, et al. *Coincidence Analysis: A New Method for Causal Inference in Implementation Science.* Research Sqaure, 2020: 1-18.

125 Squires JE, Aloisio LD, Grimshaw JM, et al. Attributes of context relevant to healthcare professionals' use of research evidence in clinical practice: a multi-study analysis. *Implement Sci* 2019;14:52. doi:10.1186/s13012-019-0900-8

126 Jones J, Wyse R, Finch M, et al. Effectiveness of an intervention to facilitate the implementation of healthy eating and physical activity policies and practices in childcare services: a randomised controlled trial. *Implement Sci* 2015;10:147. doi:10.1186/s13012-015-0340-z

127 Hamilton AB, Mittman BS, Campbell D, et al. Understanding the impact of external context on community-based implementation of an evidence-based HIV risk reduction intervention. *BMC Health Serv Res* 2018;18:11. doi:10.1186/s12913-017-2791-1

128 Hickey GL, Grant SW, Dunning J, Siepe M. Statistical primer: sample size and power calculations-why, when and how?*Eur J Cardiothorac Surg* 2018;54:4-9. doi:10.1093/ejcts/ezy169

129 Moher D, Hopewell S, Schulz KF, et al, CONSORT. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg* 2012;10:28-55. doi:10.1016/j.ijsu.2011.10.001

130 Rutterford C, Copas A, Eldridge S. Methods for sample size determination in cluster randomized trials. *Int J Epidemiol* 2015;44:1051-67. doi:10.1093/ije/dyv113

131 Goldstein CE, Weijer C, Brehaut JC, et al. Accommodating quality and service improvement research within existing ethical principles. *Trials* 2018;19:334. doi:10.1186/s13063-018-2724-2

132 DuBois JM, Prusaczyk B. Chapter 4: Ethical issues in dissemination and implementation research. Brownson RC, Colditz GA, Proctor EK, eds. In: *Dissemination and implementation research in health: translating science to practice* (2nd ed). Oxford University Press, 2018;1:63.

133 Hutton JL, Eccles MP, Grimshaw JM. Ethical issues in implementation research: a discussion of the problems in achieving informed consent. *Implement Sci* 2008;3:52. doi:10.1186/1748-5908-3-52

134 Taljaard M, Weijer C, Grimshaw JM, Eccles MP, Ottawa Ethics of Cluster Randomised Trials Consensus Group. The Ottawa Statement on the ethical design and conduct of cluster randomised trials: precis for researchers and research ethics committees. *BMJ* 2013;346:f2838. doi:10.1136/bmj.f2838

135 Straus S, Tetroe J, Graham ID. *Knowledge translation in health care: moving from evidence to practice*. John Wiley & Sons, 2013. doi:10.1002/9781118413555

136 McRae AD, Weijer C, Binik A, et al. Who is the research subject in cluster randomized trials in health research?*Trials* 2011;12:183. doi:10.1186/1745-6215-12-183

137 McRae AD, Weijer C, Binik A, et al. When is informed consent required in cluster randomized trials in health research?*Trials* 2011;12:202. doi:10.1186/1745-6215-12-202

138 Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:g1687. doi:10.1136/bmj.g1687

139 Effective Practice and Organisation of Care. EPOC taxonomy 2015. epoc.cochrane.org/epoc-taxonomy.

140 Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013;46:81-95. doi:10.1007/s12160-013-9486-6

141 Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006;185:263-7. doi:10.5694/j.1326-5377.2006.tb00557.x

Web appendix: Supplementary file 1