



OPEN ACCESS



A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance

Kathryn Skivington,¹ Lynsay Matthews,¹ Sharon Anne Simpson,¹ Peter Craig,¹ Janis Baird,² Jane M Blazeby,³ Kathleen Anne Boyd,⁴ Neil Craig,⁵ David P French,⁶ Emma McIntosh,⁴ Mark Petticrew,⁷ Jo Rycroft-Malone,⁸ Martin White,⁹ Laurence Moore¹

For numbered affiliations see end of the article.

Correspondence to:

K Skivington
Kathryn.skivington@glasgow.ac.uk
(ORCID 0000-0002-3571-1561)

Cite this as: *BMJ* 2021;374:n2061
<http://dx.doi.org/10.1136/bmj.n2061>

Accepted: 9 August 2021

The UK Medical Research Council's widely used guidance for developing and evaluating complex interventions has been replaced by a new framework, commissioned jointly by the Medical Research Council and the National Institute for Health Research, which takes account of recent developments in theory and methods and the need to maximise the efficiency, use, and impact of research.

Complex interventions are commonly used in the health and social care services, public health practice, and other areas of social and economic policy that have consequences for health. Such interventions are delivered and evaluated at different levels, from individual to societal levels. Examples include a new surgical procedure, the redesign of a healthcare programme, and a change in welfare policy. The UK Medical Research Council (MRC) published a framework for researchers and research funders on developing and evaluating complex interventions in 2000 and revised guidance in 2006.¹⁻³ Although these documents continue to be widely used and are now accompanied by a range of more detailed guidance on specific aspects of the research process,⁴⁻⁸ several important conceptual, methodological and theoretical developments have taken place since 2006. These developments have been included in a new framework commissioned by the National Institute of Health Research (NIHR) and the MRC.⁹ The framework aims to help researchers work with other stakeholders to identify the key questions about complex interventions, and to design and conduct research with a diversity of perspectives and appropriate choice of methods.

Development of the Framework for Developing and Evaluating Complex Interventions

The updated Framework for Developing and Evaluating Complex Interventions is the culmination of a process that included four stages:

- A gap analysis to identify developments in the methods and practice since the previous framework was published
- A full day expert workshop, in May 2018, of 36 participants to discuss the topics identified in the gap analysis
- An open consultation on a draft of the framework in April 2019, whereby we sought stakeholder opinion by advertising via social media, email lists and other networks for written feedback (52 detailed responses were received from stakeholders internationally)
- Redraft using findings from the previous stages, followed by a final expert review.

We also sought stakeholder views at various interactive workshops throughout the development of the framework: at the annual meetings of the Society for Social Medicine and Population Health (2018), the UK Society for Behavioural Medicine (2017, 2018), and internationally at the International Congress of Behavioural Medicine (2018). The entire process was

SUMMARY POINTS

Complex intervention research can take an efficacy, effectiveness, theory based, and/or systems perspective, the choice of which is based on what is known already and what further evidence would add most to knowledge

Complex intervention research goes beyond asking whether an intervention works in the sense of achieving its intended outcome—to asking a broader range of questions (eg, identifying what other impact it has, assessing its value relative to the resources required to deliver it, theorising how it works, taking account of how it interacts with the context in which it is implemented, how it contributes to system change, and how the evidence can be used to support real world decision making)

A trade-off exists between precise unbiased answers to narrow questions and more uncertain answers to broader, more complex questions; researchers should answer the questions that are most useful to decision makers rather than those that can be answered with greater certainty

Complex intervention research can be considered in terms of phases, although these phases are not necessarily sequential: development or identification of an intervention, assessment of feasibility of the intervention and evaluation design, evaluation of the intervention, and impactful implementation

At each phase, six core elements should be considered to answer the following questions:

- How does the intervention interact with its context?
- What is the underpinning programme theory?
- How can diverse stakeholder perspectives be included in the research?
- What are the key uncertainties?
- How can the intervention be refined?
- What are the comparative resource and outcome consequences of the intervention?

The answers to these questions should be used to decide whether the research should proceed to the next phase, return to a previous phase, repeat a phase, or stop

overseen by a scientific advisory group representing the range of relevant NIHR programmes and MRC population health investments. The framework was reviewed by the MRC-NIHR Methodology Research Programme Advisory Group and then approved by the MRC Population Health Sciences Group in March 2020 before undergoing further external peer and editorial review through the NIHR Journals Library peer review process. More detailed information and the methods used to develop this new framework are described elsewhere.⁹ This article introduces the framework and summarises the main messages for producers and users of evidence.

What are complex interventions?

An intervention might be considered complex because of properties of the intervention itself, such as the number of components involved; the range of behaviours targeted; expertise and skills required by those delivering and receiving the intervention; the number of groups, settings, or levels targeted; or the permitted level of flexibility of the intervention or its components. For example, the Links Worker Programme was an intervention in primary care in Glasgow, Scotland, that aimed to link people with community resources to help them “live well” in their communities. It targeted individual, primary care (general practitioner (GP) surgery), and community levels. The intervention was flexible in that it could differ between primary care GP surgeries. In addition, the Link Workers did not support just one specific health or wellbeing issue: bereavement, substance use, employment, and learning difficulties were all included.^{10 11} The complexity of this intervention had implications for many aspects of its evaluation, such as the choice of appropriate outcomes and processes to assess.

Flexibility in intervention delivery and adherence might be permitted to allow for variation in how, where, and by whom interventions are delivered and received. Standardisation of interventions could relate more to the underlying process and functions of the intervention than on the specific form of components delivered.¹² For example, in surgical trials, protocols can be designed with flexibility for intervention delivery.¹³ Interventions require a theoretical deconstruction into components and then agreement about permissible and prohibited variation in the delivery of those components. This approach allows implementation of a complex intervention to vary across different contexts yet maintain the integrity of the core intervention components. Drawing on this approach in the ROMIO pilot trial, core components of minimally invasive oesophagectomy were agreed and subsequently monitored during main trial delivery using photography.¹⁴

Complexity might also arise through interactions between the intervention and its context, by which we mean “any feature of the circumstances in which an intervention is conceived, developed, implemented and evaluated.”^{6 15-17} Much of the criticism of and

extensions to the existing framework and guidance have focused on the need for greater attention on understanding how and under what circumstances interventions bring about change.^{7 15 18} The importance of interactions between the intervention and its context emphasises the value of identifying mechanisms of change, where mechanisms are the causal links between intervention components and outcomes; and contextual factors, which determine and shape whether and how outcomes are generated.¹⁹

Thus, attention is given not only to the design of the intervention itself but also to the conditions needed to realise its mechanisms of change and/or the resources required to support intervention reach and impact in real world implementation. For example, in a cluster randomised trial of ASSIST (a peer led, smoking prevention intervention), researchers found that the intervention worked particularly well in cohesive communities that were served by one secondary school where peer supporters were in regular contact with their peers—a key contextual factor consistent with diffusion of innovation theory, which underpinned the intervention design.²⁰ A process evaluation conducted alongside a trial of robot assisted surgery identified key contextual factors to support effective implementation of this procedure, including engaging staff at different levels and surgeons who would not be using robot assisted surgery, whole team training, and an operating theatre of suitable size.²¹

With this framing, complex interventions can helpfully be considered as events in systems.¹⁶ Thinking about systems helps us understand the interaction between an intervention and the context in which it is implemented in a dynamic way.²² Systems can be thought of as complex and adaptive,²³ characterised by properties such as emergence, feedback, adaptation, and self-organisation (table 1).

For complex intervention research to be most useful to decision makers, it should take into account the complexity that arises both from the intervention's components and from its interaction with the context in which it is being implemented.

Research perspectives

The previous framework and guidance were based on a paradigm in which the salient question was to identify whether an intervention was effective. Complex intervention research driven primarily by this question could fail to deliver interventions that are implementable, cost effective, transferable, and scalable in real world conditions. To deliver solutions for real world practice, complex intervention research requires strong and early engagement with patients, practitioners, and policy makers, shifting the focus from the “binary question of effectiveness”²⁶ to whether and how the intervention will be acceptable, implementable, cost effective, scalable, and transferable across contexts. In line with a broader conception of complexity, the scope of complex intervention research needs to include the development, identification, and evaluation of whole system interventions and the

Table 1 | Properties and examples of complex adaptive systems

System properties	Example
Emergence	
Complex systems have emergent, often unanticipated, properties that are a feature of the system as a whole	Group based interventions that target young people at risk could be undermined by the emergence of new social relationships among the group that increase members' exposure to risk behaviours, while reducing their contact with other young people less tolerant of risk taking ²⁴
Feedback	
Where one change reinforces, promotes, balances, or diminishes another	A smoking ban in public places reduces the visibility and convenience of smoking; fewer young people start smoking, further reducing its visibility, in a reinforcing loop ²²
Adaptation	
Change of system behaviour in response to an intervention	Retailers adapted to the ban on multi-buy discounts by discounting individual alcohol products, offering them at the same price individually as they would have been if part of a multi-buy offer ²⁵
Self-organisation	
Order arising from spontaneous local interaction rather than a preconceived plan or external control	Recognising that individual treatment did not address some social aspects of alcohol dependency, recovering drinkers self-organised to form Alcoholics Anonymous

assessment of how interventions contribute to system change.²²⁻²⁷ The new framework therefore takes a pluralistic approach and identifies four perspectives that can be used to guide the design and conduct of complex intervention research: efficacy, effectiveness, theory based, and systems (table 2).

Although each research perspective prompts different types of research question, they should be thought of as overlapping rather than mutually exclusive. For example, theory based and systems perspectives to evaluation can be used in conjunction,³³ while an effectiveness evaluation can draw on a theory based or systems perspective through an embedded process evaluation to explore how and under what circumstances outcomes are achieved.³⁴⁻³⁶

Most complex health intervention research so far has taken an efficacy or effectiveness perspective and for some research questions these perspectives will continue to be the most appropriate. However, some questions equally relevant to the needs of decision makers cannot be answered by research restricted to an efficacy or effectiveness perspective. A wider range and combination of research perspectives and methods, which answer questions beyond efficacy and effectiveness, need to be used by researchers and supported by funders. Doing so will help to improve

the extent to which key questions for decision makers can be answered by complex intervention research. Example questions include:

- Will this effective intervention reproduce the effects found in the trial when implemented here?
- Is the intervention cost effective?
- What are the most important things we need to do that will collectively improve health outcomes?
- In the absence of evidence from randomised trials and the infeasibility of conducting such a trial, what does the existing evidence suggest is the best option now and how can this be evaluated?
- What wider changes will occur as a result of this intervention?
- How are the intervention effects mediated by different settings and contexts?

Phases and core elements of complex intervention research

The framework divides complex intervention research into four phases: development or identification of the intervention, feasibility, evaluation, and implementation (fig 1). A research programme might begin at any phase, depending on the key uncertainties about the intervention in question. Repeating phases

Table 2 | Research perspectives

Perspective and research question	Key points	Vaccine study example
Efficacy		
To what extent does the intervention produce the intended outcomes in experimental or ideal settings?	Conducted under idealised conditions; maximises internal validity to provide a precise, unbiased estimate of efficacy	Seeks to measure the effect of the vaccine on immune system response and report its safety ²⁸
Effectiveness		
To what extent does the intervention produce the intended outcomes in real world settings?	Intervention often compared against treatment as usual; results inform choices between an established and a novel approach to achieving the desired outcome	Seeks to determine whether the vaccination programme, implemented in a range of real world populations and settings, is effective in terms of what it set out to do (eg, prevent disease) ²⁹
Theory based		
What works in which circumstances and how?	Aims to understand how change is brought about, including the interplay of mechanisms and context; can lead to refinement of theory	Asks why effectiveness varies across contexts, and asks what this variation indicates about the conditions for a successful vaccination programme ³⁰ ; considerations that might be explored go beyond whether the vaccine works ³¹
Systems		
How do the system and intervention adapt to one another?	Treats the intervention as a disruption to a complex system ¹⁶	Seeks to understand the dynamic interdependence of vaccination rollout, population risk of infection and willingness to be vaccinated, as the vaccination programme proceeds ³²

is preferable to automatic progression if uncertainties remain unresolved. Each phase has a common set of core elements—considering context, developing and refining programme theory, engaging stakeholders, identifying key uncertainties, refining the intervention, and economic considerations. These elements should be considered early and continually revisited throughout the research process, and especially before moving between phases (for example, between feasibility testing and evaluation).

Core elements

Context

The effects of a complex intervention might often be highly dependent on context, such that an intervention that is effective in some settings could be ineffective or even harmful elsewhere.⁶ As the examples in table 1 show, interventions can modify the contexts in which they are implemented, by eliciting responses from other agents, or by changing behavioural norms or exposure to risk, so that their effects will also vary over time. Context can be considered as both dynamic and multi-dimensional. Key dimensions include physical, spatial, organisational, social, cultural, political, or economic features of the healthcare, health system, or public health contexts in which interventions are implemented. For example, the evaluation of the Breastfeeding In Groups intervention found that the context of the different localities (eg, staff morale and suitable premises) influenced policy implementation and was an explanatory factor in why breastfeeding

rates increased in some intervention localities and declined in others.³⁷

Programme theory

Programme theory describes how an intervention is expected to lead to its effects and under what conditions. It articulates the key components of the intervention and how they interact, the mechanisms of the intervention, the features of the context that are expected to influence those mechanisms, and how those mechanisms might influence the context.³⁸ Programme theory can be used to promote shared understanding of the intervention among diverse stakeholders, and to identify key uncertainties and research questions. Where an intervention (such as a policy) is developed by others, researchers still need to theorise the intervention before attempting to evaluate it.³⁹ Best practice is to develop programme theory at the beginning of the research project with involvement of diverse stakeholders, based on evidence and theory from relevant fields, and to refine it during successive phases. The EPOCH trial tested a large scale quality improvement programme aimed at improving 90 day survival rates for patients undergoing emergency abdominal surgery; it included a well articulated programme theory at the outset, which supported the tailoring of programme delivery to local contexts.⁴⁰ The development, implementation, and post-study reflection of the programme theory resulted in suggested improvements for future implementation of the quality improvement programme.

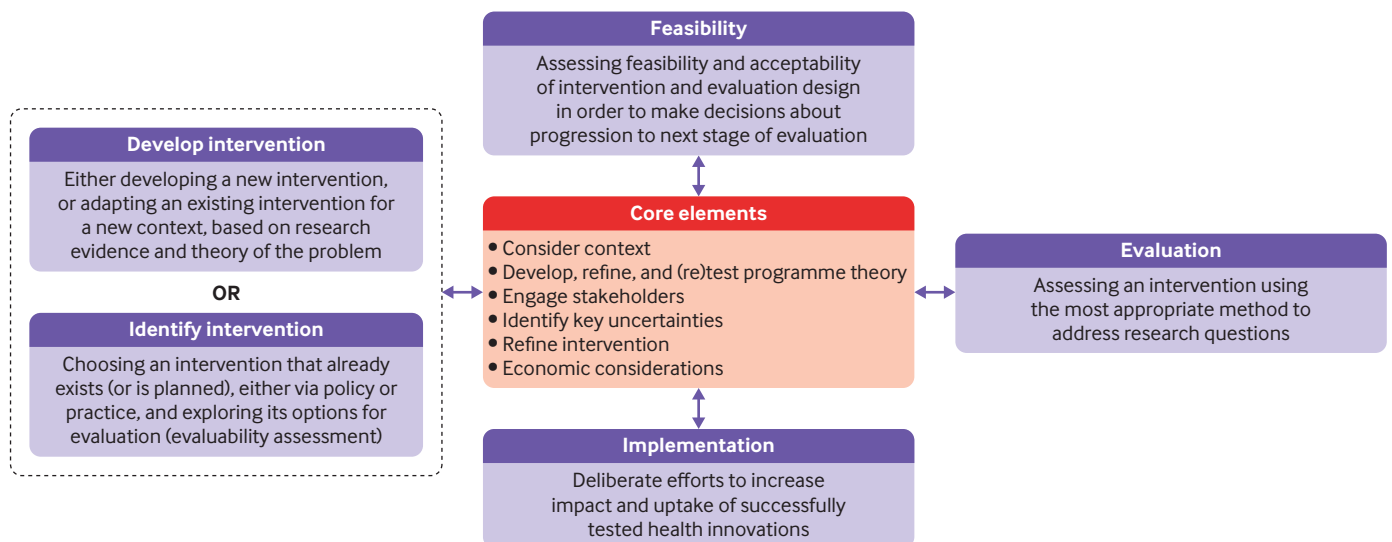


Fig 1 | Framework for developing and evaluating complex interventions. Context=any feature of the circumstances in which an intervention is conceived, developed, evaluated, and implemented; programme theory=describes how an intervention is expected to lead to its effects and under what conditions—the programme theory should be tested and refined at all stages and used to guide the identification of uncertainties and research questions; stakeholders=those who are targeted by the intervention or policy, involved in its development or delivery, or more broadly those whose personal or professional interests are affected (that is, who have a stake in the topic)—this includes patients and members of the public as well as those linked in a professional capacity; uncertainties=identifying the key uncertainties that exist, given what is already known and what the programme theory, research team, and stakeholders identify as being most important to discover—these judgments inform the framing of research questions, which in turn govern the choice of research perspective; refinement=the process of fine tuning or making changes to the intervention once a preliminary version (prototype) has been developed; economic considerations=determining the comparative resource and outcome consequences of the interventions for those people and organisations affected

A refined programme theory is an important evaluation outcome and is the principal aim where a theory based perspective is taken. Improved programme theory will help inform transferability of interventions across settings and help produce evidence and understanding that is useful to decision makers. In addition to full articulation of programme theory, it can help provide visual representations—for example, using a logic model,^{41–43} realist matrix,⁴⁴ or a system map,⁴⁵ with the choice depending on which is most appropriate for the research perspective and research questions. Although useful, any single visual representation is unlikely to sufficiently articulate the programme theory—it should always be articulated well within the text of publications, reports, and funding applications.

Stakeholders

Stakeholders include those individuals who are targeted by the intervention or policy, those involved in its development or delivery, or those whose personal or professional interests are affected (that is, all those who have a stake in the topic). Patients and the public are key stakeholders. Meaningful engagement with appropriate stakeholders at each phase of the research is needed to maximise the potential of developing or identifying an intervention that is likely to have positive impacts on health and to enhance prospects of achieving changes in policy or practice. For example, patient and public involvement⁴⁶ activities in the PARADES programme, which evaluated approaches to reduce harm and improve outcomes for people with bipolar disorder, were wide ranging and central to the project.⁴⁷ Involving service users with lived experiences of bipolar disorder had many benefits, for example, it enhanced the intervention but also improved the evaluation and dissemination methods. Service users involved in the study also had positive outcomes, including more settled employment and progression to further education. Broad thinking and consultation is needed to identify a diverse range of appropriate stakeholders.

The purpose of stakeholder engagement will differ depending on the context and phase of the research, but is essential for prioritising research questions, the co-development of programme theory, choosing the most useful research perspective, and overcoming practical obstacles to evaluation and implementation. Researchers should nevertheless be mindful of conflicts of interest among stakeholders and use transparent methods to record potential conflicts of interest. Research should not only elicit stakeholder priorities, but also consider why they are priorities. Careful consideration of the appropriateness and methods of identification and engagement of stakeholders is needed.^{46 48}

Key uncertainties

Many questions could be answered at each phase of the research process. The design and conduct of research need to engage pragmatically with the multiple uncertainties involved and offer a flexible and emergent

approach to exploring them.¹⁵ Therefore, researchers should spend time developing the programme theory, clearly identifying the remaining uncertainties, given what is already known and what the research team and stakeholders identify as being most important to determine. Judgments about the key uncertainties inform the framing of research questions, which in turn govern the choice of research perspective.

Efficacy trials of relatively uncomplicated interventions in tightly controlled conditions, where research questions are answered with great certainty, will always be important, but translation of the evidence into the diverse settings of everyday practice is often highly problematic.²⁷ For intervention research in healthcare and public health settings to take on more challenging evaluation questions, greater priority should be given to mixed methods, theory based, or systems evaluation that is sensitive to complexity and that emphasises implementation, context, and system fit. This approach could help improve understanding and identify important implications for decision makers, albeit with caveats, assumptions, and limitations.²² Rather than maintaining the established tendency to prioritise strong research designs that answer some questions with certainty but are unsuited to resolving many important evaluation questions, this more inclusive, deliberative process could place greater value on equivocal findings that nevertheless inform important decisions where evidence is sparse.

Intervention refinement

Within each phase of complex intervention research and on transition from one phase to another, the intervention might need to be refined, on the basis of data collected or development of programme theory.⁴ The feasibility and acceptability of interventions can be improved by engaging potential intervention users to inform refinements. For example, an online physical activity planner for people with diabetes mellitus was found to be difficult to use, resulting in the tool providing incorrect personalised advice. To improve usability and the advice given, several iterations of the planner were developed on the basis of interviews and observations. This iterative process led to the refined planner demonstrating greater feasibility and accuracy.⁴⁹

Refinements should be guided by the programme theory, with acceptable boundaries agreed and specified at the beginning of each research phase, and with transparent reporting of the rationale for change. Scope for refinement might also be limited by the policy or practice context. Refinement will be rare in the evaluation phase of efficacy and effectiveness research, where interventions will ideally not change or evolve within the course of the study. However, between the phases of research and within systems and theory based evaluation studies, refinement of interventions in response to accumulated data or as an adaptive and variable response to context and system change are likely to be desirable features of the intervention and a key focus of the research.

Economic considerations

Economic evaluation—the comparative analysis of alternative courses of action in terms of both costs (resource use) and consequences (outcomes, effects)—should be a core component of all phases of intervention research. Early engagement of economic expertise will help identify the scope of costs and benefits to assess in order to answer questions that matter most to decision makers.⁵⁰ Broad ranging approaches such as cost benefit analysis or cost consequence analysis, which seek to capture the full range of health and non-health costs and benefits across different sectors,⁵¹ will often be more suitable for an economic evaluation of a complex intervention than narrower approaches such as cost effectiveness or cost utility analysis. For example, evaluation of the New Orleans Intervention Model for infants entering foster care in Glasgow included short and long term economic analysis from multiple perspectives (the UK's health service and personal social services, public sector, and wider societal perspectives); and used a range of frameworks, including cost utility and cost consequence analysis, to capture changes in the intersectoral costs and outcomes associated with child maltreatment.^{52 53} The use of multiple economic evaluation frameworks provides decision makers with a comprehensive, multi-perspective guide to the cost effectiveness of the New Orleans Intervention Model.

Phases*Developing or identifying a complex intervention*

Development refers to the whole process of designing and planning an intervention, from initial conception through to feasibility, pilot, or evaluation study. Guidance on intervention development has recently been developed through the INDEX study⁴; although here we highlight that complex intervention research does not always begin with new or researcher led interventions. For example:

- A key source of intervention development might be an intervention that has been developed elsewhere and has the possibility of being adapted to a new context. Adaptation of existing interventions could include adapting to a new population, to a new setting,^{54 55} or to target other outcomes (eg, a smoking prevention intervention being adapted to tackle substance misuse and sexual health).^{20 56 57} A well developed programme theory can help identify what features of the antecedent intervention(s) need to be adapted for different applications, and the key mechanisms that should be retained even if delivered slightly differently.^{54 58}
- Policy or practice led interventions are an important focus of evaluation research. Again, uncovering the implicit theoretical basis of an intervention and developing a programme theory is essential to identifying key uncertainties and working out how the

intervention might be evaluated. This step is important, even if rollout has begun, because it supports the identification of mechanisms of change, important contextual factors, and relevant outcome measures. For example, researchers evaluating the UK soft drinks industry levy developed a bounded conceptual system map to articulate their understanding (drawing on stakeholder views and document review) of how the intervention was expected to work. This system map guided the evaluation design and helped identify data sources to support evaluation.⁴⁵ Another example is a recent analysis of the implicit theory of the NHS diabetes prevention programme, involving analysis of documentation by NHS England and four providers, showing that there was no explicit theoretical basis for the programme, and no logic model showing how the intervention was expected to work. This meant that the justification for the inclusion of intervention components was unclear.⁵⁹

Intervention identification and intervention development represent two distinct pathways of evidence generation,⁶⁰ but in both cases, the key considerations in this phase relate to the core elements described above.

Feasibility

A feasibility study should be designed to assess predefined progression criteria that relate to the evaluation design (eg, reducing uncertainty around recruitment, data collection, retention, outcomes, and analysis) or the intervention itself (eg, around optimal content and delivery, acceptability, adherence, likelihood of cost effectiveness, or capacity of providers to deliver the intervention). If the programme theory suggests that contextual or implementation factors might influence the acceptability, effectiveness, or cost effectiveness of the intervention, these questions should be considered.

Despite being overlooked or rushed in the past, the value of feasibility testing is now widely accepted with key terms and concepts well defined.^{61 62} Before initiating a feasibility study, researchers should consider conducting an evaluability assessment to determine whether and how an intervention can usefully be evaluated. Evaluability assessment involves collaboration with stakeholders to reach agreement on the expected outcomes of the intervention, the data that could be collected to assess processes and outcomes, and the options for designing the evaluation.⁶³ The end result is a recommendation on whether an evaluation is feasible, whether it can be carried out at a reasonable cost, and by which methods.⁶⁴

Economic modelling can be undertaken at the feasibility stage to assess the likelihood that the expected benefits of the intervention justify the costs (including the cost of further research), and to help decision makers decide whether proceeding to a full

scale evaluation is worthwhile.⁶⁵ Depending on the results of the feasibility study, further work might be required to progressively refine the intervention before embarking on a full scale evaluation.

Evaluation

The new framework defines evaluation as going beyond asking whether an intervention works (in the sense of achieving its intended outcome), to a broader range of questions including identifying what other impact it has, theorising how it works, taking account of how it interacts with the context in which it is implemented, how it contributes to system change, and how the evidence can be used to support decision making in the real world. This implies a shift from an exclusive focus on obtaining unbiased estimates of effectiveness⁶⁶ towards prioritising the usefulness of information for decision making in selecting the optimal research perspective and in prioritising answerable research questions.

A crucial aspect of evaluation design is the choice of outcome measures or evidence of change. Evaluators should work with stakeholders to assess which outcomes are most important, and how to deal with multiple outcomes in the analysis with due consideration of statistical power and transparent reporting. A sharp distinction between one primary outcome and several secondary outcomes is not necessarily appropriate, particularly where the programme theory identifies impacts across a range of domains. Where needed to support the research questions, prespecified subgroup analyses should be carried out and reported. Even where such analyses are underpowered, they should be included in the protocol because they might be useful for subsequent meta-analyses, or for developing hypotheses for testing in further research. Outcome measures could capture changes to a system rather than changes in individuals. Examples include changes in relationships within an organisation, the introduction of policies, changes in social norms, or normalisation of practice. Such system level outcomes include how changing the dynamics of one part of a system alters behaviours in other parts, such as the potential for displacement of smoking into the home after a public smoking ban.

A helpful illustration of the use of system level outcomes is the evaluation of the Delaware Young Health Program—an initiative to improve the health and wellbeing of young people in Delaware, USA. The intervention aimed to change underlying system dynamics, structures, and conditions, so the evaluation identified systems oriented research questions and methods. Three systems science methods were used: group model building and viable systems model assessment to identify underlying patterns and structures; and social network analysis to evaluate change in relationships over time.⁶⁷

Researchers have many study designs to choose from, and different designs are optimally suited to consider different research questions and different circumstances.⁶⁸ Extensions to standard designs of randomised controlled trials (including adaptive

designs, SMART trials (sequential multiple assignment randomised trials), n-of-1 trials, and hybrid effectiveness-implementation designs) are important areas of methods development to improve the efficiency of complex intervention research.⁶⁹⁻⁷² Non-randomised designs and modelling approaches might work best if a randomised design is not practical, for example, in natural experiments or systems evaluations.^{5 73 74} A purely quantitative approach, using an experimental design with no additional elements such as a process evaluation, is rarely adequate for complex intervention research, where qualitative and mixed methods designs might be necessary to answer questions beyond effectiveness. In many evaluations, the nature of the intervention, the programme theory, or the priorities of stakeholders could lead to a greater focus on improving theories about how to intervene. In this view, effect estimates are inherently context bound, so that average effects are not a useful guide to decision makers working in different contexts. Contextualised understandings of how an intervention induces change might be more useful, as well as details on the most important enablers and constraints on its delivery across a range of settings.⁷

Process evaluation can answer questions around fidelity and quality of implementation (eg, what is implemented and how?), mechanisms of change (eg, how does the delivered intervention produce change?), and context (eg, how does context affect implementation and outcomes?).⁷ Process evaluation can help determine why an intervention fails unexpectedly or has unanticipated consequences, or why it works and how it can be optimised. Such findings can facilitate further development of the intervention programme theory.⁷⁵ In a theory based or systems evaluation, there is not necessarily such a clear distinction between process and outcome evaluation as there is in an effectiveness study.⁷⁶ These perspectives could prioritise theory building over evidence production and use case study or simulation methods to understand how outcomes or system behaviour are generated through intervention.^{74 77}

Implementation

Early consideration of implementation increases the potential of developing an intervention that can be widely adopted and maintained in real world settings. Implementation questions should be anticipated in the intervention programme theory, and considered throughout the phases of intervention development, feasibility testing, process, and outcome evaluation. Alongside implementation specific outcomes (such as reach or uptake of services), attention to the components of the implementation strategy, and contextual factors that support or hinder the achievement of impacts, are key. Some flexibility in intervention implementation might support intervention transferability into different contexts (an important aspect of long term implementation⁷⁸), provided that the key functions of the programme are maintained, and that the adaptations made are clearly understood.⁸

In the ASSIST study,²⁰ a school based, peer led intervention for smoking prevention, researchers considered implementation at each phase. The intervention was developed to have minimal disruption on school resources; the feasibility study resulted in intervention refinements to improve acceptability and improve reach to male students; and in the evaluation (cluster randomised controlled trial), the intervention was delivered as closely as possible to real world implementation. Drawing on the process evaluation, the implementation included an intervention manual that identified critical components and other components that could be adapted or dropped to allow flexible implementation while achieving delivery of the key mechanisms of change; and a training manual for the trainers and ongoing quality assurance built into rollout for the longer term.

In a natural experimental study, evaluation takes place during or after the implementation of the intervention in a real world context. Highly pragmatic effectiveness trials or specific hybrid effectiveness-implementation designs also combine effectiveness and implementation outcomes in one study, with the aim of reducing time for translation of research on effectiveness into routine practice.^{72 79 80}

Implementation questions should be included in economic considerations during the early stages of intervention and study development. How the results of economic analyses are reported and presented to decision makers can affect whether and how they act on the results.⁸¹ A key consideration is how to deal with interventions across different sectors, where those paying for interventions and those receiving the benefits of them could differ, reducing the incentive to implement an intervention, even if shown to be beneficial and cost effective. Early engagement with appropriate stakeholders will help frame appropriate research questions and could anticipate any implementation challenges that might arise.⁸²

Conclusions

One of the motivations for developing this new framework was to answer calls for a change in research priorities, towards allocating greater effort and funding to research that can have the optimum impact on healthcare or population health outcomes. The framework challenges the view that unbiased estimates of effectiveness are the cardinal goal of evaluation. It asserts that improving theories and understanding how interventions contribute to change, including how they interact with their context and wider dynamic systems, is an equally important goal. For some complex intervention research problems, an efficacy or effectiveness perspective will be the optimal approach, and a randomised controlled trial will provide the best design to achieve an unbiased estimate. For others, alternative perspectives and designs might work better, or might be the only way to generate new knowledge to reduce decision maker uncertainty.

What is important for the future is that the scope of intervention research is not constrained by an unduly

limited set of perspectives and approaches that might be less risky to commission and more likely to produce a clear and unbiased answer to a specific question. A bolder approach is needed—to include methods and perspectives where experience is still quite limited, but where we, supported by our workshop participants and respondents to our consultations, believe there is an urgent need to make progress. This endeavour will involve mainstreaming new methods that are not yet widely used, as well as undertaking methodological innovation and development. The deliberative and flexible approach that we encourage is intended to reduce research waste,⁸³ maximise usefulness for decision makers, and increase the efficiency with which complex intervention research generates knowledge that contributes to health improvement.

Monitoring the use of the framework and evaluating its acceptability and impact is important but has been lacking in the past. We encourage research funders and journal editors to support the diversity of research perspectives and methods that are advocated here and to seek evidence that the core elements are attended to in research design and conduct. We have developed a checklist to support the preparation of funding applications, research protocols, and journal publications.⁹ This checklist offers one way to monitor impact of the guidance on researchers, funders, and journal editors.

We recommend that the guidance is continually updated, and future updates continue to adopt a broad, pluralist perspective. Given its wider scope, and the range of detailed guidance that is now available on specific methods and topics, we believe that the framework is best seen as meta-guidance. Further editions should be published in a fluid, web based format, and more frequently updated to incorporate new material, further case studies, and additional links to other new resources.

AUTHOR AFFILIATIONS

¹MRC/CSO Social and Public Health Sciences Unit, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

²Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK

³Medical Research Council ConDuCT-II Hub for Trials Methodology Research and Bristol Biomedical Research Centre, Bristol, UK

⁴Health Economics and Health Technology Assessment Unit, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

⁵Public Health Scotland, Glasgow, UK

⁶Manchester Centre for Health Psychology, University of Manchester, Manchester, UK

⁷London School of Hygiene and Tropical Medicine, London, UK

⁸Faculty of Health and Medicine, Lancaster University, Lancaster, UK

⁹Medical Research Council Epidemiology Unit, University of Cambridge, Cambridge, UK

We thank the experts who provided input at the workshop, those who responded to the consultation, and those who provided advice and review throughout the process. The many people involved are acknowledged in the full framework document.⁹ Parts of this manuscript have been reproduced (some with edits and formatting changes), with permission, from that longer framework document.

Contributors: All authors made a substantial contribution to all stages of the development of the framework—they contributed to its development, drafting, and final approval. KS and LMa led the writing of the framework, and KS wrote the first draft of this paper.

PC, SAS, and LMo provided critical insights to the development of the framework and contributed to writing both the framework and this paper. KS, LMa, SAS, PC, and LMo facilitated the expert workshop, KS and LMa developed the gap analysis and led the analysis of the consultation. KAB, NC, and EM contributed the economic components to the framework. The scientific advisory group (JB, JMB, DPF, MP, JR-M, and MW) provided feedback and edits on drafts of the framework, with particular attention to process evaluation (JB), clinical research (JMB), implementation (JR-M, DPF), systems perspective (MP), theory based perspective (JR-M), and population health (MW). LMo is senior author. KS and LMo are the guarantors of this work and accept the full responsibility for the finished article. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting authorship criteria have been omitted.

Funding: The work was funded by the National Institute for Health Research (Department of Health and Social Care 73514) and Medical Research Council (MRC). Additional time on the study was funded by grants from the MRC for KS (MC_UU_12017/11, MC_UU_00022/3), LMa, SAS, and LMo (MC_UU_12017/14, MC_UU_00022/1); PC (MC_UU_12017/15, MC_UU_00022/2); and MW (MC_UU_12015/6 and MC_UU_00006/7). Additional time on the study was also funded by grants from the Chief Scientist Office of the Scottish Government Health Directorates for KS (SPHSU11 and SPHSU18); LMa, SAS, and LMo (SPHSU14 and SPHSU16); and PC (SPHSU13 and SPHSU15). KS and SAS were also supported by an MRC Strategic Award (MC_PC_13027). JMB received funding from the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol and by the MRC ConDuCT-II Hub (Collaboration and innovation for Difficult and Complex randomised controlled Trials In Invasive procedures - MR/K025643/1). DF is funded in part by the NIHR Manchester Biomedical Research Centre (IS-BRC-1215-20007) and NIHR Applied Research Collaboration - Greater Manchester (NIHR200174). MP is funded in part as director of the NIHR's Public Health Policy Research Unit. This project was overseen by a scientific advisory group that comprised representatives of NIHR research programmes, of the MRC/NIHR Methodology Research Programme Panel, of key MRC population health research investments, and authors of the 2006 guidance. A prospectively agreed protocol, outlining the workplan, was agreed with MRC and NIHR, and signed off by the scientific advisory group. The framework was reviewed and approved by the MRC/NIHR Methodology Research Programme Advisory Group and MRC Population Health Sciences Group and completed NIHR HTA Monograph editorial and peer review processes.

Competing interests: All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/coi_disclosure.pdf and declare: support from the NIHR, MRC, and the funders listed above for the submitted work; KS has project grant funding from the Scottish Government Chief Scientist Office; SAS is a former member of the NIHR Health Technology Assessment Clinical Evaluation and Trials Programme Panel (November 2016 - November 2020) and member of the Chief Scientist Office Health HIPS Committee (since 2018) and NIHR Policy Research Programme (since November 2019), and has project grant funding from the Economic and Social Research Council, MRC, and NIHR; LMo is a former member of the MRC/NIHR Methodology Research Programme Panel (2015-19) and MRC Population Health Sciences Group (2015-20); JB is a member of the NIHR Public Health Research Funding Committee (since May 2019), and a core member (since 2016) and vice chairperson (since 2018) of a public health advisory committee of the National Institute for Health and Care Excellence; JMB is a former member of the NIHR Clinical Trials Unit Standing Advisory Committee (2015-19); DPF is a former member of the NIHR Public Health Research programme research funding board (2015-2019), the MRC/NIHR Methodology Research Programme panel member (2014-2018), and is a panel member of the Research Excellence Framework 2021, subpanel 2 (public health, health services, and primary care; November 2020 - February 2022), and has grant funding from the European Commission, NIHR, MRC, Natural Environment Research Council, Prevent Breast Cancer, Breast Cancer Now, Greater Sport, Manchester University NHS Foundation Trust, Christie Hospital NHS Trust, and BXS GP; EM is a member of the NIHR Public Health Research funding board; MP has grant funding from the MRC, UK Prevention Research Partnership, and NIHR; JR-M is programme director and chairperson of the NIHR's Health Services Delivery Research Programme (since 2014) and member of the NIHR Strategy Board (since 2014); MW received a salary as director of the NIHR PHR Programme (2014-20), has grant funding from NIHR, and is a former member of the MRC's Population Health Sciences Strategic Committee (July 2014 to June 2020). There are no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement: This project was methodological; views of patients and the public were included at the open consultation stage of the update. The open consultation, involving access to an initial draft, was promoted to our networks via email and digital channels, such as our unit Twitter account (@theSPHSU). We received five responses from people who identified as service users (rather than researchers or professionals in a relevant capacity). Their input included helpful feedback on the main complexity diagram, the different research perspectives, the challenge of moving interventions between different contexts and overall readability and accessibility of the document. Several respondents also highlighted useful signposts to include for readers. Various dissemination events are planned, but as this project is methodological we will not specifically disseminate to patients and the public beyond the planned dissemination activities.

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

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>.

- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655. doi:10.1136/bmj.a1655
- Craig P, Dieppe P, Macintyre S, et al. *Developing and evaluating complex interventions: new guidance*. Medical Research Council, 2006.
- Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;321:694-6. doi:10.1136/bmj.321.7262.694
- O' Cathain A, Croot L, Duncan E, et al. Guidance on how to develop complex interventions to improve health and healthcare. *BMJ Open* 2019;9:e029954. doi:10.1136/bmjopen-2019-029954
- Craig P, Cooper C, Gunnell D, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health* 2012;66:1182-6. doi:10.1136/jech-2011-200375
- Craig P, Ruggiero ED, Frohlich KL, et al. *Taking account of context in population health intervention research: guidance for producers, users and funders of research*. NIHR Journals Library, 2018. doi:10.3310/CIHR-NIHR-01
- 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
- Moore G, Campbell M, Copeland L, et al. Adapting interventions to new contexts-the ADAPT guidance. *BMJ* 2021;374:n1679. doi:10.1136/bmj.n1679
- Skivington K, Matthews L, Simpson SA, et al. Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. *Health Technol Assess* 2021 [forthcoming].
- Chng NR, Hawkins K, Fitzpatrick B, et al. Implementing social prescribing in primary care in areas of high socioeconomic deprivation: process evaluation of the 'Deep End' community links worker programme. *Br J Gen Pract* 2021;71:153:BJGP.2020.1153. doi:10.3399/BJGP.2020.1153
- Mercer SW, Fitzpatrick B, Grant L, et al. Effectiveness of Community-Links Practitioners in Areas of High Socioeconomic Deprivation. *Ann Fam Med* 2019;17:518-25. doi:10.1370/afm.2429
- Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ* 2004;328:1561-3. doi:10.1136/bmj.328.7455.1561
- Blencowe NS, Mills N, Cook JA, et al. Standardizing and monitoring the delivery of surgical interventions in randomized clinical trials. *Br J Surg* 2016;103:1377-84. doi:10.1002/bjs.10254
- Blencowe NS, Skilton A, Gaunt D, et al, ROMIO Study team. Protocol for developing quality assurance measures to use in surgical trials: an example from the ROMIO study. *BMJ Open* 2019;9:e026209. doi:10.1136/bmjopen-2018-026209
- Greenhalgh T, Papoutis 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
- Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. *Am J Community Psychol* 2009;43:267-76. doi:10.1007/s10464-009-9229-9
- Petticrew M. When are complex interventions 'complex'? When are simple interventions 'simple'? *Eur J Public Health* 2011;21:397-8. doi:10.1093/eurpub/ckr084
- Anderson R. New MRC guidance on evaluating complex interventions. *BMJ* 2008;337:a1937. doi:10.1136/bmj.a1937
- Pawson R, Tilley N. *Realistic Evaluation*. Sage, 1997.

- 20 Campbell R, Starkey F, Holliday J, et al. An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): a cluster randomised trial. *Lancet* 2008;371:1595-602. doi:10.1016/S0140-6736(08)60692-3
- 21 Randell R, Honey S, Hindmarsh J, et al. *A realist process evaluation of robot-assisted surgery: integration into routine practice and impacts on communication, collaboration and decision-making*. NIHR Journals Library, 2017. <https://www.ncbi.nlm.nih.gov/books/NBK447438/>.
- 22 Rutter H, Savona N, Glonti K, et al. The need for a complex systems model of evidence for public health. *Lancet* 2017;390:2602-4. doi:10.1016/S0140-6736(17)31267-9
- 23 The Health Foundation. Evidence Scan. Complex adaptive systems. Health Foundation 2010. <https://www.health.org.uk/publications/complex-adaptive-systems>.
- 24 Wiggins M, Bonell C, Sawtell M, et al. Health outcomes of youth development programme in England: prospective matched comparison study. *BMJ* 2009;339:b2534. doi:10.1136/bmj.b2534
- 25 Robinson M, Geue C, Lewsey J, et al. Evaluating the impact of the alcohol act on off-trade alcohol sales: a natural experiment in Scotland. *Addiction* 2014;109:2035-43. doi:10.1111/add.12701
- 26 Raine R, Fitzpatrick R, de Pury J. Challenges, solutions and future directions in evaluative research. *J Health Serv Res Policy* 2016;21:215-6. doi:10.1177/1355819616664495
- 27 Kessler R, Glasgow RE. A proposal to speed translation of healthcare research into practice: dramatic change is needed. *Am J Prev Med* 2011;40:637-44. doi:10.1016/j.amepre.2011.02.023
- 28 Folegatti PM, Ewer KJ, Aley PK, et al. Oxford COVID Vaccine Trial Group. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. *Lancet* 2020;396:467-78. doi:10.1016/S0140-6736(20)31604-4
- 29 Voysey M, Clemens SAC, Madhi SA, et al. Oxford COVID Vaccine Trial Group. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* 2021;397:99-111. doi:10.1016/S0140-6736(20)32661-1
- 30 Soi C, Shearer JC, Budden A, et al. How to evaluate the implementation of complex health programmes in low-income settings: the approach of the Gavi Full Country Evaluations. *Health Policy Plan* 2020;35(Supplement_2):ii35-46. doi:10.1093/heapol/ctaa127
- 31 Burgess RA, Osborne RH, Yongabi KA, et al. The COVID-19 vaccines rush: participatory community engagement matters more than ever. *Lancet* 2021;397:8-10. doi:10.1016/S0140-6736(20)32642-8
- 32 Paltiel AD, Schwartz JL, Zheng A, Walensky RP. Clinical Outcomes Of A COVID-19 Vaccine: Implementation Over Efficacy. *Health Aff (Millwood)* 2021;40:42-52. doi:10.1377/hlthaff.2020.02054
- 33 Dalkin S, Lhussier M, Williams L, et al. Exploring the use of Soft Systems Methodology with realist approaches: A novel way to map programme complexity and develop and refine programme theory. *Evaluation* 2018;24:84-97. doi:10.1177/1356389017749036
- 34 Mann C, Shaw ARG, Guthrie B, et al. Can implementation failure or intervention failure explain the result of the 3D multimorbidity trial in general practice: mixed-methods process evaluation. *BMJ Open* 2019;9:e031438. doi:10.1136/bmjopen-2019-031438
- 35 French C, Pinnock H, Forbes G, Skene I, Taylor SJC. Process evaluation within pragmatic randomised controlled trials: what is it, why is it done, and can we find it? a systematic review. *Trials* 2020;21:916. doi:10.1186/s13063-020-04762-9
- 36 Penney T, Adams J, Briggs A, et al. Evaluation of the impacts on health of the proposed UK industry levy on sugar sweetened beverages: developing a systems map and data platform, and collection of baseline and early impact data. National Institute for Health Research, 2018. <https://www.journalslibrary.nihr.ac.uk/programmes/phr/164901/#/>.
- 37 Hoddinott P, Britten J, Pill R. Why do interventions work in some places and not others: a breastfeeding support group trial. *Soc Sci Med* 2010;70:769-78. doi:10.1016/j.socscimed.2009.10.067
- 38 Funnell SC, Rogers PJ. *Purposeful program theory: effective use of theories of change and logic models*. 1st ed. Jossey-Bass, 2011.
- 39 Lawless A, Baum F, Delany-Crowe T, et al. Developing a Framework for a Program Theory-Based Approach to Evaluating Policy Processes and Outcomes: Health in All Policies in South Australia. *Int J Health Policy Manag* 2018;7:510-21. doi:10.15171/ijhpm.2017.121
- 40 Stephens TJ, Peden CJ, Pearce RM, et al. EPOCH trial group. Improving care at scale: process evaluation of a multi-component quality improvement intervention to reduce mortality after emergency abdominal surgery (EPOCH trial). *Implement Sci* 2018;13:142. doi:10.1186/s13012-018-0823-9
- 41 Bonell C, Jamal F, Melendez-Torres GJ, Cummins S. 'Dark logic': theorising the harmful consequences of public health interventions. *J Epidemiol Community Health* 2015;69:95-8. doi:10.1136/jech-2014-204671
- 42 Maini R, Mounier-Jack S, Borghi J. How to and how not to develop a theory of change to evaluate a complex intervention: reflections on an experience in the Democratic Republic of Congo. *BMJ Glob Health* 2018;3:e000617. doi:10.1136/bmjgh-2017-000617
- 43 Cook PA, Hargreaves SC, Burns EJ, et al. Communities in charge of alcohol (CICA): a protocol for a stepped-wedge randomised control trial of an alcohol health champions programme. *BMC Public Health* 2018;18:522. doi:10.1186/s12889-018-5410-0
- 44 Ebenso B, Manzano A, Uzochukwu B, et al. Dealing with context in logic model development: Reflections from a realist evaluation of a community health worker programme in Nigeria. *Eval Program Plann* 2019;73:97-110. doi:10.1016/j.evalprogplan.2018.12.002
- 45 White M, Cummins S, Raynor M, et al. Evaluation of the health impacts of the UK Treasury Soft Drinks Industry Levy (SDIL) Project Protocol. NIHR Journals Library, 2018. <https://www.journalslibrary.nihr.ac.uk/programmes/phr/1613001/#/summary-of-research>.
- 46 National Institute for Health and Care Excellence. What is public involvement in research? – INVOLVE. <https://www.invo.org.uk/find-out-more/what-is-public-involvement-in-research-2/>.
- 47 Jones S, Riste L, Barrowclough C, et al. Reducing relapse and suicide in bipolar disorder: practical clinical approaches to identifying risk, reducing harm and engaging service users in planning and delivery of care – the PARADES (Psychoeducation, Anxiety, Relapse, Advance Directive Evaluation and Suicidality) programme. *Programme Grants for Applied Research* 2018;6:1-296. doi:10.3310/pgfar06060
- 48 Moodie R, Stuckler D, Monteiro C, et al. Lancet NCD Action Group. Profits and pandemics: prevention of harmful effects of tobacco, alcohol, and ultra-processed food and drink industries. *Lancet* 2013;381:670-9. doi:10.1016/S0140-6736(12)62089-3
- 49 Yardley L, Ainsworth B, Arden-Close E, Muller I. The person-based approach to enhancing the acceptability and feasibility of interventions. *Pilot Feasibility Stud* 2015;1:37. doi:10.1186/s40814-015-0033-z
- 50 Barnett ML, Dopp AR, Klein C, Ettner SL, Powell BJ, Saldana L. Collaborating with health economists to advance implementation science: a qualitative study. *Implement Sci Commun* 2020;1:82. doi:10.1186/s43058-020-00074-w
- 51 National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. NICE, 2014. <https://www.nice.org.uk/process/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869>.
- 52 Boyd KA, Balogun MO, Minnis H, et al. Development of a radical foster care intervention in Glasgow, Scotland. *Health Promot Int* 2016;31:665-73. doi:10.1093/heapro/dav041
- 53 Deidda M, Boyd KA, Minnis H, et al. BeST study team. Protocol for the economic evaluation of a complex intervention to improve the mental health of maltreated infants and children in foster care in the UK (The BeST? services trial). *BMJ Open* 2018;8:e020066. doi:10.1136/bmjopen-2017-020066
- 54 Escoffery C, Lebow-Skelley E, Haardorfer R, et al. A systematic review of adaptations of evidence-based public health interventions globally. *Implement Sci* 2018;13:125. doi:10.1186/s13012-018-0815-9
- 55 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
- 56 Forsyth R, Purcell C, Barry S, et al. Peer-led intervention to prevent and reduce STI transmission and improve sexual health in secondary schools (STASH): protocol for a feasibility study. *Pilot Feasibility Stud* 2018;4:180. doi:10.1186/s40814-018-0354-9
- 57 White J, Hawkins J, Madden K, et al. Adapting the ASSIST model of informal peer-led intervention delivery to the Talk to FRANK drug prevention programme in UK secondary schools (ASSIST + FRANK): intervention development, refinement and a pilot cluster randomised controlled trial. *Public Health Research* 2017;5:1-98. doi:10.3310/phr05070
- 58 Evans RE, Moore G, Movsisyan A, Rehfuess E, ADAPT Panel, ADAPT Panel comprises of Laura Arnold. How can we adapt complex population health interventions for new contexts? Progressing debates and research priorities. *J Epidemiol Community Health* 2021;75:40-5. doi:10.1136/jech-2020-214468
- 59 Hawkes RE, Miles LM, French DP. The theoretical basis of a nationally implemented type 2 diabetes prevention programme: how is the programme expected to produce changes in behaviour? *Int J Behav Nutr Phys Act* 2021;18:64. doi:10.1186/s12966-021-01134-7
- 60 Ogilvie D, Adams J, Bauman A, et al. Using natural experimental studies to guide public health action: turning the evidence-based medicine paradigm on its head. *SocArXiv* 2019. doi:10.31235/osf.io/s36km
- 61 Eldridge SM, Chan CL, Campbell MJ, et al. PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ* 2016;355:i239. doi:10.1136/bmj.i239
- 62 Thabane L, Hopewell S, Lancaster GA, et al. Methods and processes for development of a CONSORT extension for reporting pilot randomized controlled trials. *Pilot Feasibility Stud* 2016;2:25. doi:10.1186/s40814-016-0065-z

- 63 Craig P, Campbell M. Evaluability Assessment: a systematic approach to deciding whether and how to evaluate programmes and policies. Evaluability Assessment working paper. 2015. <http://whatworksscotland.ac.uk/wp-content/uploads/2015/07/WWWS-Evaluability-Assessment-Working-paper-final-June-2015.pdf>
- 64 Ogilvie D, Cummins S, Petticrew M, White M, Jones A, Wheeler K. Assessing the evaluability of complex public health interventions: five questions for researchers, funders, and policymakers. *Milbank Q* 2011;89:206-25. doi:10.1111/j.1468-0009.2011.00626.x
- 65 Expected Value of Perfect Information (EVPI). YHEC - York Health Econ. Consort. <https://yhec.co.uk/glossary/expected-value-of-perfect-information-evpi/>.
- 66 Deaton A, Cartwright N. Understanding and misunderstanding randomised controlled trials. *Soc Sci Med* 2018;210:2-21. doi:10.1016/j.socscimed.2017.12.005
- 67 Rosas S, Knight E. Evaluating a complex health promotion intervention: case application of three systems methods. *Crit Public Health* 2019;29:337-52. doi:10.1080/09581596.2018.1455966
- 68 McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ* 1999;319:312-5. doi:10.1136/bmj.319.7205.312
- 69 Burnett T, Mozgunov P, Pallmann P, Villar SS, Wheeler GM, Jaki T. Adding flexibility to clinical trial designs: an example-based guide to the practical use of adaptive designs. *BMC Med* 2020;18:352. doi:10.1186/s12916-020-01808-2
- 70 Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *Am J Prev Med* 2007;32(Suppl):S112-8. doi:10.1016/j.amepre.2007.01.022
- 71 McDonald S, Quinn F, Vieira R, et al. The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview. *Health Psychol Rev* 2017;11:307-23. doi:10.1080/17437199.2017.1316672
- 72 Green BB, Coronado GD, Schwartz M, Coury J, Baldwin LM. Using a continuum of hybrid effectiveness-implementation studies to put research-tested colorectal screening interventions into practice. *Implement Sci* 2019;14:53. doi:10.1186/s13012-019-0903-5
- 73 Tugwell P, Knottnerus JA, McGowan J, Tricco A. Big-5 Quasi-Experimental designs. *J Clin Epidemiol* 2017;89:1-3. doi:10.1016/j.jclinepi.2017.09.010
- 74 Egan M, McGill E, Penney T, et al. NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. Part 1: Introducing systems thinking. NIHR School for Public Health Research, 2019. https://sphr.nihr.ac.uk/wp-content/uploads/2018/08/NIHR-SPHR-SYSTEM-GUIDANCE-PART-1-FINAL_SBnavy.pdf.
- 75 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
- 76 McGill E, Marks D, Er V, Penney T, Petticrew M, Egan M. Qualitative process evaluation from a complex systems perspective: A systematic review and framework for public health evaluators. *PLoS Med* 2020;17:e1003368. doi:10.1371/journal.pmed.1003368
- 77 Bicket M, Christie I, Gilbert N, et al. Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation. *Lond HM Treas* 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879437/Magenta_Book_supplementary_guide_-_Handling_Complexity_in_policy_evaluation.pdf
- 78 Pfadenhauer LM, Gerhardus A, Mozygemba K, et al. Making sense of complexity in context and implementation: the Context and Implementation of Complex Interventions (CICI) framework. *Implement Sci* 2017;12:21. doi:10.1186/s13012-017-0552-5
- 79 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
- 80 Landes SJ, McBain SA, Curran GM. An introduction to effectiveness-implementation hybrid designs. *Psychiatry Res* 2019;280:112513. doi:10.1016/j.psychres.2019.112513
- 81 Imison C, Curry N, Holder H, et al. Shifting the balance of care: great expectations. Research report. Nuffield Trust. <https://www.nuffieldtrust.org.uk/research/shifting-the-balance-of-care-great-expectations>
- 82 Remme M, Martinez-Alvarez M, Vassall A. Cost-Effectiveness Thresholds in Global Health: Taking a Multisectoral Perspective. *Value Health* 2017;20:699-704. doi:10.1016/j.jval.2016.11.009
- 83 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;374:86-9. doi:10.1016/S0140-6736(09)60329-9