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

Kathryn Skivington (<https://orcid.org/0000-0002-3571-1561>) Research Fellow, 1\* Lynsay Matthews (<https://orcid.org/0000-0002-5760-1080>) Research Fellow,1 Sharon Anne Simpson (<https://orcid.org/0000-0002-6219-1768>) Professor of Behavioural Sciences and Health,1 Peter Craig (<https://orcid.org/0000-0002-7653-5832>)1 Professor of Public Health Evaluation, Janis Baird (<https://orcid.org/0000-0002-4039-4361>) Professor of Public Health and Epidemiology, Honorary Consultant in Public Health,2 Jane M Blazeby (<https://orcid.org/0000-0002-3354-3330>) Professor of Surgery,3 Kathleen Anne Boyd (<https://orcid.org/0000-0002-9764-0113>) Reader in Health Economics,4 Neil Craig (<https://orcid.org/0000-0002-7370-5541>) Honorary Professor, Acting Head of the Evaluation Team within Public Health Sciences, NHS Health Scotland,5 David P French (<https://orcid.org/0000-0002-7663-7804>) Professor of Health Psychology,6 Emma McIntosh (<https://orcid.org/0000-0001-6340-3083>) Professor of Health Economics,4 Mark Petticrew (<https://orcid.org/0000-0002-6378-5517>) Professor of Public Health Evaluation,7 Jo Rycroft-Malone (<https://orcid.org/0000-0003-3858-5625>) Professor and Dean of the Faculty of Health and Medicine,8 Martin White ([https://orcid.org/0000-0002-1861-6757) Professor of Population Health Research,9](https://orcid.org/0000-0002-1861-6757%29%20Professor%20of%20Population%20Health%20Research%2C9)Laurence Moore (https://orcid.org/0000-0003-2182-823X) Director of the MRC/CSO Social & Public Health Sciences Unit.1

\* Correspondence to: Kathryn Skivington, MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, Berkeley Square, 99 Berkeley Street, Glasgow, G3 7HR. Kathryn.skivington@glasgow.ac.uk

Email addresses of authors: Kathryn.skivington@glasgow.ac.uk, Lynsay.matthews@glasgow.ac.uk, peter.craig@glasgow.ac.uk, Sharon.simpson@glasgow.ac.uk, jb@mrc.soton.ac.uk, Kathleen.Boyd@glasgow.ac.uk, J.M.Blazeby@bristol.ac.uk, neil.craig@phs.scot, david.french@manchester.ac.uk, Emma.mcintosh@glasgow.ac.uk, Mark.Petticrew@lshtm.ac.uk, j.rycroft-malone1@lancaster.ac.uk, martin.white@mrc-epid.cam.ac.uk, laurence.moore@glasgow.ac.uk

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

# Standfirst

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

# Key Messages

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 a broader range of questions including identifying what other impacts 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.

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

It is useful to think of complex intervention research in terms of phases, noting that they 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 addressed to answer the following questions: (1) how does the intervention interact with its context? (2) what is the underpinning programme theory? (3) how can diverse stakeholder perspectives be included in the research? (4) what are the key uncertainties? (5) how can the intervention be refined? and (6) 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.

# Introduction

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. Examples include a new surgical procedure, the redesign of a health care 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] 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,[2–5] there have been important conceptual, methodological and theoretical developments since 2006. These developments have been addressed in a new framework commissioned by the National Institute of Health Research (NIHR) and the MRC.[6] 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

The updated Framework for Developing and Evaluating Complex Interventions is the culmination of a process that included four stages: (1) a gap analysis to identify developments in the methods and practice since the previous framework was published; (2) a full-day expert workshop, in May 2018, of 36 participants to discuss the topics identified in the gap analysis; (3) 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); and (4) redraft using findings from the previous stages followed by a final expert review. In addition, we 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 whole process was overseen by a Scientific Advisory Group representing the range of relevant National Institute for Health Research programmes and Medical Research Council population health investments. The framework was reviewed by the NIHR-MRC 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.[6]This article introduces the framework and summarises the main messages for producers and users of evidence.

# What are complex interventions?

An intervention may 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, there was no single health or well-being issue that the Link Workers specifically supported: bereavement, substance use, employment and learning difficulties were all included.[7,8] 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 may be permitted to allow for variation in how, where and by whom interventions are delivered and received. Standardisation of interventions may relate more to the underlying process and functions of the intervention than on the specific form of components delivered.[9] For example, in surgical trials it is possible to design protocols with flexibility for intervention delivery.[10] Interventions require a theoretical deconstruction into components and then agreement about permissible and prohibited variation in the delivery of those components. This 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.[11]

Complexity may 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.”[4,12–14] Much of the criticism of and extensions to the existing framework and guidance have focussed on the need for greater attention to be placed on understanding how and under what circumstances interventions bring about change.[5,12,15] The importance of intervention-context interactions emphasises the value of identifying: (i) mechanisms of change, where mechanisms are the causal links between intervention component(s) and outcomes, and (ii) contextual factors, which determine and shape whether and how outcomes are generated.[16] Thus, attention is paid both to the design of the intervention itself and 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, a cluster randomised trial of ASSIST, a peer-led smoking prevention intervention, found that the intervention worked particularly well in tight-knit communities served by a single 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.[17] 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, surgeons who will not be using robot assisted surgery, whole-team training, and a suitably-sized operating theatre.[18]

With this framing, complex interventions can helpfully be considered as ‘events in systems’.[13] Systems thinking helps us to understand the interaction between an intervention and the context in which it is implemented in a dynamic way.[19] Systems can be thought of as complex and adaptive,[20] characterised by properties such as feedback, emergence, adaptation, and self-organisation (Table 1).

For complex intervention research to be most useful to decision makers it needs to 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 underpinned by a paradigm in which the salient question was to identify whether or not an intervention was effective. Complex intervention research driven primarily by this question may fail to deliver interventions that are implementable, cost-effective, transferable and scalable in real world conditions. For complex intervention research to deliver solutions for real world practice requires strong and early engagement with patients, practitioners and policy-makers, shifting the focus from the "binary question of effectiveness”[21 p215],  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 assessment of how interventions contribute to system change.[19,22] 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 (see 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,[23] while an effectiveness evaluation can draw upon a theory-based or systems perspective through an embedded process evaluation to explore how and under what circumstances outcomes are achieved.[24–26]

Most complex health intervention research to date has taken an efficacy or effectiveness perspective and for some research questions these perspectives will continue to be the most appropriate. However, there are questions equally relevant to the needs of decision makers, which cannot be answered by research that is restricted to an efficacy or effectiveness perspective. A wider range and combination of research perspectives and methods, that address 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. Examples of such questions include: ‘will this ‘effective intervention’ reproduce the effects found in the trial when implemented here?’;‘is the intervention cost effective?’; ‘in the absence of a silver bullet, 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 occur as a result of this intervention?’; and ‘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 (Figure 1). A research programme may begin at any phase, depending on the key uncertainties about the intervention in question. Repeating phases 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).

## The Six Core Elements

## Context

The effects of a complex intervention may often be highly dependent on context, such that an intervention that is effective in some settings may be ineffective or even harmful elsewhere.[4] 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 health care, 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, e.g. 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.[27]

## Programme theory

Programme theory describes how an intervention is expected to lead to its effects and under what conditions. It articulates (1) the key components of the intervention and how they interact; (2) the mechanisms of the intervention; (3) the features of the context that are expected to influence those mechanisms; and (4) how those mechanisms may influence the context.[28] 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.[29] 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, which tested a large-scale quality improvement programme aimed at improving 90-day survival rates for patients undergoing emergency abdominal surgery, included a well-articulated programme theory at the outset, which supported the tailoring of programme delivery to local contexts.[30] The development, implementation and post-study reflection of the programme theory resulted in suggested improvements for future implementation of the quality improvement programme.

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 be helpful to provide a visual representation of aspects of it, for example using a logic model,[31–33] realist matrix,[34] or a system map,[35] with the choice depending on which is most appropriate for the research perspective and research questions. It should be noted that 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 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, i.e., 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 Involvement (PPI)[36] 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.[37] Involving service users with lived experiences of bipolar disorder had many benefits, e.g. it enhanced the intervention but also improved the evaluation and dissemination methods. There were also positive outcomes for service users involved in the study, 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. It is important to elicit stakeholder priorities, but also to consider why they are priorities. Careful consideration of the appropriateness and methods of identification and engagement of stakeholders is needed.[36,38]

## Key uncertainties

A large number of questions *could* be addressed at each phase of the research process. The design and conduct of research needs to engage pragmatically with the multiple uncertainties involved and offer a flexible and emergent approach to exploring them.[12] It is therefore important to 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 find out. Judgements 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.[22] For intervention research in health care 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 which emphasises implementation, context and system fit. This may help improve understanding and identify important implications for decision makers, albeit with caveats, assumptions and limitations.[19] Rather than maintaining the established tendency to prioritise strong research designs that answer some questions with certainty but are unsuited to addressing many important evaluation questions, this more inclusive, deliberative process may 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, there may be good reason, based on data collected or development of programme theory, to refine the intervention.[2] 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 based on interviews and observations. This iterative process led to the refined planner demonstrating greater feasibility and accuracy.[39]

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 may 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.[40] Broad ranging cost-benefit analysis (CBA) or cost-consequence analysis (CCA) approaches, that seek to capture the full range of health and non-health costs and benefits across different sectors,[41] will often be more suitable for an economic evaluation of a complex intervention than narrower approaches such as cost-effectiveness (CEA) or cost-utility analysis (CUA). For example, the evaluation of The New Orleans Intervention Model (NIM) for infants entering foster care in Glasgow included: short- and long-term economic analysis from multiple perspectives (NHS and Personal Social Services, public sector and wider societal perspectives); and used a range of frameworks, including CUA and CCA to capture changes in the intersectoral costs and outcomes associated with child maltreatment.[42,43] The use of multiple economic evaluation frameworks provides decision-makers with a comprehensive, multi-perspective guide to the cost-effectiveness of NIM.

## The Phases

## Developing or identifying a complex intervention

We use ‘development’ to cover 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 study;[2] though here we highlight that complex intervention research does not always begin with (i) new, or (ii) researcher-led interventions. For example:

1. A key source of intervention development may be an intervention that has been developed elsewhere and there is the possibility of adapting it to a new context. Adaptation of existing interventions may include adapting to a new population, setting, ,[44,45] or to target other outcomes, e.g. a smoking prevention intervention being adapted to tackle substance misuse and sexual health.[17,46,47] 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.[44,48]
2. 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 is an important step, even if roll out has begun, as 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 guided the evaluation design and helped identify data sources to support evaluation.[35] Another example is a recent analysis of the implicit theory of the NHS Diabetes Prevention Programme, involving analysis of documentation by NHS England and 4 providers, showing that there was no explicit theoretical basis for the NHS-DPP, and that this was responsible for ‘drift’ from evidence base to provider designs and delivery.[49]

Intervention identification and intervention development represent two distinct pathways of evidence generation,[50] though 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 pre-defined progression criteria that relate to: the evaluation design, e.g. reducing uncertainty around recruitment, data collection, retention, outcomes, and analysis; and/or the intervention itself, e.g. around optimal content and delivery, acceptability, adherence, likelihood of cost-effectiveness, or capacity of providers to deliver the intervention. Where the programme theory suggests that contextual or implementation factors may influence the acceptability, effectiveness or cost-effectiveness of the intervention, these questions should be addressed.

In the past, there has been a tendency for feasibility testing to be overlooked or rushed, though its value is now widely accepted with key terms and concepts well defined.[51,52] Before initiating a feasibility study, it may be useful to conduct 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.[53] The end result is a recommendation on whether an evaluation is feasible, can be carried out at a reasonable cost and by which methods.[54]

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 it is worthwhile proceeding to a full-scale evaluation.[55] Depending on the results of the feasibility study, further work may 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 impacts 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 real-world decision making. This implies a shift from an exclusive focus on obtaining unbiased estimates of effectiveness,[56] 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 a single primary outcome and a number of 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, pre-specified subgroup analyses should be carried out and reported. Even where such analyses are underpowered they should be included in the protocol as they may be useful as input into subsequent meta-analyses, or for developing hypotheses for testing in further research. Outcome measures may 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 useful illustration of the use of system-level outcomes is the evaluation of the Delaware Young Health Program - an initiative to address the health and well-being of young people in Delaware, USA. The intervention was expected 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.[57]

There are many study designs to choose from, and different designs are optimally suited to address different research questions and different circumstances.[58] Extensions to standard randomised controlled trial designs including adaptive designs, Sequential Multiple Assignment Randomized Trial (SMART) trials, N-of-1 trials and hybrid effectiveness-implementation designs, are important areas of methods development to improve the efficiency of complex intervention research.[59–61] Non-randomised designs and modelling approaches may be optimal where a randomised design is not practical, for example in natural experiments or systems evaluations.[3,62–64] 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 may be necessary to answer questions beyond effectiveness. In many evaluations, the nature of the intervention, the programme theory and/or the priorities of stakeholders may 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. More useful may be contextualised understandings of how an intervention effects change, and details on the most important enablers and constraints on its delivery across a range of settings.[5]

Process evaluation can answer questions around fidelity and quality of implementation (e.g., what is implemented, and how?); mechanisms of change (e.g., how does the delivered intervention produce change?); and context (e.g., how does context affect implementation and outcomes?).[5] Process evaluation can provide insight into 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.[65] When taking a theory-based or systems perspective, there is not necessarily such a clear distinction between process and outcome evaluation as there is in an effectiveness study.[66] These perspectives may prioritise theory-building over evidence production and utilise case study or simulation methods to understand how outcomes and/or system behaviour are generated through intervention.[63,67]

## 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 addressed 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 may support intervention transferability into different contexts—an important aspect of long-term implementation—[68] provided the key functions of the programme are maintained, and there is understanding of exactly what adaptations have been made.

For example, the ASSIST study,[17] a school-based peer-led intervention for smoking prevention, considered implementation at each phase: it 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 RCT) the intervention was delivered as closely as possible to real-world implementation. Drawing on the process evaluation, the implementation included: (1) an intervention manual that identified *critical* components and components that could be adapted or dropped to allow flexible implementation while achieving delivery of the key mechanisms of change; and (2) a training manual for the trainers and ongoing quality assurance built into longer-term rollout.

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.[64,69,70]

Implementation questions should be included in the economic considerations addressed at 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.[71] A key issue is how to deal with multi-sector interventions, where those paying for interventions and those receiving the benefits of them may 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 pre-empt any implementation challenges that might arise.[72]

# 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 greatest impact on health care 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 may work better, or may 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 may be less risky to commission and more likely to produce a clear and unbiased answer to a specific question. A bolder approach is needed. This will 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 will involve mainstreaming new methods that are not yet widely used, as well as undertaking methodological innovation and development. The deliberative, flexible approach that we encourage is intended to reduce research waste,[73] 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.[6] 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.

# 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, LMo facilitated the expert workshop, KS and LMa developed the gap analysis and led the analysis of the consultation. KB, NC and EMc contributed the economic components to the framework. JBa, JBl, DF, MP, JRM, MW were the Scientific Advisory Group and provided feedback and edits on drafts of the framework, with particular attention to process evaluation (JBa), clinical research (JBl), implementation (JRM, DF), systems perspective (MP), theory-based perspective (JRM), 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.

# Competing interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi\_disclosure.pdf and declare: funding for the work as detailed below;

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 (Nov 2016 - Nov 2020) and member of the Chief Scientist Office Health HIPS Committee (2018 - ongoing) and NIHR Policy Research Programme (Nov 2019 - ongoing) and has project grant funding from the ESRC, MRC and NIHR; LMo is a former member of the MRC-NIHR Methodology Research Programme Panel (2015-2019) and MRC Population Health Sciences Group (2015-2020); JBa is a member of the NIHR Public Health Research Funding Committee (May 2019 - ongoing), and a core member (since 2016) and vice-chairperson (since 2018) of a NICE Public Health Advisory Committee; JBl is a former member of the NIHR Clinical Trials Unit Standing Advisory Committee (2015–19); DF is a former member of 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 (REF) 2021, sub-panel 2: Public Health, Health Services and Primary Care (Nov 2020 to Feb 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, BXS GP Limited; EMc is a member of the NIHR PHR funding board; MP has grant funding from the MRC, UKPRP and NIHR; JRM is programme director and chairperson of the NIHR’s Health Services Delivery Research Programme (2014 to present) and member of the NIHR Strategy Board (2014 to present); MW received a salary as director of the NIHR PHR Programme (2014-2020), 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.

# Acknowledgements

We thank the experts who provided input at the workshop, those who responded to the consultation, and who provided advice and review throughout the process. The many people involved are named/acknowledged in the full framework document.

# Funding

The work was funded by the NIHR (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 Kathryn Skivington (MC\_UU\_12017/11, MC\_UU\_00022/3), Lynsay Matthews, Sharon Anne Simpson, Laurence Moore (MC\_UU\_12017/14, MC\_UU\_00022/1), Peter Craig (MC\_UU\_12017/15, MC\_UU\_00022/2), Martin White (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 Kathryn Skivington (SPHSU11 and SPHSU18), Lynsay Matthews, Sharon Anne Simpson and Laurence Moore (SPHSU14 and SPHSU16), and Peter Craig (SPHSU13 and SPHSU15). Kathryn Skivington and Sharon Anne Simpson were also supported by an MRC Strategic Award (MC\_PC\_13027). JBl received funding from 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). MP is funded in part as director of the NIHR’s Public Health Policy Research Unit.

# 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 e-mail and via digital channels, such as our unit Twitter account (Twitter, Inc., San Francisco, CA, USA; [www.twitter.com](http://www.twitter.com)). 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. There will be various dissemination events, though as this project is methodological we will not specifically disseminate to patients and the public beyond the planned dissemination activities.

**Table 1: Properties of complex adaptive systems, with examples**

|  |  |
| --- | --- |
| **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 at-risk young people may 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.[74]  |
| ***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.[19]  |
| ***Adaptation*** |
| A 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.[75]  |
| ***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. |

Table 2: Research Perspectives

|  |  |  |  |
| --- | --- | --- | --- |
| **Perspective** | **Research question** | **Key points** | **Vaccine Study Example** |
| ***Efficacy*** | To what extent does the intervention produce the intended outcome(s) 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.[76] |
| ***Effectiveness*** | To what extent does the intervention produce the intended outcome(s) 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, e.g. prevent disease.[77] |
| ***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 what this variation tells us about the conditions for a successful vaccination programme.[78]Considerations that may be explored go beyond whether the vaccine ‘works’.[79]  |
| ***Systems*** | How do the system and intervention adapt to one another? | - Treats the intervention as a disruption to a complex system.[13] | Seeks to understand the dynamic interdependence of vaccination rollout, population risk of infection and willingness to be vaccinated, as the vaccination programme proceeds.[80] |

# References

1 Craig P, Dieppe P, Macintyre S, *et al.* Developing and evaluating complex interventions: new guidance. London: : Medical Research Council 2006.

2 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

3 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

4 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.

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

6 Skivington K, Matthews L, Simpson S, *et al.* Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. *Health Technol Assess*;**In press**.

7 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;:BJGP.2020.1153. doi:10.3399/BJGP.2020.1153

8 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

9 Hawe P, Shiell A, Riley T. Complex interventions: how “out of control” can a randomised controlled trial be? *BMJ* 2004;**328**:1561. doi:10.1136/bmj.328.7455.1561

10 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

11 Blencowe NS, Skilton A, Gaunt D, *et al.* 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

12 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

13 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

14 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

15 Anderson R. New MRC guidance on evaluating complex interventions. *BMJ* 2008;**337**:a1937. doi:10.1136/bmj.a1937

16 Pawson R, Tilley N. *Realistic Evaluation*. SAGE 1997.

17 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. *The Lancet* 2008;**371**:1595–602. doi:10.1016/S0140-6736(08)60692-3

18 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*. Southampton (UK): : NIHR Journals Library 2017. http://www.ncbi.nlm.nih.gov/books/NBK447438/ (accessed 22 Jul 2021).

19 Rutter H, Savona N, Glonti K, *et al.* The need for a complex systems model of evidence for public health. *The Lancet* 2017;**390**:2602–4. doi:10.1016/S0140-6736(17)31267-9

20 The Health Foundation. Evidence Scan. Complex adaptive systems. London: : The Health Foundation 2010. https://www.health.org.uk/publications/complex-adaptive-systems (accessed 22 Jul 2021).

21 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

22 Kessler R, Glasgow RE. A Proposal to Speed Translation of Healthcare Research Into Practice. *Am J Prev Med* 2011;**40**:637–44. doi:10.1016/j.amepre.2011.02.023

23 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

24 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

25 French C, Pinnock H, Forbes G, *et al.* 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

26 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/#/ (accessed 22 Jul 2021).

27 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

28 Funnell SC, Rogers PJ. *Purposeful program theory: effective use of theories of change and logic models*. 1st ed. San Francisco, CA: : Jossey-Bass 2011.

29 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* 2017;**7**:510–21. doi:10.15171/ijhpm.2017.121

30 Stephens TJ, Peden CJ, Pearse RM, *et al.* 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

31 Bonell C, Jamal F, Melendez-Torres GJ, *et al.* ‘Dark logic’: theorising the harmful consequences of public health interventions. *J Epidemiol Community Health* 2015;**69**:95–8. doi:10.1136/jech-2014-204671

32 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

33 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

34 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

35 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 (accessed 22 Jul 2021).

36 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/ (accessed 22 Jul 2021).

37 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 Appl Res* 2018;**6**:1–296. doi:10.3310/pgfar06060

38 Moodie R, Stuckler D, Monteiro C, *et al.* Profits and pandemics: prevention of harmful effects of tobacco, alcohol, and ultra-processed food and drink industries. *The Lancet* 2013;**381**:670–9. doi:10.1016/S0140-6736(12)62089-3

39 Yardley L, Ainsworth B, Arden-Close E, *et al.* 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

40 Barnett ML, Dopp AR, Klein C, *et al.* Collaborating with health economists to advance implementation science: a qualitative study. *Implement Sci Commun* 2020;**1**:82. doi:10.1186/s43058-020-00074-w

41 National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. London: : NICE 2014. https://www.nice.org.uk/process/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869 (accessed 22 Jul 2021).

42 Boyd KA, Balogun MO, Minnis H. Development of a radical foster care intervention in Glasgow, Scotland. *Health Promot Int* 2016;**31**:665–73. doi:10.1093/heapro/dav041

43 Deidda M, Boyd KA, Minnis H, *et al.* 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

44 Escoffery C, Lebow-Skelley E, Haardoerfer 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

45 Stirman SW, Miller CJ, Toder K, *et al.* 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

46 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

47 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 Res* 2017;**5**:1–98. doi:10.3310/phr05070

48 Evans RE, Moore G, Movsisyan A, *et al.* How can we adapt complex population health interventions for new contexts? Progressing debates and research priorities. *J Epidemiol Community Health* 2020;:jech-2020-214468. doi:10.1136/jech-2020-214468

49 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

50 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

51 Eldridge SM, Chan CL, Campbell MJ, *et al.* CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ* 2016;:i5239. doi:10.1136/bmj.i5239

52 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

53 Craig P, Campbell M. Evaluability Assessment: a systematic approach to deciding whether and how to evaluate programmes and policies. *Evaluability Assess*;:18.

54 Ogilvie D, Cummins S, Petticrew M, *et al.* 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

55 Expected Value of Perfect Information (EVPI). YHEC - York Health Econ. Consort. http://yhec.co.uk/glossary/expected-value-of-perfect-information-evpi/ (accessed 22 Jul 2021).

56 Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. *Soc Sci Med* 2018;**210**:2–21. doi:10.1016/j.socscimed.2017.12.005

57 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

58 McKee M, Britton A, Black N, *et al.* 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

59 Burnett T, Mozgunov P, Pallmann P, *et al.* 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

60 Collins LM, Murphy SA, Strecher V. The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART). *Am J Prev Med* 2007;**32**:S112–8. doi:10.1016/j.amepre.2007.01.022

61 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

62 Tugwell P, Knottnerus JA, McGowan J, *et al.* Big-5 Quasi-Experimental designs. *J Clin Epidemiol* 2017;**89**:1–3. doi:10.1016/j.jclinepi.2017.09.010

63 Egan M, McGill E, Penney T, *et al.* NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. Part 1: Introducing systems thinking. London: : National Institute for Health Research 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 (accessed 22 Jul 2021).

64 Green BB, Coronado GD, Schwartz M, *et al.* 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

65 Bonell C, Fletcher A, Morton M, *et al.* 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

66 McGill E, Marks D, Er V, *et al.* 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

67 Bicket M, Christie I, Gilbert N, *et al.* Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation. *Lond HM Treas* 2020.

68 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

69 Curran GM, Bauer M, Mittman B, *et al.* 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

70 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

71 Imison C, Curry N, Holder H, *et al.* Shifting the balance of care: great expectations. Research report. Nuffield Trust

72 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

73 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *The Lancet* 2009;**374**:86–9. doi:10.1016/S0140-6736(09)60329-9

74 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

75 Robinson M, Geue C, Lewsey J, *et al.* Evaluating the impact of the alcohol act on off‐trade alcohol sales: a natural experiment in S cotland. *Addiction* 2014;**109**:2035–43. doi:10.1111/add.12701

76 Folegatti PM, Ewer KJ, Aley PK, *et al.* 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. *The Lancet* 2020;**396**:467–78. doi:10.1016/S0140-6736(20)31604-4

77 Voysey M, Clemens SAC, Madhi SA, *et al.* 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. *The Lancet* 2021;**397**:99–111. doi:10.1016/S0140-6736(20)32661-1

78 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**:ii35–46. doi:10.1093/heapol/czaa127

79 Burgess RA, Osborne RH, Yongabi KA, *et al.* The COVID-19 vaccines rush: participatory community engagement matters more than ever. *The Lancet* 2021;**397**:8–10. doi:10.1016/S0140-6736(20)32642-8

80 Paltiel AD, Schwartz JL, Zheng A, *et al.* Clinical Outcomes Of A COVID-19 Vaccine: Implementation Over Efficacy: Study examines how definitions and thresholds of vaccine efficacy, coupled with different levels of implementation effectiveness and background epidemic severity, translate into outcomes. *Health Aff (Millwood)* 2021;**40**:42–52. doi:10.1377/hlthaff.2020.02054