BMJ Global Health

Taking account of context in systematic reviews and guidelines considering a complexity perspective

Andrew Booth,[1] Graham Moore,[2] Kate Flemming,[3] Ruth Garside,[4] Nigel Rollins,[5] Özge Tunçalp,[6] Jane Noyes[7]

ABSTRACT

Systematic review teams and guideline development groups face considerable challenges when considering context within the evidence production process. Many complex interventions are context-dependent and are frequently evaluated within considerable contextual variation and change. This paper considers the extent to which current tools used within systematic reviews and guideline development are suitable in meeting these challenges. The paper briefly reviews strengths and weaknesses of existing approaches to specifying context. Illustrative tools are mapped to corresponding stages of the systematic review process. Collectively, systematic review and guideline production reveals a rich diversity of frameworks and tools for handling context. However, current approaches address only specific elements of context, are derived from primary studies which lack information or have not been tested within systematic reviews. A hypothetical example is used to illustrate how context could be integrated throughout the guideline development process. Guideline developers and evidence synthesis organisations should select an appropriate level of contextual detail for their specific guideline that is parsimonious and yet sensitive to health systems contexts and the values, preferences and needs of their target populations.

INTRODUCTION

Complex interventions, whether clinical, health service, public health or from beyond the health sector, include multiple components that interact along non-linear causal pathways within a complex adaptive system.[1 2] Interventions may share a common label and yet contain diverse content and variable components and use diverse modes of delivery, making effects unpredictable. By intervention, we include diverse procedures, practices, programmes and policies operating at different levels of system change. Yet the context or setting within which the intervention is introduced and interacts, 'a dynamic multidimensional environment'[3] constitutes the most significant aspect of this complexity.[4]

Summary box

► In systematic reviews and guidelines about complex interventions, considering context is challenging because it is diversely defined and dynamic.
► Context is typically caricatured as a location in time, in an organisational level or within a geographical space. Reviews and guidelines rarely provide sufficient detail of the context of an intervention in terms of its location in social, cultural, political or economic space.
► The extent to which reviews and guidelines of complex interventions consider context is limited by inadequate information on context in included primary studies.
► Collectively, existing tools help systematic reviewers to explore or report on context holistically. However, individually the tools differ in use of terms such as context, setting, environment and location and in what they consider as relevant.
► Reporting standards and templates, as well as improved reporting of context in primary studies, are needed to better account for context in systematic reviews and guidelines taking a complexity perspective.

Decision makers need to know when, why, how and in what circumstances complex interventions work well.[5 6] If a decision-maker does not understand the interplay between context and an intervention or policy (or more accurately the mechanisms by which the intervention or policy achieves its effect), then they can have only limited confidence in whether they would see the same effect in their own context.[7] Such a complexity perspective poses particular challenges in a global context when producing and implementing statements that include 'recommendations intended to optimise patient care or population health, informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options'.[8] Local decision makers need to decide whether a particular guideline is relevant at all, whether it is likely

to achieve a better or suboptimal effect and whether the intervention needs to be modified for contextual sensitivity.

Complex interventions interact with and modify the context within which they are introduced. Sometimes complex interventions adapt to the context in which they are implemented while some interventions explicitly aim to change the contexts within which they are implemented.[2] Complex interventions may alter a discrete aspect of practice (eg, how doctors interact with patients) or something more radical (eg, smoke-free legislation).[8] Key decisions when producing a particular guideline are whether it is helpful and informative to view an intervention as complex.[2 9 10] Not every topic justifies the added time and resource use required to explore complexity;[2 11] guideline developers may conclude that contextual variation is less critical to intervention effects than intervention complexity. Where complexity is to be considered, guideline developers must decide which characteristics or sources of complexity are to be prioritised. Evidence to Decision (EtD) frameworks offer a potential menu but not all aspects need to be addressed in every review or guideline.[10] Having decided whether a review will gravitate towards a complex intervention perspective or towards a complex system perspective, a review team can develop specific review questions to consider complexity and select evidence, synthesis methods and review tools accordingly.[12] Qualitative[13] and other forms of evidence[14] may help in understanding how an intervention interacts with, reshapes and is reshaped by its target context.

This paper reviews a selection of tools for considering context and evaluates their suitability for use in reviews and guidelines that take a complexity perspective. The paper offers guidance on good practice when considering context in systematic reviews and guidelines. The paper addresses the following questions: How is context defined and understood?; Why is context important in reviews and guidelines of complex interventions?; What tools and methods are available to handle context in systematic reviews? and What might represent good practice for considering context in reviews and guidelines?

HOW IS CONTEXT DEFINED AND UNDERSTOOD?

A plethora of definitions exists for 'context', each varying in the extent to which they adopt an 'intervention perspective' or a 'system perspective';[2] making it problematic to privilege a single frame of reference.[15] UK Medical Research Council guidance defines context as 'factors external to the intervention which may influence its implementation, or whether its mechanisms of impact act as intended'.[16] Further guidance defines context as 'any feature of the circumstances in which an intervention is conceived, developed, implemented and evaluated'.[17] Methodologically, a tension exists between the technical advantages of deciding on a single standardised definition of context and the conceptual richness of diverse and competing definitions. A systematic review

should clearly articulate its internal definition of context, justify its selection and apply this definition consistently through the review process.

Context and setting, although used interchangeably, are not synonymous. The former typically carries an overarching sense while the latter is often locally bounded (eg, an acute hospital setting in a low-income or middle-income development context). Research and evaluation, broadly speaking, seek either to ignore, control, document or explore contextual variation. Many intervention and evaluation designs seek to ensure that the context for the intervention group is the same as that in the control group, so that the intervention is the only meaningful difference between groups. Certain contexts are more easily explored by randomised controlled trials (eg, secondary or tertiary care)[18] but become less amenable to controlled studies on implementation. Current design and statistical approaches do not adequately deal with contextual variations within source studies and imperfectly represent 'real world' conditions within the target context.[19] Demarcation between the outer setting (the economic, political and social context within which an organisation resides) and the inner setting (features of structural, political and cultural contexts through which an implementation process will proceed) are 'not always clear', given that the interface is 'dynamic and sometimes precarious'.[20] One concept analysis states: 'Context reflects a set of characteristics and circumstances that consist of active and unique factors, within which the implementation is embedded'.[21] Subsequently, the concept analysis unpacks context into eight domains (ie, locational, geographical, epidemiological, sociocultural, socioeconomic, ethical, legal and political). Context typically incorporates the institutional context (setting), the wider environment (eg, epidemiological, genetic, geographical, sociocultural, health system specific, political and others[22]) or a composite of both.

Authors of primary and secondary research find context challenging to operationalise, research and evaluate.[23] Primary studies are criticised for failing to characterise context,[17] yet many guidance documents recommend that contextual factors are examined within systematic reviews and syntheses.[24] Review teams face particular challenges resulting from (1) the lack of contextual detail in included studies and (2) unsophisticated approaches to handling context within existing review tools.

The Preferred Reporting Items for Systematic reviews and Meta-analyses (PRISMA) reporting standards,[25] an evidence-based minimum set of items to help authors improve reporting of systematic reviews and meta-analyses, refer to 'context' as the circumstances that explain why a review is currently needed (ie, the need for decision support and the prevalent state of knowledge). PRISMA does not even reference 'context' in its common usage as 'the population targeted by implementation'. The PRISMA-CI extension for Complex Interventions, a supplemental set of items for reporting systematic reviews of complex interventions, partially addresses

such limitations,[26 27] extending to include the elements of Time and Setting. Nevertheless, this 'intervention perspective' proves unable to accommodate a complex systems perspective.[2]

Adding an additional component to a complex intervention (eg, a practice, behaviour or a way of organising care) almost invariably displaces something else, whether planned or not, thereby challenging conventional evaluation.[4] While the context of an intervention is the focus for change, it is important to acknowledge that 'usual care' equally operates within a complex adaptive system. Interventions may not transfer to other contexts because usual care in those contexts is already more effective than the intervention or at least more responsive to local needs. Furthermore, studies frequently offer only a sketchy picture of what constitutes 'usual care'.

WHY IS CONTEXT IMPORTANT IN REVIEWS AND GUIDELINES OF COMPLEX INTERVENTIONS?

Key tools and methods used in the systematic review process similarly lack clarity and consensus surrounding the term 'context'. Context is typically either reported in insufficient detail or omitted altogether from many systematic reviews. Systematic reviews are frequently characterised as seeking to 'strip away' context.[28] Transferability of review findings, particularly to Low and Middle Income Countries (LMICs), depends on resources, health systems, cultural norms and the physical environment.[29] Wider considerations include geographic, seasonal, environmental, cultural, economic, sociopolitical, regulatory and organisational inner (eg, payment policies) and outer (eg, organisational cultures) settings.[30] Differences between national health services and insurance-based health systems are similarly important within high income countries.[30] Above all, those examining context should seek to extend analysis beyond the caricaturisation of time and place to explore other, less salient, features of contextual variation.

Table 1 characterises three different evaluation perspectives ranging from an intervention perspective, through a more hybrid 'partial system' approach, to the system perspective. The table reveals increasing complexity with subsequent rows increasingly being likely to require, and to reward, examination of contextual variation.

Systematic review and guideline production seeks to 'globalise the evidence, localise the decisions'.[31] This globalise-localise strategy distinguishes three 'contexts' of evidence production: (1) the context of the review question (the review context); (2) the contexts of included studies (the source contexts) and (3) the implementation context into which recommendations are introduced (the target context). So, a review focused on low-income and middle-income countries (the review context) may access and include studies conducted in Africa and South America (source contexts) to inform practice in South East Asia (the target context). The context within which a study is conducted may differ considerably from the target context within which study results are later used.[32]

WHAT TOOLS AND METHODS ARE AVAILABLE TO HANDLE CONTEXT IN SYSTEMATIC REVIEWS?

Context is currently factored into systematic review processes in two main ways:

1. Contextual items included within existing review tools (eg, specific tools for describing interventions and for assessing complexity).[33] Many review tools already include one or more items for context.
2. Use of a supplementary framework to explore contextual issues.[30] A guideline group can use a context framework when planning reviews (eg, when formulating questions within a review) or while conducting reviews (eg, when representing subgroups within the synthesis).

Table 2 provides examples of known tools and frameworks for taking account of context. Full text searching of Google Scholar for references to 'context' within articles

Table 1 Perspectives mapped to role of context and evaluation activities

Perspective	Role of context	Characteristic evaluation activity	Important elements of context to consider
Intervention perspective	Context tends to have a 'disruptive' role in enhancing or impairing intervention effects	Health Technology Assessment*	Characteristics of usual care Existing skills, training and expertise Resources acceptability and feasibility
'Partial' system perspective	Contextual variation potentially explains differential effects across organisations	Health Services Delivery Research	Organisational context, culture and values Acceptability and feasibility
System perspective	Contextual variation represents a form of 'chaos' against which evaluators identify common mechanisms	Health Systems and Public Health Research	As above, but extended to include economic, cultural, social, gender and political factors

The above perspectives lie on a continuum; a 'partial system perspective' characterises a scenario that acknowledges selective aspects of a wider system but does not engage with the full range of system factors.
*Increasingly initiatives such as EUnetHTA or INTEGRATE-HTA endorse a societal perspective of health technology assessment (HTA). Organizations such as the WHO are increasingly moving from clinical recommendations to clinical/health system (hybrid) guidelines, further blurring such distinctions.

Table 2 Tools and frameworks for handling context, mapped to stages of a systematic review process

Stage of the review	Tool or framework	Strengths	Limitations
Throughout	Realist evaluation[7]/Realist synthesis[34]	Offers a method for integrating quantitative and qualitative data. Focuses on context as one line of inquiry, specifically addressed when assessing programme outcomes.[15]	Resource intensive; requires reviewers to prioritise which explanations would benefit from further exploration.
	Systems thinking[4]	Offers holistic picture of complex phenomena and interactive elements including context.	No specific review or synthesis tools or methodologies have been designed for this perspective.
	Evidence to Decision frameworks (eg, WHO-INTEGRATE[10] or GRADE EtD[38] framework)	Seek to identify elements of context throughout entire evidence production and guideline development process.	Comprehensive coverage makes it difficult to use entire tool within a tight systematic review time window. Reviewer must prioritise items relevant to specific review.
	Logic models[2]	Provide a graphical representation of context surrounding an intervention.	Offer simplistic perspective with a focus on inputs, processes, outputs and outcomes. Context is only identified as a backdrop to programme activities.
Planning the review			
Question formulation and defining eligibility criteria	Question structures:[11] PICOC,[41] CIMO,[42] SPICE,[43] PerSPE©TIF[11]	Include important elements of context/setting when compared with intervention-based PICO formulation.	Adopt a simple linear framework for representing review question.
Protocol development	PROSPERO (International prospective register of systematic reviews) template[45]	Item #23 (context) requires summary details of the setting.	Focuses on context as it relates to study eligibility.
	Use of frameworks (see itemised list below)	Offer prompts to ensure appropriate issues are considered when conducting the review.	Few available frameworks for systematic consideration of context.
	PRISMA-P[44]	Reporting standard for all elements of a protocol.	Superficial approach to context (setting and time frame).
Conducting the review			

Continued

Table 2 Continued

Stage of the review	Tool or framework	Strengths	Limitations
Data extraction and/ or analysis	Framework for the design and reporting of cluster RCTs.[46]	Requires careful description of context to help interpret findings.	Designed as a framework for primary research. Requires evaluation within a synthesis context.
	CFIR[20]	Itemises 39 factors influencing implementation of interventions offering taxonomy for classifying factors.[20] Terminology reported as clear and coherent.[91]	Complex framework that may include too many factors to be practical for data extraction reflecting that it has been designed for primary studies.[91]
	CICI framework[22]	Based on concept analysis therefore offers broad coverage of context issues.	Limited conceptual thinking on the inter-relationships of framework elements of context and implementation.
	Diffusion of innovations[40]	Considers temporal issues of context and mechanisms of sustainability and spread.	Does not offer practical framework for most aspects of context.
	iCAT-SR[33]	Requires review team to make judgements on the extent to which the effects of the intervention depend on the context or setting within which it is implemented.	Limited specific detail on context although supplemented by embedded items with contextual implications, for example, organisational level.
	Normalisation process theory[49]	Acknowledges wider societal and contextual barriers and organisational and social factors. Defines context at early stages of intervention; implementation and evaluation.	While recognising context as a dynamic environment, and as a process, not a place; challenging to operationalise within a synthesis context.
	PARIHS[47]	Context is core construct and interacts with other constructs.	Focuses on implementation setting, not wider social, political or economic environment. Originally designed for primary studies.[47]
	PROGRESS-Plus[53]	Includes population-based, social and cultural factors to explore/explain variation in intervention feasibility, acceptability, meaningfulness and effectiveness.	Broad population focus at expense of less common but equally important factors resulting in individual variation.
	RE-AIM framework[48]	Uses stakeholders to prioritise aspects for particular focus.	Primary study focus. Extensive resource use when exploring all aspects within a synthesis. Focuses mainly on number individuals adopting evidence as main contextual factor.[16]
	Theoretical Domains Framework[50]	Specifically designed for implementation contexts. Includes Environmental context and resources and Social influences. Extensive use in systematic reviews.	Focuses on organisational level context with passing acknowledgement of external factors in the environment that cause stress.
	TIDieR[55]	Includes a WHERE question to elicit different levels of contextual detail.	Focuses on intervention description; offers exemplar for how context might be developed. Focuses on spatial aspects of context.

Continued

Table 2 Continued

Stage of the review	Tool or framework	Strengths	Limitations
Literature searching	Citations, Lead authors, Unpublished materials, (Google) Scholar, Theories, Early examples, Related projects (CLUSTER) technique[58]	Identifies additional 'sibling studies' to expand detail of context.	Potentially time-consuming and dependent on linked publications.
	Web searches for supplementary data on context not present in published reports	Offers accompanying detail to complement published studies.	Requires opportunistic searching; challenging to document.
Data extraction	Context templates	Prompts review team to compile proforma assessments to standardise reporting of context.	Missing data leads to inconsistent reporting and impaired comparison.
	iCAT_SR tool[33]	Includes context among 10 dimensions that examine interplay of factors contributing to complexity.[2]	Context dimension is optional, not required.[92] Other elements implicitly relate to context.
	TIDieR[55]	As discussed above.	As discussed above.
Making use of review findings			
Producing the report	Reporting Standards (eg, CONSORT;[62] PRISMA;[25] PRISMA-CI;[26] ENTREQ[63])	Prompts to identify and elicit missing data on context.	Context or setting typically constitutes a single (or limited) number of items.
Grading of findings	GRADE[65]	Global influence supported by established methods group. Recently addressed contextualisation for (1) guidelines and (2) HTAs/ systematic reviews.[93]	Lacks structured approach for incorporating non-epidemiological evidence on context and implementation.[39]
	GRADE-CERQual[66]	Includes relevance element and broadens beyond time and place. Work in progress.	Requires further development to extend beyond less salient aspects of context.
Research recommendations	EPICOT-dts[84]	Optional elements include Disease burden or relevance and Time aspect of core EPICOT elements.	Little evidence of uptake and utilisation (excepting NICE and isolated WHO guidelines).
Implementation	Oxford Implementation Index[77]	*Contextual factors* is largest of four domains: *intervention design*, actual *delivery* by trial practitioners, *uptake* of intervention by participants.[77]	Sheer number of contextual elements requires prioritisation, particularly within context of evidence synthesis.
	Case studies, profiles, vignettes	Designed for use within local implementation context and to permit local tailoring.	Only focus on isolated simplistic aspects of context.
	Spreadsheets, calculators (eg, NICE Physical Activity Return on Investment Tool)[75]	Enables user to evaluate a portfolio of interventions in their geographical area (eg, region, county or local authority) and model different payback returns. Interventions are mixed and matched to identify best 'value for money'.	Underpinning assumptions behind model may change. Requires periodic updating.

Continued

Table 2 Continued

Stage of the review	Tool or framework	Strengths	Limitations
Adaptation	SUPPORT tools[59]	Offer a comprehensive and systematic guide for considering local context throughout the decision-making process.	Limited number of examples. May oversimplify decision making process—lack a complex system perspective.
	ADAPTE and other adaptation guidelines[94]	Offer alternatives to developing de novo guidelines (eg, ADAPTE covers organisational and cultural considerations).	Frameworks need to be evaluated for rigour, efficiency and transparency of process.

CFIR, Consolidated Framework for Implementation Research; CICI, Context and Implementation of Complex Interventions; CIMO, Context-Intervention-Mechanisms-Outcome;CONSORT, Consolidated Standards of Reporting Trials; ENTREQ, Enhancingtransparency in reporting the synthesis of qualitative research; EPICOT,Evidence, Population, Intervention, Comparison, Outcomes, and Time Stamp; EtD, Evidenceto Decision; GRADE, Grading of Recommendations, Assessment, Development,and Evaluation; GRADE-CERQual, Grading of Recommendations, Assessment, Development,and Evaluation-Confidence in Evidence from Reviews of Qualitative research; HTA, Health Technology Assessment; iCAT-SR, Intervention Complexity Assessment Tool for Systematic Reviews; LMIC, NICE,National Institute for Health and Clinical Excellence; PARIHS, PromotingAction on Research Implementation in Health Services; PerSPE©TIF, Perspective-Setting-Phenomenon-Environment-Comparison-Timing-Findings; PICOC, Population-Intervention-Comparison-Outcome-Context; PRISMA-P, PreferredReporting Items for Systematic Reviews and Meta-Analyses for Protocols; PRISMA-CI,Preferred Reporting Items for Systematic Reviews and Meta-Analyses for ComplexIntervention; PROGRESS-Plus, Place of residence, Race/ethnicity/culture/language, Occupation, Gender/sex, Religion, Education, Socioeconomic status, Social capital;RCT, randomised controlledtrial; RE-AIM, Reach, Effectiveness,Adoption, Implementation, and Maintenance; SPICE, Setting-Perspective-Interest,-Phenomenon-of-Comparison-Evaluation; TIDieR, Template for Intervention Descriptionand Replication.

Box 1 Factoring in context within systematic reviews for complex interventions

Throughout the review
0. Use a review method (eg, realist synthesis or meta-ethnography) that allows exploration of context and/or

Planning the review
1. Consider contextual details during question formulation and when defining eligibility criteria.
2. Describe contextual considerations in the review protocol.

Conducting the review
3. Extract contextual details as part of data extraction.
4. Seek missing contextual details.
5. Consider contextual characteristics during statistical and qualitative analyses and explore contextual heterogeneity when appropriate.

Making use of review findings
6. Report contextual details in supplementary materials.
7. Develop support tools to facilitate implementation and adaptation.
8. Describe implications for future research.

---

reporting systematic review methodology was used to identify candidate tools. This paper seeks to illustrate well-known or recently developed approaches to be used at different stages of the systematic review process and to summarise their potential strengths and weaknesses. While not seeking to be exhaustive, the table can be used to identify composite approaches that explore complementary perspectives of context.

How to consider context when planning, conducting and implementing reviews

Review authors have numerous opportunities to consider contextual details when planning and conducting the review and, subsequently when making recommendations (box 1).

Throughout the review

One approach to handling context within systematic reviews and guidelines is to use a review method that explicitly articulates context within its programme logic. Realist synthesis[34] may identify contextual variation to be explored through complementary use of quantitative and qualitative data sources.[16] Realist synthesis seeks to explain why interventions may or may not work, within which contexts, how and in what circumstances'.[35] Specific techniques for juxtaposing quantitative and qualitative evidence are elaborated elsewhere in this series.[13 14] Synthesis methods that specifically examine context, as with realist review, or that seek to preserve context when 'translating' across studies (as with a meta-ethnography alongside an effects review) offer a holistic approach. However, they struggle to deliver within the constrained time and resource timetable[36] that characterises guideline production.

The value of systems thinking[4] is explored elsewhere. However, although applications of system thinking in public health are emerging,[37] few systematic reviews have encapsulated systems thinking within their review methods. EtD frameworks, such as the GRADE EtD[38] framework and the WHO-INTEGRATE framework,[10] itemise many contextual factors impacting on decision-making. Nevertheless, exhaustive frameworks may prove resource-intensive when harnessed to a systematic review. Guideline developers and the review team should consider which, if any, elements from this expansive menu will prove informative for their particular guideline/review.

Logic models are graphic descriptions of a system and are designed to identify important elements and relationships within that system.[2] In seeking to portray multiple interacting elements, they represent a light touch response to context; the focus is on the intervention or programme itself, together with associated inputs and activities, and context offers a frame for intervention delivery. So a logic model may be useful for initiating discussion of context but will probably require further, more granular, specification.

Planning the review
Consider contextual details during question formulation and when defining eligibility criteria
Methodologists emphasise that a review team should consider context from initiation of the review.[11 39] Assumptions about context first surface within the Population, Intervention and Comparison elements of a review question; a review may define its context as 'low and middle-income settings', as 'Africa' or as a named country. However, a review question may favour geopolitical aspects of context over equally important considerations such as rural-urban differences, ethnicity, infrastructure and policy. Health systems frequently characterise context as primary, secondary, tertiary and community but this demarcation may run counter to a systems perspective.

Question formulation frameworks represent flat, linear representations of context[11] that do not accommodate a systems perspective. In contrast, a systems perspective, where macrolevel environmental factors and microlevel factors of setting dynamically interplay, may require a comprehensive question framework.[40] Few question formulation frameworks specify context (or associated subconcepts, eg, setting or environment). Nevertheless, question structures such as Population-Intervention-Comparison-Outcome-Context (PICOC),[41] Context-Intervention-Mechanisms-Outcomes (CIMO)[42] and Setting-Perspective-Interest, phenomenon of-Comparison-Evaluation (SPICE)[43] offer alternatives to the PICO structure, which omits context (environment or setting) altogether and may prompt guideline developers or review authors to articulate context concerns. A review team may incorporate question structures that include environment/setting within their inclusion and

exclusion criteria and the data extraction form. An innovative, yet untried, question formulation, PerSPE©TiF[11] may prompt a review team to identify elements of context not typically articulated within an intervention perspective.

Describe how context is to be defined within the review protocol
Other aspects of the review protocol have an implicit role in defining context, for example, inclusion and exclusion criteria, the choice of data sources and search terms as well as language and date restrictions. However, two principal structures for codifying the review protocol, PRISMA-P[44] and PROSPERO[45] noticeably lack detail of context. PRISMA-P, a reporting standard for review protocols, simply lists setting and time frame alongside the PICO elements. No detail is given on specific details to be included. The template used to register reviews in PROSPERO, an international registry of systematic reviews, requires contextual details 'if these form part of the review's eligibility criteria but are not reported elsewhere in the PROSPERO record'. Examples include: 'studies in hospital accident and emergency departments' or 'research in low-income and middle-income countries.' Detailed reporting in a protocol template can help in agreeing the scope of the review and to surface assumptions of similarity and difference across contexts (eg, combining studies with broadly comparable generalisable study features, eg, low-income and middle-income countries) and splitting (only combining studies with highly comparable populations for example, East Africa).[45] However, such templates require the review team to preidentify contextual differences before starting to undertake the review.

Conducting the review
Extract contextual details as part of data extraction
Numerous purpose-specific frameworks exist that offer a potential structure for data extraction forms. Generic frameworks, designed for primary research, are suggested in the Medical Research Council Process Evaluation Guidance.[16] Examples include the Cluster randomised controlled trials[46] framework, PARIHS[47] and RE-AIM.[48] However, frameworks designed for use within primary studies may include too much detail to sustain data extraction for a review. Other frameworks remain experimental such as the Context and Implementation of Complex Interventions (CICI) framework, produced for a European Union-funded project as a generic framework for exploring for the context of any health technology.[22] Similarly, the WHO-INTEGRATE framework enumerates the 'Health system and broader context'.[10] When conducting a systematic review of a complex intervention, a team must prioritise important contextual elements considered important for data extraction rather than seeking to use frameworks in their entirety.

Implementation frameworks (such as Normalisation Process Theory[49] and the Theoretical Domains Framework[50]) offer an alternative lens on context. A systematic

review of facilitators and barriers to team working in primary care used Normalisation Process Theory to identify the importance of contextual variation in funding models and colocation.[51] The Theoretical Domains Framework was used in a systematic review on barriers and enablers to acute stroke management in the emergency department revealing important contextual differences in social influences.[52] Frameworks must optimise usability and detail; the Diffusion of Innovations framework[40] does not appear sufficiently nuanced for data extraction whereas the meta-framework, Consolidated Framework for Implementation Research,[20] appears overly detailed. Craig and colleagues have combined and consolidated a comprehensive list of features of context.[17]

Potential is offered by PROGRESS-Plus, devised by the Cochrane Equity Methods Group to examine any intervention or programme, where a review team identifies concerns of equity.[53] The PROGRESS-Plus framework includes contextually sensitive details such as race, religion and ethnicity and may therefore help in isolating important variations that impact on the effects of a complex intervention.[54]

The Intervention Complexity Assessment Tool for Systematic Reviews (iCAT-SR)[33] and the Template for Intervention Description and Replication (TIDieR)[55] specify elements of context alongside salient details of the target intervention. The iCAT_SR tool,[33] which seeks to locate an intervention on a simple-complex continuum, references context as how 'the effects of an intervention may be dependent on the societal, political, economic, health system or environmental context in which the intervention is delivered'. Such an approach helps to 'embed' context alongside intervention complexity.

The TIDieR[55] checklist is used for describing interventions in sufficient detail for replication and can help to extract information on WHERE an intervention was delivered, which is important in considering context. The TIDieR framework encourages a detailed description of location including any necessary infrastructure or relevant features.[55] The importance of specific factors depends on the type of health decision and the decision-making context, but the TIDieR framework may be helpful when extracting detailed information about spatial aspects of context.[55] A subsequent paper evaluating TIDieR[56] contrasts full reporting of interventions (to enhance replicability) with detailed reporting of context (to surface implementation issues). TIDieR, as the first major reporting tool specifically to examine implications when reporting context, signals an opportunity to recalibrate away from seeking to replicate context to the more realistic aspiration of seeking to understand and interpret its effects.

While these tools and frameworks present possible structures for extracting data on context, they tend to artificially separate intervention characteristics and contextual considerations, which is unhelpful when reviews and guidelines specifically need to identify and take account of sources of complexity.

Seek missing contextual details

Where the review team considers that important details about context are missing from the included studies, there are three principal options for locating this information.[57] The CLUSTER technique[58] has been devised for searching systematically for associated ('sibling') reports, such as pilot, feasibility studies, doctoral theses, book chapters, process evaluations and qualitative research, that offer supplementary detail beyond that retrieved from a systematic subject search.

Alternatively, a team can conduct supplementary web searches to build up context templates from routine government data and from policy sources for the study context of each 'index study'. Policy and sociodemographic characteristics appear as national profiles on websites of government agencies or of international organisations such as the World Bank or WHO. Institutional websites also host organisational details, for example, the number of hospital beds or the catchment area for a health service. Finally, the team can harness the collective knowledge and expertise from an expert advisory group, but such approaches pose challenges to the need to be both systematic and explicit.

Consider contextual characteristics during statistical and qualitative analyses and explore contextual heterogeneity when appropriate

Where appropriate and feasible, a review team should consider context, as identified a priori in the protocol, when grouping studies, conducting analyses and exploring heterogeneity.[59] Higgins and colleagues describe statistical and non-statistical methods available for synthesising quantitative data,[19] including subgroup analyses and meta-regression for examining sources of heterogeneity including context. Flemming and colleagues describe approaches for qualitative evidence synthesis, particularly helpful for exploring variation in context and explaining why certain interventions work in some contexts and not others.[13] Noyes and colleagues address methods for mixed method synthesis and how such evidence can help a decision-maker to understand how context may affect feasibility and acceptability.[14] All these methods possess advantages and disadvantages; essentially, though, the choice of methods centres on the extent to which a review team aims to gain a greater understanding of the study context (by examining potential explanatory factors within the study data) or the extent to which the team seeks to focus on transferability from the study context to the target context (by exploring shared contextual characteristics).

Making use of the review findings
Report contextual details in supplementary materials

Reporting standards for primary studies seek to encourage more complete reporting of context; however, they prescribe diverse levels of suggested reporting detail.[60] Craig and colleagues affirm that 'indicating and describing clearly the contexts in which the research

was carried out, how they affect implementation and variation in impacts and the limits to the applicability of the results should make reports of primary studies and evidence syntheses more useful'[17] to researchers and decision-makers alike. Shoveller and colleagues report from a snapshot of current journals that context was typically treated as a 'black box' or problematised as something to be 'controlled for'.[61] A systematic review of a complex intervention can only report contextual details if they are available in the index primary trial reports or in supplementary publications. CONSORT (2010) suggests reporting 'other information about setting or location that could have influenced the observed results'[62] yet offers limited guidance on how context should be reported.

Currently, PRISMA[25] demonstrates superficial engagement with the need to report details of context. The specific PRISMA-CI extension for complex interventions does, however, extend coverage of contextual detail and hint at potential future directions.[26 27] Thus, the PICO structure becomes a six-element PICOTS mnemonic to include Time and Setting.[27] Furthermore, PRISMA-CI recognises organisational features, geographic location; financial setting and clinical setting and, optionally, 'rival activities'[27]. In contrast, the Enhancing Transparency in Reporting the synthesis of Qualitative research (ENTREQ) reporting standard[63] does not include any item specifically targeted at context and, indeed only refers to population and country. ENTREQ does, however, advocate the need to specify context when formulating the review question and also alludes to concerns of transferability.

The WHO has recently sought to enhance reporting of programme level detail, recognising that the impact of many social and behavioural programmes is 'tied to the local context (eg, sociocultural, socioeconomic, geographical, legal, political, health system) and to the processes of implementation, which may not be easily described'.[64]

Tools for assessing the quality (or certainty) of evidence contained in systematic reviews, that is the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach for quantitative outcomes[65] and the Confidence in the Evidence from Reviews of Qualitative Research (GRADE-CERQual) approach for synthesised qualitative findings[66] facilitate post hoc consideration of context when moving from evidence to recommendations. GRADE assessment criteria recognise contextual variation (eg, geographic setting, culture, subcultures and resources) and also differences in perspective (patients, providers, public health and society and of individual vs population based evaluation frames).[67] Both the GRADE domain of indirectness[68] and the GRADE-CERQual component of 'relevance' require a reviewer to assess the extent to which study contexts match the target contexts.[69]

Grading systems are constrained by their reliance on the level of detail in the source primary studies; detail that is often absent. Furthermore, they depend on the ability of the reviewer to identify salient contextual detail a priori and to make this detail the focus for analysis. Finally, summary of findings[65 66] are limited in the contextual detail they can accommodate, necessitating difficult choices.

Develop support tools to facilitate implementation and adaptation of systematic review findings

Applicability or generalisability checklists, of which more than 25 frameworks exist,[70] help when translating systematic review findings into specific contexts. The SUPPORT tools (https://health-policy-systems.biomedcentral.com/articles/supplements/volume-7-supplement-1), written for decision makers and those who support these decision makers, include tools on assessing the applicability of review findings[71] and using local contextual evidence.[72]

Tools to support implementation include case studies, country profiles and vignettes which prompt the user to consider potential variation in context. Individual country case studies can help decision-makers to visualise differences between a source context and a target context or to benchmark across countries.[73] However, numbers of case studies are necessarily limited and they may lack the depth required to unearth critical differences. Country profiles require a study author to make an a priori 'editorial' judgement on what is 'relevant'. Vignettes use narrative techniques to characterise complex circumstances and to trigger critical thinking about similarities and differences with local practice. Typically vignettes simplify population characteristics (eg, age or severity of a condition) but they could equally capture contextual characteristics. For example, Gauld *et al,* (2012) use vignettes to characterise qualitatively the primary care systems of 12 different countries, supported by routine quantitative data on key indicators.[74]

Spreadsheets and calculators require users to add local contextual data to be factored within resultant calculations and projections. So, for example, the National Institute for Health and Clinical Excellence in the UK offers return on investment tools that allow local health providers to configure optimal packages of interventions for physical activity and alcohol or tobacco control for their populations.[75] The NICE Guidelines programme routinely produces three types of implementation support tool to accompany each guideline:[76]

► a baseline assessment tool,
► clinical audit tools,
► a costing report and costing template, or a costing statement.

Such tools use common software, such as Microsoft Excel, to improve uptake.

The Oxford Implementation Index offers a rare example of a tool to help reviewers identify and extract contextual information associated with implementation from studies included in a systematic review.[77] Other tools include implementation workbooks and worksheets.[78 79]

The ADAPTE toolkit[80] highlights criteria for assessing whether a recommendation is applicable to a local context.

Considering contextual factors on completion of a review, using applicability checklists and decision-making frameworks, helps to make review findings more useful to guideline developers. Alternatively, strengthening the conduct of the systematic review[19 29] can lead a review team to consider a complexity perspective a priori using an EtD-making framework.[10]

Context is pivotal when localising guideline recommendations. Guidance from the GRADE Working Group directs guideline developers in deciding when to adopt (use a guideline 'as is'), adapt (introduce local evidence and expertise to rework guidelines) or contextualise (reframe recommendations in light of local resources and constraints) and when to initiate guideline development de novo.[81 82]

Describe implications for future research

An important outcome of any systematic review, or guideline, is articulating gaps in knowledge or evidence on the topic. However, often reports of reviews and guidelines provide only general observations and miss the opportunity to identify important gaps relating to different contexts. Where important contexts are missing, resulting in 'absolute gaps' or 'synthesis gaps'[83] the review team can identify implications for future research. The Evidence Population Intervention Comparison Outcome Timestamp (EPICOT) framework, developed by organisations involved in commissioning or summarising research, is recommended for formulating research recommendations.[84] Optional elements in the framework (current state of evidence (E), appropriate study type (s), disease burden and relevance (d) and timeliness (t)) help to ensure more detailed definition of context.[84] The WHO Guidelines on Indoor Air Quality[85] used this EPICOT reporting structure. However, the EPICOT framework shares limitations of the PICO question framework in providing a static linear framework and it does not take into account other important characteristics of the immediate setting or the wider environment.

WHAT MIGHT REPRESENT GOOD PRACTICE FOR CONSIDERING CONTEXT IN REVIEWS AND GUIDELINES?

This paper reaffirms that, where possible, systematic review authors and guideline developers should identify relevant aspects of context from the beginning of the process, when the scope of the review or guideline is determined and the questions identified.[86] Sandelowski offers some useful 'rules of thumb' on how to anchor text to context in restatements of findings.[87] Little consensus exists on how best to identify and select relevant aspects of context that may be important in reviews and guidelines considering a complexity perspective. Other papers[2 10] in this series offer practical suggestions on how to achieve this.

A review team should recognise that different types of evidence can contribute to an improved understanding of context. For example, Craig and colleagues illustrate how inclusion of qualitative evidence, in the form of a meta-ethnography can help in exploring context.[17] Natural experiments and uncontrolled studies can present opportunities for understanding context and yet are frequently excluded from systematic reviews.[17] Systematic reviews can further be designed to incorporate implementation and outcome evidence, rarely reported within a single primary study but usefully juxtaposed within implementation reviews. Ideally, a review team will plan a priori analyses based on anticipated differences in context.[88] Nevertheless, if they have not preplanned such analyses, exploratory post hoc analyses, clearly identified as such, may be appropriate in generating further hypotheses.[88]

From initiation of a review, the review team should look out for studies reporting interventions from highly disparate contexts, whether differences are openly apparent or masked by incomplete reporting.[77] Variations in context impact on the review question, the search strategy, the inclusion and exclusion criteria, the data extraction forms and any subgroup analyses and the resultant recommendations. Inevitably, further contextual features emerge throughout the review; a review team should document these, together with implications for the final report. Ongoing interaction with stakeholders is one way to ensure that contextual factors do not surface unexpectedly during the course of a review.

Box 2 illustrates how a joined-up approach to handling context might be built within an existing guideline process. The chosen example of a complex intervention, antenatal care, is delivered in diverse ways across different countries and settings. Therefore, results from individual trials may not generalise to other types of population and setting.[89] Box 2 illustrates multiple points at which context might be introduced within a systematic review while illustrating the diversity of approaches to context employed by each tool or framework.

AVOIDING INTERPRETATION ERRORS IN CONNECTION WITH CONTEXT

Based on the above scenario (box 2), we have identified and labelled three ways in which readers may misinterpret context within systematic review findings; variously termed the 'universal fallacy', the 'focal fallacy' and the 'evidence availability fallacy' (also see online supplementary appendix):

1. A reader may think that a review finding that is not qualified by contextual descriptors (eg, time, place, socioeconomic factors, religious beliefs and so on), is a universal finding (ie, applying to all contexts) (the universal fallacy).

Box 2   Factoring in Context—a hypothetical example

Planning the review
A guideline development group on antenatal care selects the WHO-INTEGRATE framework[10] to identify and discuss the importance of context throughout the development process. The group uses the PerSPE©TiF[11] question formulation (From the *Per*spective of a Pregnant Woman in the *Se*tting of Rural Communities how does the *Ph*enomenon of Facility Based Care within an *E*nvironment of Poor Transport Infrastructure and Distantly-Located Facilities (*C*ompare with Traditional Birth Attendants at Home) up to and including the *Ti*me of Delivery in terms of *F*indings related to the Woman's Perceptions and Experiences?) to define the scope of the review. Using the PerSPE©TiF structure, they craft a review protocol together with inclusion and exclusion criteria to define and anticipate salient aspects of context.

Conducting the review
The review team takes details from the protocol and selects an appropriate data extraction framework (the Context and Implementation of Complex Interventions (CICI) framework,[22] deciding that broader environmental concerns are more critical to this population-based guideline than the specific setting. They access websites to elicit extra contextual data and identify 'clusters' of associated studies using the CLUSTER technique.[58] Where critical details are missing, the review team contacts authors of included studies requesting additional information, setting a deadline for response. The review team juxtaposes results from a quantitative meta-analysis and a qualitative synthesis using a logic model.[95] Important details of context inform plans for statistical and qualitative analyses and the statistical team explore contextual heterogeneity across predefined subgroups.[19]

Making recommendations
The review team provides appendices detailing the context described in each included study. The guideline development group pays particular attention to factors relating to Indirectness and Relevance when formulating their GRADE[65] and GRADE-CERQual[66] recommendations. The group uses an adaptation framework (ADAPTE[94]) to provide contextually-sensitive recommendations for major regions of the world, allowing each country to tailor country-specific recommendations, where feasible. The guideline developers provide accompanying spreadsheets and vignettes to facilitate roll-out and the group frames recommendations or using the EPICOT-dts framework.[84]

---

2. A reader may think that a finding only applies to the context that is the focus of the review question (when it is actually more generalisable) (the focal fallacy).

3. A reader may think that a finding only applies to contexts studied within included studies (when this simply characterises where research has been undertaken) (the evidence availability fallacy).

To avoid these interpretative errors, review authors and guideline developers can better articulate context-sensitivity when framing review findings and subsequent recommendations:

► **By Exception**—A review team could highlight that exceptions (divergent study contexts) are different from all the other study contexts or from all other contexts encompassed by a review question, for

example, 'This finding seems to apply irrespective of context, except in X, Y, Z where primary care systems continue to be disrupted by conflict and migration'.

► **By Inclusion**—A review team could explicitly state contexts to which findings apply. This flags the distinction between 'We have evidence that this only applies to these contexts' and 'We only have evidence that this applies to these contexts'.

► **By Exclusion**—A review team could explicitly state contexts to which findings do not apply.

► **By Assumption**—A review team could define relevance according to the presence or absence of common mechanisms or contextual factors, for example, 'This finding is likely to apply in those contexts with a developed method for priority setting for nutrition'.

Supporting recommendations by assumption should be treated with caution, comparable to circumspect use of indirect comparison in quantitative research. Nevertheless, the value of explanatory theory is increasingly recognised and its contribution undoubtedly extends beyond the theories of change or behavioural and social theories that are the focus of this paper.[90]

LIMITATIONS OF THIS ANALYSIS

Our brief selection of high-profile or recently developed frameworks and tools for taking account of context in systematic reviews and guidelines reveals that such tools remain limited in their facility to accommodate a complex systems perspective.[2] This selection is necessarily arbitrary given that information retrieval strategies to identify review tools that target context systematically do not currently exist. Highlighted tools and frameworks are illustrative requiring that the example given in box 2 is hypothetical and, indeed aspirational.

CONCLUSION

Approaches to context within systematic reviews typically share the following limitations:

► Tools typically capture context at a single timepoint and are barely able to capture prevailing trends.

► Many tools derive from use in primary research and either have limited utility within synthesis or have not been fully evaluated.

► Particular tools focus on a single specific level of context, particularly true of implementation tools that target the direct setting within which the intervention is delivered.

We have also identified a potential tension given that tools that are best equipped for considering context may, by their very nature, not prove practical as a structure or framework for data extraction or analysis. In the absence of empirical evidence on features of context that impact most on intervention success, we observe that tools tend to become increasingly expansive. This, in turn, makes data extraction and analysis more time consuming and resource intensive.

For reviewers and guideline developers, this paper emphasises the practical need to select from this toolkit judiciously. Context should be addressed early in guideline development and then revisited throughout the review process. Frameworks offer a useful structure for ensuring that data on context are extracted for the review. Checklists prompt decision-makers to consider important factors and reporting standards ensure that, where reported, requisite details are present in the report or in supplementary material accompanying each report. Knowledge translation tools such as vignettes, country profiles, calculators and spreadsheets can help in translating between the study contexts and the target context. With such diverse approaches, the current way forward requires composite use of different tools at appropriate stages of the process. Our brief scan of illustrative tools reveals a need for a coherent approach that runs throughout the review process and that underpins the individual tools used at each stage. Approaches to context to date suggests that the components already exist; the outstanding methodological challenge is to agree a coherent approach within which existing tools might be integrated.

Recent years have seen collective recognition of the importance of context when selecting, evaluating and implementing complex interventions. This paper consolidates collective learning in identifying and using tools for handling context and meeting the ever more complex demands of guideline developers and other decision-makers. This overview signals how context might possibly be addressed within systematic reviews and guidelines. For the present, review teams are encouraged to extract contextual details when available, elicit them when feasible and reconstruct them when available data allows. Review teams and guideline producers should seek to navigate a path that is parsimonious in terms of resource and data extraction and yet sensitive to their specific contexts and the values, preferences and needs of their target populations.

Author affiliations
[1]School of Health and Related Research, University of Sheffield, Sheffield, UK
[2]School of Social Sciences, Cardiff University, Cardiff, UK
[3]Department of Health Sciences, The University of York, York, UK
[4]European Centre for Environment and Human Health, University of Exeter, Truro, UK
[5]Department of Maternal, Newborn, Child and Adolescent Health, World Health Organization, Genève, Switzerland
[6]Department of Reproductive Health and Research including UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), World Health Organization, Geneva, Switzerland
[7]School of Social Sciences, Bangor University, Wales, UK

REFERENCES

1. Petticrew M. When are complex interventions 'complex'? When are simple interventions 'simple'? *Eur J Public Health* 2011;21:397–8.
2. Petticrew M, Knai C, Thomas J, *et al*. Implications of a complexity perspective for systematic reviews and guideline development in health decision making. *BMJ Global Health* 2019;0:e000899. doi:10.1136/bmjgh-2018-000899.
3. Kelly MP, Noyes J, Kane RL, *et al*. AHRQ series on complex intervention systematic reviews-paper 2: defining complexity, formulating scope, and questions. *J Clin Epidemiol* 2017;90:11–18.
4. Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. *Am J Community Psychol* 2009;43–267–76.
5. Webster J, Chandramohan D, Hanson K. Methods for evaluating delivery systems for scaling-up malaria control intervention. *BMC Health Serv Res* 2010;10(Suppl 1):S8.
6. Mills A, Gilson L, Hanson K, *et al*. What do we mean by rigorous health-systems research? *Lancet* 2008;372:1527–9.
7. Pawson R, Tilley N. *Realistic evaluation*. London: Sage, 1997.
8. Porter S, O'Halloran P. The use and limitation of realistic evaluation as a tool for evidence-based practice: a critical realist perspective. *Nurs Inq* 2012;19:18–28.
9. Moore GF, Evans RE. What theory, for whom and in which context? Reflections on the application of theory in the development and evaluation of complex population health interventions. *SSM Popul Health* 2017;3:132–5.
10. Rehfuess EA, Stratil JM, Scheel IB, *et al*. The WHO-INTEGRATE evidence to decision framework version 1.0: integrating who norms and values and a complexity perspective. *BMJ Global Health* 2019;0:e000844. doi:10.1136/bmjgh-2018-000844.
11. Booth A, Noyes J, Flemming K, *et al*. Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Global Health* 2019;0:e001107. doi:10.1136/bmjgh-2018-001107.
12. Petticrew M, Rehfuess E, Noyes J, *et al*. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. *J Clin Epidemiol* 2013;66:1230–43.
13. Flemming K, Booth A, Garside R, *et al*. Qualitative evidence synthesis for complex interventions and guideline development: clarification of the purpose, designs and relevant methods. *BMJ Global Health* 2019;0:e000882. doi:10.1136/bmjgh-2018-000882.
14. Noyes J, Booth A, Moore G, *et al*. Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: Clarifying the purpose, designs and outlining some methods. *BMJ Global Health* 2019;0:e000893. doi:10.1136/bmjgh-2018-000893.
15. McKillop A, Shaw J, Sheridan N, *et al*. Understanding the attributes of implementation frameworks to guide the implementation of a model of community-based integrated health care for older adults with complex chronic conditions: a metanarrative review. *Int J Integr Care* 2017;17:10.

16. Moore GF, Audrey S, Barker M, *et al*. Process evaluation of complex interventions: medical research council guidance. *BMJ* 2015;350:h1258.

17. Craig P, Di Ruggiero E, Frohlich KL. *Taking account of context in population health intervention research: guidance for producers, users and funders of research*. Southampton: NIHR Evaluation, Trials and Studies Coordinating Centre, 2018.

18. Poland B, Krupa G, McCall D. Settings for health promotion: an analytic framework to guide intervention design and implementation. *Health Promot Pract* 2009;10:505–16.

19. Higgins JPT, López-López JA, Becker BJ, *et al*. Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Global Health* 2019;0:e000858. doi:10.1136/bmjgh-2018-000858.

20. Damschroder LJ, Aron DC, Keith RE, *et al*. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009;4:50.

21. Pfadenhauer LM, Mozygemba K, Gerhardus A, *et al*. Context and implementation: a concept analysis towards conceptual maturity. *Z Evid Fortbild Qual Gesundhwes* 2015;109:103–14.

22. 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.

23. Howarth E, Devers K, Moore G. Contextual issues and qualitative research. *NIHR Journals Library* 2016.

24. Mathes T, Antoine SL, Prengel P, *et al*. Health technology assessment of public health interventions: a synthesis of methodological guidance. *Int J Technol Assess Health Care* 2017;33:135–46.

25. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.

26. Guise JM, Butler ME, Chang C, *et al*. AHRQ series on complex intervention systematic reviews-paper 6: PRISMA-CI extension statement and checklist. *J Clin Epidemiol* 2017;90:43–50.

27. Guise JM, Butler M, Chang C, *et al*. AHRQ series on complex intervention systematic review-paper 7: PRISMA-CI elaboration and explanation. *J Clin Epidemiol* 2017;90:51–8.

28. Greenhalgh T. Outside the box: Why are Cochrane reviews so boring? *Br J Gen Pract* 2012;62:371.

29. Gruen RL, Morris PS, McDonald EL, *et al*. Making systematic reviews more useful for policy-makers. *Bull World Health Organ* 2005;83:480.

30. Butler M, Epstein RA, Totten A, *et al*. AHRQ series on complex intervention systematic reviews-paper 3: adapting frameworks to develop protocols. *J Clin Epidemiol* 2017;90:19–27.

31. Eisenberg JM. Globalize the evidence, localize the decision: evidence-based medicine and international diversity. *Health Aff* 2002;21:166–8.

32. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* 2005;365:82–93.

33. Lewin S, Hendry M, Chandler J, *et al*. Assessing the complexity of interventions within systematic reviews: development, content and use of a new tool (iCAT_SR). *BMC Med Res Methodol* 2017;17:76.

34. Brennan C, Greenhalgh J, Pawson R. Guidance on guidelines: Understanding the evidence on the uptake of health care guidelines. *J Eval Clin Pract* 2018;24:105–16.

35. Rycroft-Malone J, McCormack B, Hutchinson AM, *et al*. Realist synthesis: illustrating the method for implementation research. *Implement Sci* 2012;7:33.

36. Booth A, Noyes J, Flemming K, *et al*. *Guidance on choosing qualitative evidence synthesis methods for use in health technology assessments of complex interventions*. In: Integrate-HTA, 2016.

37. Carey G, Malbon E, Carey N, *et al*. Systems science and systems thinking for public health: a systematic review of the field. *BMJ Open* 2015;5:e009002.

38. Alonso-Coello P, Schünemann HJ, Moberg J, *et al*. Grade Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016;353:i2016.

39. Rehfuess EA, Akl EA. Current experience with applying the GRADE approach to public health interventions: an empirical study. *BMC Public Health* 2013;13:9.

40. Rogers EM. *The diffusion of innovation*. 5th edn. New York, 2003.

41. Petticrew M, Roberts H. *Systematic reviews in the social sciences: a practical guide*. Blackwell Publishing: Oxford, 2006.

42. Denyer D, Tranfield D, van Aken JE. Developing design propositions through research synthesis. *Organization Studies* 2008;29:393–413.

43. Booth A. Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech* 2006;24:355–68.

44. Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.

45. Sideri S, Papageorgiou SN, Eliades T. Registration in the international Prospective Register of Systematic Reviews (PROSPERO) of systematic review protocols was associated with increased review quality. *J Clin Epidemiol* 2018;100:103–10.

46. Grant A, Treweek S, Dreischulte T, *et al*. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials* 2013;14:15.

47. Helfrich CD, Damschroder LJ, Hagedorn HJ, *et al*. A critical synthesis of literature on the promoting action on research implementation in health services (PARIHS) framework. *Implement Sci* 2010;5:82.

48. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: a systematic review of use over time. *Am J Public Health* 2013;103:e38–e46.

49. Murray E, Treweek S, Pope C, *et al*. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. *BMC Med* 2010;8:63.

50. Mosavianpour M, Sarmast HH, Kissoon N, *et al*. Theoretical domains framework to assess barriers to change for planning health care quality interventions: a systematic literature review. *J Multidiscip Healthc* 2016;9:303–10.

51. O'Reilly P, Lee SH, O'Sullivan M, *et al*. Assessing the facilitators and barriers of interdisciplinary team working in primary care using normalisation process theory: an integrative review. *PLoS One* 2017;12:e0177026.

52. Craig LE, McInnes E, Taylor N, *et al*. Identifying the barriers and enablers for a triage, treatment, and transfer clinical intervention to manage acute stroke patients in the emergency department: a systematic review using the theoretical domains framework (TDF). *Implement Sci* 2016;11:157.

53. O'Neill J, Tabish H, Welch V, *et al*. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol* 2014;67:56–64.

54. Attwood S, van Sluijs E, Sutton S. Exploring equity in primary-care-based physical activity interventions using PROGRESS-Plus: a systematic review and evidence synthesis. *Int J Behav Nutr Phys Act* 2016;13:60.

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

56. Cotterill S, Knowles S, Martindale AM, *et al*. Getting messier with TIDieR: embracing context and complexity in intervention reporting. *BMC Med Res Methodol* 2018;18:12.

57. Noyes J, Hendry M, Lewin S, *et al*. Qualitative "trial-sibling" studies and "unrelated" qualitative studies contributed to complex intervention reviews. *J Clin Epidemiol* 2016;74:133–43.

58. Booth A, Harris J, Croot E, *et al*. Towards a methodology for cluster searching to provide conceptual and contextual "richness" for systematic reviews of complex interventions: case study (CLUSTER). *BMC Med Res Methodol* 2013;13:118.

59. Lavis JN, Oxman AD, Lewin S, *et al*. SUPPORT Tools for evidence-informed health Policymaking (STP). *Health Res Policy Syst* 2009;7(Suppl 1):I1.

60. Altman DG, Simera I. A history of the evolution of guidelines for reporting medical research: the long road to the EQUATOR Network. *J R Soc Med* 2016;109:67–77.

61. Shoveller J, Viehbeck S, Di Ruggiero E, *et al*. A critical examination of representations of context within research on population health interventions. *Critical Public Health* 2016;26:487–500.

62. Schulz KF, Altman DG, Moher D, *et al*. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332.

63. Tong A, Flemming K, McInnes E, *et al*. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Med Res Methodol* 2012;12:181.

64. Kågesten AE, Tunçalp Ö, Portela A, *et al*. Programme Reporting Standards (PRS) for improving the reporting of sexual, reproductive, maternal, newborn, child and adolescent health programmes. *BMC Med Res Methodol* 2017;17:117.

65. Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.

66. Lewin S, Glenton C, Munthe-Kaas H, *et al*. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative

evidence syntheses (GRADE-CERQual). *PLoS Med* 2015;12:e1001895.

67. Andrews JC, Schünemann HJ, Oxman AD, *et al*. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol* 2013;66:726–35.

68. Guyatt GH, Oxman AD, Kunz R, *et al*. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol* 2011;64:1303–10.

69. Noyes J, Booth A, Lewin S, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 6: how to assess relevance of the data. *Implement Sci* 2018;13(Suppl 1):4.

70. Burchett H, Umoquit M, Dobrow M. How do we know when research from one setting can be useful in another? A review of external validity, applicability and transferability frameworks. *J Health Serv Res Policy* 2011;16:238–44.

71. Lavis JN, Oxman AD, Souza NM, *et al*. SUPPORT Tools for evidence-informed health Policymaking (STP) 9: assessing the applicability of the findings of a systematic review. *Health Res Policy Syst* 2009;7(Suppl 1):S9.

72. Lewin S, Oxman AD, Lavis JN, *et al*. SUPPORT tools for evidence-informed policymaking in health 11: Finding and using evidence about local conditions. *Health Res Policy Syst* 2009;7(Suppl 1):S11.

73. Hunter BM, Requejo JH, Pope I, *et al*. National health policy-makers' views on the clarity and utility of Countdown to 2015 country profiles and reports: findings from two exploratory qualitative studies. *Health Res Policy Syst* 2014;12:40.

74. Gauld R, Blank R, Burgers J. The world health report 2008 - primary healthcare: how wide is the gap between its agenda and implementation in 12 high-income health systems? *Healthc Policy* 2012;7:38–58.

75. Pokhrel S, Owen L, Coyle K. Suppor*ting decision making*. London: Palgrave, 2017.

76. National Institute for Health and Care Excellence (NICE). *Assessing resource impact process manual: technology appraisals and highly specialised technologies*. London: National institute for Health and Care Excellence (NICE), 2017.

77. Montgomery P, Underhill K, Gardner F, *et al*. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *J Clin Epidemiol* 2013;66:874–82.

78. Alvarez E, Lavis J. *Implementing health systems guidance: a workbook to support the contextualization of recommendations at the national or subnational level*. Hamilton, Canada: McMaster Health Forum, 2016.

79. Alvarez E, Lavis JN, Brouwers M, *et al*. Developing a workbook to support the contextualisation of global health systems guidance: a case study identifying steps and critical factors for success in this process at WHO. *Health Res Policy Syst* 2018;16:19.

80. Adapte Collaboration. *The ADAPTE process: resource toolkit for guideline adaptation Version 2.0*, 2009.

81. Tugwell P, Knottnerus JA. Adolopment - a new term added to the Clinical Epidemiology Lexicon. *J Clin Epidemiol* 2017;81:2.

82. Schünemann HJ, Wiercioch W, Brozek J, *et al*. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. *J Clin Epidemiol* 2017;81:101–10.

83. Snilstveit B, Vojtkova M, Bhavsar A, *et al*. Evidence & Gap Maps: a tool for promoting evidence informed policy and strategic research agendas. *J Clin Epidemiol* 2016;79:120–9.

84. Brown P, Brunnhuber K, Chalkidou K, *et al*. How to formulate research recommendations. *BMJ* 2006;333:804–6.

85. World Health Organization. Guidelines approved by the guidelines review committee. In: *Who indoor air quality guidelines: household fuel combus*tion. Geneva: World Health Organization, 2014.

86. Burford B, Lewin S, Welch V, *et al*. Assessing the applicability of findings in systematic reviews of complex interventions can enhance the utility of reviews for decision making. *J Clin Epidemiol* 2013;66:1251–61.

87. Sandelowski M, Leeman J, Knafl K, *et al*. Text-in-context: a method for extracting findings in mixed-methods mixed research synthesis studies. *J Adv Nurs* 2013;69:1428–37.

88. Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992;116:78–84.

89. Dowswell T, Carroli G, Duley L, *et al*. Alternative versus standard packages of antenatal care for low-risk pregnancy. *Cochrane Database Syst Rev* 2015;(7):CD000934.

90. Noyes J, Hendry M, Booth A, *et al*. Current use was established and Cochrane guidance on selection of social theories for systematic reviews of complex interventions was developed. *J Clin Epidemiol* 2016;75:78–92.

91. Kirk MA, Kelley C, Yankey N, *et al*. A systematic review of the use of the consolidated framework for implementation research. *Implement Sci* 2016;11:72.

92. Moore GF, Evans RE, Hawkins J, *et al*. All interventions are complex, but some are more complex than others: using iCAT_SR to assess complexity. *Cochrane Database Syst Rev* 2017;7:Ed000122.

93. Hultcrantz M, Rind D, Akl EA, *et al*. The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol* 2017;87:4–13.

94. Darzi A, Abou-Jaoude EA, Agarwal A, *et al*. A methodological survey identified eight proposed frameworks for the adaptation of health related guidelines. *J Clin Epidemiol* 2017;86:3–10.

95. Rehfuess EA, Booth A, Brereton L, *et al*. Towards a taxonomy of logic models in systematic reviews and health technology assessments: a priori, staged, and iterative approaches. *Res Synth Methods* 2018;9:13–24.