Bringing together realist and economic approaches in the evaluation of health and social care interventions: a scoping review of theoretical, methodological and practical implications

Andrew Fletcher [a,*] [ID], Sonia Dalkin [a], Rob Anderson [b], Rachel M. Baker [c], Cam Donaldson [c], Vivienne Hibberd [i], Meghan Bruce Kumar [a,d], Felicity Shenton [h], Gill Westhorp [e], Geoff Wong [f], Judy Wright [g], Angela Bate [a]

[a] Northumbria University, UK
[b] University of Exeter, UK
[c] Glasgow Caledonian University, UK
[d] KEMRI-Wellcome Trust, Nairobi, Kenya
[e] Charles Darwin University, Australia
[f] University of Oxford, UK
[g] University of Leeds, UK
[h] NIHR ARC NENC (PPIE Lead), UK
[i] PPIE, UK

ABSTRACT

Background: In the evaluation of complex interventions, economic evaluations aim to determine the relative cost-effectiveness of interventions but generate little explanation of how or why contexts and underlying causal mechanisms impact this. Conversely, realist approaches aim to explain 'what works, for whom, in which circumstances and why' but rarely capture the economic costs and consequences of interventions. As a result, many evaluations remain partial.

Objective: To identify past attempts to integrate realist and economic evaluation approaches and summarise the recent developments in realist and economic evaluation approaches in the evaluation of complex health and social care interventions.

Methods: We conducted a series of scoping reviews using online academic databases, personal libraries and expert stakeholder workshops, to identify the theoretical, methodological, and practical challenges and developments in bringing together realist and economic evaluation approaches.

Findings and recommendations: Although increasing, there remain relatively few examples of evaluations that have attempted to integrate realist and economic evaluation approaches, and challenges for their integration mean that further guidance is required. The wider literature indicated challenges in the theoretical (e.g. ontology, causality), methodological (e.g. accounting for context, study design, mixing methods) and practical (e.g. terminology, scale and scope) domains, for which we have developed recommendations.

Conclusion: To deliver services that are both effective and efficient, evaluations must synthesise relevant explanatory evidence with cost and outcome data to enable policymakers and commissioners to make informed decisions. Findings and recommendations from this review were used to inform the development of guidance for the integration of realist and economic evaluation approaches.

Table of acronyms

ARC	Applied Research Collaboration
ASSIA	Applied Social Sciences Index and Abstracts
AU	Australia
CCA	Cost Consequence Analysis
CEA	Cost Effectiveness Analysis
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CIMO	Context Intervention Mechanism Outcome

(continued on next column)

(continued)

CMMO	Context Mechanism (resources) Mechanism (reasoning) Outcome
CMO(C)	Context Mechanism Outcome (Configuration)
CSM	Complex Systems Modelling
EE	Economic Evaluation
HRQoL	Health Related Quality of Life
HTA	Health Technology Assessment
ICAMO	Intervention Context Actor Mechanism Outcome
IIAG	International Interdisciplinary Advisory Group

(continued on next page)

(*continued*)

IPT	Initial Programme Theory
KEMRI	Kenya Medical Research Institute
MEDLINE	Medical Literature Analysis and Retrieval System Online
MRC	Medical Research Council
NENC	North East and North Cumbria
NHS	National Health Service
NIHR	National Institute for Health and Care Research
PHEM-B	Public Health Economic Models (PHEM) - Behaviour
PPIE	Patient and Public Involvement and Engagement
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PT	Programme Theory
QALY	Quality Adjusted Life Year
QCA	Qualitative Comparative Analysis
RAMESES	Realist And Meta-narrative Evidence Syntheses: Evolving Standards
RCT	Randomised Controlled Trial
RE	Realist Evaluation
REE	Realist Economic Evaluation
REEM	Realist Economic Evaluation Methods
S/ICMO	Strategy/Intervention Context Mechanism Outcome
SEM	Structural Equation Modelling
(TD)ABC	(Time-Driven) Activity-Based Costing
ToC	Theory of Change
UK	United Kingdom

1. Introduction

Achieving and sustaining improvements in health and social care requires well-designed interventions, programmes, or policies that are often inherently complex or implemented within complex systems; and robust evaluation that is also often complex (Shiell et al., 2008; Pettigrew, 2013; Rutter et al., 2017; Skivington et al., 2021). Interventions are considered complex for several reasons including but not limited to: the number of components of the intervention (including people, skills, processes, organisations and systems), the scale of the intervention (numbers of groups or settings), and the interaction and dynamic nature of these (Skivington et al., 2021). Complexity also arises from external factors and the fact that interventions exist and interact with in a complex world. This is reflected in the recent update to the UK Medical Research Council (MRC) guidance for developing and evaluating complex interventions, which outlines six core elements that appear to align closely with approaches adopted in realist and economic evaluation (Skivington et al., 2021).

Realist evaluation, although commonly associated with evaluating complex interventions, is an approach that can be applied to range of interventions, depending on the research question and the level of complexity within which a change or intervention is implemented. While realist evaluation (RE) seeks to explain how, for whom and in which circumstances an intervention works, it omits the key question, *at what cost relative to benefits?* Ignoring this question implicitly denies the existence of resource scarcity without offering a decision-informing alternative. Conversely, while economic evaluation (EE) can determine the relative cost-effectiveness, on average, of an intervention, it provides little or no explanatory evidence about how or why interventions might be cost-effective, or whether they are relatively more or less cost-effective in some contexts and not others. As such, there is a powerful argument for integrating economic methods and realist approaches. However, where REs and EEs have been undertaken in the same study, they often run as separate processes, either in parallel (without explicit integration) or sequentially, limiting opportunities to share valuable learning between the evaluations (for example, Davies et al. (2024), who note, *"Generally, the estimates of resource use and of benefits are separate domains of the evaluation"* (p. 84)).

Our research seeks to integrate RE and EE approaches to enable evaluators to establish what works, for whom, in which circumstances, while integrating better understanding of costs and consequences (including opportunity cost) (Dalkin and Bate, 2022). The specific

potential of combining RE with EE was first suggested by Anderson in his 2004 review of Pawson and Tilley's (1997) book, *Realistic Evaluation* (Anderson, 2004). The idea was later more fully elaborated using an evidence synthesis of the cost-effectiveness of shared care for chronic disease (Anderson and Hardwick, 2016). This paper represents the first attempt to bring together realist evaluators and economic evaluators to answer the following questions.

- What can we learn from past attempts to integrate (implicitly or explicitly) RE and EE?
- What are the recent developments in RE and EE approaches in the evaluation of complex health and social care interventions?

2. Methods

Initially four searches were designed to identify.

- past attempts to integrate realist approaches (explicitly or implicitly) in EE
- past attempts to integrate economic concepts (explicitly or implicitly) in RE
- recent developments in EE in the evaluation of complex health and social care interventions
- recent developments in RE in the evaluation of complex health and social care interventions

These were supplemented by four expert stakeholder workshops, which were used to discuss the scoping review results and identify points of conflicts and connections between the two approaches.

2.1. Scoping reviews

We followed guidance on reviews of methods papers (Gentles et al., 2016; Tricco et al., 2018; McBride et al., 2022) to gather documents efficiently. Table 1 presents a summary of the search strategies, databases, and inclusion and exclusion criteria. The final protocol was developed through discussion with the REEM research team and registered with the Open Science Framework on July 14, 2022 (https://osf.io/x7eg6).

Searches were conducted by an information specialist (JW), initially between November 2022 and August 2023, and repeated and updated in March 2025. Searches were limited to peer reviewed publications from 2008 onwards to capture advances since the previous MRC complex intervention guidance update (Craig et al., 2008). The bibliographic databases searched are listed in Table 1 and the final search strategy for MEDLINE can be found in Appendix 1. Google Scholar and ASSIA gave more new hits on the second search, because some indexing had changed within ASSIA, and Google Scholar updates its algorithms frequently.

The search strategy was developed to avoid retrieving large numbers of studies that applied realist and/or EE approaches with no methodological development. Subject headings and text words were generated by the authors from analysis of 69 core 'seed papers' identified from i) their personal libraries and ii) a call for papers disseminated through social media. Search strategies were tested for retrieval of the seed papers and were peer-reviewed by a second information specialist using the PRESS checklist (McGowan et al., 2016).

The combined searches (1–4) and selection of studies from the initial search period, and the additional studies from the updated search period are presented in the following PRISMA diagrams (Figs. 1 and 2 respectively). References and abstracts were managed in Endnote X9 (https://endnote.com/) and uploaded to Rayyan (https://www.rayyan.ai/). A mixed-discipline team of realist and economic evaluators jointly screened 100 abstracts for each search, initially using simple criteria (such as whether the paper addressed one or more of the scoping questions) to refine the inclusion/exclusion criteria (Table 1). The

Table 1
Scoping reviews inclusion/exclusion criteria.

Review	Aim	Databases searched	INCLUSION Peer reviewed academic articles, in English, 2008-present	EXCLUSION
1	Identify EEs that use realist or similar theory-driven approaches.	ASSIA (ProQuest), EconLit (EBSCOhost), Ovid MEDLINE ® ALL and Google Scholar (via Harzing's Publish or Perish).	**Full economic evaluation** AND **Explicitly reference any of:** Realism, role of context in programme or intervention outcomes, mechanisms, programme or intervention theory, theory testing	**Reference to logic models only, unless realist logic models or innovation in logic modelling** AND/OR **Protocols**
2	Identify REs that incorporate resource requirements and impacts.	Assia (ProQuest), Core Collection, Ovid MEDLINE ® ALL, Health Management Information Consortium (Ovid) and Google Scholar (via Harzing's Publish or Perish).	**Realist evaluation** AND **Explicitly references any of:** Cost-benefit analysis, resource allocation, economic concepts	**Realist reviews or syntheses** AND/OR **Protocols** AND/OR **Only acknowledge economic concepts in passing** (e.g. 'economic evaluation is needed') **or a realist evaluation of economic intervention/incentive** (e.g. a fiscal policy)
3	Identify recent developments in methods/guidance recommended for EE of complex interventions in health and social care.	ASSIA (ProQuest), EconLit (EBSCOhost), Ovid MEDLINE ® ALL and Google Scholar (via Harzing's Publish or Perish).	**Propose a framework/guideline/approach, or recommend novel methods or combinations of methods** OR **Explicitly include methods or methodological issues for a) economic evaluation or b) estimating value/costs/or resource use** AND **Evaluate complex interventions** or include equivalent terms that intrinsically imply complexity, e.g. 'public health interventions', 'multi-component interventions', 'service development', 'health policy programmes', etc.	**No detailed discussion/study of methods** OR **Macro/high-level evaluations** not focussed on a specific intervention. OR **Not sufficiently complex.** For example, simple or complicated interventions that are simply vulnerable to broader variables.
4	Identify recent developments in methods/guidance recommended for REs of complex health and social care interventions.	Assia (ProQuest), Core Collection, Ovid MEDLINE ® ALL, WorldCat and Google Scholar (via Harzing's Publish or Perish).	**Use Pawson and Tilley's (1997) Realistic Evaluation** or meaningfully cite its descendants. AND **Combine data collection methods in a novel way** or combine RE with other methodological approaches. OR **Propose guidelines or recommendations** for novel approaches in RE at any level, e.g. • Pragmatic level of 'doing' realist research (such as realist interviewing) • Methodological level (such as combining approaches)	**Realist reviews or protocols for realist reviews** OR **Protocols for realist evaluations** unless they specifically propose methods development or guidance OR **No detailed discussion/study of methods or methodological nuances** OR **Different meanings of 'realist'**, e.g. realism in political science or critical realist studies

remaining abstracts were screened by mixed-discipline teams of two researchers (a health economic evaluator and a realist evaluator). Additional references were derived after the initial searches were completed through citation searching, personal recommendations, email alerts and social media, and are included in Fig. 1.

Full-text screening, review and data extraction were undertaken by the same mixed discipline teams, as above. Extraction of information focussed on: where the methodological contribution lay (e.g. programme theory development, mixing methods); whether the article was primarily philosophical, methodological or practical; and whether it proposed recommendations or guidance. Articles that reported conducting a separate RE and EE, either sequentially or in parallel, with no methodological development, were excluded at this stage, as this review is concerned primarily with the integration of the two approaches and/or developments in the approaches. One exemplar full extraction table is provided in Appendix 2.

2.2. Expert stakeholder workshops

Expert workshop participants were purposively selected based on their scholarly contribution in realist or economic evaluation. They were identified by members of the core research team and from early literature scoping. Equal numbers of realist and economic evaluators were selected, with careful attention paid the ratio of male to female scholars, and their relative seniority, to cultivate a good balance of established

perspectives and newer voices.

A mix of online and in person workshops (n = 4) took place between April 2023 and May 2024. The workshops were organised in advance to discuss specific topics emerging from the scoping reviews. These included terminology, meanings, and definitions in RE and EE; the significance and implications of ontology and theory in RE and EE; and identifying existing and novel methods of data collection and analysis. Ground rules of respectful communication and the need to explain technical terms where necessary were established. All workshops were audio recorded (or recorded via Microsoft Teams for online meetings). Artefacts, including Miro boards, flipchart sheets and Post-it notes were gathered at the end of each workshop. Content analysis (Bryman, 2016) was used to analyse these data and produce summaries to inform key emerging issues in this scoping review. The aim was not to reach consensus, but to surface all views, perspectives and sources in the literature. Where there was disagreement, this was noted and reported in the findings.

In addition, a Patient and Public Involvement and Engagement (PPIE) group (n = 6) was involved from the outset. Members were selected based on having previous involvement in REs and/or EEs. The group had regular PPIE-specific meetings and were represented by the REEM PPIE co-applicant (VH) at the monthly REEM research project team meetings. These PPIE meetings explored challenges and approaches to involving PPIE in a primarily methodological project; ensured draft outputs from the REEM research project were accessible to
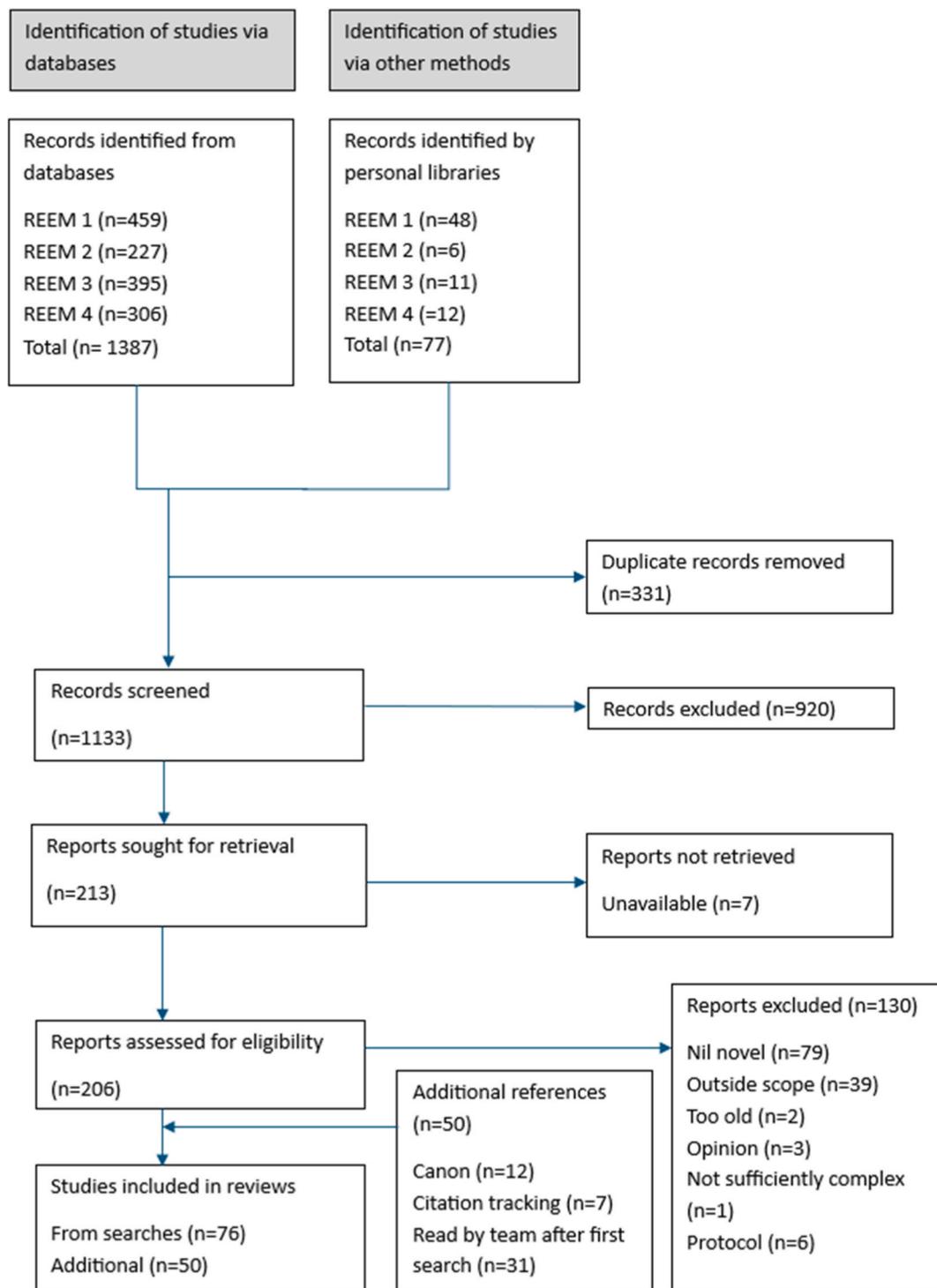
Fig. 1. PRISMA diagrams REEM 1–4 2022 combined.

lay audiences; developed a glossary of shared or disputed project-relevant terms; and commented on the process and findings of the project.

3. Findings

The PRISMA diagrams for the 2023 searches (n = 4) and the 2025 searches (n = 4) are shown in Figs. 1 and 2 respectively. The searches identified 1461 abstracts after removal of duplicates. Of these, 1199 were excluded, resulting in 284 full texts to retrieve. Of these, 182 were excluded for reasons including 'nil novel', 'unavailable', 'not sufficiently

complexity oriented', 'opinion' or 'synthesis claims do not materialise'. Another 50 texts were retrieved through citation tracking; were relevant to specific points in this article (e.g. canonical texts); or were discovered by the team between the 2023 search and the 2025 search. The total number of references in this article is 152.

Table 2 presents the outputs from the expert stakeholder workshops.

Findings from the scoping review searches are reported together with the outputs from the discussions at the expert stakeholder workshops, as one narrative, to synthesise learning applicable to the integration of RE and EE. These are presented in two sections. The first reports previous attempts to bring together RE and EE approaches,
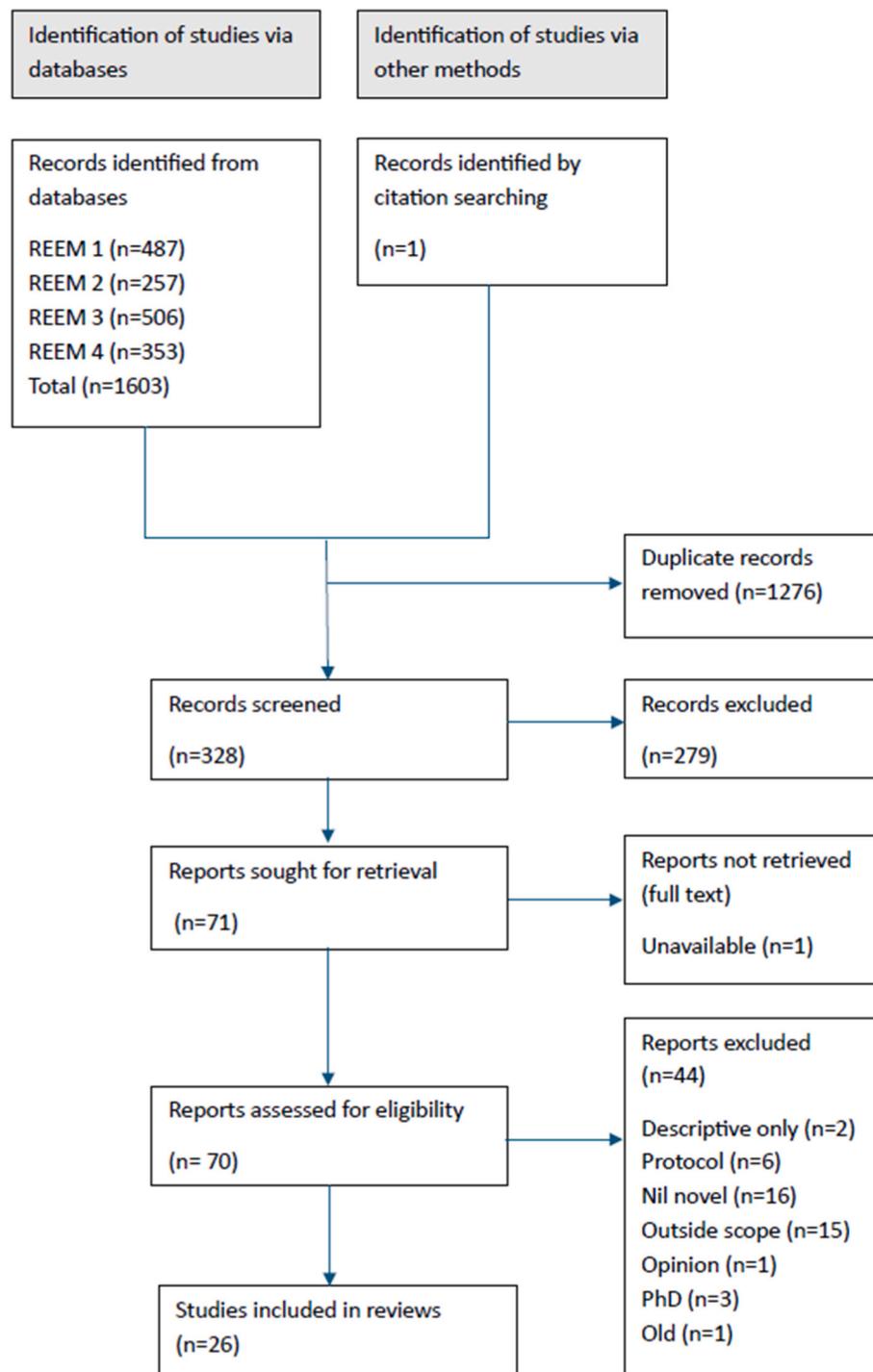
Fig. 2. PRISMA diagrams REEM 1–4 2025 combined.

either explicitly or implicitly. The second reports the theoretical, methodological, and practical challenges, developments and implications for integrating RE and EE approaches. Implications of these findings are presented in the discussion.

3.1. Section 1: previous attempts to bring together realist and economic evaluation

Longstanding calls from some health economists have a) identified the conceptual problems inherent in evaluating the cost-effectiveness of complex interventions (Godber et al., 1997; Coast et al., 2000), b) suggested the need for more theory-informed approaches and explanatory EE (Coast et al., 2000; Birch and Gafni, 2003) and c) led to an increasing focus on theory-driven and realist approaches (Deaton, 2010). However, it was Anderson (2004) who first suggested the potential of combining RE with EE. This was later elaborated by Hardwick et al. (2013) and Anderson and Hardwick (2016), who developed programme theories (PTs) about the effectiveness and cost-effectiveness of shared care for chronic conditions. Their latter paper described the key similarities and differences between RE and EE and, while they observed that a) cost-effectiveness-specific PTs are possible for some programmes and b) these can sometimes be expressed in terms of contexts, mechanisms and outcomes, they acknowledged more work is needed by realists to understand resource implications at different points in a programme

Table 2
Outputs from expert stakeholder workshops.

Workshop	Date	Format	No. of attendees/discipline		Data capture method	Themes emerging from meeting discussions
1	25-Apr-23	Online	Realists	8	Meeting recording/notes/Teams chat file/ Miro (online collaboration	Ontological/epistemological assumptions; essential tenets or requirements for integrating RE and EE
			Health economists	8		
			PPIE, information specialist	3		
2	06-Jun-23	Hybrid	Realists	8	Meeting recording (+4 breakout rooms)/ Initial survey for discussion	Definitions and terms used in RE and EE; comparators, causality, PTs, ontology and epistemology
			Health economists	6		
			PPIE	1		
3	4/5-Oct-23	F2F	Realists	8	Meeting notes and breakout group audio recordings	Technical definitions, counterfactuals, theory development, evidence hierarchies
			Health economists	10		
4	08-May-24	Online	Realists	6	Meeting recording and transcript	Comments and feedback on draft principles for integrating RE and EE; glossary terms
			Health economists	8		
			Information specialist	1		

theory (PT).

These ideas were revisited by Anderson et al. (2018, p. 108), who suggested two extensions to existing guidance on the design and conduct of REs:

"1) articulating the costs of the resources offered by the intervention and the cost implications of changes to stakeholder reasoning and behaviour in response to the intervention; and 2) surfacing the ways in which new interventions seek to be more cost-effective than existing care, thus identifying 'cost-effectiveness programme theories'". He notes, *"all three of the main components of realist explanations – contexts, mechanisms and outcomes – can either require additional resources in order to exist or impact on resource use."*

Further recent examples include a framework for synthesising realist and economic evidence in health and criminal justice (Johnson et al., 2015; Croci, 2024), and the development of economically optimised PTs in education (Rees et al., 2022a,b). The few studies that have explicitly combined RE and EE approaches could be categorised into: a) those that used RE to contextualise EE findings (i.e. *ex-post* theorising) – for example Witter et al. (2016), who drew on RE to understand how the cost and effects of national policies for improving financial access to emergency obstetric care in four African countries might be linked to differences in implementation; and b) those that used RE to theorise the parameters for a costing analysis or EE (i.e. *ex-ante* theorising).

Most studies belong in this latter category. Goodman et al. (2017) and, in a related study, Gordon et al. (2018) conducted a RE to explore the use of different NHS service delivery models in care homes to identify the key features/mechanisms associated with promoting continuity of care (as an outcome). This enabled them to develop a more nuanced micro-costing strategy to identify the costs (both to the NHS and from a societal perspective) associated with these key features/mechanisms. Latterly, Butler et al. (2022) used RE to assess different models of hospice at home services, to identify what works (and does not) and to determine the relative cost-effectiveness of different models. Similarly, in their exploration of the sustainability of outcomes in international development, Feeny et al. (2023) proposed a framework that used RE to identify the pathways through which an intervention works to identify resource use implications. Likewise, in education, Rees et al. (2022a) conducted a realist synthesis to evaluate different supervision training models and develop a micro-costing analysis to produce 'economically optimised programme models'. In a related study, Rees et al. (2022b) used RE to link outcomes with 'cost-sensitive mechanisms and contexts' and, despite the small dataset, claimed that their approach responds to common criticisms that *"cost-related research is both context-blind and atheoretical"*, although it generates challenges in terms of *"staying internally coherent with the principles of realist rather than economic evaluation"* (Rees et al., 2022a, p. 414). Moving beyond costing analyses, Beeston et al. (2020) adopted a Theory of Change (ToC) approach, using mixed methods and data to evaluate minimum unit pricing for alcohol in Scotland to underpin their cost consequence analysis (CCA). However, while ToC shares some common ground with RE, the causal explanations used in ToC are not usually aligned with notions of generative causation as in RE. On the other hand, Walsh et al. (2024) used a mixed methods RE in first contact physiotherapy for musculoskeletal disorders in primary care, to inform the scope of a cost-effectiveness analysis (CEA), specifically to determine the costs and range of benefits beyond the intervention itself.

There have been several attempts to integrate RE and EE. While we recognise the categories in the three sections above are subjective and potentially overlapping, they assist our understanding of how others have sought to integrate these approaches to date. However, due to the small number of studies that explicitly claim to integrate RE and EE, further searches were undertaken to explore recent developments in RE and EE, to identify theoretical, methodological and practical challenges and implications for integrating RE and EE approaches.

3.2. Section 2: the theoretical, methodological and practical challenges, developments and implications for integrating RE and EE approaches

The emergent themes from the scoping review and expert stakeholder workshops, detailing the conflicts and connections between RE and EE, are summarised in Table 3. These are discussed under the following headings, outlining the theoretical, methodological and practical challenges and potential developments for integrating RE and EE approaches.

3.3. Theoretical challenges and developments

3.3.1. Ontology and epistemology

While realist evaluators are explicit about their ontological and epistemological choices, adopting a realist philosophy of science, Coast and De Allegri (2018, p. 2) observed that economic evaluators *"do not tend to spend much time considering issues of ontology and epistemology. Questions of the nature of being and theories of knowledge are largely taken for granted by economists, with realist, positivist and deductive approaches seen as standard".* It was agreed at workshop #1 that EE, as applied to health, is pragmatic and decision-informing, and less concerned with defining or clarifying explicit ontological assumptions. However, Connections between pragmatism and ontology are not straightforward (Elder-Vass, 2022). Gilmore et al. (2019, p. 1) note, *"Striking a balance between theory and pragmatism, while adhering to realist ontological underpinnings of generative causation and retroduction, is no easy task".* Furthermore, while Törnberg (2018) agrees that ontological assumptions are often implicit in economics (as opposed to EE *per se*), he argues that the emergence of new pluralist approaches can be delineated into two broad categories, those that are implicitly ontologically aligned with complexity theory (e.g. those that use computational models, agent-based economics, etc.) and those that are aligned with critical realism (e.g. those that draw on social, structural and institutional perspectives). He therefore proposes that complex realism (a realist

Table 3

Themes emerging from scoping review results and expert stakeholder workshops.

Heading	Theme	Summary
Theoretical	**Ontology and epistemology** **Causality** **(Programme) Theory**	Theoretical challenges and developments in RE and EE highlight several considerations. Given the relative ontological flexibility of EE approaches, alongside the need for pragmatism and adherence to generative causation, any integration of RE and EE should be situated within a realist philosophy of science. This will involve the use of retroductive reasoning and both programme and substantive theory, potentially using heuristics, to demonstrate generative causality.
Methodological	**Accounting for context** **Identifying, measuring and valuing mechanism outcomes and costs** **Study design** **Integrating and triangulating data collection and analysis**	Context in RE and EE is fundamental to defining generalisability and transferability. Outcome measurement takes a more central role in EE, which often uses proxies – indicators of the 'real'. Attention should therefore be given to the definition, measurement and explanation of all components of the realist configuration and resources, commensurate with EE. This would also potentially allow for micro costing across the entire PT. Overlaps between RE and EE designs could enable better understanding of: the critical parameters of the evaluation, the decision question and what types of data are needed, including the necessity for a comparator to enable the incremental measurement of changes. The range of innovation and/or mixed method study designs in the review highlighted some valuable potential integrative approaches, although any such analysis would need to produce realist causal explanations.
Practical	**Terminology** **Scale and scope of evaluations**	There is a clear need for evaluators to understand where shared terms may cause confusion and ensure that any terms are transparently agreed and defined. In addition, developing and testing more comprehensive PTs for an integrated RE and EE is likely to require significant time, to allow for more responsive use of appropriate quantitative methods and enable more iterative analysis between multiple datasets. There will likely therefore be a trade-off between navigating the practical and funding difficulties of conducting such research and developing sufficiently nuanced and sophisticated studies.

philosophy of science) offers a unifying ontology.

3.3.2. Causality

In RE, outcomes are *"brought about by the action of some underlying 'mechanism'. Mechanisms are not variables or attributes and thus not always directly measurable. […] Causal explanation is thus a matter of producing theories of the mechanisms"* (Pawson, 2008, p. 1). While RE is underpinned by this generative understanding of causality, it can be argued that EE is rooted in a logic of successionist, linear causation, which *"abounds in the heartlands of survey and experimental work"* (idem., p. 3) and causal inference to estimate causal impacts. It is patterns in quantitative data that allow researchers to attribute effect to cause. As with ontology, assumptions around causation in EE are often implicit. Only a few EE papers in our scoping reviews explicitly stated their approach to causality, including examples where the authors referred to generative causation (Anderson and Hardwick, 2016; Walker et al., 2022). Most study designs and analyses implied a successionist causality. However, it was acknowledged in the stakeholder workshops that while the focus of EE is on the impact of interventions and thus causal inference, interrogation of how and why has not been explicit. Instead, economics has, over recent decades, developed increasingly sophisticated econometric techniques to estimate causal inference as well as focusing more on measuring *how much* an intervention works rather than *how* it works (i.e. measuring magnitude rather than mechanisms).

RE uses realist PTs to generate causal explanations in the form of Context Mechanism Outcome (CMO) configurations (Pawson and Tilley, 1997). The CMO has undergone several elaborations, including CMMO (CMO with explicit identification of resources and reasoning within the mechanism) (Dalkin et al., 2015), CIMO (CMO plus intervention) (Mazzocato et al., 2010), ICMAO (adding intervention and agency) (Higgins et al., 2012), S/ICMO (strategy/intervention) (Willis et al., 2014), and ICAMO (intervention and actor) (Mukumbang et al., 2018). In a review of three heuristics, De Weger et al. (2020) conclude, while different heuristics can add value, *"any configuration type must explain the causal link between context, mechanism and outcome, and any additional explanatory factors must adhere to that rule of generative causation"* (p. 7). However, discussions in workshop #3 highlighted that adding extra components, such as 'actor' or 'intervention', into CMOs may help to surface more detail/parameters relevant to EE, helping economic evaluators to identify potential comparators, opportunity costs, benefits or resource implications. Therefore, more work is required to understand the added benefit and limitation of adding components to the CMO and how terms can be conflated (see terminology).

3.3.3. (Programme) theory (PT)

RE is theory-driven. Realist PTs are causal explanations of what works (i.e. what changes in outcomes are produced), for whom, under what circumstances and how (Pawson and Tilley, 1997), for a particular type of intervention or policy. The language, logic and development of PTs are critical elements of rigour and there have been calls for greater transparency in PT development in RE (Welch and Tricco, 2016). Dalkin et al. (2021) suggested that many PT development discussions happen 'behind closed doors', so foregrounding and recording both the process and the reasoning of PT development can increase transparency. The RAMESES II reporting standards (Wong et al., 2016) were developed partially to make realists accountable when using the approach, including conveying this transparency.

As RE has evolved, so too has the range of approaches to recording explicitly how evaluators have journeyed from initial PTs to refined PTs, including the reasoning they used. Different types of reasoning differ in how they use information to arrive at conclusions. PTs in RE are developed primarily using retroductive reasoning, *"a mechanism-focused approach to inference making, which seeks to clarify the basic prerequisite or conditions of a phenomenon"* (Kabongo et al., 2020, p. 4). This may also incorporate other forms of reasoning, such as abduction, *"inventive thinking required to imagine the existence of mechanisms"* (ibid.); the

'zigzagging' model, which combines induction, abduction, deduction and retroduction (Cooper et al., 2020); or a 'realist dialogic approach' in which mixed-discipline research teams discuss IPTs with senior academics using a mix of abduction and reflection (Francis-Auton et al., 2022). Different forms of reasoning, particularly retroductive reasoning, are needed in RE to combine different types of data about the causal relationships between C, M and O. This involves an iterative approach to data collection and analysis, which can be seen to be at odds with more traditional EE, which often adopts a more linear process, drawing on inductive and deductive reasoning only, not retroductive reasoning.

Alongside the use of more explicit reasoning in RE, transparency in the process of testing and refining theory has also been elaborated in recent publications. Mukumbang et al. (2016) and Gilmore et al. (2019) outline five development phases for refining PTs, placing data analysis and PT development in relation to substantive theory. Dalkin et al. (2021) provided a worked example of moving through this cycle, recording PT development using qualitative data analysis software, to help reveal the 'back and forth' process of retroductive reasoning throughout a realist study. Similarly, McEwan et al. (2024) developed a framework for IPT development, allowing researchers to show their 'working out', *"while avoiding pitfalls of over-explanation and recording"* (p. 418). Other analytical approaches to enhancing transparency are also cited in the literature, for example, using a deliberative analytical approach (Nurjono et al., 2018) or journey maps in addition to documentary analysis (Price et al., 2021) to *"provide a chronological depiction of the interventions and [were used] to develop and refine initial program theories and inform the subsequent project activities"* (p. 3).

PT testing in RE can also be augmented by drawing on other theory-driven approaches such as Qualitative Comparative Analysis (QCA), which proposes to analyse the causal contribution of different conditions, or Theory of Change (ToC), which draws on available evidence and causal analysis to explain how interventions create change. For Kabongo et al. (2020), the combination of ToC and RE elucidated different levels of abstraction. They note, *"The ToC was used to explicate the implementation theory for program planning, improvement, and the development of robust monitoring systems at a macro program level. The RE approach was then used on a more micro-level (user-oriented understanding) of how and why the intervention would work or not"* (p. 9).

The reporting of theory, specifically substantive theory, in EE is somewhat different to that in RE. EE papers included in this review rarely explicitly described how traditional economic theories impacted on the choices made in the evaluation or implications for the analyses undertaken. Exceptions included Byford and Sefton (2003), who discuss the implication of welfare economics for evaluating complex interventions. They surmised that, rooted in the theoretical traditions of neoclassical and welfare economics, economic evaluation *"should be comparative in nature, societal in scope and concerned with the resulting wellbeing of the individuals involved"* but that *"the evaluation of a simple intervention is likely to have the potential to come much closer to the ideal than the evaluation of a complex one"* (p. 100). Likewise, Premji and Griffin (2024) further discuss the implications of welfare theory and the emergence of extra-welfarism for methods in EE. Specifically, that approaches to capturing benefits, including non-health and health consequences, should identify the total resource impacts in a system and make potential trade-offs clearer for decision makers. They advocate methods that recognise *"the institutional arrangements and opportunity costs associated with public sector budgets"* and *"the use of general frameworks to consider multiple outcomes and nest within them both traditional cost-benefit and cost-effectiveness approaches"* (p. 1480).

Furthermore, the role of substantive theory in EE was a topic of discussion in workshops #1 and #3, which explored how the use of theory in EE has informed the methods of data collection and analysis that dominate today. Specifically, discussion focussed on the relative merits of welfarism and extra-welfarism, recognising that: a) while welfarism advocates a broad societal perspective (whereby the impact of an intervention or initiative on all resources and on all aspects of

wellbeing are accounted for, whether quantifiable or not), the application of the theory tended to be narrower, focusing on a concept of efficiency based on distribution of resources through maximising welfare (utility) through aggregation of individual preferences; b) 'extra-welfarism', as relatively unique to EE in health, emerged to overcome the perceived narrowness of the application of welfare, to allow for consideration of extra measures of wellbeing in addition to utility, including health and non-health gains; c) broader utility measures were developed, among other things, including health-related quality of life (HRQoL), quality adjusted life years (QALYs) and the extension of the capability approach; and d) that the utility measures used with extra-welfarist approaches are also constructed from individual preferences, and replace those traditionally used in welfare economics rather than supplementing them.

As well as substantive theory, EE also draws on ToC and PT (implicitly or explicitly) when defining the parameters of an evaluation. However, beyond this, the use of PT in EE is limited. Responding to the original MRC Framework for the design and evaluation of complex interventions to improve health, Byford and Sefton (2003) advocated exploring the theory *behind* an intervention to help generate appropriate hypotheses regarding its likely effectiveness and cost-effectiveness, and to provide information on potential sources of variability and confounding factors. They suggested identifying the intervention components, to explore and determine the mechanisms by which outcomes are influenced, prior to conducting an EE. Furthermore, King (2016, 2021) observed, *"Evaluations of policies and programs often use a theory of change to articulate how the intervention is intended to function and the mechanisms by which it is supposed to generate outcomes. When an evaluation includes cost and efficiency considerations, economic and other concepts can be added to a theory of change to articulate a theory of value creation that articulates the mechanisms by which the intervention should use resources efficiently, effectively and create sufficient value to justify the resource use"* (p. 1). Using theory in the ways described by Byford and Sefton (2003) and King (2021), to understand and identify what to measure and test to build robust EEs, may be seen as moving towards more explicit use of PT and potentially realist approaches.

In summary, the theoretical challenges and developments in RE and EE highlight several important considerations for the integration of these approaches. Given the need for pragmatism and the relative ontological flexibility of EE approaches, alongside adherence to generative causation, any integration of RE and EE should be situated within a realist philosophy of science. This will involve the use of retroductive reasoning and both programme and substantive theory; the choice, development and testing of which must be transparent, including the use of heuristics to demonstrate generative causality. While this may be at odds with the ways in which EE uses information to arrive at conclusions, and the understanding of causation, recent EE literature has highlighted the added benefit theory driven approaches can bring. Indeed, the integration of RE and EE offers opportunities to bridge economic welfarist and extra-welfarist approaches, given the emphasis in RE on developing PTs for everything that counts, and the adoption of a broad range of methods for testing PTs.

3.4. Methodological challenges and developments

3.4.1. Accounting for context

The roles and uses of the term 'context' in social sciences have a history of being under-conceptualised or overlooked, leading to confusion and inconsistent use (Greenhalgh and Manzano, 2022). For RE, mechanisms will only operate in the presence of specific contextual circumstances (Pawson and Tilley, 1997).

In RE, 'context' is only discernible when considered as a unique component in a CMO configuration, i.e. a context is only a context if it triggers a mechanism. Citing Boudon (2014), Greenhalgh and Manzano (2022, p. 584) emphasised, *"in the social sciences, it is impossible to talk about context in general terms, since context is always defined specifically."*

For this reason, realist context tends to be teased out through refinement of PTs. Acquiring data about context, with a view to determining which specific elements are actively triggering a mechanism, can be done in various ways. While realist evaluators do not adhere to any particular method, some typical approaches include: researcher embeddedness (Ebenso et al., 2019; Gilmore, 2019; Karnon et al., 2024); stakeholder involvement (Douthwaite et al., 2017; Dalkin et al., 2018; Cooper et al., 2020; Bhattacharya et al., 2022; Ataman et al., 2024); and the use of participatory and longitudinal approaches (Taylor et al., 2021). These approaches can significantly benefit theory gleaning (Wong et al., 2016) and IPT development.

Coast et al. (2000, p. 43) observed that EEs are *"extremely context dependent"*. The importance and significance of context in EE, particularly in the evaluation of complex interventions, is at the heart of tensions around the generalisability (or transferability) of EE findings beyond the evaluation. García-Mochón et al. (2021) argue that EE must be contextualised and adjusted for key local conditions to account for generalisability and transferability. Here, they use 'generalisability' to refer to assumptions around whether findings from one setting can apply in a different setting, and 'transferability' to refer to whether the methods for data collection and analysis used in one setting might be similarly valid in another. In their scoping review on social care EEs, Weatherly et al. (2017) noted that evaluations in social contexts were generally more complex, with less clearly defined pathways and found, *"typically, authors suggested that the study findings had restricted relevance due to the specific context and heterogeneity of the intervention* e.g. *that there were different models of a service, which might not be easily replicable in other parts of the jurisdiction, or reflect current practice elsewhere"* (p. 16). Context in EE therefore has implications for the choice of comparator; accounting for existing resources (e.g. the same resources might have different costs in different places) or different opportunity costs (e.g. different budget constraints); the impact of people in and on the intervention; and the impact of external factors (Birch and Gafni, 2003; Byford and Sefton, 2003).

To enhance generalisability, EE strives to account for context, although context is inherently and inevitably restricted to quantifiable elements rather than the broader aspects considered in RE. EE therefore tends to stratify for context in its design or analyses, reducing context to tangible variables (defined in terms of geographical or people/patient subgroups or organisational characteristics; see 'terminology' subsection). Examples used in EE that attempt to adjust for or account for context in these ways include.

a) Adopting as broad a perspective as possible, or 'the societal perspective' (Hammitt, 2012; Garrison et al., 2018; Haacker et al., 2020; Breslau et al., 2023), to identify all costs and wider societal impacts (Premji and Griffin, 2024) or benefits related to the delivery and implementation of an intervention, beyond the patient or sector, irrespective of who pays, while explicitly understanding who bears the cost, acknowledging that these will differ between interventions, geographies, and people/patients.

b) Identifying subgroups to explore and explain how costs and outcomes vary between groups including using: 'net health benefit' and 'value of information analysis' approaches to identify which subgroups will benefit most (Espinoza et al., 2014); sub-group and regression analysis to adjust for potential confounders or identify key subgroups (Hoch and Dewa, 2014); and Latent Class Analysis to identify subgroups that share certain qualitative characteristics (inferred from unobserved latent variables that may explain differences in context) (Sorensen et al., 2018).

c) The use of objective approaches to describe and analyse the extent to which different interventions will impact differentially on different groups. For example, distributional CEA (Asaria et al., 2015; Cookson et al., 2017; Kwon et al., 2023), equity informed CEA (Avancena and Prosser, 2021), extended CEA (Verguet et al., 2016) and 'Generalised Cost-Effectiveness Analysis' (Bertram et al., 2021;

Shafrin et al., 2024, p. 97). And the use of normative approaches to assign differential weights to different group gains to adjust for this, for example equity weights (Weatherly et al., 2009; Culyer and Bombard, 2012; Asaria et al., 2015; Cookson et al., 2017, 2021; Round and Paulden, 2018; Avancena and Prosser, 2021; Paulden and McCabe, 2021; Schultz et al., 2024).

d) Adopting qualitative methods alongside EE to identify and understand contextual impacts including the use of qualitative health economics (Coast and De Allegri, 2018) and qualitative approaches alongside EE (Rogers et al., 2009; Swank et al., 2015), as well as participatory methods (Kumar et al., 2023) to *"generate context-specific and applied evaluation outputs [to] help meet demand from policymakers for locally relevant evidence that captures both resources and mechanisms and centre the service user within a network of decision makers"* (p. e1001).

The concept of transferability was first introduced by Lincoln and Guba (1985) as *"an evaluative criterion suitable for constructivist qualitative research to replace the (post-)positivist criterion of external validity, or generalisability, that describes the extent to which causal relationships identified in one study hold true in other populations, settings, or times"* (Stalmeijer et al., 2024, p. 1). RE focuses specifically on causal processes (configurations of CMO) and therefore in RE the analysis enables learning that may be transferable. However, this is specifically related to the transferability of the potential actions of the identified mechanism(s) under similar context conditions, i.e. mechanisms can be common across different settings (Wong, 2018). Mukumbang and Wong (2025) recently revisited the Firestone (1993) typology of generalisation (sample to population generalisation; analytic or theoretical generalisation; case to case transfer) to understand how generalisation could apply in RE. They argued that analytic generalisation is possible in mechanism-based theory-driven research methods or approaches like realist methodologies, as the premise is that the same causal forces theorised in the research setting are also found in other settings. This is used alongside contextualisation (considering the relevant contextual features in relation to the mechanisms of action) and retroduction (a theory-building approach to uncovering the underlying causal mechanisms responsible for an outcome) (Mukumbang, 2023; Mukumbang and Wong, 2025).

3.4.2. Identifying, measuring and valuing mechanism outcomes and costs

In the evaluation of complex interventions, the measurement and valuation of costs and outcomes for EE is complicated due to the range of costs and outcomes that can be considered and where these fall within the system (Weatherly et al., 2009), as well as 'spillovers' or 'externalities' (where outcomes and costs can extend to informal caregivers, family members, communities, other sectors or wider society). With respect to costs, developments in micro-costing methods have sought to increase the accuracy with which costs can be identified, measured and attributed, including the use of methods from accounting such as ingredients-based approaches (Adesina and Bollinger, 2013), activity-based costing (ABC) (Cunnama et al., 2016), and time-Driven ABC (da Silva Etges et al., 2019; Cidav et al., 2020) to produce more granular cost estimates.

However, considerably more attention in the EE literature has been given to outcome measurement and valuation in the evaluation of complex interventions. In particular it has been argued that Quality Adjusted Life Years (QALYs) (as a generic preference based outcome measure), which are common to EEs, do not: 1) sufficiently capture the breadth of outcomes resulting from public health or complex interventions, especially non-health; 2) capture the outcomes of relevance or importance to different stakeholders; and are not 3) sensitive enough to measure changes in these interventions over time (Byford and Sefton, 2003; Kelly et al., 2005; Weatherly et al., 2009; Payne et al., 2013; Whitaker et al., 2016). In response, developments in EE included: 1) weighting outcomes using a multiplier effect to account for outcome spillovers (Al-Janabi et al., 2016), such as outcomes to relatives and

carers (the R-QALY (Davidson and Levin, 2010)); 2) extending outcome measures to capture cross-sector outcomes, such as the wellbeing index (Weatherly et al., 2009); the 'wellbeing-adjusted life-year' (Brazier and Tsuchiya, 2015), Carer QALY (Peasgood et al., 2021) or education QALY (Weatherly et al., 2009); 3) developing alternative outcome measures to capture broader outcomes such as capabilities (Lorgelly et al., 2010) and social impact and value (Fujiwara, 2010; Ashton et al., 2024); or 4) considering multiple outcomes collectively using multi-criteria approaches (Wildman and Wildman, 2019) or value for money rubrics (Peterson and Skolits, 2020). Nevertheless, these alternative responses have yet to achieve widespread adoption in health and social care decision-making, which is synonymous with the QALY. In addition, disaggregated reporting of costs and outcomes was advocated to improve transparency and detail, and to enhance decision-making by meeting the needs of policymakers. This included the presentation of qualitative data alongside EE findings (Frew and Breheny, 2019); the use of CCA (Mauskopf et al., 1998; Sanders et al., 2016); and 'extended impact inventory' (Walker et al., 2019).

In RE, Pawson and Tilley (1997) argue that the main objective is to improve policy and programmes and therefore a clear understanding of outcome patterns and why they vary is required. To interrogate these outcome patterns, a generative approach to causation (see 'Causality' subsection) is required, to uncover CMO configurations. According to Bhaskar's (2013) stratified ontology (which sets out three overlapping domains: the real, the actual and the empirical), outcome patterns exist in the empirical domain, while mechanisms are located in the real. This means they are unobservable and exist independently of our knowledge or experience. Due to their location in the domain of the real, there remains debate within RE as to whether mechanisms can be measured. Williams (2018, p. 26) notes, there will always *"be a conceptual gap between mechanisms formulated by the scientist (within programme theories) and the 'real' mechanisms out there […] Researchers should be eclectic about the use of theories and methods, and use these as heuristic tools to propose possible mechanisms. However, our theories should always be testable, so that theory is always linked to method"*. This suggests that all data collection and theories should or could be considered as 'proxies' or 'indicators' in RE, particularly those that are unobservable, such as realist mechanisms.

3.4.3. Study design

Due to its inherently comparative nature and need for a counterfactual, EE has sometimes been equated with experimental or quasi-experimental study designs to estimate effect sizes and the incremental cost per unit of effect, in addition to attributing links between inputs and outputs. However, experimental designs are often less well-suited to the evaluation of complex interventions, where exposure cannot be controlled, making evidencing causal attribution more difficult. Furthermore, Hawkins (2016) argues that RCTs' use of counterfactuals, while valid when investigating the effects of an intervention in a simple or complicated system, becomes meaningless in a complex adaptive system where such a counterfactual is also inherently complex and adaptive, meaning there is no straightforward falsifiable scenario. Similarly, Baghbanian et al. (2025, p. 2) noted, *"One critical gap in Health Technology Assessment (HTA) is its limited ability to fully capture the dynamic interactions, contextual factors and long-term impacts associated with complex interventions. As such, these multi-component, context-sensitive interventions require broader, adaptive methods to assess their effectiveness in real-world settings – an area where traditional HTA may struggle"*. This has led to a growing interest in the use of EE alongside natural experiments (Deidda et al., 2019). The recent framework for using natural experiments to evaluate population health interventions (Craig et al., 2025) provides an integrated guide to the evaluation of policies and other large scale population health interventions as natural experiments. Similar to the guidance for developing and evaluating complex interventions (Skivington et al., 2021), it explicitly advocates for the use of mixed methods, but stops short of suggesting an integration of RE and natural experiments. Instead, it is focused on *"evaluations whose primary*

aim is to estimate effect sizes of interventions, rather than studies whose main goal is to address questions such as the contexts in which interventions work" (Craig et al., 2025, p. 7).

Alternatively, there is growing use of models and economic modelling for evaluating complex interventions in complex systems, where any number of hypothetical comparators can be simulated. Squires et al. (2012) argue that complex systems modelling (CSM) of public health interventions provides a deeper understanding of the impact of changing factors that affect the system. In developing guidance for the use of CSMs for public health evaluation, Breeze et al. (2023) noted that researchers need to access a diffuse and diverse literature from across multiple disciplines, and must pay close attention to defining CSMs and their characteristics; when to use them; what model types should be used (based on the study's dimensions); and what processes are important (including stakeholder engagement, understanding the problem, setting model boundaries, uncertainty, validation, and how to use data and evidence). Recent developments in the use of economic modelling in public health evaluation have proposed and explored the use of theory informed models that bear some similarity to the configurational approach adopted in RE, including potentially considering realist mechanisms. Brown et al. (2021), explored the feasibility of analysing realist programme theories alongside Structural Equation Modelling (SEM), concluding that the approaches have the potential to be combined in mixed methods research but that their integration requires well-developed theories and large datasets, and may oversimplify regularities in order to manage complex reality. Similarly, Söling et al. (2023) used SEM alongside RE to explore and test contextual (behavioural) factors influencing physicians' adoption of digital health technologies. They claimed that the modified methodological approach enabled them to integrate interdisciplinary evidence of contextual influences on change processes (using RE), with quantitative theory-based testing, to examine the statistical relationships between context, latent mediating variables, and outcomes (SEM). Most recently, Squires et al. (2024) also proposed the 'Public Health Economic Models – Behaviour' (PHEM-B) toolbox, which outlines potential methods and tools for incorporating theories of behaviour change into public health economic models. This helps to ensure models are more theory informed, to improve their accuracy and relevance when evaluating the cost-effectiveness of public health programmes.

On the other hand, RE has become increasingly equated with qualitative methods and thus designs, although there are instances where more typically quantitative or experimental approaches (Dyer and Williams, 2021; Bhattacharya et al., 2022; Smeets et al., 2022; Westhorp and Feeny, 2024) or quasi-experimental approaches (Parr and Churchill, 2020) have been used. Significant attention has been given to the prospect of 'Realist RCTs', experimental studies that incorporate RE to determine the contexts and mechanisms by which an intervention works. Bonell et al. (2012) note, *"longitudinal quantitative data are then also required to test the wider relevance of findings from qualitative research. For example, a realist RCT might collect qualitative data in the period between baseline and follow-up questionnaire surveys and use emergent findings from the qualitative research to inform additional quantitative measures to be included in follow up surveys"* (p. 2304). However, some argue that realist and positivist ontologies are incompatible and suggest that these should be termed 'theory-informed RCTs' (Marchal et al., 2013). Others have outlined guidance on how realist RCTs might be conducted; for example, Jamal et al. (2015) suggested that realist RCTs could enable the use of interim analysis of qualitative data to identify the need for any further quantitative data. Furthermore, Singleton et al. (2023) conclude that critical realist trials are possible using Partial Least Squares SEM, to move beyond correlational analysis, to support realist identification of the mechanisms that generate correlations. Most recently, Bonell et al. (2024) elaborated several ways in which the incorporation of realist enquiry methods into RCTs can *"generate evidence which is both more scientific and more useful"* (p. 101).

3.4.4. Integrating and triangulating methods of data collection and analysis

Despite its claimed 'methods agnosticism', REs most commonly use qualitative methods, especially interviews. The mapping review by Renmans and Castellano Pleguezuelo (2023) found that 97% of all REs use some kind of interviews (with low use of realist interviews); surveys are less common and lack a clear function; and relatively few studies used innovative data collections methods. Dyer and Williams (2021) note, *"Realism needs quantitative methods to be relevant, but equally we believe quantitative methods are enriched by realism"* (p. 110). Indeed, Westhorp and Feeny (2024) recently argued that the use of both quantitative and qualitative surveys can considerably strengthen standard approaches to realist analysis, so long as the PTs are well-developed.

Realist evaluators often endorse the use of mixed qualitative and quantitative methods (Fletcher et al., 2016; Gilmore et al., 2019; Butler et al., 2022). Lukersmith et al. (2023) advocated using multi-layered data analysis for large and diverse mixed methods datasets, and using machine learning to assist in the coding of complex variable information in the CMO configuration, which could then be analysed for patterns and associations. Furthermore, the need for mixed methods and novel approaches was highlighted in previous attempts to integrate RE and EE approaches. In their integration of RE and EE approaches, Rees et al., 2022b, p. 984) conclude, *"Ultimately, further realist economic evaluation should consider mixed methods approaches including more objective quantitative measures (e.g. knowledge tests), alongside larger sample sizes in order to make better sense of patterns in the data."*

EE, while predominantly reliant on quantitative methods of data collection and analysis, embraces qualitative methods, to: identify and understand the parameters of the EE (including theorising how the intervention works); or contextualise EE findings (see Findings subsection one). However, the application of qualitative methods alongside EE should be done so critically and discriminately. As Törnberg (2018) notes, adopting complex realism in economics *"rejects the idea of universal formal models, pointing instead to an epistemology where a plurality of methods is employed – deep qualitative study, quantitative techniques, inferential statistics, as well as modelling and simulation – while being mindful with respect to their epistemological limitations, conscious of the ontological nature of the system under study, and hence drawing on social theory to frame the research and make sense of the findings"* (p. 528). He therefore argues that such integration calls for *"a question-driven and methodologically pluralist approach, that uses modelling tools from both neoclassical economics and complexity science to explore stasis and dynamics, but does so within an epistemological framing that enables economists to draw insights that are situated, reflexive and meaningful"* (ibid.)

In summary, the methodological challenges and developments in RE and EE highlight several important considerations for the integration of these approaches. Despite conceptualising context in different ways, context in RE and EE is fundamental to defining the generalisability and transferability. Outcome measurement takes a more central role in EE compared to RE. EE is mainly quantitative, and comparing outcomes and costs requires measurement and quantification; RE requires mixed methods and some realist evaluations have been qualitative. The underpinning philosophical stance of RE will mean that while PT measurement can be attempted in the integration of RE ad EE, often these measures will only be proxies – indicators of the 'real'. Attention should therefore be given to the definition, measurement and explanation of all components of the realist configuration and resources, commensurate with EE. This would also potentially allow for micro costing across the entire PT. Together, developments in design and methods of evaluations of complex interventions in both RE and EE suggest some clear overlaps, if not convergence. Drawing on both realist and economic evaluators' perspectives could allow better understanding of: the critical parameters of the evaluation, the decision question and what types of data are needed, and to create a more comprehensive study design. This includes the necessity for a comparator to enable the incremental measurement of changes (in both costs and outcomes). The range of innovation and/or mixed method study designs in the review highlighted some valuable potential approaches when integrating RE and EE. However, evaluators must exercise caution when integrating RE and EE, to remain true to the ontological and epistemological underpinnings of the method. Any such analysis would be required to produce realist causal explanations.

3.5. Practical challenges and developments

3.5.1. Terminology

Differences in meanings of terms shared between RE and health EE became evident during the scoping reviews and expert stakeholder workshops, presenting challenges for both discussing and integrating the approaches. In some cases, terms were used in a method-specific way beyond their standard dictionary definition. The PPIE group also found jargon from both approaches difficult to understand, with terms with different meanings further complicating matters. Particularly problematic terms identified during the expert stakeholder workshops and reviews included a) context, b) mechanisms and c) resources.

a) In RE, context refers to *"those features of a situation into which a programme is introduced that affect the operation of [particular] programme mechanisms"* (Greenhalgh et al., 2017, p. 1). In their examination of 40 realist studies' definitions and operationalisations of 'context', Greenhalgh and Manzano (2022) identified two key 'narratives': *"Context as 1) observable features (space, place, people, things) that triggered or blocked the intervention; assuming that context operates at one moment in time and sets in motion a chain reaction of events. Or 2) the relational and dynamic features that shaped the mechanisms through which the intervention works; assuming that context operates in a dynamic, emergent way over time at multiple different levels of the social system"* (p. 583). The first definition (observable features) might be easier to measure and value but narrows the range of what might be considered to function as context, potentially restricting the range of causal explanations. The second definition (relational and dynamic features) is flexible and more in line with open, adaptive and complex systems, but may be harder to measure or cost. Sheaff et al. (2021) developed a taxonomy of contexts in RE, which they claim helps to *"differentiate contexts according to the ways in which they impact upon policy mechanisms (and so help disambiguate a key term in evaluations), aid the pooling and comparison of findings across realist evaluations, indeed across policy research generally and facilitate future realist evaluations"* (p. 191). Of note here is the aim to disambiguate the term and to aid comparison, however, definitions of context and mechanism used in the taxonomy could be contested due to further developments in the terms (e.g. Greenhalgh and Manzano, 2022).

In EE, 'context' is mainly limited to referencing the setting (often location), person or subgroup (individual attributes) or sector, for example, the 'legislative context' (Pedersen and Rieper, 2008); it does not have the same specific meaning as it does in RE. The current Consolidated Health Economic Evaluation Reporting Standards (CHEERS) states, EEs should *"provide relevant contextual information that may influence findings"* (Husereau et al., 2022, p. 4). However, it does not state what is meant by context and, while the word 'relevant' is important here, it does not define context as a component in a hypothesised causal configuration, as in RE.

b) The definition of mechanisms is contested across disciplines (Mahoney, 2001; Marchal et al., 2012). In realist philosophy, mechanisms are hidden causal processes that are context sensitive and cause outcomes (Westhorp, 2014). Programmes (interventions) work through 'mechanisms', i.e. by providing resources that influence the decisions made by programme actors, including beneficiaries. The interaction between resource and reasoning, is known as a programme mechanism (Pawson and Tilley, 1997). In EE, mechanisms tend to be defined (if at all) as a tangible mediator(s) and/or

moderator(s) associated with the intervention that may lead to differential outcome effects (Baron and Kenny, 1986; Kraemer, 2012).

c) EEs are inherently concerned with the efficient allocation of resources, which are generally classified as human, land, capital and consumable items, all of which have alternative uses and therefore, potentially, opportunity costs. It is not the resource itself that is of interest in EE but rather the opportunity costs attached to that resource, i.e. the potential opportunity to use a scarce resource in another (potentially more efficient) way. Importantly, things that can be consumed, such as a service or product, or a care professional's time, carry an opportunity cost regardless of whether funding is used to obtain it. Realists characterise resources more broadly, as being provided or enabled by an intervention and altering the reasoning of participants (Pawson and Tilley, 1997). This was elaborated by Dalkin et al. (2015), who further disaggregated realist mechanisms into 'resources' and 'reasoning', thereby enabling researchers to interpret which elements of an intervention or its context might be considered as providing resources of relevance in a PT. However, definitions of resources are still contested within RE.

Due to differences and overlaps in terminology, the need for lexical and conceptual clarity is common in interdisciplinary research, and can drive the evolution of terms, sometimes leading to the development of a shared glossary (Butler et al., 2022; Rees et al., 2022a,b).

3.5.2. Scale and scope of evaluations

Workshop #1 highlighted the need for pragmatism in the evaluation of complex interventions (Billings et al., 2020); sufficient time for longitudinal data collection and analysis (Rolfe et al., 2020; Taylor et al., 2021) to measure changes over time (Kelly et al., 2005; Weatherly et al., 2009; Payne et al., 2013; Whitaker et al., 2016); and to allow for designs to be iterative and adapted based on accumulative learning from the evaluation over time (Douthwaite et al., 2017). Indeed Coast and De Allegri (2018) suggested that *"the linear approach to research associated with deducing theory and then testing associated hypotheses that is common in quantitative health economics, is replaced in qualitative health economics with a much more iterative approach to research, where questions are explored with initial data collection and early analysis, and this early analysis then drives further sampling and data collection decisions, in a process that continues until theory is fully developed or understanding is complete."* (p. 2). Jamal et al. (2015, p. 8) suggest that protocols for such studies *"should allow for some iterative adaptation, allowing for some secondary and exploratory analyses to be determined from interim analysis of emerging qualitative process evaluation data"*, which may require extra time and resources. Our expert stakeholder group acknowledged not only the time required to assemble a well-integrated multidisciplinary team but also to sustain effective integration over long periods.

In summary, the practical challenges and developments in RE and EE highlight several important considerations for the integration of these approaches. The review highlighted differences in meanings in terms across the approaches. In the integration of RE and EE there is therefore a need for evaluators to understand where shared terms may cause confusion and ensure that any terms are transparently agreed and defined. In addition, developing and testing more comprehensive PTs for an integrated and RE and EE is likely to require significant time, to allow for more responsive use of appropriate quantitative methods and enable more iterative analysis between multiple datasets. There will likely therefore be a trade-off between navigating the practical and funding challenges of conducting such research and developing sufficiently nuanced and sophisticated studies. The advantages of longitudinal approaches, which can reveal change over time must however be balanced against the implications for cost and duration of the study.

4. Recommendations

The findings from the scoping review and expert stakeholder workshops suggest several theoretical, methodological and practical recommendations for the integration and RE and EE.

4.1. Theoretical recommendations

- REE should be theory driven. Involving economic and realist evaluators from the outset will enable both realist and economic logic to be appropriately embedded into theory development.
- REE approaches should be grounded in a realist philosophy of science, adopting a generative understanding of causation as opposed to successionist or attributional causation. Of particular importance is the idea that REE findings can be transferable, or analytically or theoretically generalisable, based on causal forces (mechanisms).

4.2. Methodological recommendations

- REE must adopt a broad understanding of context in line with RE, while also drawing on EE perspectives. A more nuanced understanding of context will allow evaluators to understand and map resources in more detail across the PT. This also suggests the adoption of as broad a perspective as possible when costing.
- REE, like EE, requires a comparator, in which the differential (incremental) effects and costs can be measured.
- REE should take an iterative approach to data collection and analysis, and remain open to emergent new theories as data collection and analysis progresses.
- REE is inherently interdisciplinary and will therefore likely use mixed methods in an integrated way to test PTs that aim to explain if, how, for whom and in which circumstances interventions work, and at what costs. This will involve drawing on existing methods of data collection and analysis, and other relevant expertise, guidance and reporting standards.
- REE should draw on methods commonly used in RE or EE, as long as the generative causal configuration can be maintained within the analysis, or different analyses can be used to form the configuration (e.g. combining linear statistical tests of differences in outcomes with qualitative findings).
- Conducting a REE is likely to be complicated, making use of multiple data sources. Transparency in documenting analyses and reporting will therefore be key. Reports should appropriately identify decisions and assumptions made throughout the study and (where appropriate) draw on sensitivity analysis to strengthen findings.

4.3. Practical recommendations

- REE would benefit from a glossary of terms to ensure these are defined and agreed. If agreement cannot be reached on specific terms, having multiple meanings about the same terms is possible, but this must be explicit within the team and explained in all reporting.
- REE will require an iterative and developmental approach, using pragmatic study designs that may change as theories are refined. Designs that are sensitive to changes in outcomes over time will inevitably take more time and may require significant resources to fund and deliver.

These recommendations were used to inform the development of guidance for the integration of RE and EE (redacted for author anonymity), retaining the decision-informing goals of EE, while also achieving the generalisable knowledge-producing evidence-based programme theory inherent to RE.

5. Strengths and limitations

This is the first paper to scope the literature on integrating RE and EE, and we recognise there is potential for refinements. For example, since most academic papers contain some reference to 'methodology', methodological scoping reviews can generate large numbers of results. An information specialist determined search strategies that would yield a reasonable number of useable articles, but we acknowledge there may be articles missing. While the number of studies that explicitly combined RE and EE approaches was comparatively small, the searches on recent developments in methods generated more results. Only studies that clearly demonstrated methodological innovation in RE and EE were included. The integration of RE and EE has both national and international relevance, with the potential to usefully inform both UK and international government approaches to evaluation (e.g. the UK Magenta Book). We believe this paper therefore makes a significant contribution to the social sciences and health services research methodological literature.

6. Conclusion

The integration of RE and EE speaks to the recommendations made in the updated MRC guidance for developing and evaluating complex interventions (Skivington et al., 2021). Our scoping review identified and critiqued the few articles that have attempted to integrate these approaches, before widening out to explore recent developments in RE and EE in the evaluation of complex interventions. We identified the theoretical, methodological and practical similarities and differences between the two approaches, to produce recommendations to inform the development of guidance for their integration.

Ethical approval

Ethics not required.

Funding statement

CRediT authorship contribution statement

Andrew Fletcher: Formal analysis, Writing – original draft, Writing – review & editing. **Sonia Dalkin:** Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing – review & editing. **Rob Anderson:** Formal analysis, Writing – review & editing. **Rachel M. Baker:** Formal analysis, Writing – review & editing. **Cam Donaldson:** Formal analysis, Writing – review & editing. **Vivienne Hibberd:** Writing – review & editing. **Meghan Bruce Kumar:** Formal analysis, Writing – review & editing. **Felicity Shenton:** Writing – review & editing. **Gill Westhorp:** Formal analysis, Writing – review & editing. **Geoff Wong:** Formal analysis, Writing – review & editing. **Judy Wright:** Data curation, Methodology, Writing – review & editing. **Angela Bate:** Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2026.119050.

Data availability

Data will be made available on request.

References

Adesina, A., Bollinger, L.A., 2013. Estimating the cost-savings associated with bundling maternal and child health interventions: a proposed methodology. BMC Public Health 13, 1–7.

Al-Janabi, H., van Exel, J., Brouwer, W., Coast, J., 2016. A framework for including family health spillovers in economic evaluation. Med. Decis. Mak. 36 (2), 176–186. https://doi.org/10.1177/0272989X15605094.

Anderson, R., 2004. Realistic evaluation [book review]. Appl. Health Econ. Health Pol. 2 (1), 84–86.

Anderson, R., Hardwick, R., Pearson, M., Byng, R., 2018. Using realist approaches to explain the costs and cost-effectiveness of programmes. In: Doing Realist Research. SAGE Publications Ltd, London.

Anderson, R., Hardwick, R., 2016. Realism and resources: towards more explanatory economic evaluation. Evaluation 22 (3), 323–341. https://doi.org/10.1177/1356389016652742.

Asaria, M., Griffin, S., Cookson, R., Whyte, S., Tappenden, P., 2015. Distributional cost-effectiveness analysis of health care programmes–a methodological case study of the UK bowel cancer screening programme. Health Econ. 24 (6), 742–754. https://doi.org/10.1002/hec.3058.

Ashton, K., Cotter-Roberts, A., Clemens, T., Green, L., Dyakova, M., 2024. Advancing the social return on investment framework to capture the social value of public health interventions: semistructured interviews and a review of scoping reviews. Public Health 226, 122–127.

Ataman, R., Ahmed, S., Berta, W., Thomas, A., 2024. Sustainability of an outcome measure in outpatient stroke rehabilitation: a realist evaluation. J. Eval. Clin. Pract. 30 (4), 559–574.

Avancena, A.L.V., Prosser, L.A., 2021. Examining equity effects of health interventions in cost-effectiveness analysis: a systematic review. Value Health 24 (1), 136–143. https://doi.org/10.1016/j.jval.2020.10.010.

Baghbanian, A., Carter, D., Merlin, T., 2025. Methods for the health technology assessment of complex interventions: a scoping review. PLoS One 20 (3), e0315381.

Baron, R.M., Kenny, D.A., 1986. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J. Personality Soc. Psychol. 51 (6), 1173.

Beeston, C., Robinson, M., Giles, L., Dickie, E., Ford, J., MacPherson, M., McAdams, R., Mellor, R., Shipton, D., Craig, N., 2020. Evaluation of minimum unit pricing of alcohol: a mixed method natural experiment in Scotland. Int. J. Environ. Res. Publ. Health 17 (10). https://doi.org/10.3390/ijerph17103394.

Bertram, M.Y., Lauer, J.A., Stenberg, K., Edejer, T.T.T., 2021. Methods for the economic evaluation of health care interventions for priority setting in the health system: an update from WHO CHOICE. Int. J. Health Pol. Manag. 10 (11), 673–677. https://doi.org/10.34172/ijhpm.2020.244.

Bhaskar, R., 2013. A Realist Theory of Science. Routledge.

Bhattacharya, D., Whiteside, H., Tang, E., Kantilal, K., Loke, Y., Atkins, B., Hill, C., 2022. A review of trial and real-world data applying elements of a realist approach to identify behavioural mechanisms supporting practitioners to taper opioids. Br. J. Clin. Pharmacol. 88 (9), 4019–4042. https://doi.org/10.1111/bcp.15379.

Billings, J., de Bruin, S.R., Baan, C., Nijpels, G., 2020. Advancing integrated care evaluation in shifting contexts: blending implementation research with case study design in project SUSTAIN. BMC Health Serv. Res. 20 (1), 971. https://doi.org/10.1186/s12913-020-05775-5.

Birch, S., Gafni, A., 2003. Economics and the evaluation of health care programmes: generalisability of methods and implications for generalisability of results. Health Policy 64 (2), 207–219. https://doi.org/10.1016/s0168-8510(02)00182-3.

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., Moore, L., 2012. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. Soc. Sci. Med. 75 (12), 2299–2306. https://doi.org/10.1016/j.socscimed.2012.08.032.

Bonell, C., Melendez-Torres, G.J., Warren, E., 2024. Realist Trials and Systematic Reviews: Rigorous, Useful Evidence to Inform Health Policy. Cambridge University Press.

Boudon, R., 2014. What is context? KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie 66 (Suppl. 1), 17–45.

Brazier, J., Tsuchiya, A., 2015. Improving cross-sector comparisons: going beyond the health-related QALY. Appl. Health Econ. Health Pol. 13, 557–565.

Breeze, P.R., Squires, H., Ennis, K., Meier, P., Hayes, K., Lomax, N., Shiell, A., Kee, F., de Vocht, F., O'Flaherty, M., 2023. Guidance on the use of complex systems models for economic evaluations of public health interventions. Health Econ. 32 (7), 1603–1625.

Breslau, R.M., Cohen, J.T., Diaz, J., Malcolm, B., Neumann, P.J., 2023. A review of HTA guidelines on societal and novel value elements. Int. J. Technol. Assess. Health Care 39 (1), e31.

Brown, A., Hecker, K.G., Bok, H., Ellaway, R.H., 2021. Strange bedfellows: exploring methodological intersections between realist inquiry and structural equation modeling. J. Mix. Methods Res. 15 (4), 485–506.

Bryman, A., 2016. Social Research Methods. Oxford university press.

Butler, C., Wilson, P., Abrahamson, V., Mikelyte, R., Gage, H., Williams, P., Brigden, C., Swash, B., Rees-Roberts, M., Silsbury, G., Goodwin, M., Greene, K., Wee, B., Barclay, S., 2022. In: Optimum Models of Hospice at Home Services for end-of-life Care in England: a realist-informed mixed-methods Evaluation. https://doi.org/10.3310/MSAY4464.

Byford, S., Sefton, T., 2003. Economic evaluation of complex health and social care interventions. Natl. Inst. Econ. Rev. 186 (1), 98–108.

Cidav, Z., Mandell, D., Pyne, J., Beidas, R., Curran, G., Marcus, S., 2020. A pragmatic method for costing implementation strategies using time-driven activity-based costing. Implement. Sci. 15 (1), 1–15.

Coast, J., Hensher, M., Mulligan, J.-A., Shepperd, S., Jones, J., 2000. Conceptual and practical difficulties with the economic evaluation of health services developments. J. Health Serv. Res. Policy 5 (1), 42–48.

Coast, J., De Allegri, M., 2018. Qualitative Methods in Health Economics. Oxford University Press.

Cookson, R., Mirelman, A.J., Griffin, S., Asaria, M., Dawkins, B., Norheim, O.F., Verguet, S., A, J.C., 2017. Using cost-effectiveness analysis to address health equity concerns. Value Health 20 (2), 206–212. https://doi.org/10.1016/j.jval.2016.11.027.

Cookson, R., Robson, M., Skarda, I., Doran, T., 2021. Equity-informative methods of health services research. J. Health Organisat. Manag. https://doi.org/10.1108/JHOM-07-2020-0275. *ahead-of-print*(ahead-of-print).

Cooper, C., Lhussier, M., Carr, S., 2020. Blurring the boundaries between synthesis and evaluation. A customized realist evaluative synthesis into adolescent risk behavior prevention. Res. Synth. Methods 11 (3), 457–470. https://doi.org/10.1002/jrsm.1407.

Craig, P., Campbell, M., Deidda, M., Dundas, R., Green, J., Katikireddi, S.V., Lewsey, J., Ogilvie, D., de Vocht, F., White, M., 2025. Using natural experiments to evaluate population health and health system interventions: new framework for producers and users of evidence. Br. Med. J. 388.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., 2008. Developing and evaluating complex interventions: the new medical research council guidance. Br. Med. J. 337.

Croci, G., 2024. A realistic evaluation of homicide reduction programmes. Int. Crim. Justice Rev., 10575677241271121

Culyer, A.J., Bombard, Y., 2012. An equity framework for health technology assessments. Med. Decis. Mak. 32 (3), 428–441. https://doi.org/10.1177/0272989X11426484.

Cunnama, L., Sinanovic, E., Ramma, L., Foster, N., Berrie, L., Stevens, W., Molapo, S., Marokane, P., McCarthy, K., Churchyard, G., 2016. Using top-down and bottom-up costing approaches in LMICs: the case for using both to assess the incremental costs of new technologies at scale. Health Econ. 25, 53–66.

da Silva Etges, A.P.B., Cruz, L.N., Notti, R.K., Neyeloff, J.L., Schlatter, R.P., Astigarraga, C.C., Falavigna, M., Polanczyk, C.A., 2019. An 8-step framework for implementing time-driven activity-based costing in healthcare studies. Eur. J. Health Econ. 20 (8), 1133–1145.

Dalkin, S., Lhussier, M., Williams, L., Burton, C.R., Rycroft-Malone, J., 2018. Exploring the use of soft systems methodology with realist approaches: a novel way to map programme complexity and develop and refine programme theory. Evaluation 24 (1), 84–97. https://doi.org/10.1177/1356389017749036.

Dalkin, S.M., Greenhalgh, J., Jones, D., Cunningham, B., Lhussier, M., 2015. What's in a mechanism? Development of a key concept in realist evaluation. Implement. Sci. 10, 49. https://doi.org/10.1186/s13012-015-0237-x.

Dalkin, S.M., Hardwick, R.J.L., Haighton, C.A., Finch, T.L., 2021. Combining realist approaches and normalization process theory to understand implementation: a systematic review. Implement Sci. Commun. 2 (1), 68. https://doi.org/10.1186/s43058-021-00172-3.

Dalkin, S., Bate, A., 2022. Developing Realist Economic Evaluation Methods (REEM) and Guidance to Evaluate the Impact, Costs, and Consequences of Complex Interventions. NIHR. Retrieved 1 December from. https://fundingawards.nihr.ac.uk/award/NIHR135102.

Davidson, T., Levin, L.-Å., 2010. Is the societal approach wide enough to include relatives? Incorporating relatives' costs and effects in a cost-effectiveness analysis. Appl. Health Econ. Health Pol. 8, 25–35.

Davies, F., Edwards, M., Price, D., Anderson, P., Carson-Stevens, A., Choudhry, M., Cooke, M., Dale, J., Donaldson, L., Evans, B.A., 2024. Evaluation of different models of general practitioners working in or alongside emergency departments: a mixed-methods realist evaluation. Health Social Care Deliv. Res. 12 (10), 1–152.

De Weger, E., Van Vooren, N., Wong, G., Dalkin, S., Marchal, B., Drewes, H., Baan, C., 2020. What's in a realist configuration? Deciding which causal configurations to use, how, and why. J. Qual. Methods 19, 1609406920938577.

Deaton, A., 2010. Instruments, randomization, and learning about development. J. Econ. Lit. 48 (2), 424–455.

Deidda, M., Geue, C., Kreif, N., Dundas, R., McIntosh, E., 2019. A framework for conducting economic evaluations alongside natural experiments. Soc. Sci. Med. 220, 353–361. https://doi.org/10.1016/j.socscimed.2018.11.032.

Douthwaite, B., Mayne, J., McDougall, C., Paz-Ybarnegaray, R., 2017. Evaluating complex interventions: a theory-driven realist-informed approach. Evaluation 23 (3), 294–311. https://doi.org/10.1177/1356389017714382.

Dyer, W., Williams, M., 2021. Quantification and realist methodologies. Int. J. Soc. Res. Methodol. 24 (1), 109–121.

Ebenso, B., Manzano, A., Uzochukwu, B., Etiaba, E., Huss, R., Ensor, T., Newell, J., Onwujekwe, O., Ezumah, N., Hicks, J., 2019. Dealing with context in logic model development: reflections from a realist evaluation of a community health worker programme in Nigeria. Eval. Progr. Plann. 73, 97–110.

Elder-Vass, D., 2022. Pragmatism, critical realism and the study of value. J. Crit. Realism 21 (3), 261–287.

Espinoza, M.A., Manca, A., Claxton, K., Sculpher, M.J., 2014. The value of heterogeneity for cost-effectiveness subgroup analysis: conceptual framework and application. Med. Decis. Mak. 34 (8), 951–964. https://doi.org/10.1177/0272989X14538705.

Feeny, S., Westhorp, G., Williams, E., 2023. Understanding sustainable outcomes in international development: towards a realist evaluation framework. J. Int. Dev. 35 (1), 21–42.

Firestone, W.A., 1993. Alternative arguments for generalizing from data as applied to qualitative research. Educ. Res. 22 (4), 16–23.

Fletcher, A., Jamal, F., Moore, G., Evans, R.E., Murphy, S., Bonell, C., 2016. Realist complex intervention science: applying realist principles across all phases of the medical research council framework for developing and evaluating complex interventions. Evaluation 22 (3), 286–303. https://doi.org/10.1177/1356389016652743.

Francis-Auton, E., Sarkies, M.N., Pomare, C., Long, J.C., Hardwick, R., Nguyen, H.M., Braithwaite, J., 2022. Real talk: a realist dialogic approach in a realist evaluation. Int. J. Qual. Methods 21. https://doi.org/10.1177/16094069221120748.

Frew, E., Breheny, K., 2019. Methods for public health economic evaluation: a Delphi survey of decision makers in English and Welsh local government. Health Econ. 28 (8), 1052–1063. https://doi.org/10.1002/hec.3916.

Fujiwara, D., 2010. The department for work and pensions social cost-benefit analysis framework. Department Work Pens. Work. Pap. Ser. 86.

García-Mochón, L., Rovira Forns, J., Espin, J., 2021. Cost transferability problems in economic evaluation as a framework for an European health care and social costs database. Cost Eff. Resour. Allocation 19 (1), 43.

Garrison Jr., L.P., Pauly, M.V., Willke, R.J., Neumann, P.J., 2018. An overview of value, perspective, and decision Context-A health economics approach: an ISPOR special task force report [2]. Value Health 21 (2), 124–130. https://doi.org/10.1016/j.jval.2017.12.006.

Gentles, S.J., Charles, C., Nicholas, D.B., Ploeg, J., McKibbon, K.A., 2016. Reviewing the research methods literature: principles and strategies illustrated by a systematic overview of sampling in qualitative research. Syst. Rev. 5, 1–11.

Gilmore, B., 2019. Realist evaluations in low- and middle-income countries: reflections and recommendations from the experiences of a foreign researcher. BMJ Glob. Health 4 (5), e001638. https://doi.org/10.1136/bmjgh-2019-001638.

Gilmore, B., McAuliffe, E., Power, J., Vallières, F., 2019. Data analysis and synthesis within a realist evaluation: toward more transparent methodological approaches. Int. J. Qual. Methods 18. https://doi.org/10.1177/1609406919859754.

Godber, E., Robinson, R., Steiner, A., 1997. Economic evaluation and the shifting balance towards primary care: definitions, evidence and methodological issues. Health Econ. 6 (3), 275–294. https://doi.org/10.1002/(sici)1099-1050(199705)6:3<275::Aid-hec255>3.0.Co;2-2.

Goodman, C., Davies, S.L., Gordon, A.L., Dening, T., Gage, H., Meyer, J., Schneider, J., Bell, B., Jordan, J., Martin, F., Iliffe, S., Bowman, C., Gladman, J.R.F., Victor, C., Mayrhofer, A., Handley, M., Zubair, M., 2017. In *optimal NHS service delivery to care homes: a realist evaluation of the features and mechanisms that support effective working for the continuing care of older people in residential settings*. https://doi.org/10.3310/hsdr05290.

Gordon, A.L., Goodman, C., Davies, S.L., Dening, T., Gage, H., Meyer, J., Schneider, J., Bell, B., Jordan, J., Martin, F.C., Iliffe, S., Bowman, C., Gladman, J.R.F., Victor, C., Mayrhofer, A., Handley, M., Zubair, M., 2018. Optimal healthcare delivery to care homes in the UK: a realist evaluation of what supports effective working to improve healthcare outcomes. Age Ageing 47 (4), 595–603. https://doi.org/10.1093/ageing/afx195.

Greenhalgh, T., Pawson, R., Wong, G., Westhorp, G., Greenhalgh, J., Manzano, A., Jagosh, J., 2017. What Realists Mean by Context; Or, Why nothing Works Everywhere or for everyone.

Greenhalgh, J., Manzano, A., 2022. Understanding 'context' in realist evaluation and synthesis. Int. J. Soc. Res. Methodol. 25 (5), 583–595.

Haacker, M., Hallett, T.B., Atun, R., 2020. On discount rates for economic evaluations in global health. Health Pol. Plann. 35 (1), 107–114. https://doi.org/10.1093/heapol/czz127.

Hammitt, J.K., 2012. Discounting health and cost-effectiveness analysis: a response to Nord. Health Econ. 21 (7), 878–882. https://doi.org/10.1002/hec.1782.

Hardwick, R., Pearson, M., Byng, R., Anderson, R., 2013. The effectiveness and cost-effectiveness of shared care: protocol for a realist review. Syst. Rev. 2, 1–7.

Hawkins, A.J., 2016. Realist evaluation and randomised controlled trials for testing program theory in complex social systems. Evaluation 22 (3), 270–285.

Higgins, A., O'Halloran, P., Porter, S., 2012. Management of long term sickness absence: a systematic realist review. J. Occup. Rehabil. 22, 322–332.

Hoch, J.S., Dewa, C.S., 2014. Advantages of the net benefit regression framework for economic evaluations of interventions in the workplace: a case study of the cost-effectiveness of a collaborative mental health care program for people receiving short-term disability benefits for psychiatric disorders. J. Occup. Environ. Med. 56 (4), 441–445. https://doi.org/10.1097/JOM.0000000000000130.

Husereau, D., Drummond, M., Augustovski, F., de Bekker-Grob, E., Briggs, A.H., Carswell, C., Caulley, L., Chaiyakunapruk, N., Greenberg, D., Loder, E., Mauskopf, J., Mullins, C.D., Petrou, S., Pwu, R.F., Staniszewska, S., Force, C., 2022. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. Int. J. Technol. Assess. Health Care 38 (1), e13. https://doi.org/10.1017/S0266462321001732.

Jamal, F., Fletcher, A., Shackleton, N., Elbourne, D., Viner, R., Bonell, C., 2015. The three stages of building and testing mid-level theories in a realist RCT: a theoretical and methodological case-example. Trials 16, 466. https://doi.org/10.1186/s13063-015-0980-y.

Johnson, S.D., Tilley, N., Bowers, K.J., 2015. Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews. J. Exp. Criminol. 11, 459–473.

Kabongo, E.M., Mukumbang, F.C., Delobelle, P., Nicol, E., 2020. Combining the theory of change and realist evaluation approaches to elicit an initial program theory of the MomConnect program in South Africa. BMC Med. Res. Methodol. 20 (1), 282. https://doi.org/10.1186/s12874-020-01164-y.

Karnon, J., Partington, A., Gray, J., Pincombe, A., Schultz, T., 2024. Local level economic evaluation: what is it? What is its value? Is it sustainable? Appl. Health Econ. Health Pol. 22 (3), 273–281.

Kelly, M.P., McDaid, D., Ludbrook, A., Powell, J., 2005. Economic Appraisal of Public Health Interventions. Health Development Agency, London.

King, J., 2016. Using economic methods evaluatively. Am. J. Eval. 38 (1), 101–113. https://doi.org/10.1177/1098214016641211.

King, J., 2021. Expanding theory-based evaluation: incorporating value creation in a theory of change. Eval. Progr. Plann. 89, 101963. https://doi.org/10.1016/j.evalprogplan.2021.101963.

Kraemer, H.C., 2012. Determining gene moderation of environmental risk factors for a mental disorder: a "perfect storm" of methodological problems. Int. J. Methods Psychiatr. Res. 21 (3), 185–194.

Kumar, M.B., Roder-DeWan, S., Nyondo-Mipando, A.L., Mirzoev, T., Marston, C., 2023. Participatory economic approaches in global health evaluations. Lancet Global Health 11 (7), e1001–e1002.

Kwon, J., Squires, H., Young, T., 2023. Economic model of community-based falls prevention: seeking methodological solutions in evaluating the efficiency and equity of UK guideline recommendations. BMC Geriatr. 23 (1), 187.

Lincoln, Y.S., Guba, E.G., 1985. Naturalistic Inquiry, vol. 75. sage.

Lorgelly, P.K., Lawson, K.D., Fenwick, E.A., Briggs, A.H., 2010. Outcome measurement in economic evaluations of public health interventions: a role for the capability approach? Int. J. Environ. Res. Publ. Health 7 (5), 2274–2289. https://doi.org/10.3390/ijerph7052274.

Lukersmith, S., Salvador-Carulla, L., Chung, Y., Du, W., Sarkissian, A., Millington, M., 2023. A realist evaluation of case management models for people with complex health conditions using novel methods and tools—what works, for whom, and under what circumstances? Int. J. Environ. Res. Publ. Health 20 (5), 4362.

Mahoney, J., 2001. Beyond correlational analysis: recent innovations in theory and method. Sociol. Forum.

Marchal, B., Van Belle, S., Van Olmen, J., Hoerée, T., Kegels, G., 2012. Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. Evaluation 18 (2), 192–212.

Marchal, B., Westhorp, G., Wong, G., Van Belle, S., Greenhalgh, T., Kegels, G., Pawson, R., 2013. Realist RCTs of complex interventions–an oxymoron. Soc. Sci. Med. 94, 124–128.

Mauskopf, J.A., Paul, J.E., Grant, D.M., Stergachis, A., 1998. The role of cost—consequence analysis in healthcare decision—making. Pharmacoeconomics 13, 277–288.

Mazzocato, P., Savage, C., Brommels, M., Aronsson, H., Thor, J., 2010. Lean thinking in healthcare: a realist review of the literature. BMJ Qual. Saf. 19 (5), 376–382.

McBride, O., Bunting, E., Harkin, O., Butter, S., Shevlin, M., Murphy, J., Mason, L., Hartman, T.K., McKay, R., Hyland, P., Levita, L., Bennett, K.M., Stocks, T.V.A., Gibson-Miller, J., Martinez, A.P., Vallieres, F., Bentall, R.P., 2022. Testing both affordability-availability and psychological-coping mechanisms underlying changes in alcohol use during the COVID-19 pandemic. PLoS One 17 (3), e0265145. https://doi.org/10.1371/journal.pone.0265145.

McEwan, K., Girling, M., Bate, A., Atkinson, J., Clarke, A., Dalkin, S., 2024. 'For want of a nail': developing a transparent approach to retroduction and early initial programme theory development in a realist evaluation of community end of life care services. Int. J. Soc. Res. Methodol. 27 (4), 417–430.

McGowan, J., Sampson, M., Salzwedel, D.M., Cogo, E., Foerster, V., Lefebvre, C., 2016. PRESS peer review of electronic search strategies: 2015 guideline statement. J. Clin. Epidemiol. 75, 40–46.

Mukumbang, F.C., 2023. Retroductive theorizing: a contribution of critical realism to mixed methods research. J. Mix. Methods Res. 17 (1), 93–114.

Mukumbang, F.C., Marchal, B., Van Belle, S., van Wyk, B., 2018. Unearthing how, why, for whom and under what health system conditions the antiretroviral treatment adherence club intervention in South Africa works: a realist theory refining approach. BMC Health Serv. Res. 18 (1), 1–15.

Mukumbang, F.C., Van Belle, S., Marchal, B., Van Wyk, B., 2016. Realist evaluation of the antiretroviral treatment adherence club programme in selected primary healthcare facilities in the metropolitan area of Western Cape Province, South Africa: a study protocol. BMJ Open 6 (4), e009977.

Mukumbang, F.C., Wong, G., 2025. Mechanism-based middle-range theories: using realist syntheses to reconcile specificity to context and generalizability. Qual. Health Res., 10497323251316401

Nurjono, M., Shrestha, P., Lee, A., Lim, X.Y., Shiraz, F., Tan, S., Wong, S.H., Foo, K.M., Wee, T., Toh, S.A., Yoong, J., Maria Vrijhoef, H.J., 2018. Realist evaluation of a

complex integrated care programme: protocol for a mixed methods study. BMJ Open 8 (3), e017111. https://doi.org/10.1136/bmjopen-2017-017111.

Parr, S., Churchill, H., 2020. The troubled families programme: learning about policy impact through realist case study research. Soc. Pol. Adm. 54 (1), 134–147.

Paulden, M., McCabe, C., 2021. Modifying nice's approach to equity weighting. Pharmacoeconomics 39 (2), 147–160. https://doi.org/10.1007/s40273-020-00988-2.

Pawson, R., 2008. *Causality for Beginners* NCRM Research Methods Festival 2008. Oxford. https://eprints.ncrm.ac.uk/id/eprint/245/.

Pawson, R., Tilley, N., 1997. Realistic Evaluation. sage.

Payne, K., McAllister, M., Davies, L.M., 2013. Valuing the economic benefits of complex interventions: when maximising health is not sufficient. Health Econ. 22 (3), 258–271. https://doi.org/10.1002/hec.2795.

Peasgood, T., Mukuria, C., Carlton, J., Connell, J., Devlin, N., Jones, K., Lovett, R., Naidoo, B., Rand, S., Rejon-Parrilla, J.C., Rowen, D., Tsuchiya, A., Brazier, J., 2021. What is the best approach to adopt for identifying the domains for a new measure of health, social care and carer-related quality of life to measure quality-adjusted life years? Application to the development of the EQ-HWB? Eur. J. Health Econ. 22 (7), 1067–1081. https://doi.org/10.1007/s10198-021-01306-z.

Pedersen, L.H., Rieper, O., 2008. Is realist evaluation a realistic approach for complex reforms? Evaluation 14 (3), 271–293. https://doi.org/10.1177/1356389008090856.

Peterson, C., Skolits, G., 2020. Value for money: a utilization-focused approach to extending the foundation and contribution of economic evaluation. Eval. Progr. Plann. 80, 101799. https://doi.org/10.1016/j.evalplan.2020.101799.

Petticrew, M., 2013. Public health evaluation: epistemological challenges to evidence production and use. Evidence Policy 9 (1), 87–95. https://doi.org/10.1332/174426413x663742.

Premji, S., Griffin, S., 2024. Assessing the health and welfare benefits of interventions using the wider societal impacts framework. Value Health 27 (11), 1479–1487.

Price, H., Domoney, J., Ariss, S., Hughes, E., Trevillion, K., 2021. Documentary analysis within a realist evaluation: recommendations from a study of sexual assault referral centres. Int. J. Qual. Methods 20. https://doi.org/10.1177/16094069211047818.

Rees, C.E., Foo, J., Nguyen, V.N.B., Edouard, V., Maloney, S., Ottrey, E., Palermo, C., 2022a. Unpacking economic programme theory for supervision training: preliminary steps towards realist economic evaluation. Med. Educ. 56 (4), 407–417. https://doi.org/10.1111/medu.14701.

Rees, C.E., Nguyen, V.N.B., Foo, J., Edouard, V., Maloney, S., Palermo, C., 2022b. Balancing the effectiveness and cost of online education: a preliminary realist economic evaluation. Med. Teach. 44 (9), 977–985. https://doi.org/10.1080/0142159X.2022.2051463.

Renmans, D., Castellano Pleguezuelo, V., 2023. Methods in realist evaluation: a mapping review. Eval. Progr. Plann. 97, 102209. https://doi.org/10.1016/j.evalplan.2022.102209.

Rogers, P.J., Stevens, K., Boymal, J., 2009. Qualitative cost-benefit evaluation of complex, emergent programs. Eval. Progr. Plann. 32 (1), 83–90. https://doi.org/10.1016/j.evalplan.2008.08.005.

Rolfe, S., Garnham, L., Godwin, J., Anderson, I., Seaman, P., Donaldson, C., 2020. Housing as a social determinant of health and wellbeing: developing an empirically-informed realist theoretical framework. BMC Public Health 20 (1), 1138. https://doi.org/10.1186/s12889-020-09224-0.

Round, J., Paulden, M., 2018. Incorporating equity in economic evaluations: a multi-attribute equity state approach. Eur. J. Health Econ. 19 (4), 489–498. https://doi.org/10.1007/s10198-017-0897-3.

Rutter, H., Savona, N., Glonti, K., Bibby, J., Cummins, S., Finegood, D.T., Greaves, F., Harper, L., Hawe, P., Moore, L., Petticrew, M., Rehfuess, E., Shiell, A., Thomas, J., White, M., 2017. The need for a complex systems model of evidence for public health. Lancet 390 (10112), 2602–2604. https://doi.org/10.1016/S0140-6736(17)31267-9.

Sanders, G.D., Neumann, P.J., Basu, A., Brock, D.W., Feeny, D., Krahn, M., Kuntz, K.M., Meltzer, D.O., Owens, D.K., Prosser, L.A., Salomon, J.A., Sculpher, M.J., Trikalinos, T.A., Russell, L.B., Siegel, J.E., Ganiats, T.G., 2016. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. JAMA 316 (10), 1093–1103. https://doi.org/10.1001/jama.2016.12195.

Schultz, S., Beissmann, F., Zorbas, C., Yoong, S., Peeters, A., Backholer, K., 2024. A realist impact evaluation of a tool to strengthen equity in local government policy-making. Int. J. Equity Health 23 (1), 179.

Shafrin, J., Kim, J., Cohen, J.T., Garrison, L.P., Goldman, D.A., Doshi, J.A., Krieger, J., Lakdawalla, D.N., Neumann, P.J., Phelps, C.E., 2024. Valuing the societal impact of medicines and other health technologies: a user guide to current best practices. Forum Health Econ. Pol.

Sheaff, R., Doran, N., Harris, M., Lang, I., Medina-Lara, A., Fornasiero, M., Ball, S., McGregor-Harper, J., Bethune, R., 2021. Categories of context in realist evaluation. Evaluation 27 (2), 184–209.

Shiell, A., Hawe, P., Gold, L., 2008. Complex interventions or complex systems? Implications for health economic evaluation. Br. Med. J. 336 (7656), 1281–1283.

Singleton, H., Porter, S., Beavis, J., Falconer, L., Hernandez, J.P., Holley, D., 2023. Accounting for complexity in critical realist trials: the promise of PLS-SEM. J. Crit. Realism 1–20.

Skivington, K., Matthews, L., Simpson, S.A., Craig, P., Baird, J., Blazeby, J.M., Boyd, K.A., Craig, N., French, D.P., McIntosh, E., 2021. A new framework for developing and evaluating complex interventions: update of medical research council guidance. Br. Med. J. 374.

Söling, S., Demirer, I., Köberlein-Neu, J., Hower, K.I., Müller, B.S., Pfaff, H., Karbach, U., Group, A.S., 2023. Complex implementation mechanisms in primary care: do

physicians' beliefs about the effectiveness of innovation play a mediating role? Applying a realist inquiry and structural equation modeling approach in a formative evaluation study. BMC Primary Care 24 (1), 131.

Smeets, R.G.M., Hertroijs, D.F.L., Mukumbang, F.C., Kroese, M., Ruwaard, D., Elissen, A. M.J., 2022. First things first: how to elicit the initial program theory for a realist evaluation of complex integrated care programs. Milbank Q. 100 (1), 151–189. https://doi.org/10.1111/1468-0009.12543.

Sorensen, S.S., Jensen, M.B., Pedersen, K.M., Ehlers, L., 2018. Examining the heterogeneity and cost effectiveness of a complex intervention by segmentation of patients with chronic obstructive pulmonary disease. Value Health 21 (2), 239–247. https://doi.org/10.1016/j.jval.2017.08.001.

Squires, H., Chilcott, J., Akehurst, R., Burr, J., 2012. Guidance for modellers for developing public health economic models. Lancet 380, S68.

Squires, H., Kelly, M.P., Gilbert, N., Sniehotta, F., Purshouse, R.C., Garcia, L., Breeze, P., Brennan, A., Gardner, B., Bright, S., 2024. The PHEM-B toolbox of methods for incorporating the influences on behaviour into public health economic models. BMC Public Health 24 (1), 1–25.

Stalmeijer, R.E., Brown, M.E., O'Brien, B.C., 2024. How to discuss transferability of qualitative research in health professions education. Clin. Teach. 21 (6), e13762.

Swank, M.F., Brennan, L.K., Gentry, D., Kemner, A.L., 2015. Using frameworks to diagram value in complex policy and environmental interventions to prevent childhood obesity. J. Publ. Health Manag. Pract. 21 (Suppl. 3), S116–S120. https://doi.org/10.1097/PHH.0000000000000210.

Taylor, C.N., Mackay, M., Perkins, H.C., 2021. Social impact assessment and (realist) evaluation: meeting of the methods. Impact Assess. Proj. Apprais. 39 (6), 450–462.

Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., Moher, D., Peters, M.D., Horsley, T., Weeks, L., 2018. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann. Intern. Med. 169 (7), 467–473.

Törnberg, P., 2018. Complex realist economics: toward an ontology for an interested pluralism. Rev. Soc. Econ. 76 (4), 509–534. https://doi.org/10.1080/00346764.2018.1480796.

Verguet, S., Kim, J.J., Jamison, D.T., 2016. Extended cost-effectiveness analysis for health policy assessment: a tutorial. Pharmacoeconomics 34 (9), 913–923.

Walker, S., Fox, A., Altunkaya, J., Colbourn, T., Drummond, M., Griffin, S., Gutacker, N., Revill, P., Sculpher, M., 2022. Program evaluation of Population- and system-level policies: evidence for decision making. Med. Decis. Mak. 42 (1), 17–27. https://doi.org/10.1177/0272989X211016427.

Walker, S., Griffin, S., Asaria, M., Tsuchiya, A., Sculpher, M., 2019. Striving for a societal perspective: a framework for economic evaluations when costs and effects fall on multiple sectors and decision makers. Appl. Health Econ. Health Pol. 17 (5), 577–590. https://doi.org/10.1007/s40258-019-00481-8.

Walsh, N.E., Berry, A., Halls, S., Thomas, R., Stott, H., Liddiard, C., Anchors, Z., Cramp, F., Cupples, M.E., Williams, P., 2024. Clinical and cost-effectiveness of first contact physiotherapy for musculoskeletal disorders in primary care: the FRONTIER, mixed method realist evaluation. Health Social Care Deliv. Res. 12 (49).

Weatherly, H., Drummond, M., Claxton, K., Cookson, R., Ferguson, B., Godfrey, C., Rice, N., Sculpher, M., Sowden, A., 2009. Methods for assessing the cost-effectiveness of public health interventions: key challenges and recommendations. Health Policy 93 (2–3), 85–92. https://doi.org/10.1016/j.healthpol.2009.07.012.

Weatherly, H.L.A., Neves De Faria, R.I., Van den Berg, B., Sculpher, M.J., O'Neill, P., Nolan, K., Glanville, J., Isojarvi, J., Baragula, E., Edwards, M., 2017. Scoping Review on Social Care Economic Evaluation Methods.

Welch, V.A., Tricco, A.C., 2016. Reporting guidelines for realist evaluations seek to improve clarity and transparency. BMC Med. 14, 1–2.

Westhorp, G., 2014. Realist impact evaluation: an introduction. www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9138.pdf.

Westhorp, G., Feeny, S., 2024. Using surveys in realist evaluation. Eval. J. Australas., 1035719X241292083

Whitaker, R., Hendry, M., Aslam, R., Booth, A., Carter, B., Charles, J.M., Craine, N., Tudor Edwards, R., Noyes, J., Ives Ntambwe, L., Pasterfield, D., Rycroft-Malone, J., Williams, N., 2016. Intervention now to eliminate repeat unintended pregnancy in teenagers (INTERRUPT): a systematic review of intervention effectiveness and cost-effectiveness, and qualitative and realist synthesis of implementation factors and user engagement. Health Technol. Assess. 20 (16), 1–214. https://doi.org/10.3310/hta20160.

Wildman, J., Wildman, J.M., 2019. Combining health and outcomes beyond health in complex evaluations of complex interventions: suggestions for economic evaluation. Value Health 22 (5), 511–517. https://doi.org/10.1016/j.jval.2019.01.002.

Williams, M., 2018. Making up Mechanisms in Realist Research. Doing realist research, pp. 25–40.

Willis, C., Saul, J., Bitz, J., Pompu, K., Best, A., Jackson, B., 2014. Improving organizational capacity to address health literacy in public health: a rapid realist review. Public Health 128 (6), 515–524.

Witter, S., Boukhalfa, C., Cresswell, J.A., Daou, Z., Filippi, V., Ganaba, R., Goufodji, S., Lange, I.L., Marchal, B., Richard, F., team, F.E., 2016. Cost and impact of policies to remove and reduce fees for obstetric care in Benin, Burkina Faso, Mali and Morocco. Int. J. Equity Health 15 (1), 123. https://doi.org/10.1186/s12939-016-0412-y.

Wong, G., 2018. Realist reviews in health policy and systems research. In: Evidence Synthesis for Health Policy and Systems: a Methods Guide. World Health Organization.

Wong, G., Westhorp, G., Manzano, A., Greenhalgh, J., Jagosh, J., Greenhalgh, T., 2016. RAMESES II reporting standards for realist evaluations. BMC Med. 14 (1), 1–18.