Quality standards for realist syntheses and meta-narrative reviews

Geoff Wong[1], Trish Greenhalgh[1], Gill Westhrop[2] and Ray Pawson[3]

[1] Centre for Primary Care and Public Health, Queen Mary University of London, 58 Turner Street, London E1 2AB, UK

[2] Community Matters, P.O. Box 443, Mount Torrens, SA 5244, Australia

[3] Department of Social Research Methodology, University of Leeds, Leeds LS2 9JT, UK

Corresponding author
Dr Geoff Wong
Senior Lecturer in Primary Care
Global Health Innovation and Policy Unit
Centre for Primary Care and Public Health
Blizard Institute
Barts and The London School of Medicine and Dentistry
Queen Mary, University of London
Yvonne Carter Building
58 Turner Street
London E1 2AB
(Tel):   +44 (0)20 7882 2483
(Fax):   +44 (0)20 7882 2200
(Email): grckwong@gmail.com

Introduction

Our work in the field of realist syntheses and/or meta narrative reviews clearly indicated to us that guidance on what might be considered as high quality in the execution of these two review methods were needed. We were funded as part of the RAMESES (**R**ealist **A**nd **M**eta-narrative **E**vidence **S**yntheses: **E**volving **S**tandards) project to produce these quality standards.

In these notes, we explain how we have set about producing the quality standards for:
1) Researchers and peer-reviewers using these methods
2) Funders/commissioners of research

Whilst the core components of the quality standards we have developed are the same for each of the two 'versions' listed above, we have adapted each one in an attempt to make them more focussed and useful for the intended users.

As part of the RAMESES project we produced reporting standards for realist syntheses (1;2) and meta-narrative reviews (3;4). These quality standards are intended primarily to provide researchers, peer-reviewers or funders/commissioners of research with guidance on what might constitute methodological rigour when a realist synthesis or meta-narrative review is planned or has been undertaken. When judgements about the *quality of reporting* are needed, then the following documents should be used:

1) RAMESES publication standards: realist syntheses
> http://www.biomedcentral.com/1741-7015/11/21
> or
> http://onlinelibrary.wiley.com/doi/10.1111/jan.12095/pdf

2) RAMESES publication standards: meta-narrative reviews
> http://www.biomedcentral.com/1741-7015/11/20
> or
> http://onlinelibrary.wiley.com/doi/10.1111/jan.12092/pdf

These quality standards are not training materials. If researchers wish to learn about how to undertake these types of reviews, other resources exist. As part of the RAMESES project we also developed training materials for realist syntheses and meta-narrative reviews. The training materials we developed focus specifically on the main review processes that reviewers have found the most challenging. However, within these training materials we have provided advice on the resources reviewers may wish to consult when learning to undertake these review types.

The training materials for realist syntheses and meta-narrative reviews are freely available online at:
http://www.ramesesproject.org/index.php?pr=Project_outputs#train

Finally, in the spirit of the RAMESES project, we have developed these quality standards with the expectation that they are likely to need to be updated and revised as more and more realist synthesis and meta-narrative reviews are undertaken and methodological lessons learnt. These quality standards thus act more as a starting point rather than 'rules written in stone'. We thus

invite interested researchers to contact us directly or join the RAMESES JISCMail (www.jiscmail.ac.uk/RAMESES) to help us in improving these standards.

Development methods

These quality standards have been developed as part of the RAMESES project (5). The protocol for this project may be found online at: http://www.biomedcentral.com/1471-2288/11/115 In brief, for these quality standards we drew on the following sources of data to inform the content of the quality standards:

1) personal expertise as researchers and trainers in the field
2) comments for participants in the Delphi panels
3) feedback from participants at workshops and training sessions run by members of the project team
4) comments made on RAMESES JISCMail

The data from the sources above were channelled and collated contemporaneously and used to develop the initial quality standards for researchers using the realist or meta-narrative method. The initial drafts were circulated with the review team and they were iteratively refined for content and clarity. Box 1 provides an illustration of how we drew on the data sources to produce the quality standards using an example for realist syntheses.

Box 1: Illustration of the type data we drew on to identify the need for and develop quality standards for realist syntheses

Quality standard: Programme theories
Identification of need
As researchers and trainers in realist synthesis we had noted that there was some confusion amongst researchers about the nature, need and role of realist programme theory (theories) in realist syntheses. To develop the briefing materials and initial drafts of the reporting standards for realist syntheses, we searched for and analysed a number of published syntheses and noted that our impressions were well founded.
When we one day conference in March 2011 the topic of the nature, need and role of realist programme theory (theories) in realist syntheses emerged again. In our Delphi process we encouraged participants to provide free text comments. These closely reflected the comments we received from our one day conference.

Development of the quality criteria
We drew on our content expertise of the topic area and published methodological literature to develop the quality criteria. In addition we found that some of our Delphi panel participants provided us with clear indications that support the criteria we set. For example we suggested that a realist synthesis should develop a programme theory and one that did not was 'Inadequate'. Delphi panel participants' free text comment echoed our suggestion:
"How could identification of programme theories not be appropriate…"
"…it cannot be an RS [realist synthesis] without candidate [programme] theories."

We were also able to draw on the discussions that took place on JISCMail to support some of our criteria. For example, in under 'Excellent' we suggested that, "The relationship between the programme theory and relevant substantive theory is identified."
As illustration, a comment from JISCMail that we drew upon to support this criterion was: "In a review, one focus[es] first on what is reported but one can – and probably should, in order to produce some added value -- reflect the findings and outcomes of the study under review against the theories and/or best practice that already exist."

The Quality standards for realist syntheses and meta-narrative reviews

Quality standards for researchers using the methods and peer-reviewers

The quality standards for these user groups are set out using rubrics. By peer-reviewers here, we specifically refer to individuals who have been asked to appraise the quality of completed reviews. For each review process that requires a judgement about its quality, we have provided a brief description of why the process is important and also descriptors of criteria against which a decision about quality might be arrived at. The quality standards for:

- Realist syntheses for researchers are set out in Table 1
- Meta-narrative reviews are in Table 2.

As an illustrative example to explain how the layout of these quality standards, in the quality standard for 'Focussing reviews' for realist syntheses, this aspect of the review could be judged as being adequate if, "Attempts are made by the review team to progressively focus the review topic in a way that takes account of the priorities of the review and the realities of time and resource constraints." For this aspect of a review to be judged as 'Good' we recommend that as well as fulfilling the criteria for adequate (hence our use of the terms "Adequate plus:"), reviews would need to ensure (amongst others) that, "The focussing process is iterative."

Table 1: Quality standards for realist reviews for researchers and peer-reviewers

QUALITY STANDARDS FOR REALIST SYNTHESIS (for researchers and peer-reviewers)				
1. The research problem				
Realist synthesis is a theory-driven method that is firmly rooted in a realist philosophy of science and places particular emphasis on understanding causation and how causal mechanisms are shaped and constrained by social context. This makes it particularly suitable for reviews of certain topics and questions – for example, complex social programmes that involve human decisions and actions. A realist research question contains some or all of the elements of 'What works, how, why, for whom, to what extent and in what circumstances, in what respect and over what duration?' and applies realist logic to address the question. Above all a realist research seeks to answer the 'why?' question. Realist synthesis always has explanatory ambitions. It assumes that programme effectiveness will always be partial and conditional and seeks to improve understanding of the key contributions and caveats.				
Criterion	**Inadequate**	**Adequate**	**Good**	**Excellent**
The research topic is appropriate for a realist approach	The research topic is: • not appropriate for secondary research; and/or • does not require understanding of how and why outcomes are generated.	The research topic is appropriate for secondary research. It requires understanding of how and why outcomes are generated and why they vary across contexts.	Adequate plus: Framing of the research topic reflects a thorough understanding of a realist philosophy of science (generative causation in contexts; mechanisms operating at other levels of reality than the outcomes they generate).	Good plus: There is a coherent argument as to why a realist approach is more appropriate for the topic than other approaches, including other theory based approaches.
The research question is constructed in such a way as to be suitable for a realist synthesis	The research question is not structured to reflect the elements of realist explanation. For example, it: • only requires description; and/or • only requires a numerical aggregation of outcomes; and/or • only requires summary of processes; and/or • specifies methods that are inadequate to generate realist understanding (e.g. 'a thematic analysis of …')	The research question includes a focus on how and why the intervention, or programme (or similar classes of interventions or programmes - where relevant) generates its outcomes, and contains at least some of the additional elements, "for whom, in what contexts, in what respects, to what extent and over what durations".	Adequate plus: The rationale for excluding any elements of 'the realist question' from the research question is explicit. The question has a narrow enough focus to be managed within a realist review.	Good plus: The research question is a model of clarity and as simple as possible.
2. Understanding and applying the underpinning principles of realist reviews				
Realist syntheses apply realist philosophy and a realist logic of enquiry. This influences everything from the type of research question to a review's processes (e.g. the construction of a realist programme theory, search, data extraction, analysis and synthesis to recommendations).				

The key analytic process in realist review involves iterative testing and refinement of theoretically based explanations using empirical findings in data sources. The pertinence and effectiveness of each constituent idea is then tested using relevant evidence (qualitative, quantitative, comparative, administrative, and so on) from the primary literature on that class of programmes. In this testing, the ideas within a programme theory are re-cast and conceptualised in realist terms. Reviewers may draw on any appropriate analytic techniques to undertake this testing.

	Inadequate	Adequate	Good	Excellent
The review demonstrates understanding and application of realist philosophy and realist logic which underpins a realist analysis.	Significant misunderstandings of realist philosophy and/or logic of analysis are evident. Common examples include: • programme/intervention activities or strategies are confused with mechanisms • no attempts are made to uncover mechanisms • outcomes are assumed to be caused by the programme/intervention • relationship(s) between an outcome, its causal mechanism(s) and context(s) are not explained • some theory is provided but this is not explicitly linked to outcome(s)	Some misunderstandings of realist philosophy and/or logic of analysis exist, but the overall approach is consistent enough that a recognisably realist analysis results from the process.	The review's assumptions and analytic approach are consistent with a realist philosophy at all stages of the review. Where necessary a realist programme theory is developed and tested.	Good plus: Review methods, strategies or innovations used to address problems or difficulties within the review are consistent with a realist philosophy of science.

3. Focussing the review

Because a realist review may generate a large number of avenues that might be explored and explained, and because resources and timescale are invariably finite, it may be necessary to 'contain' a review by progressively focusing both its breadth (how wide an area?) and depth (how much detail?). This important process needs to be considered from the start and may involve iterative rounds of discussion and negotiation with (for example) content experts, funders and/or users. It is typical and legitimate for the review's objectives, question and/or the breadth and depth of the review to evolve as the review progresses.

	Inadequate	Adequate	Good	Excellent
The review question is sufficiently and appropriately focussed.	The review question is too broad to be answerable within the time and resources allocated. There is no evidence that progressive focussing occurred as the review was undertaken.	Attempts are made by the review team to progressively focus the review topic in a way that takes account of the priorities of the review and the realities of time and resource constraints. Attempts are documented so that	Adequate plus: The focussing process is iterative. Commissioners of the review are involved in decision-making about focussing. Decisions made about which avenues are pursued and which	Good plus: The review team draws on external stakeholder expertise to drive the focussing process in order to achieve maximal end-user relevance.

		they can be described in publications as appropriate.	are left open for further inquiry are recorded and made available to users of the review.	

4. Constructing and refining a realist programme theory

Early in the review, the main ideas that went into the making of a class of interventions (the programme theory – which may or may not be realist in nature) are elicited. This initial programme theory sets out how and why a class of intervention is thought to 'work' to generate the outcome(s) of interest. This initial programme theory then needs to be 're-cast' in realist terms (a rough outline of the contexts in which, populations for which, and main mechanisms by which, particular outcomes are expected to be achieved.) This initial tentative theory will be progressively refined over the course of the review.

	Inadequate	Adequate	Good	Excellent
An initial realist programme theory is identified and developed.	A realist programme theory is not offered or; A program theory is offered but is not converted to a realist program theory at any stage of the review.	An initial program theory is identified and described in realist terms (that is, in terms of the relationship between contexts, mechanisms and outcomes). The refined theory is consistent with the evidence provided.	Adequate plus: An initial realist programme theory is set out at the outset. The theory is refined iteratively as the review progresses.	Good plus: The relationship between the programme theory and relevant substantive theory is identified. Implications of the final theory for practice, and for refinements to substantive theory where appropriate, are described. The final realist program theory comprises multiple context-mechanism-outcome configurations (describing the ways different mechanisms fire in different contexts to generate different outcomes) and an explanation of the pattern of CMOs.

5. Developing a search strategy

Searching in a realist review is guided by the objectives and focus of the review, and revised iteratively in the light of emerging data. Searching is directed at finding data that can be used to test theory, and may lie in a broad range of sources that may cross traditional disciplinary, programme and sector boundaries. The search phase is thus likely to involve searching for different sorts of data, or studies from different domains, with which to test different aspects of any provisional theory.

	Inadequate	Adequate	Good	Excellent
The search process is such that it would identify data to enable the review team to develop, refine and test programme theory or theories.	The search is incapable of supporting a rigorous realist review. Common errors include: • The search is driven by a	Searches are driven by the objectives and focus of the review. The search strategy is piloted and	Adequate plus: further searches are undertaken in light of greater understanding of the topic area. These searches are designed to	Good plus: the searching deliberately seeks out data from situations outside the program under study where it can be

Inadequate	Adequate	Good	Excellent
methodological hierarchy of evidence (e.g. privileging RCTs) rather than the need to identify data to develop, refine or test program theory/ies • The search process is not informed by the objectives and focus of the review • The database(s) selected are narrow in the subject matter that they contain (e.g. limited to specific topics rather than extending to social science, psychology etc.) • Searching is undertaken once only at the outset of the review and there is no iterative component	refined to check that it is fit for purpose. Documents are sought from a wide range of sources which are likely to contain relevant data for theory development, refinement and testing. There is no restriction on the study or documentation type that is searched for.	find additional data that would enable further theory development, refinement or testing.	reasonably inferred that the same mechanisms(s) might be in operation.

6.   Selection and appraisal of documents

Realist review requires a series of judgements about the relevance and robustness of particular data for the purposes of answering specific questions within the overall review question.

An appraisal of the contribution of any section of data (within a document) should be made on two criteria:
- *Relevance* – whether it can contribute to theory building and/or testing; and
- *Rigour* – whether the method used to generate that particular piece of data is credible and trustworthy.

The selection and appraisal stage may need to run in parallel with the analysis stage.

	Inadequate	Adequate	Good	Excellent
The selection and appraisal process ensures that sources relevant to the review containing material of sufficient rigour to be included are identified. In particular, the sources identified allow the reviewers to make sense of the topic area; to develop, refine and test theories; and to support inferences about mechanisms.	The selection and appraisal process does not support a rigorous and complete realist review. For example: • Selection is overly driven by methodological hierarchies (e.g. the restriction of the sources to RCTs to the exclusion of other forms of evidence) • Sources are appraised using a technical checklist for a particular method (e.g. assessment of quality for an	Selection of a document for inclusion into the review is based on what it can contribute to the process of theory development, refinement and/or testing (i.e. relevance). Appraisals of rigour judge the plausibility and coherence of the method used to generate data.	Adequate plus: During the appraisal process limitations of the method used to generate data are identified and taken into consideration during analysis and synthesis.	Good plus: Selection and appraisal demonstrate sophisticated judgements of relevance and rigour within the domain.

	RCT) rather than by making a defensible judgement on the relevance and rigour of the source • Selection and appraisal processes are overly restrictive and exclude materials that may be useful for a realist analysis • Selection and appraisal processes are not sensitive enough to exclude irrelevant materials			

7. Data extraction

In a review, data extraction assists analysis and synthesis. Of particular interest to the realist reviewer are data that support the use of realist logic to answer the review's question(s) – e.g. data on context, mechanisms, and outcome configurations, demi-regularities, middle-range and/or programme theories.

	Inadequate	Adequate	Good	Excellent
The data extraction process captures the necessary data to enable a realist review.	The data extraction process does not capture the necessary data to enable a realist review. For example: • Data extraction is undertaken mechanically and with no attention to how the data informs the review • No or very limited piloting has been undertaken to test aspects of the data extraction process and improve it	Data extraction focuses on identification and elucidation of context-mechanism outcome configurations and refinement of program theory. Piloting and refinement of the data extraction process has been undertaken where appropriate. Quality control processes are in place to check that all review team members apply common processes and standards in data extraction.	Adequate plus: Data extraction processes support later processes of analysis (e.g. by organising data into sets relevant for later analysis). The data extracted is comprehensive enough to identify main CMO patterns.	Good plus: The data extraction process is continually refined as the review progresses, so as to capture relevant data as the review question is focussed and/or program theory is refined.

8. Reporting

Realist reviews may be reported in multiple formats – lengthy reports, summary reports, articles, websites and so on. Reports should be consistent with the publication standards for realist synthesis. (See RAMESES publication standards: Realist syntheses at: http://onlinelibrary.wiley.com/doi/10.1111/jan.12095/full or http://www.biomedcentral.com/1741-7015/11/21).

	Inadequate	Adequate	Good	Excellent
The realist synthesis is reported using the items listed in the	Key items are missing. For example	Most items reported. In particular the following items should be	All items are reported clearly and in sufficient detail for an external	Good plus: The report is well written and easy to understand.

9

RAMESES Reporting standard for realist syntheses.	• No defined research question • Limited or no reporting of the review's processes (i.e. methods used) • Limited or no explanations and justifications provided for any adaptations made on the realist review process • Insufficient detail is reported to enable readers to judge the plausibility and coherence of the findings	reported: • Rationale for review • Objectives and focus of review • All method section items (i.e. items 5 to 11 in the RAMESES publication standards: Realist syntheses)	reader to understand and to judge the methods used and the plausibility and coherence of the findings.	Additional materials are made available for external readers to investigate aspects of the review in more detail.

Table 2: Quality standards for meta-narrative reviews for researchers and peer-reviewers

QUALITY STANDARDS FOR META-NARRATIVE REVIEWS (for researchers and peer-reviewers)				
1. The research problem				
Meta-narrative review is a relatively new method of systematic review, designed for topics that have been differently conceptualized and studied by different groups of researchers. To understand the many approaches, reviewers have to consciously and reflexively step out of their own world-view, learn some new vocabulary and methods, and try to view a topic through multiple different sets of eyes. An over-arching narrative of the different perspectives, based on an increased understanding of them, is produced which highlights what different research teams might learn from one another's approaches.				
Criterion	**Inadequate**	**Adequate**	**Good**	**Excellent**
The research topic is appropriate for a meta-narrative approach	The research topic is: • not appropriate for secondary research; and/or • does not require understanding of how a topic has been conceptualised and studied differently by different groups.	The research topic is appropriate for secondary research. It would benefit from illumination of how a topic has been conceptualised and studied differently by different groups.	Adequate plus: Framing of the research topic reflects a thorough understanding of the value, importance and implications of different approaches on research practice and findings.	Good plus: There is a coherent argument as to why a meta-narrative review is more appropriate for the topic than potential alternatives.
The research question is constructed in such a way as to be suitable for a meta-narrative review	The research question is not structured to reflect the elements of meta-narrative explanation. For example, it: • requires only description; and/or • requires only a numerical aggregation of outcomes; and/or • requires only a summary of processes; and/or • specifies methods that are inadequate to generate meta-narrative understanding (e.g. 'a thematic analysis of …')	The research question includes a focus on how a topic has been conceptualised and studied differently by different groups.	Adequate plus: The research question includes an element that addresses the implications of different conceptualisations and approaches to a topic on research findings.	Good plus: The research question is a model of clarity and as simple as possible.
2. Understanding and applying the purpose and underpinning principles of meta-narrative reviews				
Meta-narrative review (which is rooted in a constructivist philosophy of science), is inspired by the work of Thomas Kuhn, who observed that science progresses in paradigms. Meta-narrative reviews often look historically at how particular research traditions or epistemic traditions have unfolded over time and shaped the 'normal science' of a topic area. The review seeks first to identify and understand as many as possible of the potentially important different research traditions which have a bearing on the topic. In the synthesis phase, by means of an over-arching narrative, the findings from these different traditions are compared and contrasted to build a rich				

picture of the topic area from multiple perspectives. The goal of meta-narrative review is sense-making of a complex (and perhaps contested) topic area. During analysis and synthesis, six guiding principles (pragmatism, pluralism, historicity, contestation, reflexivity and peer review) should be used.

Criterion	Inadequate	Adequate	Good	Excellent
The review demonstrates understanding and application of the purpose and principles underpinning a meta-narrative review.	Significant misunderstandings of purpose and principles underpinning a meta-narrative review. Common examples include: • Analysing only one paradigm/epistemic tradition • No application of the six underlying principles	Some misunderstandings of purpose and principles underpinning a meta-narrative review, but the overall approach is consistent enough that a recognisable set of distinct meta-narratives together with a higher-order synthesis of these results from the process.	The review's assumptions and analytic approach are consistent with the purpose and underpinning principles of a meta-narrative review. In particular, the philosophical position is explicitly constructivist. A sufficient range of paradigms/epistemic traditions has been included to make sense of an unfolding and complex topic area from multiple perspectives and to use contrasts between these as higher-order data.	Good plus: Review methods, strategies or innovations used to address problems or difficulties within the review are philosophically coherent and make a clear and illuminative contribution to the knowledge base on the topic area.

3. Focussing the review

A meta-narrative review asks some or all of the following questions: (1) Which research (or epistemic) traditions have considered this broad topic area?; (2) How has each tradition conceptualized the topic?; (3) What theoretical approaches and methods did they use?; (4) What are the main empirical findings?; and (5) What insights can be drawn by combining and comparing findings from different traditions?'

Because a meta-narrative review may generate a large number of avenues that might be explored and explained, and because resources and timescale are invariably finite, it may be necessary to 'contain' a review by progressively focusing both its breadth (how wide an area?) and depth (how much detail?). This important process needs to be considered from the start and may involve iterative rounds of discussion and negotiation with (for example) content experts, funders and/or users. It is typical and legitimate for the review's objectives, question and/or the breadth and depth of the review to evolve as the review progresses.

Criterion	Inadequate	Adequate	Good	Excellent
The review question is sufficiently and appropriately focussed.	The review question is too broad to be answerable within the time and resources allocated. There is no evidence that progressive focussing occurred as the review was undertaken.	Attempts were made by the review team to progressively focus the review topic in a way that takes account of the priorities of the review and the realities of time and resource constraints.	Adequate plus: There is evidence that the focussing process was iterative. Commissioners of the review were involved in decision-making about focussing. Decisions made about which avenues were pursued and which left open for further inquiry are clearly documented and made	Good plus: The review team draws on external stakeholder expertise to drive the focussing process in order to achieve maximal end-user relevance.

		available to users of the review.	

4. Scoping the literature

An important process in a meta-narrative review is to identify a sufficiently broad range of sources to be able to build as comprehensive a map as possible of research undertaken on the topic. This scoping step is used to identify in broad terms the different research traditions, situated in different literatures, which have addressed the topic of interest. Initial attempts to make sense of a topic area may involve not just informal 'browsing' of the literature but also consulting with experts and stakeholders.

Criterion	Inadequate	Adequate	Good	Excellent
The scoping of the literature has been sufficiently and appropriately undertaken	The scoping of the literature has been limited and cursory (e.g. only a single source is used – perhaps the Medline database – and/or the review has inappropriately concentrated on a single research tradition – for example 'evidence based medicine')	Attempts made to utilise a broad range of relevant sources and to build as comprehensive a map as possible of the research traditions on the topic.	Adequate plus: A coherent and through search strategy, deliberately including exploratory methods such as browsing and modified in the light of emerging findings, is used to identify research traditions.	Good plus: Systematic use is made of experts and stakeholders in identifying research traditions.

5. Developing a search strategy

Searching in a meta-narrative review is guided by the objectives and focus of the review, and revised iteratively in the light of emerging data. Searching is directed at finding sufficient data to develop and more sense of the relevant research traditions that have been identified, and may lie in a broad range of sources that may cross traditional disciplinary, programme and sector boundaries. This stage is likely to involve searching for different kinds of data in different ways.

Criterion	Inadequate	Adequate	Good	Excellent
The search process is such that it would identify data to enable the review team to develop and refine the map of seminal papers and primary research studies.	The search is incapable of supporting the development of a rigorous meta-narrative review. Errors may include: • The search is driven by a methodological hierarchy of evidence (e.g. privileging RCTs) rather than the need to identify the range of research paradigms (concepts, theories, methods and instruments) that have been brought to bear on a topic • The search process is not informed by the objectives and focus of the review • The database(s) selected are narrow in the subject matter that they contain (e.g. limited	Searches are driven by the objectives and focus of the review and are piloted and refined to check that they are fit for purpose. Documents are sought from wide range of sources which are likely to contain relevant data on research traditions. There is no predefined restriction on the study or documentation type that is searched for	Adequate plus: further searches are undertaken in light of greater understanding of the topic area, particularly through the use of citation-tracking of seminal papers. These searches are designed to find additional data that would allow greater sense to be made of component research traditions and/or draw higher order insights from contrasts between traditions.	Good plus: The search reflects a high degree of scholarly insight into the key research traditions of the review.

	to biomedical topics and approaches rather than extending to social science, psychology etc.) • Searching is undertaken once only at the outset of the review and there is no iterative component			

6. Selection and appraisal of documents

Meta-narrative review is not a technical process, rather, it is a process of sense-making of the literature, selecting and combining data from primary sources to produce an account of how a research tradition unfolded and why, and then (in the synthesis phase) comparing and contrasting findings from these different traditions to build a rich picture of the topic area from multiple perspectives. This process requires a series of judgements about the unfolding of research in particular traditions, and about the relevance and robustness of particular data within that tradition.

Meta-narrative review takes its quality criteria from the traditions included in the review. Studies in these separate traditions should be appraised using the quality criteria that a competent peer-reviewer in that tradition would choose to use.

The description of the selection and appraisal process should be sufficiently detailed to enable a reader to judge how likely it is that researchers inadvertently excluded data that may have significantly altered the findings of the review.

Criterion	Inadequate	Adequate	Good	Excellent
The selection and appraisal process ensures that sources relevant to the review containing material likely to help identify, develop and refine understanding of research traditions are included.	The selection and appraisal process does not support a rigorous and complete meta-narrative review. For example: • Selection is overly driven by methodological hierarchies (in particular the restriction of the sources to RCTs to the exclusion of other forms of evidence) • Sources are appraised using a technical checklist focused on methodological procedure rather than by making a defensible judgement on the contribution that a source might make. • Selection and appraisal processes are overly restrictive and exclude materials that may help sense-making of research	Selection of a document for inclusion into the review is based on what it can contribute to making sense of research traditions. All the key high-quality sources are identified and included in the review and the poor-quality ones accurately excluded.	Adequate plus: During the appraisal process studies in the separate traditions are appraised competently using the quality criteria acceptable to that tradition.	Good plus: The judgements made when appraising papers are a model of good scholarship in the relevant tradition.

	traditions. • Selection and appraisal processes are not sensitive enough to exclude irrelevant materials			

7. Data extraction

In a review, data extraction assists analysis and synthesis. Of particular interest to the meta-narrative reviewer are data elements that would contribute to constructing a story of how research on a topic unfolded over time in a particular tradition.

Criterion	Inadequate	Adequate	Good	Excellent
The data extraction process captures the necessary data to enable a meta-narrative review	The data extraction process does not capture the necessary data to enable a meta-narrative review. For example: • Data extraction is undertaken mechanically and with no attention to how the data informs the review • No or very limited piloting is undertaken to test aspects of the data extraction process and improve it	Data extraction focuses on identification and elucidation of data that informs how research on a topic unfolded over time in a particular tradition. Piloting and refinement of the data extraction process is undertaken where appropriate. Quality control processes are in place to check that all review team members apply common processes and standards in data extraction.	Adequate plus: Data extraction processes support later processes of analysis (e.g. by organising data into sets relevant for later analysis). The data extracted are comprehensive enough to identify important topics that concern a research tradition, for example: • upstream (antecedent) traditions from which these emerged; background philosophical assumptions; • research questions and how they were framed; • key conceptual and theoretical issues; • preferred methodologies, study designs, and quality criteria; • key actors and events in the unfolding of the tradition; • landmark empirical or theoretical studies; • significant findings and how these shaped subsequent work; and • key debates and areas	Good plus: The data extraction process is continually refined as the review progresses, so as to capture relevant data as the review question is focussed and/or research traditions identified and elucidated.

| | | | of dispute within the tradition, including links with or breaches from other traditions. | |

8. Synthesis phase

Having identified the individual meta-narratives, the next phase in a meta-narrative review is to compare and contrast these to generate higher-order data (e.g. to identify and explain 'conflicting' findings).

Criterion	Inadequate	Adequate	Good	Excellent
The meta-narrative should include a synthesis phase where philosophical, conceptual, methodological and empirical differences between traditions are discussed and explained.	The synthesis phase is missing or fails to engage with the underlying philosophical, conceptual or theoretical contrasts between traditions.	Some attempt is made to show how different groups of researchers produced different findings as a result of different philosophical assumptions, different ways of conceptualising the topic, different theoretical explanations or different study designs and methods.	Adequate plus: The contrasting accounts of different traditions are synthesised in a way that generates robust higher-order data (for example, about the contestation between different research storylines at policy level).	Good plus: The review generates additional philosophical, conceptual, theoretical or methodological insights that inform innovations in research.

9. Reporting

Meta-narrative reviews may be reported in multiple formats – lengthy reports, summary reports, articles, websites and so on. Reports should be consistent with the publication standards for meta-narrative reviews. (See RAMESES publication standards: meta-narrative reviews at: http://onlinelibrary.wiley.com/doi/10.1111/jan.12092/full or http://www.biomedcentral.com/1741-7015/11/20).

Criterion	Inadequate	Adequate	Good	Excellent
The meta-narrative review is reported using the items listed in the relevant RAMESES Reporting standard	Key items are missing. For example • No defined research question • Limited or no reporting of the review's processes (i.e. methods used) • Limited or no explanations and justifications provided for any adaptations made on the meta-narrative review process • Insufficient detail is reported to enable readers to judge the plausibility and coherence of the findings	Most items reported. In particular the following items should be reported: • Rationale for review • Objectives and focus of review • All method section items (i.e. items 5 to 12 in the RAMESES publication standards: meta-narrative reviews)	All items are reported clearly and in sufficient detail for an external reader to understand and to judge the methods used and the plausibility and coherence of the findings.	Good plus: The report is well written and easy to understand. Additional materials are made available for external readers to investigate aspects of the review in more detail.

Quality standards for funders/commissioners of research

As more and more realist syntheses and meta-narrative reviews are being funded/commissioned, peer-reviewers at this stage need to make judgements on two broad areas – proposed review processes and methodological expertise. We appreciate that many funding bodies and commissioners will already have processes in place to guide the peer-reviewers they appoint. As such we see this guidance we have produced not as replacement for, but as supplementation to any existing organisational peer-review processes and guidance.

The quality standards for:

- Realist syntheses for funders/commissioners of research are set out in Table 3
- Meta-narrative reviews are in Table 4.

These have been abridged and adapted from their respective counterparts in Tables 1 and 2 to better suit the needs of this user group.

Table 3: Quality standards for realist reviews for funders/commissioners of research

QUALITY STANDARDS FOR REALIST SYNTHESIS (for funders/commissioners of research)

1. The research problem

Criterion	Inadequate	Adequate	Good	Excellent
Is the research topic is appropriate for a realist approach?	Research topic: • Is not appropriate for secondary research; and/or • Does not require understanding of how and why outcomes are generated.	Research topic: • is appropriate for secondary research. • Requires understanding of how and why outcomes are generated and why they vary across contexts.	Adequate plus: Framing of the research topic reflects a thorough understanding of a realist philosophy of science.	Good plus: There is a coherent argument as to why a realist approach is more appropriate for the topic than other approaches.
Is the research question is constructed in such a way as to be suitable for a realist synthesis?	The research question is not structured to reflect the elements of realist explanation.	The research question includes a focus on how and why the intervention, or programme generates its outcomes, and contains at least some of the additional elements, "for whom, in what contexts, in what respects, to what extent and over what durations".	Adequate plus: • The rationale for excluding any elements of 'the realist question' from the research question is explicit. • The question has a narrow enough focus to be managed within a realist review.	Good plus: The research question is a model of clarity and as simple as possible.

2. Understanding and applying the underpinning principles of realist reviews

Criterion	Inadequate	Adequate	Good	Excellent
Does the review team demonstrate understanding and application of realist philosophy and realist logic which underpins a realist analysis?	Significant misunderstandings of realist philosophy and/or logic of analysis are evident.	Some misunderstandings of realist philosophy and/or logic of analysis exist, but the overall approach is consistent enough that a recognisably realist analysis results from the process.	• The review's assumptions and analytic approach are consistent with a realist philosophy at all stages of the review. • Where necessary a realist programme theory is developed and tested.	Good plus: Proposed review methods, strategies or innovations planned to address problems or difficulties within the review are consistent with a realist philosophy of science.

3. Focussing the review

Criterion	Inadequate	Adequate	Good	Excellent

Is, or will, the review question be sufficiently and appropriately focussed?	• The review question is too broad to be answerable within the time and resources allocated. • There is no evidence that progressive focussing will occur as the review progresses.	Process proposed enables the review team to progressively focus the review topic in a way that takes account of the priorities of the review and the realities of time and resource constraints.	Adequate plus: • The focussing process is iterative. • Commissioners of the review are involved in decision-making about focussing.	Good plus: The review team draws on external stakeholder expertise to drive the focussing process in order to achieve maximal end-user relevance.

4. Constructing and refining a realist programme theory

Criterion	Inadequate	Adequate	Good	Excellent
Does the review team plan to identify, develop and refine their initial realist programme theory?	There are no plans to identify, develop and refine a realist programme theory.	There are plans to identify, develop and refine a realist programme theory.	Adequate plus: The initial realist programme theory is set out at the outset and will be refined iteratively as the review team's understanding of the topic grows.	Good plus – there are plans to: • Identify and explain the relationship between the programme theory and relevant substantive theory. • Draw on, where necessary, external expertise to develop their programme theory.

5. Developing a search strategy

Criterion	Inadequate	Adequate	Good	Excellent
Is the proposed search process such that it would identify data to enable the review team to develop, refine and test programme theory or theories?	The search is incapable of supporting a rigorous realist review.	The proposed searches will: • Be driven by the objectives and focus of the review. • Be piloted and refined. • Seek out documents from wide range of sources likely to contain relevant data. • Not be restricted by study or documentation type.	Adequate plus: Further searches will be undertaken in light of greater understanding of the topic area.	Good plus The searching will deliberately seeks out data from situations where it can be reasonably inferred that the same mechanisms(s) might be in operation.

6. Selection and appraisal of documents

Criterion	Inadequate	Adequate	Good	Excellent
Will the selection and appraisal process ensure that documents of **relevance** to the review containing material of sufficient **rigour** to be included are	The proposed selection and appraisal process does not support a rigorous and complete realist review	Selection of a document for inclusion will be based on: • Relevance – i.e. what it can contribute to the process of theory development,	Adequate plus: During the appraisal process limitations of the method used to generate data will be identified and taken into consideration during	As for 'Good'

| identified? | | refinement and/or testing.
• Rigour – judgements will be made based on the plausibility and coherence of the method used to generate data. | analysis and synthesis. | |

7. Data extraction

Criterion	Inadequate	Adequate	Good	Excellent
Will the data extraction process capture the necessary data to enable a realist review?	• The data extraction process does not capture the necessary data to enable a realist review. • No piloting of the data extraction process is planned.	Data extraction processes will: • Focus on identification and elucidation of context-mechanism outcome configurations and refinement of program theory. • Be piloted and refined where appropriate. • Include quality control processes to ensure uniformity of processes and standards.	Adequate plus: Data extraction processes will: • Support later processes of analysis (e.g. by organising data into sets relevant for later analysis). • Be comprehensive enough to identify main CMO patterns.	Good plus: There are plans to continually refine the data extraction process as the review progresses, so as to capture relevant data as the review question is focussed and/or program theory is refined.

8. Reporting

Criterion	Inadequate	Adequate	Good	Excellent
Will the review team use the items listed in the RAMESES Reporting standard for realist syntheses when reporting their realist synthesis?	No information provided	RAMESES Reporting standard for realist syntheses will be used for reporting.	Adequate plus: Firm commitment made to adhere to all items within the RAMESES Reporting standard for realist syntheses.	As for 'Good'

Table 6: Quality standards for meta-narrative reviews for funders/commissioners of research

QUALITY STANDARDS FOR META-NARRATIVE REVIEWS (for funders/commissioners of research)

1. The research problem

Criterion	Inadequate	Adequate	Good	Excellent
Is the research topic is appropriate for a meta-narrative approach?	Research topic: • Is not appropriate for secondary research; and/or • Does not require understanding of how a topic has been conceptualised and studied differently by different groups.	Research topic: • Is appropriate for secondary research. • Would benefit from illumination of how a topic has been conceptualised and studied differently by different groups.	Adequate plus: Framing of the research topic reflects a thorough understanding of the value, importance and implications of different approaches on research practice and findings.	Good plus: There is a coherent argument as to why a meta-narrative review is more appropriate for the topic than potential alternatives.
Is the research question is constructed in such a way as to be suitable for a meta-narrative review?	• The research question is not structured to reflect the elements of meta-narrative explanation.	The research question includes a focus on how a topic has been conceptualised and studied differently by different groups.	Adequate plus: The research question includes an element that addresses the implications of different conceptualisations and approaches to a topic on research findings.	Good plus: The research question is a model of clarity and as simple as possible.

2. Understanding and applying the purpose and underpinning principles of meta-narrative reviews

Criterion	Inadequate	Adequate	Good	Excellent
Does the review team demonstrate an understanding and application of the purpose and principles underpinning a meta-narrative review?	Significant misunderstandings of purpose and principles underpinning a meta-narrative review.	Some misunderstandings of purpose and principles underpinning a meta-narrative review, but the overall planned approach is consistent enough that a recognisable set of distinct meta-narratives together with a higher-order synthesis of these is likely to results from the process.	• The review's assumptions and planned analytic approach are consistent with the purpose and underpinning principles of a meta-narrative review. • The philosophical position is explicitly constructivist. • A sufficient range of paradigms/epistemic traditions is likely to be included for sense-making and use made of contrasts between these as higher-order data.	Good plus: Review methods, strategies or innovations planned to address problems or difficulties within the review are philosophically coherent and make a clear and illuminative contribution to the knowledge base on the topic area.

3. Focussing the review

Criterion	Inadequate	Adequate	Good	Excellent
Is, or will, the review question be sufficiently and appropriately focussed?	• The review question is too broad to be answerable within the time and resources allocated. • There is no evidence that progressive focussing will occur as the review progresses.	Attempts will be made by the review team to progressively focus the review topic in a way that takes account of the priorities of the review and the realities of time and resource constraints.	Adequate plus: • The focussing process will be iterative and reflexive. • Commissioners of the review will be involved in decision-making about focussing.	Good plus: The review team will draw on external stakeholder expertise to drive the focussing process in order to achieve maximal end-user relevance.

4. Scoping the literature

Criterion	Inadequate	Adequate	Good	Excellent
Has sufficient and appropriate scoping of the literature been planned?	The planned scoping of the literature appears to be limited and cursory.	Attempts will be made to utilise a broad range of relevant sources and to build as comprehensive a map as possible of the research traditions on the topic.	Adequate plus: A coherent and through search strategy will be used, deliberately including exploratory methods such as browsing and will be modified in the light of emerging findings.	Good plus: Systematic use will be made of experts and stakeholders in identifying research/epistemic traditions.

5. Developing a search strategy

Criterion	Inadequate	Adequate	Good	Excellent
Is the proposed search process such that it would identify data to enable the review team to develop and refine the map of seminal papers and primary research studies?	The planned search is incapable of supporting the development of a rigorous meta-narrative review.	The proposed searches will: • Be driven by the objectives and focus of the review. • Be piloted and refined. • Seek out documents from a wide range of sources likely to contain relevant data on research traditions. • Not be restricted by study or documentation type.	Adequate plus: Further searches will be undertaken in light of greater understanding of the topic area, particularly through the use of citation-tracking of seminal papers.	As for 'Good'

6. Selection and appraisal of documents

Criterion	Inadequate	Adequate	Good	Excellent
Will the selection and appraisal process ensure that sources relevant to the review containing material likely to help identify, develop and refine understanding of research traditions be included?	• The selection and appraisal process will not support a rigorous and complete meta-narrative review.	Selection of a document for inclusion into the review will: • Be based on what it can contribute to making sense of research traditions. • Accurately include all the key high-quality sources identified and exclude the poor-quality	Adequate plus: During the appraisal process studies in the separate traditions will be appraised using the quality criteria acceptable to that tradition.	As for 'Good'

		ones.		

7. Data extraction

Criterion	Inadequate	Adequate	Good	Excellent
Will the data extraction process capture the necessary data to enable a meta-narrative review?	The data extraction process will not capture the necessary data to enable a meta-narrative review.	Data extraction processes will: • Focus on identification and elucidation of data that informs how research on a topic unfolded over time in a particular tradition. • Be Piloted and refined where appropriate. • Include quality control processes to ensure uniformity of processes and standards.	Adequate plus: Data extraction processes will: • Support later processes of analysis (e.g. by organising data into sets relevant for later analysis). • Be comprehensive enough to identify important topics that concern a research tradition.	Good plus: The data extraction process will be continually refined as the review progresses, so as to capture relevant data as the review question is focussed and/or research traditions identified and elucidated.

8. Synthesis phase

Criterion	Inadequate	Adequate	Good	Excellent
Will synthesis of the meta-narratives include discussion and explanation of the philosophical, conceptual, methodological and empirical differences between traditions?	A synthesis phase: • Is not planned, or • Is planned in such a way that it fails to engage with the underlying philosophical, conceptual or theoretical contrasts between traditions.	The planned synthesis phase will attempt to show how different groups of researchers produced different findings as a result of different philosophical assumptions, ways of conceptualising the topic, theoretical explanations or study designs and methods.	Adequate plus: Contrasting accounts of different traditions will be sought out and synthesised in a way that generates robust higher-order data.	As for 'Good'

9. Reporting

Criterion	Inadequate	Adequate	Good	Excellent
Will the review team use the items listed in the RAMESES Reporting standard for meta-narrative reviews when reporting their meta-narrative review?	No information provided	RAMESES Reporting standard for meta-narrative reviews will be used for reporting.	Adequate plus: Firm commitment made to adhere to all items within the RAMESES Reporting standard for meta-narrative reviews.	As for 'Good'

Conclusion

Guidance on what might be considered as high quality in the execution of realist syntheses and meta-narrative reviews are needed. This document set out how quality should be judged for realist syntheses and meta-narrative reviews. We have provided quality standards for: a) researchers and peer-reviewers; and b) Funders/commissioners of research. These quality standards complement the training materials and RAMESES publication standards we have developed for realist syntheses and meta-narrative reviews.

Our expectation is that these quality standards are likely to need to be updated and revised as more and more realist synthesis and meta-narrative reviews are undertaken and methodological lessons learnt. These quality standards thus act more as a starting point rather than 'rules written in stone'. We thus invite interested researchers to contact us directly or join the RAMESES JISCM@il (www.jiscmail.ac.uk/RAMESES) to help us in improving these standards.

Acknowledgement

References

(1)   Wong G, Greenhalgh T, Westhorp G, Pawson R. RAMESES publication standards: realist syntheses. BMC Medicine 2013;11:21.

(2)   Wong G, Greenhalgh T, Westhorp G, Pawson R. RAMESES publication standards: realist syntheses. Journal of Advanced Nursing 2013;69(5):1005-22.

(3)   Wong G, Greenhalgh T, Westhorp G, Pawson R. RAMESES publication standards: meta-narrative reviews. BMC Medicine 2013;11:20.

(4)   Wong G, Greenhalgh T, Westhorp G, Pawson R. RAMESES publication standards: meta-narrative reviews. Journal of Advanced Nursing 2013;69(5):987-1004.

(5)   Greenhalgh T, Wong G, Westhorp G, Pawson R. Protocol - Realist and Meta-narrative Evidence Synthesis: Evolving Standards(RAMESES). BMC Medical Research Methodology 2011;11:115.