

Producing policy relevant systematic reviews: navigating the policy-research interface

Oliver S, Bangpan M, Dickson K.

Accepted version of paper published in *Evidence and Policy*:

Oliver S, Bangpan M, Dickson K. (2017) Producing policy relevant systematic reviews: navigating the policy-research interface. *Evidence and Policy*.
<http://www.ingentaconnect.com/content/tpp/ep/pre-prints/content-ppevidpol1600048r2>

Abstract

This study employed insider research and reflective practice to investigate exchanges across the research-policy interface to understand the practice of producing policy-relevant systematic reviews. Interviewees came from 11 systematic reviews or review programmes which spanned four models of policy relevant reviews and between them provided evidence for understanding policy problems, comparing policy options, or implementing policy decisions. No review methodology was found to be uniquely appropriate for policy-relevant systematic reviews. It was the mutual engagement across the research-policy interface that made the reviews policy relevant. This involved thinking about the issues and seeing them from multiple viewpoints to identify and shape questions; this prompted implicit or explicit value driven debates. The intellectual work to shape a policy-relevant systematic review is an iterative, collective endeavour that requires partners from either side of the policy-research interface to engage with the unfamiliar, listen, challenge and co-construct questions and answers.

Background

Policy-relevant systematic reviews 'present findings clearly for policy audiences to: illuminate policy problems; challenge or develop policy assumptions; or offer evidence about the impact or implementation of policy options; and take into account diversity of people and contexts (Oliver and Dickson 2016). Evidence informed policy making depends on the capacity of policy makers to use such reviews and the capacity of researchers to produce them. That each of these capacities needs the other is apparent for research about both health (Cooke 2005) and international development (Bates et al 2006); and has been shown to be applicable to the production of systematic reviews (Oliver et al 2015a). These capacities need to be matched by timeliness and relevance of evidence to inform policy dilemmas (Oliver K et al 2014; Whitty 2015). Our earlier study of institutional mechanisms to produce policy relevant systematic reviews identified consensus development methods or knowledge brokers (skilled communicators bringing extensive experience of the worlds of policy and research) as potential solutions for navigating the policy-research interface as they both help with clarifying key concepts that are pertinent to policy problems in order for them to be addressed by systematic reviews (Oliver and Dickson 2016). The current study investigates in more detail exchanges across the research-policy interface to understand the practice of producing policy-relevant systematic reviews.

Methods

When reporting our initial investigations of systematic reviews at the research policy interface (Oliver and Dickson 2016) we acknowledged our direct experience of working in and with policy and research organisations, and systematic reviewing. In this subsequent phase of our research, in addressing detailed processes of making systematic reviews policy relevant, we now recognise our investigations as 'insider research', where the researcher has a direct involvement or connection with the research setting (Robson 2002), even over many years, as in 'deep insider' research (Edwards 2002). The advantages are easier access and understanding of the context, cultures and history. Such research can also be constrained by researchers occupying two roles (interviewer and colleague with formal or informal relationships). In conducting this study, we protected our working relationships by inviting interviewees to comment on how we incorporated their interview data (and subsequent thoughts) into our findings, and amending the text, when necessary, to accurately reflect their expertise and insights before wider circulation.

What the insider research literature does not address is how the researcher role may influence not only our relationships with colleagues, but also how we think about our routine work through the research encouraging a more critical engagement with familiar practices.

The issues raised by insider interviews during the course of this research prompted fresh thinking during the course of our day-to-day work. In other words, conducting this study prompted us to become more reflective about our current systematic review practice. Rather than ignoring this additional critical thinking we have incorporated it into our analysis, making this a report of both insider research and reflective practice. We have tried to distinguish clearly where 'data' and 'analysis' came from insider interviews, where they came from reflective practice, and where the two approaches to investigating policy relevant reviews influenced each other.

Sampling systematic reviews for investigation

We sought a diversity of systematic reviews/ review programmes that were initiated either side of the policy-research interface, and that between them included all four models of policy-relevant reviews identified by our earlier study (Oliver and Dickson 2016). Although the aim was to inform policy for health systems, reviews were also eligible from international development and environmental science when they also addressed health.

Insider interviews

We collected data between May 2015 and April 2016. For each systematic review (or review series) we sought the person responsible for ensuring the policy relevance. They held roles conducting reviews, commissioning reviews or providing methodological guidance. We began by approaching individuals in our own networks and in the course of the interviews they recommended others they worked with for seeking additional insights. To understand how interviewees sought to make reviews policy relevant, we adopted narrative interviewing whereby each interviewee was invited to tell the story of their review, from its origin to the current time. Some reviews were still on-going and some were recently completed. After the narrative had unfolded, sometimes with the help of further prompts for details, we asked for their views about issues that had been raised by earlier interviewees or by our own direct experience. Rather than unbiased data collection, this approach made the interviews a conversation (in this case methodological debates) amongst equals, of value to interviewers and interviewees (Oakley 1991).

We shared our draft paper with each interviewee not as 'respondent validation' but as an opportunity for them to reflect again on our conversation, and develop ideas further if they wished. An opportunity for further testing of the emerging ideas came from a meeting of reviewers and policy specialists at a conference [Cochrane Colloquium 2016]. In this way we sought their thoughtful insights and responses to issues emerging from the field, and we present the findings here as the result of co-construction of knowledge.

Reflective practice

While formally investigating how colleagues in our networks approach the preparation of systematic reviews for policy, we were also, as part of other work in 2015 and early 2016, actively engaged in

preparing such reviews ourselves, discussing with funders how to commission such reviews, supporting other research teams prepare such reviews, or supervising postgraduate dissertations. Prompts for reflection, and efforts to achieve greater clarity about working across the research-policy interface, came from discussions in the course of our own work: with commissioners about their priorities for particular reviews; with researchers seeking our advice for how to make their systematic reviews policy-relevant; with postgraduate students about their own motivations for policy relevant systematic reviews; with the audience at a student-led conference about evidence-based international development; and interdisciplinary discussions to develop procedures for delivering evidence on demand to policy teams.

Analysis

Our earlier constant comparative analysis identified institutional mechanisms grouped within higher level themes (demand and supply of systematic reviews, and mutual understanding between the producers and users of reviews) which were then framed in terms of two overlapping social worlds which shared dimensions of motivations for activities, engagement with each other, supportive structures, procedures and impact (Oliver and Dickson 2016). The new interview data collected for this paper were analysed using the same framework (figure 2). Recurrent themes were recognised from interview notes and repeated listening and partial transcriptions of sound recordings. Subsequently the detailed themes populating the framework in our original study were inspected to prompt recognition of themes overlooked in our current data. Data collection continued after the framework had been populated to check the degree of thematic saturation.

Ethics

The study was approved by the UCL Institute of Education Research Ethics Committee. Participants were informed of the study and gave their written consent. Efforts were made, in discussion with participants, to anonymise individuals. Members of organisations that are recognisable were involved in planning and advising the study in the knowledge that their organisation may be named. Interviewees checked our partial transcripts of their interview and a draft version of the report. Where data originated from other individuals and were not in the public domain, they were not used without that individual also giving informed consent to be interviewed. Insider interviewees had the option to be acknowledged publicly.

Findings

Diversity of approaches to systematic reviewing

We recruited the project leads from 11 systematic reviews or review programmes. (Their ID numbers indicate model number (first digit), review number (second digit); and, later in the text, when multiple

interviewees were engaged in a single review they are distinguished by a unique following letter.) This sample included reviews that were:

- Driven by academic interests in health systems and planned for widespread use (ID1.3);
- Driven by specific organisational or individual professional interests in making or implementing policy (ID4.1, ID4.2, ID2.1, ID4.3); or
- The result of a priority setting exercise (ID1.1, ID1.2, ID3.1, ID3.2, ID3.3, ID3.4).

Nine were registered with a review facility where experienced reviewers support those not specialised in reviewing to conduct a review. The two exceptions were either conducted by experienced reviewers independently (ID2.1) or as part of an accredited postgraduate qualification (ID4.2).

These varied in terms of the geographic location of the reviewers and the jurisdiction of the related policy makers (table 1); the substantive focus, who initiated the review and whether evidence was needed for understanding policy problems, evaluating policy options, or implementing policy decisions (Lavis 2009) (table 2). Figure 1 shows these systematic reviews between them spanning all four models for producing policy relevant systematic reviews (Oliver and Dickson 2016), which are distinguished by two key dimensions where:

- The review is produced either for a general audience as a 'public good' for enduring problems (models 1 and 3) or with context specific concerns in mind (national or local) (models 2 and 4)
- Prior consensus on key concepts and definitions is either strong (models 1 and 2) or weak (models 3 and 4).

[Insert figure 1 here]

The reviewers varied in their experience of systematic reviewing, their experience of the worlds of policy and research, and their experience of spanning the interface.

Additional evidence products within the portfolio of our own work became the focus of reflection. These included: an evidence/ gap analysis to inform the commissioning of a review; a commissioned rapid review; a policy review for selected countries; three dissertations by postgraduate students with relevant professional backgrounds; and an evidence brief. Evidence briefs are designed for particular health systems and synthesise the best available global research evidence together with relevant local data and studies to clarify policy problems, and describe what is known about the choice of policy responses and their associate implementation issues (Moat et al 2014).

How these diverse reviews were developed is described below in terms of an analytical framework (figure 1) which encompassed 'two overlapping social worlds which shared dimensions of motivations, engagement methods, structures and procedures to support the demand and supply of reviews' (Oliver and Dickson 2016).

Motivation for systematic reviews

Each review was influenced by both policy and academic motivations. Five were inspired by priority setting exercises with stakeholders. Two other systematic reviews were commissioned to inform a specific decision or project where there was an urgent need for implementation or cost saving. Two series of reviews were prompted by national interests; one of these drew on global data and the other on developing country data. One review was conducted as part of postgraduate studies; its focus related to the professional experience of the student who was employed in an organisation implementing policy [ID4.2a, ID4.2b]. One review was prompted by academic interest, a sense of enjoyment and satisfaction from contributing to the field; this academic motivation was supplemented by a link to international priorities once it was registered with a review facility [ID1.3].

The focus of some reviews was influenced by the experience of the reviewers, with some reviewers choosing to focus on the familiar rather than over extending themselves:

[In broader public policy there were lots of potentially relevant studies] and the issue is... it was beyond our expertise, and we would have to invest more time and need more expertise to go beyond health, so we are keeping it to health to match our resources – human and financial resources. [ID3.1b]

The effectiveness focus of two reviews was chosen for several reasons, including the opportunity for reviewers to develop their research skills [ID1.1a, ID3.3a].

I wanted my first review to be an effectiveness review, ideally an EPOC [Cochrane Effective Practice and Organisation of Care] review... because that is the best way to learn to do a review, most supported, most guided. [ID3.3a]

Systematic reviewers had ideological and practical motivations. One interviewee emphasised the methodological criteria for a Cochrane Review Group, implying a motivation to populate the Cochrane Library with reviews of effectiveness [ID1b]. Others felt a strong accountability towards stakeholders who had participated in an exercise that prioritised their review [ID3.3a, ID3.3b]. One of these explained:

In a country like South Africa there is a moral obligation to contribute... not necessarily to policy... but to the uplift of people... to retain a human rights perspective... Lots of researchers of my generation and older come from an anti-apartheid background... [For] my generation and above, a lot of the work only has meaning if it's applied [ID3.4]

Generally there was a strong sense of responsibility to deliver rigorous, useful research. The weight of this responsibility, and the commitment to develop methods for policy relevant reviews, was complementary [ID4.3a].

I find doing the rapid reviews, doing the service work [for specific policy customers] informs the methods work [methodological development], and vice versa. I can sleep at night by saying I'm doing the very best I can in the short space of time. [ID4.3a]

More prosaically, some systematic reviews were influenced by opportunities for academic for publications or financial consultancies [ID2.1]. One programme of reviews combined financial and academic rewards by providing a committee of academics with consultancy funds, a secretariat to support them producing a systematic review, and the opportunity to meet in comfortable surroundings to participate in intellectual debate with other leaders in their field [ID1.2]. Conversely, the lack of funding was demotivating. Slowest progress was made where funding was either absent [ID4.2a] or available only part way through the work [ID1.3].

Engagement of the policy-research worlds

Engagement with policy issues:

Systematic reviews that were commissioned for a specific decision or project required reviewers to engage with questions posed by policy makers (in government [ID4.3a, ID4.1] and in an international NGO [ID2.1]). In circumstances where

there is a specific user for a systematic review (say... for a particular programming decision) as opposed to a more broadly felt need, it's all the more important to engage with this user (if not individually, perhaps institutionally) during the process. Of course there are practical limitations – the 'user' may move on to a different role, or priorities may change. [ID4.1]

Shaping the questions, whether posed by commissioners, stakeholders or inspired by horizon scanning required thinking skills, a familiarity with the research literature and a commitment to focus on policy priorities.

How will it respond to national priorities and national goals and to the community at needs, and policy makers working on this in the region, what about challenges of implementation?... One of the policy makers said I really simply want to know how I go about developing and implementing [an intervention like this]. [It's a] straightforward request in the sense that this is important for us to really understand about going about reform. [ID3.2]

We reformulated most of the questions... we couldn't figure out how to translate [the original questions] into systematic review parameters... [and they were] not well aligned with the funder's action plan [ID2.1]

Where policy concern was initially on containing rising costs, a researcher raised concerns about such a narrow focus ignoring concomitant patient safety issues [ID4.3b]. It was a researcher's familiarity with the wider system and historical context from living within the country that prompted this lateral thinking and influenced the focus of the review. In this instance, broader knowledge was assimilated

and the review ultimately addressed safety issues alongside the original policy focus of cost containment.

Our own experience of lateral thinking came during the preparation of an evidence brief. When seeking systematic reviews to inform plans for setting up a new programme, we recognised that a systematic review on 'taskshifting' within existing services provided relevant international evidence. Similarly, when seeking systematic reviews to map the evidence and gaps about 'problematic masculinities' we found relevant evidence in a review about youth and peacebuilding. In each case, the relevant research was framed differently: as workforce management rather than programme implementation; and framing young men positively rather than negatively.

Developing an appropriate question to drive a systematic review includes identifying existing reviews.

One tool we use [to guide our discussion asks about the question], is there any previous systematic review, can it be answered by an existing study? [ID3.2]

These discussions comprised a major part of the review teams' work, either directly with policy makers or indirectly through a contract manager from the commissioning organisation who liaised with both the research team and the policy makers.

However, interest in policy relevant systematic reviews was not always initiated by the policy makers themselves [ID1.1, ID1.2, ID3.1, ID3.2, ID3.3, ID3.4].

[There is] no blue print for how to make it policy relevant, and whole idea of stakeholder engagement... People are just trying to run the country... It's our agenda how might systematic reviews help, we're making them stop to think about it a different way [ID3.3b]

Sometimes reviewers purposely looked outside academia for questions, such as horizon scanning for emerging issues [ID3.1A]. Some reached out to stakeholders to identify questions through formal priority setting exercises [ID1.1, ID1.2, ID3.1, ID3.2, ID3.3, ID3.4] or through consultations [ID2.1]. However, priority setting surveys revealed topics of stakeholder interest rather than questions suitable for a systematic review:

It was a bit of a broken telephone... Can't throw questions away, but have to translate them, they come as localised questions not systematic review questions. There is some lost in translation because you have to slightly change it. I don't know what the best way to do this.... Ask what are the problems? Just to name key areas we should be looking at? But lots of responses... from policy makers, and the public came up with things that were of interest to them... eg we have a problem with delivering cancer services in the community... is there anything you can suggest? [ID3.3b]

Engaging across and beyond disciplines: Policy driven questions did not fit neatly into academic disciplines. They needed expert input in various combinations: from clinical and legal experts, and

ethics and information science [ID4.3a]; from health policy and media experts [ID3.1A]; from systems thinkers [ID3.3a] and someone critically engaging with a familiar area.

The danger is reading into it more than is there, but when doing reviews like this, you need people who know a lot about the area to do them, [who have] worked in the area so [they] could recognise the phenomenon for what it was, [who] know what changing [a] culture is when you see it. The localism of language is critical. What is a driver, a policy driver in one context may not be seen as such in another political system... [ID4.2b]

Structures to support the supply of policy relevant reviews

Editorial support: Reviews registered with review facilities had access to their editorial and peer review processes. Interviewees raised the mismatch between policy timelines and review timelines encountered above, and the mismatch between support offered by review facilities (focused on methodology) and support needed (stakeholder engagement as well as methodology).

High expectations from policy makers and stakeholders on delivery time. [ID3.2]

It would benefit from editorial support... [but] I have to balance the added value of having editorial support and the need to deliver on time [ID2.1]

[I] would have liked efficient ways to involve stakeholders. Maybe an advisory group. Not part of our normal way of working. Should be routine like doing research. A more proactive way of involving stakeholders and linking to dissemination and implementation – not our natural way of working. I have no idea personally about setting up advisory groups, role and contribution. Guidance here and support would be helpful. [ID3.1A]

One solution suggested for attracting critical feedback to reviews not registered with review facilities was *Peerage for Science*, a free service for scientific peer review and publishing (www.peerageofscience.org/). [ID1.2]

Knowledge brokers: Our earlier study identified knowledge brokers as an important institutional mechanism to support policy relevant reviews addressing urgent local needs (Oliver and Dickson 2016). In the current study, although no interviewees used the term ‘knowledge broker’, they did describe individuals mediating interactions between the worlds of policy and research that aligned with the growing literature about knowledge brokers generally (Bornbaum et al 2015). For instance, focusing one review question involved multiple interactions with an independent intermediary from an international agency [ID4.3a]. Elsewhere a review funder liaised with policy teams and a review facility to translate policy questions into review questions [ID4.1]:

A challenge is that we need to remain receptive to [policy maker] needs and clear about expectations of what is or is not feasible [with systematic reviews] [ID4.1]

Guidance when formulating another review question took into account local knowledge of policy controversies [ID4.3b]. In another instance a reviewer brought their own experience spanning the worlds of research and policy [ID4.2a].

When working with policy teams we have encountered the funding organisation providing a knowledge broker. In these circumstances the knowledge brokers have either encouraged us to challenge policy teams' assumptions (when they were shaping questions to commission reviews), or to align our efforts with their assumptions, ideologies or goals in others (when we were conducting systematic reviews addressing questions they had developed). Maintaining our independence while working closely with policy teams was made easier when we could also call on academic peers to act as critical friends, listen to our analysis, interrogate our understanding and help us keep our intellectual distance.

Academic supervision: Postdoctoral study suited the production of a policy relevant review where a part time postgraduate student whose paid employment was in a policy organisation could address familiar policy issues. [ID4.2b]. However, this Masters in Public Health programme, like many others, only allocated a single supervisor to each student, an arrangement which offers less support than review facilities. Our own experience confirms postgraduate programmes as suitable for the production of policy relevant systematic reviews, particularly when students bring relevant professional experience.

Employment structures: Typical research contracts for university employees, aligned with external funding contracts, take time to agree but time is short when systematic reviews are commissioned for urgent, unpredictable policy decisions with short deadlines. A particularly successful rapid review unit, when deadlines are tight, employs casual staff who are paid by the hour. They are based in numerous countries in different time zones, progressing the work while the project lead is sleeping [ID4.3a]. This approach works well for the task of applying standardised tools to multiple studies, with minimal supervision, but not for building capacity to conduct more complex reviews.

That's why [for more complex reviews] it is important to retain systematic reviewers. Once experienced reviewers, they are more productive over time, and once you lose them, much of the capacity is wasted. [ID3.2]

Procedures for spanning the worlds of policy and research

Peer review: When a review is prompted by academic interest and supported by academic systems emphasising methodology, the opportunity for policy input is limited to peer review. Even here blind peer review means systematic review authors may not know if comments come from a policy perspective [ID1.3]. In contrast, when a review is prompted by policy interests, there are opportunities for input throughout the review process.

Developing review questions: Typical effectiveness reviews are framed with the mnemonic PICOT: Population, Intervention, Comparison, Outcomes and Time. In environmental science reviews, when PICOT could not be applied, an alternative was PECO: Population, Exposure (eg grazing of land), Comparison, Outcome (ID1.2). However, not all questions were amenable to this standardised framework.

Where the question doesn't easily split into PICO, or if definitions are difficult to explain, we put it at the bottom of priority list. If it isn't sufficiently narrow, we go back to stakeholders to narrow it more [ID1.2].

However, issues of interest to policy clients do not necessarily follow a PICOT structure. An alternative mnemonic was offered for policy relevant reviews, namely TOPIC – Timely, Organisation, Problem driven, Implementation and Context – because:

'Clients' have 'problems' (affecting many components of the health system) rather than 'research questions'. Issues are not only about "what works" but about "what is the problem" (or how important is it, who is affected), or "how to implement" (or how much will it costs, will it be appropriate to a given context). Translating an 'issue' into a research question to be addressed by a systematic review is not a simple task; ideally it requires an analytical framework which can 'make sense' of what may be important to ask, what other issues to consider. [ID2.1]

Discussions were often difficult and not merely intellectual when they tapped into differing drivers for what review question should be addressed or how. Researchers addressing health systems have the opportunity to embrace the complexity, or to apply review methods to a complex intervention as if it were a black box. The first approach accommodates systems thinking, the second is simpler to apply. A review of key governance issues presented a difficult choice for the lead author who was both a skilled systems thinker and a novice reviewer hoping for a manageable effectiveness review that addressed a well-defined meaningful core concept. A black box effectiveness review was expected to be empty of eligible studies, which concerned some members of the team, but not others [ID3.3a]. The co-authors repeatedly remembered their commitment to addressing issues raised by stakeholders, which were rarely reviewable questions nor well conceptualised [ID3.3a, ID3.3b]. Resolving these competing drivers took months of reading, with 'a lot of arguments', challenge and resistance within the team, and 'quite dramatic... negotiation' with mediation about the feasibility of the question. Developing the review protocol took a year, six months of which was spent struggling to get to a question. A member of the editorial team remembered this process as similar to other reviews of complex questions, although considered it further complicated by having a topic coming from consultation process, leaving the team feeling obliged to deliver something to meet that identified need [ID3.3b]. The mediator acknowledged that the lead author, who was going through the process for the first time, was surprised how long it took to go from issue to reviewable question, but considered it not that unusual [ID3.3b]. 'A stronger steer right at the start' from the editorial team could

have been helpful, but the editorial team was also learning about this particularly challenging field through the review team, and had yet to clarify a policy about how to scale reviews in new areas (by either lumping or splitting questions) that could have guided the reviewers [ID3.3b]. The mediator considered the discussions worthwhile as they provided a valuable learning process in going from policy issue to reviewable question, but hadn't anticipated how long it would take, and with the benefit of hindsight recognises that review teams would appreciate more support in managing these tensions.

We don't have mechanisms for 'containment' how do you manage inevitable anxiety about a limited time frame to deliver to funders. We don't see that as our role, our role is more technical. Although capacity strengthening at [the review facility] brings that other stuff into play, but we don't talk about that process. We don't have that bit of the review on our review guidance. 'Don't be surprised if you experience these things'. I guess because I know [the review authors] well and I'm in [their local city], that stuff just happened... Thinking doesn't appear in the protocol – the big discussion that happened isn't there. (ID3.3b)

Where reviews were not registered with a review facility, the teams managed these tensions themselves.

[With] all the team around the table... We discussed all those issues [policy interest, existing studies, our capacity]. Then confirmed the topic. Then formed the team, then the timeline. Early discussion was in the team, and with some stakeholders – we bounced [around] ideas with them. [There was] no one route to really get a question... Very different from a clinical review. Iterative process, requiring dialogue, framing, shaping until a specific review question, until it can be addressed by a systematic review. There is a drawback of too much standardising. You need flexibility and engagement until you have a good review question. For some studies it is very fast. This one longer. There were some misconceptions, a lack of clarity of [the type of intervention] and how it links to [the whole system]. [ID3.2]

Even for a rapid review of nine weeks it took

At least three weeks to decide on the research question – going back and forth quite a bit, with the librarian, calling [the knowledge broker] to discuss how we can make sense of the topic because of how the question evolved, and [go] through the PICO criteria. [ID4.3a]

Developing conceptual frameworks: The complexity of policy concerns and the process of stakeholder engagement raise challenges for developing conceptual frameworks to shape systematic reviews. Some systematic reviews about health systems sought frameworks that were available in the literature, for instance about access to medicines [ID3.4], packages for reform [ID3.2], or frameworks for regulation [ID4.3a]. More commonly, a clear conceptual framework as an early product was a first step in the review process.

A lot of [effectiveness] reviews treat 'health systems strengthening' like a drug... [An advisor said]... most valuable was mapping out the pathway [whereas, for drug reviews this step appears in papers reporting biochemical causal pathways] [ID3.3a]

Many interviewees mentioned identifying a conceptual framework to clarify or understand the focus of interest as a key step in getting started, although this is rarely considered in depth in systematic review guidance.

Different review teams made use of conceptual frameworks developed by other experts in the field.

First is to identify a framework for interaction, involve [a content expert], that totally changed our thinking about the intervention [ID3.1A]

Where data are sparse for qualitative synthesis, we use frameworks structured to organise the content. We used a classification of service users' concerns – an old framework [ID2.1]

After the review question, when starting working on scoping, and when we are trying to look at the protocol, and the framework, this is when we go and consult on the framework...does it capture what we're looking for or anything missing? [ID3.2]

One participant compared this stage in two reviews [ID3.2]. Where the review team or their stakeholders related to the issues on a daily basis, identifying or developing a framework proved relatively easy and was achieved within a week or two of recognising the issues as a priority. However, where the concept of a reform package was unfamiliar, scoping a review required much deliberation. It took another four to six weeks to develop a framework and criteria for eligible studies.

Tensions in choosing a question for the governance review mentioned above were resolved once an appropriate conceptual framework was identified which could clarify the structures for accountability through routine information systems and the potential interventions for improving them [ID3.3a].

We had half a day together and a white board. [The lead author] had prepared how [the key governance focus] is conceptualised – in different ways. We talked about it. We realised it was not possible to have a review looking at both community based mechanisms and national governance level interventions... [The lead author] went away to think about it. She then suggested [a sub-question] where she'd done work in this area [which would make an important contribution to governance. We discussed it. It made sense. [This sub-question] was on the EPOC priority list and was not already addressed... We had many iterations, getting narrow and narrower... One of the strengths is that they are interested in engaging conceptually with the material, wanted to understand how field of [governance] and [the sub system] fits into it. They did more than many reviewers would. It took a lot of time, and was very confusing initially. It will pay off eventually. When they write up they'll be able to contextualise findings appropriately... [ID3.3b]

Insider interviewees, when talking about developing questions and conceptual frameworks, repeatedly mentioned how it takes time, thinking skills, and familiarity with the issues and the relevant literature.

There is a danger of

moving very quickly to the technical from the conceptual. There is an element of serendipity about what knowledge you have access to at the time and draw in, you may have just read something, and that's anathema to the whole idea of systematic review, but in the conceptualisation of a review, it is so important to be able to think laterally. [ID3.3b]

Gauging the literature: Many interviewees commented on the nature and scale of the literature available for their review, and the implications that had for their review. Where the focus is primarily on effectiveness there is an option to broaden the substantive focus of a review in order to include more studies.

Where an area includes a well understood range of interventions and is reasonably clear conceptually, splitting [it into several reviews] may be a good approach, but this doesn't apply to many areas within [health systems research]. Lumping is a better start, like doing a scoping review. Also because there are not many studies, so burden on authors not too much. (ID1.1b)

However, broadening the substantive eligibility criteria raises the likelihood of heterogeneous studies which precludes statistical meta-analysis. (ID2.1)

Another option is to broaden the type of studies for drawing conclusions about effectiveness to include quasi-experimental studies. This approach 'won't miss any data [where experimental studies are lacking], but may get biased results' [ID1.3]. The alternative is to emphasise the internal rigour of studies for drawing conclusions about effectiveness. A rise in standards for the internal rigour of studies to be included in systematic reviews of effectiveness has not been matched by a rise in the number of such studies conducted in health systems research. Consequently updating a review resulted in less evidence being presented to readers than in the original review [ID1.1b].

Other reviewers were inherently interested in a broader literature for addressing other questions.

We wanted to look at effectiveness studies and to have idea about the size of literature to decide what we would go for, focus only on effectiveness or consider data for mixed methods. We want to make sure this is feasible, based on a scoping review we decided to extend the search to go for quantitative and qualitative study designs. Because quantitative literature is limited— two or three studies in scoping review looking at effectiveness. We are also interested in intrinsic value of qualitative literature. But we wanted to know whether it is manageable. We want to understand how things work [in qualitative studies], different strategies used. (ID3.1A)

For some teams, the goal was to review existing research rather than highlight a lack of evidence and report an empty review [ID1.1a, ID2.1, ID3.1b].

Challenges we faced were mainly to refine the question and focus of the review. So we did a scoping review to look at the lay of the land, what sorts of studies exist. We got help from contact expert... to help define the question and where to focus... We searched two databases. It was a preliminary search strategy to see what is out there. We didn't want to go for systematic review and have no idea what nature of literature, focus, what types of study designs, what research questions are out there. [ID3.1b]

[For this review we] couldn't have an empty systematic review, so worked in a patchy way until we found something. [For another review we] only found biomedical markers, not mortality and morbidity. Had to find modelling articles [and work with] covariance.[ID2.1]

Many requests from policy makers requires in large part qualitative work with only a small fraction being answered by quantitative analysis. This is challenging because there is less consensus about what constitutes a meaning nuanced qualitative synthesis [ID4.1], and fewer teams offer such reviews [responses to tender].

A systematic review seeking international studies addressing policy change, particular outside conventional practices of health systems, is likely to find little research, and that driven by research interests which may necessarily be narrowly framed to pass peer review [ID4.2b]. When facing an immature literature an alternative approach is to focus, not on research literature, but on

policy literature, with its broad frameworks, memos, committees and political drivers [which may contain] tiny fragments describing elements of practice, some of which are impact, as described by individuals' [ID4.2b]

This is possible with the help of a tool for assessing grey literature in terms of authority, accuracy, coverage, objectivity, date and significance (Tyndall 2008) [ID4.2a].

Reflecting on our own work, we have seen how this critical step of gauging the literature has implications for providing editorial support because the appropriate review method has not always been immediately apparent. In such circumstances we have avoided being prescriptive about methods, and sometimes allowed the methods to evolve iteratively, taking into account not only the literature available, but also the team's experience of research generally, and of synthesis methodology specifically.

Impact of navigating research-policy interface

Timely, relevant products: One of the desired outputs is a rapid product.

[policy makers] use knowledge they already have, so if this is only have now in the literature they will use it, especially if they are translated or communicated in effective ways. Other challenges... We realise more and more, [policy makers] want something quick, they don't have the patience to wait for a systematic review [ID3.2].'

There are tensions between this commitment to providing evidence fast and a commitment to providing evidence that is sound,

Doing the [low budget] reviews makes demands on my own time and my staff time. We have other projects to do too. We try to finalise a review in two months, but not full time. It's very tight, quite stressful. [ID2.1]

Our expertise was in doing it fast... We are asked to do many more rapid reviews for policy makers. They are very challenging... We're purists we like everything to be double checked especially for policy makers. But if we don't give them anything, may be they'd use a single paper or expert opinion... It's very stressful. I've definitely aged. [ID4.3a]

When making policy decisions, some policy makers are seeking policy focused reviews that draw on the existing literature [ID2.1] to offer solutions, interventions, strategies [ID3.2], where the reviewing process is summarised very concisely and a clear report also offers direct access to details in an appendix when required [ID4.3a]. One series of reviews presented their findings through an interactive geographical map. This allowed readers to identify studies geographically close, and to interrogate a database that coded studies in terms of their contexts [ID1.2]. When commissioning new research, policy makers are seeking comprehensive information about evidence and gaps.

When conducting our own systematic reviews we have frequently been encouraged to develop recommendations directly relevant to the funders. The later challenge was to guide reviewers in framing 'context specific, actionable messages' that were justified by the research.

It would be useful to be more conscious of this aspect [contextual variation]... [asking] what are the meaningful subgroups at the beginning, and check with practitioners at the beginning. In future it would be useful to have someone... who could specifically look at the issue of variation... while questions being finalised or during protocol development [to ask] what are the meaningful subgroups? [ID4.1]

As policy teams and knowledge brokers we have worked with have become more familiar with the principles of systematic reviewing their expectations of scientific rigour have been raised alongside their expectations for policy relevance. Tensions have resulted from policy relevant studies being excluded because their design was inadequate to answer the question; and narrative methods have met with disappointment when statistical meta-analysis was not appropriate.

Developing capacity in navigating the interface: Researchers engaged in policy relevant systematic reviews appreciate the opportunity to influence policy decisions with rigorous research [ID1.1a, ID1.1b, ID4.2b], and to have developed research methods that makes this possible [ID4.3a]. An interim step is developing policy makers' interest.

How [do] we continue to increase the demand and to sustain engagement of policy makers and stakeholders throughout the process of systematic reviews? [ID3.2]

[We have a] small group of stakeholders, because [it] is a small country, and they all know each other. We return to same stakeholders for later meetings with new projects, so they accrue experience of systematic reviews. We try to keep them engaged along the way, there is not much turnover in decision makers, we don't lose people along the way over two years, they have opportunities to chat about the draft, so they are aware of the project by release event. [ID1.2]

However, some policy interests have been too challenging for current reviewing capacity.

With limited resources and lots of questions we have to be strategic, have to acknowledge some ideas from policy makers... were too difficult for us and not wise or strategic for us. If we had a larger team, we could venture into broader topics. [ID3.2]

Discussion

Summary of findings

The diversity of reviews identified as policy relevant confirmed that no research methodology is uniquely appropriate for policy-relevant systematic reviews. It is the mutual engagement across the research-policy interface that makes the reviews policy relevant. This involves an intellectual task of thinking about the issues and seeing them from multiple viewpoints to identify and shape questions; this necessarily prompts implicit or explicit value driven debates. Spanning the policy-research interface requires engaging with unfamiliar issues, mutual challenge and tensions between technical and political issues.

Strengths and limitations of the study

This study combined the strengths of insider research with reflective practice. In doing so it offers an understanding of policy relevant systematic reviews co-constructed with leaders in the field. We conducted 12 interviews, with a total of 13 participants, related to eight systematic reviews and three series of systematic reviews. Despite the sample being very varied in terms of geography, jurisdictions of policy partners, and review topics, they are all linked indirectly through five partially overlapping international networks covering health practice, health systems, social policy, international development and environmental science. By the last interview no new themes emerged, although the last interviewee offered novel details within themes, and it is possible that participants from other unconnected networks, if interviewed, might introduce other themes. Our understanding of each domain spanning the research-policy interface (diversity, motivation, engagement, structures, procedures and impact) was supported by three or more interviewees in this study, and was consistent with, while developing further, our understanding developed from our prior studies (Oliver and Dickson 2016; Oliver et al 2015) and with reflective practice in our own work which was analysed using the same framework. We triangulated our data sources: interviewees' oral accounts of review

processes were compared with contemporaneous records (emails), and written accounts of review processes (published, conference presentations). Our interpretation of the findings was checked both with participants and through peer debriefing.

Discussion of the findings in relation to the wider literature

Gauging the literature: Although some reviews ‘map’ (Oakley et al 2005) or ‘scope’ key concepts, types of evidence and gaps (Colquhoun et al 2014), our insider interviewees talked of an informal scoping of the literature, as part of initiating a review, to gauge the nature and scale of the literature to choose the most appropriate tools to advance knowledge within the available resources.. Rather than asking whether there are sufficient studies to answer a pre-specified question, the alternative is to ask what questions is this literature capable of answering, and what approach is appropriate. In the natural sciences, just as a biologist chooses a microscope to learn about microbes, an astronomer uses a telescope to learn about stars. Similarly, in synthesis, sociologists employ thematic synthesis to reveal concepts and meta-ethnography to develop theory while epidemiologists employ statistical meta-analysis to measure difference and programme evaluators employ case studies and process evaluations to assess how services are delivered and received. To maximise learning from the available literature, instead of the question determining the review methods, a prior step in some policy relevant reviews is the literature available informing the choice of question.

Rapid, responsive reviews: With lack of timeliness and relevance being most commonly cited barriers to policy use of evidence (Oliver et al 2014) the emergence of rapid reviews has emphasised both methodological and interpersonal solutions for more timely relevant products (Thomas et al 2013). Rapid responsive review services have appeared with key roles for knowledge brokers (Wilson et al 2015; Campbell et al 2011). The wider research literature offers evidence of knowledge brokers working with stakeholders to define research questions and priorities from clinical/management questions or policy gaps/ concerns (Bornbaum et al 2015). The current study confirms knowledge brokering activities, whoever undertakes them, as central to addressing policy questions with systematic reviews. Elsewhere knowledge brokers are seen to increase the perceived clarity of rapid review proposals and the confidence of reviewers that they could meet policy makers’ needs (Moore et al 2017).

Slow research: Timeliness with reviews is achieved by using or adapting off the shelf products (Model 1), or reviewing rapidly for urgent decisions (Model 4). However, if reviews are to last, they need early investment in thinking. The emphasis given by our interviewees to the time required to shape the review question and the conceptual framework, the responsiveness and iteration, and valuing relevance over data are all features of slow research recently recommended for addressing global health (Adams et al 2014).

Constructive conflict: For policy-relevant reviews, this time is required for systematic reviewers (or knowledge brokers) to explore the issues with policy makers while keeping a critical distance. The challenging nature of this activity prompted this question from one of our peers: ‘were you ever

brought to the edge of your integrity?’ We consider the risk of a researcher or knowledge broker losing critical distance in debates spanning the research-policy interface is analogous to the risk of an ethnographer losing appropriate detachment towards the people they study by ‘going native’ (Hammersley and Atkinson 1995, p. 10). These tensions are recognised in a recent critical review of the ‘dark side’ of knowledge brokering (Kislov et al 2017). The aim instead is mutual challenge and constructive conflict, where different ideas, opinions and perspectives are examined, compared and reconciled (Amason et al 1995).

Developing review questions: When prompted, during peer debriefing, to describe in procedural terms how, in our own collaborative practice, we move from a problem requiring attention to a review question, we recognised the parallels with both qualitative analysis and non-directive counselling. These analogies, which were confirmed by one of our interviewees recognising the value of ‘lateral thinking’ when framing a review question, raises the possibility of turning to these literatures for practical guidance.

Analysing stakeholders’ knowledge and uncertainties corresponds to asking analytical questions of existing data, in other words qualitative analysis or primary data (Strauss and Corbin 1998) or synthesising qualitative studies (Noblit and Hare 1988).

Similarly, conversations with stakeholders to analyse problems and evidence gaps, like non-directive counselling, focus on learning and implications for action (Egan 1990; Rogers 2004). Just as the detailed steps in qualitative analysis and non-directive counselling have been recognised and shared to advance practice (See box 1), we now have the same opportunity to clarify not only what knowledge brokers do, but how they do it to develop questions and conceptual frameworks for policy-relevant reviews. The next step is to develop and test guidance based on these insights.

Conclusions

No review methodology was uniquely appropriate for policy-relevant systematic reviews. It was the mutual engagement across the research-policy interface that made the reviews policy relevant, within the constraints of the available literature. This involved thinking about the issues and seeing them from multiple viewpoints to identify and shape questions; this prompted implicit or explicit value driven debates. The intellectual work to shape a policy-relevant systematic review is an iterative, collective endeavour that requires partners from either side of the policy-research interface to engage with the unfamiliar, listen, challenge and co-construct questions and answers.

Acknowledgements

We were privileged to interview very thoughtful people who shared and reflected on particularly challenging aspects of collaborative work for producing policy relevant systematic reviews. With their permission we name them here: Elie Akl, Lama Bou-Karroum, Xavier Bosch-Capblanch, Karen Daniels, Fadi El Jardali, Anirban Ganguly, Neal Haddaway, Sophie Hill, Natalie Leon, Simon Lewin, Willem Odendaal, Ania Sieracka, Chandrashekar Sreeramareddy and Andrea Tricco. We also benefited from thought-provoking discussions with Morel Fourman.

We are very grateful to the two anonymous reviewers who offered thoughtful and constructive comments.

The authors are very grateful to the World Health Organization for its provision of financial support for this investigation through the Alliance for Health Policy and Systems Research (research grant no. 2015/532540-0; PO-No. 201256565). The study benefited from discussions with the Advisory Group on Health Systems Research Synthesis convened by the funders.

Additional reflections were on our own work funded by the British Council, the UK Department for International Development, the National Institute for Health and Care Excellence (NICE) and UNICEF.

References

Adams, V, Burke, NJ, and Whitmarsh, I, 2014, Slow Research: Thoughts for a Movement in Global Health. *Medical Anthropology*, 33:3, 179-197, DOI: 10.1080/01459740.2013.858335

Amason, AC, Hochwarter, WA., Thompson, KR, & Harrison, AW, 1995, Conflict: An important dimension in successful management teams. *Organizational Dynamics*, 24, 20-35.

Cooke, J, 2005, A framework to evaluate research capacity building in health care. *BMC Fam Pract.* 6:44.

Bates, I, Akoto, AY, Ansong, D, Karikari, P, Bedu-Addo, G, Critchley, J, et al., 2006, Evaluating health research capacity building: an evidence-based tool. *PLoS Med.* 3(8):e299.

Bornbaum CC, Kornas K, Peirson L, Rosella LC (2015) Exploring the function and effectiveness of knowledge brokers as facilitators of knowledge translation in health-related settings: a systematic review and thematic analysis. *Implementation Science* 10: 162.

Campbell D, Donald B, Moore G, Frew D. (2011) Evidence check: knowledge brokering to commission research reviews for policy. *Evidence and Policy.* 7:97–107.

- Colquhoun, HL, Levac, D, O'Brien, KK, Straus, S, Tricco, AC, Perrier, L, Kastner, M, Moher, D, 2014, Scoping reviews: time for clarity in definition, methods, and reporting. *Journal of Clinical Epidemiology* 67 (2014) 1291e1294.
- Edwards, B, 2002, Deep insider research. *Qualitative Research Journal*, 2(1), 71-84.
- Egan, G, 1990, *The Skilled Helper: A Systematic Approach to Effective Helping*. Brooks/ Cole, Belmont California, USA. 4th Edition.
- Hammersley, M, and Atkinson, P, 1995, *Ethnography: Principles in Practice*, London, Routledge.
- Kislov, R, Wilson, P, Boaden R, 2017, The 'dark side' of knowledge brokering. *Journal of Health Services Research & Policy* 22(2) 107–112.
- Lavis, J, 2009, How Can We Support the Use of Systematic Reviews in Policymaking? *Plos Medicine* Nov; 6(11): e1000141.
- Moat KA, Lavis JN, Clancy SJ, El-Jardali F, Pantoja T and for the Knowledge Translation Platform Evaluation study team (2014) Evidence briefs and deliberative dialogues: perceptions and intentions to act on what was learnt. *Bulletin of the World Health Organization* 92(1): 20–28.
- Noblit, GW, Hare, RD, 1998, *Meta-Ethnography: Synthesizing Qualitative Studies*. Newbury Park: Sage.
- Oakley, A, 1981, Interviewing women: A contradiction in terms? In: Roberts H (ed.) *Doing Feminist Research*. London: Routledge and Kegan Paul, 30-61
- Oakley, A, Gough, D, Oliver, S, Thomas, J, 2005, The politics of evidence and methodology: lessons from the EPPI-Centre. *Evidence & Policy: A Journal of Research, Debate and Practice* 1 (1) 5 – 31.
- Oliver, S, Bangpan, M, Stansfield, C, Stewart R, 2015, Capacity for conducting systematic reviews in Low and Middle Income Countries: a rapid appraisal. *Health Research Policy and Systems* 13:23 doi:10.1186/s12961-015-0012-0
- Oliver, S, Dickson, K, 2016, Policy-relevant systematic reviews to strengthen health systems: models and mechanisms to support their production. *Evidence and Policy*. 12 (2) 235-259.
- Oliver, K, Innvar, S, Lorenc, T, Woodman, J, Thomas, J, 2014, A systematic review of barriers to and facilitators of the use of evidence by policymakers, *BMC Health Services Research* 14, 2.
- Robson, C, 2002, *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*, Massachusetts: Blackwell Publishers.
- Rogers, J, 2004, *Coaching Skills: a Handbook*. Open University, Maidenhead, Berkshire, UK.
- Strauss, A, Corbin, J, 2004, *Basics of Qualitative Research: Techniques and procedures for developing grounded theory*. Sage; London.

Thomas J, Newman M, Oliver S. (2013) Rapid evidence assessments of research to inform social policy: taking stock and moving forward. *Evidence & Policy* 9 (1) 5–27.

Tyndall, J, 2008, How low can you go? Towards a hierarchy of grey literature. Presented at Dreaming 08: Australian Library and Information Association Biennial Conference, 2-5 September 2008, Alice Springs. <http://hdl.handle.net/2328/3326>

Wilson, M, Lavis, J, Gauvin FP (2015) Developing a rapid-response program for health system decision-makers in Canada: findings from an issue brief and stakeholder dialogue. *Systematic Reviews* 4: 25.

Whitty, C, 2016, What makes an academic paper useful for health policy? *BMC Medicine* (2015) 13:301 DOI 10.1186/s12916-015-0544-8

Table 1: Characteristics of research and policy teams

Income country	WHO Region hosting review activity	Jurisdiction of policy partners
HIC	Americas	National ²
	Europe	National ¹
		Global
Western Pacific	Local	
Upper MIC	Africa	National and Local ¹
	Eastern Mediterranean	National
	South and East Asia	Anonymous peer review ¹
Regional ¹		

¹Reviews registered with one of four review facilities

²Policy partner in Africa

Table 2: Substantive focus of reviews

		Nature or scale of phenomenon	Policy options	Implementation
Health systems building blocks	Leadership/ governance	Public involvement	Contracting out ¹ Malpractice policies ²	
	Health care financing			Health benefit package ¹
	Health workforce			
	Access to essential medicines		Access to medicines ¹ Vaccine programmes ²	
	Health information systems		Social media ¹ Information for management ¹	

	Service delivery		Integrated programme management ³	
International development			Various ²	
Environmental science			Various ¹	

¹Aligned with priority setting exercise; ²Initiated by policy makers; ³Initiated by researcher

Box 1: Thinking and communication processes analogous to developing a question or conceptual framework

<p>Qualitative analysis</p> <p>Analysing primary data or reports of qualitative research involves asking questions:</p> <ul style="list-style-type: none">• that sensitise the researchers to the landscape of interest – what is going on here, who is involved, how do they define the situation, what does it mean to them, are their definitions and meanings the same or different, what are they all doing (the same or differently) and why?• that explore recurring themes as stakeholders talk;• about processes, variation, connections (or assumptions) about key concepts, changes over time and pertinent structural influence;• about exceptions or contradictions; and• about where to look for evidence and how to recognise it in different contexts.	<p>Non-directive counselling</p> <p>Questions focused on learning and implications for action (Egan 1990; Rogers 2004) involve:</p> <ul style="list-style-type: none">• asking open ended questions to encourage talk and reflection on specific examples;• adopting the stakeholders' own language;• asking future oriented questions about how stakeholders would use the evidence;• provoking thinking, demanding clarification and challenging assumptions;• summarising responses to confirm understanding, invite correction and introduce language that links with wider understandings;• interrupting repetition or vague assertions;• moving the conversation on; and• getting to the crux of the matter and articulating the main focus.
---	--

		Evidence for generic use	
		✓ (matching consensus priority interests)	✗ (matching specific policy interests)
Key concepts clear & widely agreed	✓	<p>1. Update of an effectiveness review</p> <ul style="list-style-type: none"> • Interviewed lead author (ID1.1a)* & review facilitator (ID1.1b) • Document: previous version of review <p>2. Series of effectiveness reviews</p> <ul style="list-style-type: none"> • Interviewed lead author (ID1.2) • Documents: website with reviews, completed/ on-going <p>3. Effectiveness review*</p> <ul style="list-style-type: none"> • Interviewed lead author (ID1.3) • Documents: title registration and protocol <p>Gender analysis of effectiveness reviews</p> <ul style="list-style-type: none"> • Reflections 	<p>4. Series of rapid reviews with multiple questions</p> <ul style="list-style-type: none"> • Interviewed lead author (ID2.1) • Documents: Call for reviews; one published paper; notification of award; PowerPoint presentation. <p>Rapid review with multiple questions</p> <ul style="list-style-type: none"> • Reflections <p>Re-analysis of effectiveness reviews for specific location</p> <ul style="list-style-type: none"> • Reflections
	✗	<p>5. Effectiveness & economic analysis</p> <ul style="list-style-type: none"> • Interviewed lead author (ID3.1a) and co-author (ID3.1b) • Document: protocol on PROSPERO <p>6. Implementation review</p> <ul style="list-style-type: none"> • Interviewed lead (ID3.2) • Document: protocol on PROSPERO <p>7. Effectiveness review</p> <ul style="list-style-type: none"> • Interviewed lead (ID3.3a) and senior author (ID3.4) & review facilitator (ID3.3c) <p>8. Review of models and frameworks</p> <ul style="list-style-type: none"> • Interviewed senior author (ID3.4) <p>Theory building review</p> <ul style="list-style-type: none"> • Reflections 	<p>9. Two country specific review programmes</p> <ul style="list-style-type: none"> • Interviewed funder (ID4.1 = two stage reviews) • Document: feedback from policy teams on questions <p>10. Review of models</p> <ul style="list-style-type: none"> • Interviewed lead author (ID4.2a) and supervisor (ID4.2b) <p>11. Review of models and frameworks</p> <ul style="list-style-type: none"> • Interviewed lead author (ID4.3a) • Documents: Contemporaneous written input from stakeholders through review process (ID4,3b) <p>Evidence brief for a specific country</p> <ul style="list-style-type: none"> • Reflections <p>Policy review of specific countries</p> <ul style="list-style-type: none"> • Reflections <p>Briefing to commission effectiveness review</p> <ul style="list-style-type: none"> • Reflections

Figure 1: Diversity of systematic reviews (numbered if subject to interview, not if subject to reflection)

*ID numbers relate to each review-interviewee combination: first digit indicates model number, second digit indicates review number, and letter indicates one of multiple interviewees for this review. Some interviewees spoke about two or more reviews, therefore the number of unique ID numbers is higher than the number of interviewees.

Figure 2: analytical framework for policy relevant systematic reviews

