Secure Livelihoods Research Consortium

Making systematic reviews work for international development research

Key messages

- Using systematic review principles can help researchers improve the rigour and breadth of literature reviews
- Conducting a full systematic review is a resource intensive process and involves a number of practical challenges
- Systematic reviews should be viewed as a means to finding a robust and sensible answer to a focused research question

Briefing Paper 1
January 2012

Systematic reviews can explore the impacts of cash transfers on women's livelihoods
©Save the Children

The question of 'what works' in international development policy and practice is becoming ever more important against a backdrop of accountability and austerity. Donors are under increasing pressure to adopt spending practices that not only generate positive development and humanitarian outcomes, but that also represent value for money. As part of this drive towards achieving greater (cost) effectiveness, there has been a surge of interest in 'evidence-informed policy making' – the careful use of empirical evidence in the design and implementation of externally-funded policies and programmes in developing countries (DFID, 2011).

Systematic reviews (SRs) are a rigorous and transparent form of literature review. Described by Petrosino et al. (cited in van der Knaap et al., 2008: 49) as 'the most reliable and comprehensive statement about what works', SRs involve identifying, synthesising and assessing all available evidence, quantitative and/or qualitative, in order to generate a robust, empirically derived answer to a focused research question (Box 1).

Increasingly considered a key tool for evidence-informed policy making, a number of donors – most notably the UK Department for International Development (DFID) and AusAid – are focusing attention and resources on testing the appropriateness of SRs in assessing the impacts of development and humanitarian interventions. With the second round of DFID-funded SRs now under way, this briefing paper reflects upon the use of SRs in international development research.

SLRC Briefing Papers present information, analysis and key policy recommendations on issues relating to livelihoods, basic services and social protection in conflict-affected situations. This and other SLRC Briefing Papers are available from **www.odi.org.uk/slrc**. Funded by DFID.

Secure Livelihoods Research Consortium
Overseas Development Institute (ODI)
203 Blackfriars Road, London SE1 8NJ
United Kingdom

T +44 (0)20 7922 8249
F +44 (0)20 7922 0399
E slrc@odi.org.uk
www.odi.org.uk/slrc

It draws on the authors' shared experience of conducting eight SRs[1] on the respective impacts of:

- microfinance programmes (Duvendack et al., 2011)
- cash transfers and employment guarantee schemes (Hagen-Zanker et al., 2011)
- employment creation programmes (Holmes et al., forthcoming)
- 'Markets for the Poor' (M4P) programmes (SLRC, forthcoming)
- school feeding programmes (ibid.)
- seeds-and-tools interventions (ibid.)
- social funds (ibid.)
- water committees (ibid.)

The briefing paper identifies where an SR approach adds value to development research and where it becomes problematic. While the findings are valid across a broader development context, six of the reviews focus specifically on Fragile and Conflict-Affected Situations (FCAS).

Box 1: What are systematic reviews (SRs)?

Used widely and for many years in medical research and the natural sciences, SRs are considered a 'rigorous method to map the evidence base in an [as] unbiased way as possible, and to assess the quality of the evidence and synthesize it' (DFID, 2011). SRs rely upon the use of an objective and transparent approach for the entire process in order to minimise bias and ensure future replicability.

The following steps are often included in SRs:

1. Deconstruct the research question by population, intervention, outcome and comparator – these form the basis of search strings.
2. Produce a protocol that describes definitions, search strings, search strategy, inclusion and exclusion criteria, and approach to synthesis.
3. Pilot the protocol and, if necessary, revise the search strategy.
4. Conduct systematic searching (i.e. study retrieval) of academic databases and perform hand-searching of institutional websites.
5. Screen retrieved studies, using pre-defined inclusion and exclusion criteria on relevance by title, abstract and full text. Often the screening process is piloted to ensure each researcher screens consistently.
6. Characterise studies that have been included in the final analysis by intervention, study quality, outcomes, and/or by research design and type of analysis.
7. Extract relevant quantitative and/or qualitative data, synthesise evidence, and, if possible, perform meta-analysis.

While most SRs apply these steps in a fixed and rigid fashion, some of the SRs referred to in this paper adopted a more flexible approach by continuing to comply with the core principles of SR methodology (rigour, transparency, replicability), while tailoring the protocol as and when required. This reflects the fact that SRs do not constitute a homogenous approach: there are different 'levels' of SR (see http://www.matrixknowledge.com/evidence/wp-content/uploads/2009/05/evidence-review-comparison.gif), and researchers working in various disciplines have previously attempted to make SRs more useful by combining them with other methodological approaches (e.g. van der Knaap et al., 2008).

How SR principles can improve literature reviews

Our shared experience of conducting SRs suggests that adhering to SR principles can improve the quality and strength of literature reviews in a number of ways.

Increasing breadth, retaining focus

Through the adoption of broad search strategies, pre-defined search strings, and uniform in/exclusion criteria (Box 1), SRs effectively force researchers to search for studies beyond their own knowledge and networks, thereby reducing researcher bias. At the same time, the careful deconstruction of the research question at the outset in terms of population, intervention, comparator and outcome, ensures that the review process remains tightly focused. In theory, this improves the likelihood of generating a clearer, more objective answer to the research question.

Empirical evidence, not preconceived knowledge

In comparison to other literature reviews in international development research, SRs focus much more strongly on evidence, impact, validity, and causality. By extracting information on research design (sampling strategy and data collection methods), analytical methods and causal chains, SRs are effective at gauging the robustness of evidence. Classifying the quality and characteristics of impact studies against standardised criteria also enables the possibility of producing cross-study comparisons and meta-analyses, which are valuable for evidence-informed policy making. In other words, SRs encourage researchers to engage with studies more critically and to be consistent in prioritising empirical evidence over preconceived knowledge.

Transparency and replicability

The use of a clear SR protocol is effective not only in guiding researchers throughout the process – keeping them 'on track' – but also in improving the methodological transparency of the review and enabling future replication. SRs are, therefore, able to produce a relatively objective baseline against which future research and evidence on certain interventions can be assessed. This might prove particularly useful for 'measuring' the knowledge contribution of a research programme over a number of years.

Practical challenges of SRs

Despite the added value of an SR approach, we encountered a number of practical problems throughout the process:

- **Searching academic databases.** SRs require access to a wide range of databases and peer-reviewed journals, which can be problematic for non-academic researchers and those based in southern research organisations. Promoting SRs as best practice, therefore, sits uneasily alongside donors' interests in developing southern research capacity and in encouraging a more inclusive process of evidence building.
- **Searching institutional websites.** Searching institutional websites – essential as relevant research is often located outside the formal peer-reviewed channels (for example, of the nine studies included in the social funds SR, just two were retrieved from academic journals) – undermines the objectivity of the search and retrieval process. This happens for a number of reasons: differences in websites' search functions mean

Table 1: Number of studies included in our SRs

Systematic review topic	Cash transfers & employment guarantee schemes	Employment creation	Microfinance	School feeding programmes	Water committees	Social funds	Seeds-and-tools	Markets for the Poor
FCAS-focused		x		x	x	x	x	x
Studies retrieved	35,991	46,177	3,620	128	1,225	76	2,372	482
Studies screened on title/abstract (after removing duplicates)	24,263	9,558	2,643	116	1,192	72	2,325	464
Studies screened on full text	356	191	201	45	189	29	118	55
Studies included in final analysis	37	8	58	16	6	9	9	3

that search strings have to be either adapted or discarded altogether; and relevant websites may be excluded, whether unintentionally (lack of knowledge) or otherwise (time/resource constraints), meaning that potentially high numbers of pertinent studies can be missed.

- **Screening.** There is inevitable subjectivity in the screening process, particularly when high numbers of researchers are involved, as each member of the research team interprets inclusion criteria slightly differently.
- **Classifying included studies.** Data and methodology are, in general, poorly described, and researchers often have to rely on authors' self-proclaimed research design and results, which introduces another source of bias. In addition, information on statistical significance is missing from many studies.
- **Synthesis.** Meta-analysis is rarely possible because of the non-availability of data as well as methodological diversity. This makes it difficult to draw meaningful conclusions. Moreover, complex interventions tend to generate multiple outcomes which SRs may not be able to capture (Boaz et al., 2002).
- **Generating useful policy recommendations.** Findings are often too broad, too incomparable, and too research-oriented (and therefore of greater interest to an academic rather than a policy audience).
- **The SR process is resource intensive.** Using a rigid SR procedure is an extremely demanding and time consuming process, in part because of the high number of studies that are often assessed (Table 1).

What are the fundamental concerns?

In addition to the practical difficulties outlined above, the use of SRs in international development research throws up a series of deeper dilemmas.

What's good for an SR may not be so good for a peer-reviewed journal

Empirical impact studies in development studies are not written in a uniform fashion, unlike in the natural and medical sciences. This is problematic from a practical perspective: unclear titles and vague, unstructured abstracts make it more difficult to accurately assess the relevance of a study on the basis of a title or abstract alone. But there is also a more fundamental concern: the attributes that get research published in a peer-reviewed development journal are very different to those required for inclusion in an SR.

SR inclusion criteria demand a high level of detail on method, data and impact that many peer-reviewed articles either do not contain for lack of space or forego in favour of, say, deeper explorations of historical context. Peer-reviewed journals, therefore, may not be the most appropriate sources for SR study retrieval.

Assessing evidence

SRs were pioneered in the natural sciences where the predominant methodologies are quantitative. These can be assessed in a relatively straightforward (although not entirely unproblematic) fashion using pre-existing methodological quality scales, such as the Maryland Scientific Methods Scale. In international development research, however, qualitative methodologies are just as common. This poses a challenge because quality appraisal techniques for assessing qualitative studies are underdeveloped, and it is not yet clear how well SRs are able to compare qualitative with quantitative methodologies and findings.

There are many research questions of a qualitative nature for which an SR approach would be inappropriate. But the problems of assessing qualitative evidence mean that SRs lean even further towards quantitative studies and measurable outcomes than they perhaps otherwise would. Given the rising importance and prominence of SRs in policy making, this may have serious long-term policy implications if donors become unwilling to fund interventions that generate less tangible, more difficult-to-measure outcomes (such as those that aim to strengthen community cohesion or build state-citizen relations). Therefore, while efforts have been made outside the field of international development to make SRs more inclusive of qualitative evidence (for example, the Cochrane Collaboration's qualitative methods network; see also Petticrew and Roberts, 2006), this remains a challenging area that requires greater attention.

What do SRs miss? The importance of context and process
Quantitative methodologies adopted in the natural sciences aim to measure impact and causality by controlling for confounding factors. However, 'cutting out the noise' probably misses the point in international development research (and the social sciences more broadly), where context is everything.

It is only through considerations of political economy, social relations and institutions that we gain a fuller understanding of why particular interventions work in particular environments at particular times. Similarly, while investigations into causality and impact are undeniably vital, understanding process and the internal dynamics of interventions is just as important.

Outcomes are ultimately shaped by programme design and delivery, as well as by context (see Pawson and Tilley, 1997 on context-mechanism-outcome configurations), and SRs do not necessarily help us understand these dimensions. In other words, the question of *why* things work is just as policy relevant as whether or not they do in the first place.

Conclusions and recommendations

SRs are a new tool in international development research and have the potential to enhance and promote evidence-informed policy making. But they may not be as objective as they appear, and their strengths must be balanced against a number of practical and fundamental limitations.

There is a need to adapt the methodology to make SRs work for international development and humanitarian research, and finding ways to achieve this will only happen through experimentation with the process. Ultimately, SRs should be seen as a means to an end – helping to get a robust and sensible answer to a focused research question – and not an end in themselves (Lichtenstein et al., 2008 drew a similar conclusion on the use of SRs in the field of nutrition). Rather than following a rigid SR methodology, our shared experience suggests that a more useful approach for development researchers might involve a mixture of compliance and flexibility: compliance with the broad SR principles (rigour, transparency, replicability) and flexibility to tailor the process towards improving the quality of the overall findings, particularly if time and budgets are constrained. In short, we should be focusing on the utility that can be gained from an SR approach rather than its rigid application.

We offer five specific conclusions:

- Applying SR principles to a literature review is highly valuable as it increases breadth, improves transparency and emphasises the importance of empirical evidence over preconceived knowledge.
- SRs can be used to identify knowledge gaps and highlight methodological inconsistencies and weaknesses; they are therefore useful in identifying future research priorities.
- Full SRs are expensive: researchers need to consider whether the full application of a rigid SR approach is justified in relation to the time and resources required.
- SR methodology can be adjusted or developed (see, for example, the work of van der Knaap et al., 2008) if it helps to get a more useful answer to the research question.
- More work is needed to find better ways to assess qualitative research and compare it with quantitative work.

And one reflection:

- As researchers, we should pay more attention to the way we write titles and abstracts. By including information on methodology, data and key findings, we can make our work more user-friendly for other researchers and policy makers.

[1] A number of the SRs may eventually be published as 'systematic maps' due to the low number of relevant studies included for final analysis.

Written by Jessica Hagen-Zanker, Maren Duvendack, Richard Mallett and Rachel Slater based at the ODI, with Samuel Carpenter and Mathieu Tromme. For futher information please contact Jessica Hagen-Zanker (j.hagen-zanker@odi.org.uk).

References

Boaz, A., Ashby, A. and Young, K. (2002) 'Systematic Reviews: What have they got to offer evidence based policy and practice?' Working Paper 2. London: ESRC UK Centre for Evidence Based Policy and Practice

DFID (2011) 'Feature: Systematic Reviews in International Development: An Initiative to Strengthen Evidence-Informed Policy Making'. Accessed 22 September 2011 [http://www.dfid.gov.uk/r4d/SystematicReviewFeature.asp]

Duvendack, M., Palmer-Jones, R., Copestake, J. G., Hooper, L., Loke, Y. and Rao, N. (2011) 'What is the evidence of the impact of microfinance on the well-being of poor people?' London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London

Hagen-Zanker, J., McCord, A. and Holmes, R. (2011) 'The impact of employment guarantee schemes and cash transfers on the poor'. ODI systematic review

Holmes, R., McCord, A. and Hagen-Zanker, J. (forthcoming) 'What is the evidence of the impact of employment creation on a) stability and b) poverty reduction in fragile states?' ODI systematic review

Lichtenstein, A. H., Yetley, E. A., and Lau, J. (2008) 'Application of Systematic Review Methodology to the Field of Nutrition'. *Journal of Nutrition* 138: 2297-306

Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*. London: Sage

Petticrew, M. and Roberts, H. (2006) *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford: Blackwell Publishing

SLRC (Secure Livelihoods Research Consortium) (forthcoming) 'The impacts of five development interventions in fragile and conflict-affected situations'. SLRC systematic review

van der Knaap, L. M., Leeuw, F. L., Bogaerts, S., and Nijssen, L. T. J. (2008) 'Combining Campbell Standard and the Realist Evaluation Approach: The Best of Two Worlds?'. *American Journal of Evaluation* 29(1): 48-57