



### Systematic reviews: Let's Work out the Kinks

*Birte Snilstveit and Hugh Waddington*

A recent [briefing paper by ODI on systematic reviews](#) in international development is a welcome contribution to the debate on systematic reviews in international development, triggering commentary by high profile commentators such as Ben Goldacre who wrote extensively on the subject in ['Bad science'](#).

At 3ie we have conducted, funded and quality assured systematic reviews in international development since 2008 as part of an [overall strategy](#) to promote evidence-based policy in international development, and ultimately improve the lives of people in low and middle income countries. Our experience has highlighted a number of challenges in making systematic reviews work for international development. But we think the analysis offered in the brief conflates different issues — namely what the authors struggled with when conducting their own SRs, and the inherent constraints of the SR methodology. An analysis of the former informs areas for the SR community to work together to address while the latter should lead to a better understanding of the place of SRs among other forms of evidence.

### Conducting Systematic Reviews is hard!

Lack of access to academic databases and difficulties with searching institutional sites does not point to inherent issues with the SRs as a tool for evidence based policy. Rather, it underscores a need to work to make both published and unpublished impact evaluations more readily available, particularly for researchers in low and middle income countries. To this end, 3ie is compiling a searchable [database of impact evaluations](#), in addition to establishing a prospective registry of impact evaluations in international development. The British Library for Development Studies at IDS has also recently launched an online [digital library](#) and researchers from low and middle income countries can also access journal articles for free from the Global Development Network's [online services](#).

A key lesson to emerge from recent applications of systematic review methodology to international development is the importance of a good question. Even with plenty of data available, making systematic reviews work for international development requires applying the methodology to clearly defined research questions on issues where a review seems sensible. A review in medicine will often ask a narrow question such as the [Cochrane Collaboration's](#) recent review on the [efficacy of oseltamivir \(tamiflu\) for preventing and treating influenza](#). Many of the review questions development researchers have attempted to answer in recent systematic reviews are too broad, which inevitably leads to challenges. There is a trade-off between depth and breath, but if our goal is to build a sustainable community of practice around credible, high quality reviews we should be favouring depth of analysis where a trade-off needs to be made. In contrast, any benefits



## International Initiative for Impact Evaluation (3ie)

---

from trading off depth for breadth in terms of broadening support for reviews in the short term will ultimately undermine their credibility among users in the long term. The solution is to conduct more systematic reviews, addressing clearly defined research questions, and bringing findings together in overviews of reviews. 3ie's [database of systematic reviews](#) facilitates this by gathering all systematic reviews in international development in one place. As more systematic reviews are being produced the database will provide gapmaps and thematic collections of systematic reviews across a range of sectors.

The ODI paper argues most systematic reviews apply the methodology in a fixed and rigid fashion, noting they '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'. However, this statement is missing the point. The protocol should state a priori what evidence is eligible for inclusion, where it will be located, and how it will be assessed. This prevents cherry-picking of evidence, data-mining or lowering or raising the bar arbitrarily to provide evidence supporting or refuting a preconceived conclusion. While medical reviewers are going up against big pharmaceuticals, both development researchers and commissioners have biases and vested interests in the outcomes of a review. We need to start becoming more aware of bias in social science research, including in international development. A researcher might favour a conclusion based on their ideological perspective or to improve their chances of publication. Similarly, if an implementing agency has spent large sums of money on an intervention, it can be inconvenient if the evidence suggests the intervention is not effective. For this reason a published study protocol with explicit inclusion criteria and description of review methods is a defining feature of systematic reviews, and avoid vested interests driving conclusions.

Systematic reviews are often criticised for being too rigid and throwing out most of the evidence. However, this is in part a misunderstanding. When results from a typical search process indicate that 10,000 records were identified, of which only 30 are included in the review - the point is that most of the 10,000 studies are excluded because they are irrelevant to the review question, not because of issues with study design and quality. Developing search strategies in collaboration with an information specialist, and piloting and refining the search as part of developing the study protocol can reduce the number of irrelevant hits.

However, we agree that there is a need to include a broader range of evidence in systematic reviews. Reviews have traditionally focused on synthesising quantitative evidence, typically from randomised control trials. This narrow focus is rightly criticised, especially for its limitations in informing policy and practice. In our view, evidence hierarchies and the quantitative versus qualitative schism are unhelpful. Use of evidence appropriate to the research question is more important (see Petticrew and Roberts on ['Horses for Courses'](#)). A research question which asks 'what works' should draw on evidence from high quality impact evaluation studies which are capable of addressing attribution. But answering other questions such as 'why it works' requires a broader range



of evidence, including qualitative. This informs 3ie's approach to systematic reviews. We do not restrict our reviews to RCTs, but include a broader range of quantitative impact evaluation designs (including studies using propensity score matching, regression discontinuity and instrumental variables) to address 'what works'. When appropriate we also promote incorporating a broader range of evidence to address other questions of relevance to policy and practice, such as 'how does it work' and 'what are the barriers and facilitators of intervention effectiveness' in a separate review component. While there is less agreement on how to appraise qualitative evidence, this is not a challenge which is particular to international development. Colleagues in health and social welfare have been working on methods for systematic reviewing of qualitative evidence for at least a decade and there is a range of existing tools and [guidance](#) which authors can draw on.

The systematic reviewing field is still nascent in development, and has so far largely been demand driven. The explosion in demand for systematic reviews has far outstripped the supply of development researchers with experience using the methodology. To build capacity and ensure high quality products, teams should ideally include at least one experienced reviewer on the team, which may mean bringing in technical support from a field in which reviewing is better established, such as health. Teams should also engage with an experienced group for training, support and peer review. 3ie has partnered with the Campbell Collaboration and to set up a group dedicated to helping build capacity for production of high quality systematic reviews in international development - the [International Development Coordinating Group \(IDCG\)](#). The IDCG provides support and quality assurance to review authors and systematic review commissioners.

### **The Place of Systematic Reviews in the Evidence Ecosystem**

The authors rightly raise the importance of contextualising reviews, and we agree that this is something systematic reviews need to be able to grapple with better. But if one holds the view that 'context is everything', as argued in the ODI piece, then the differences with the philosophy of science underpinning systematic reviews are likely so fundamental that one will inevitably have objections to the methodology of SR itself. We do think lessons can be learnt across contexts, and argue it is better to generalise findings from a collection of evidence from multiple similar programmes across different contexts, than from single studies. Indeed, a major purpose of the SR approach should be to provide policy makers with evidence on whether an approach has been shown to work, and is worth applying in a different context.

On the subject of resource management, one of the conclusions offered in the ODI brief is that systematic reviews are expensive, and that 'researchers need to consider whether the full application of a rigid SR approach is justified'. But while more time consuming and expensive than literature reviews, SRs are still cheaper than impact evaluations and are particularly valuable in areas where there is an emerging evidence base, as well as a way of highlighting evidence gaps and methodological issues to consider for future studies.



## International Initiative for Impact Evaluation (3ie)

---

Better process management, relatively large cross-disciplinary teams with substantive and methodological expertise and efforts to make evidence more easily available provide opportunities to speed up the review process and reduce costs.

In this vein, it is important to recognise that systematic review is only one part of the puzzle for evidence based policy. It would be naive to think that policy will be based on only one source of evidence. [Guidelines for public health interventions](#) for instance are developed in consultation with stakeholders and draw on a range of different types of evidence and information, including systematic reviews, but also more context specific research and expert opinion. Users and producers of research need to collaborate to develop similar systems and communities of practice for evidence based policy making in the field of international development.

*Birte Snilstveit is an Evaluation Officer and Hugh Waddington is a Senior Evaluation Officer at 3ie. They manage 3ie's Systematic Reviews Programme and the Campbell Collaboration International Development Coordinating Group and are based in 3ie's London office.*