Using and commissioning systematic evidence in international development

Hugh Waddington
Anna Heard

USAID Evidence and Innovation
Washington DC
28 September 2017
Partners in systematic evidence production
Objectives

1. Introduction to systematic evidence: what, why and where to find it

2. How to ask answerable questions for systematic reviews

3. Processes needed for rigor and relevance
Someone asks you to prepare a brief in three hours on the effectiveness of a particular intervention. Where would you go for sound evidence?
There’s a lot of evidence out there

Sheer amount of available evidence “Beyond the capacity of the human mind”
Why single studies can be useful

Scale up what works

Inform policy discussions

Close down or re-design what doesn’t
But single studies give a *biased* view of the overall evidence

- Misrepresent the *balance* of research evidence
- Illuminate only one part of a policy issue
- Are sample-specific, time-specific, context-specific
- Are often of poor quality.
Findings are often contradictory

Single study evidence: “....the program did not have significant impacts on the performance of graduates and their neighbors”

Are farmer field schools effective?

“Results demonstrated remarkable, widespread and lasting developmental impacts”
What does the systematic evidence say?

FFS effective in small-scale project settings

- 25% pesticides reduction
- 10% yields increase
- 17% revenues increase

Source: Waddington and White (2014)
No diffusion to neighbour farmers

Source: Waddington and White (2014)
FFS doesn’t work at scale

Source: Waddington and White (2014)
What are systematic reviews?

A way of establishing the overall balance of evidence on a topic or policy

And separating higher quality from lower quality evidence

And identifying what is generalisable and what is context specific

A way of addressing known problems in available evidence (e.g. publication bias)
“Drawing on systematic, critically appraised evidence is a no-brainer” (Marie Gaarder, 3ie)

Formal land titling

Effects on land productivity (output per ha)

- Peru 1993-2004 (Torrero & Field, 2005)
- Cambodia 1989-2004 (Markussen, 2008)
- Vietnam 1993-2006 (Van den Broeck et al., 2007)
- Ethiopia 1998-2006 (Holden et al., 2009)
- Madagascar 1961-2005 (Jacoby & Minten, 2007)
- Madagascar 1961-2002 (Bellemare, 2013)*
- Malawi 2004-2006 (Chiwra, 2008)*
- Zambia 1995-2001 (Smith, 2004)*

Less improvement in Africa

Source: Steve Lawry, Cyrus Samii et al.
What makes a review systematic?

- Systematic **searching** for studies and study inclusion criteria as determined by study **protocol**
- Systematic **critical appraisal** of studies (separating the ‘wheat’ from the ‘chaff’)
- Systematic and transparent **data extraction**
- Systematic **analysis** (e.g. meta-analysis, qualitative synthesis)
- Systematic **reporting of findings and implications**
  + Systematic **engagement**
Sources of systematic evidence

700 records, including summary and critical appraisal
Objectives

1. Introduction to systematic evidence: what, why and where to find it

2. How to ask answerable questions for systematic reviews

3. Processes needed for rigor and relevance
## Types of evaluation questions

<table>
<thead>
<tr>
<th>Type of evaluation question</th>
<th>Example question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First generation question</strong></td>
<td>“What is the effect of [intervention] on [outcomes] among [participants]?”</td>
</tr>
<tr>
<td>‘does it work compared to doing nothing?’</td>
<td></td>
</tr>
<tr>
<td><strong>Second generation question</strong></td>
<td>“What are the relative effects of [intervention a] and [intervention b, c, d, etc.]?”</td>
</tr>
<tr>
<td>‘what works best?’</td>
<td>“What intensities or levels of intervention are most effective”</td>
</tr>
<tr>
<td></td>
<td>“What approaches to implementation are most (cost-) effective?”</td>
</tr>
</tbody>
</table>
Corresponding to two main types of reviews

**Women’s Self-Help Groups Systematic Review**

- Full-text articles assessed for eligibility \(n = 362\)
- 354 articles excluded on relevance and 74 excluded on methodology

**What are impacts of economic self-help groups on women’s individual-level empowerment?**

- Impact evaluations included \(n = 23\)
- Qualitative studies included \(n = 11\)

**Education Effectiveness Systematic Review**

- Full-text articles assessed for eligibility \(n = 2042\)
- 1,600 articles excluded on relevance and methodology

**What are impacts of different education interventions on access and learning outcomes for primary and secondary students?**

- Studies included \(n = 238\) from 420 papers

Source: Carinne Brody et al. (2015)  
Source: Birte Snilstveit et al. (2015)
How to ask answerable questions: PICO

- **POPULATION**: Which regions, countries and groups of people?

- **INTERVENTION**: One intervention? Many interventions?

- **COMPARISON**: compared to what – nothing, ‘business as usual’, other intervention?

- **OUTCOME**: single and/or multiple outcomes? Intermediate and/or final outcomes? (+process/implementaion?)

+ **STUDY DESIGN**: quantitative and/or qualitative, causal and/or correlational?
The question setting process

- Consultation with funder
- Needs assessment
- Clarification of question
- Agreement of question
- Question and PICO articulated clearly

Do public schools that receive computer-assisted learning get better test scores than public schools which don’t in L&MICs?
Use of Evidence Gap Map as a scoping exercise
Main types of reviews supported

• **Review of effects** (single intervention, multiple outcomes)
  “The impact of daycare on children’s nutritional and cognitive development in developing countries”

• **Comparative review of effects** (multiple interventions)
  “The impacts of education programmes on attendance and learning in low- and middle-income countries”

• **Review of economic appraisal**
  “Cost-effectiveness of cash-based approaches in humanitarian emergencies”
  “Willingness to pay for clean water in sub-Saharan Africa”

• **Effectiveness review** (‘full causal chain’ synthesis)
  “Targeting, barriers and enablers, impacts and cost-effectiveness of farmer field schools”
Question setting for systematic evidence

What question related to development effectiveness would you like answered?

For example:

• How do we end the AIDS epidemic?
• How do we get more kids immunized?
• Does keeping girls in school improve their economic outcomes?
Discuss in groups how to make your question answerable using PICO (15 mins)

Question 1:

Question 2:

Question 3:

Question 4:

Question 5:
Objectives

1. Introduction to systematic evidence: what, why and where to find it

2. How to ask answerable questions for systematic reviews

3. Processes needed for rigor and relevance
What makes a review rigorous and relevant?

- Systematic **stakeholder engagement**
- Answering **relevant questions** using mixed methods and intervention theory
- Systematic **synthesis** (including meta-analysis where appropriate)
- Systematic **reporting** of findings and implications
- **User-friendly** knowledge translation and dissemination
- Results provided in **timely** way within decision-making cycle
## Timelines and costs of full SRs

<table>
<thead>
<tr>
<th>Review type</th>
<th>Typical cost</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review of effects</td>
<td>$100,000</td>
<td>12-18 months</td>
</tr>
<tr>
<td>Comparative review of effects</td>
<td>$100-250,000</td>
<td>12-24 months</td>
</tr>
<tr>
<td>‘Full causal chain’ effectiveness review</td>
<td>$150-250,000</td>
<td>12-24 months</td>
</tr>
</tbody>
</table>

- So reviews need to be built into known processes (e.g. major policy development), or commissioned by communities of practice
- Technical support from recognised agency (e.g. 3ie, Campbell Collaboration, EPPI-centre)
Updates, reviews of reviews and EGMs

- **Review update (update searches, engagement, quality)**
  
  “Services for street-connected children in L&MICs: systematic review update”

- **Review of reviews (evidence syntheses)**
  
  “The impacts of microfinance in developing countries: review of credit, savings, insurance and financial inclusion”

<table>
<thead>
<tr>
<th>Review type</th>
<th>Typical cost</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Update of systematic review</td>
<td>$30-100,000</td>
<td>3-9 months</td>
</tr>
<tr>
<td>Review of reviews</td>
<td>$30-100,000</td>
<td>6-12 months</td>
</tr>
<tr>
<td>Evidence gap map</td>
<td>$50-150,000</td>
<td>3-9 months</td>
</tr>
</tbody>
</table>
Step 1: Determine **scope** and **review questions consultatively**

Step 2: **Collect all relevant studies** on a particular topic

Step 3: **Assess biases** in the evidence base

Step 4: **Synthesise evidence** on programme impacts and/or process and implementation evidence

Step 6: **Communicate findings and implications transparently**
Collect studies

- We usually want to collect studies systematically.
- If we do not, we cannot say anything generalisable about development effectiveness.
- And we cannot assess important problems like publication bias.
Assess biases in studies

We assess bias to account for:

- Factors other than the intervention which *confound* causal estimates
- *Measurement error*
- *Reporting bias (including publication bias)*
“You shouldn’t trust *everything* you read in a peer-review journal” – Howard White, Campbell Collaboration

- Structured questions
- Decision rules leading to different categories of bias
- Done by two reviewers working independently
- Transparently reported

<table>
<thead>
<tr>
<th></th>
<th>Random sequence generation</th>
<th>Allocation concealment</th>
<th>Blinding</th>
<th>Incomplete outcome data</th>
<th>Selective reporting</th>
<th>Other biases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miguel 2004 (Cluster)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Nga 2009</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Olds 1999</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
What results look like before and after control for bias

Pooled effect in percent

- Contract farming
- Farmer field schools
- Microcredit

- Biased estimate
- Unbiased estimate
Synthesising evidence

Synthesis can be narrative or statistical:

- *Narrative along the causal chain:* immediate outcomes $\rightarrow$ intermediate outcomes $\rightarrow$ impacts
- *Statistical pooling across studies* for particular outcomes
- *Statistical pooling within and across studies*
Many SRs use outdated methods to synthesise evidence

<table>
<thead>
<tr>
<th>Country</th>
<th>Effect on empowerment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>Positive</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Positive</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Positive</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Positive</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Positive</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Positive</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Positive</td>
</tr>
<tr>
<td>India</td>
<td>Positive</td>
</tr>
<tr>
<td>India</td>
<td>Positive</td>
</tr>
<tr>
<td>India</td>
<td>Positive</td>
</tr>
<tr>
<td>Morocco</td>
<td>Not significant</td>
</tr>
<tr>
<td>Pakistan</td>
<td>Not significant</td>
</tr>
<tr>
<td>South Africa</td>
<td>Positive</td>
</tr>
<tr>
<td>South Africa</td>
<td>Positive</td>
</tr>
<tr>
<td>Uganda</td>
<td>Positive</td>
</tr>
</tbody>
</table>

14 > 7 so microcredit empowers women?
The conclusion from meta-analysis is very different

<table>
<thead>
<tr>
<th>Country</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>Garikipati (2008)</td>
</tr>
<tr>
<td>Morocco</td>
<td>Crepon et al. (2011)</td>
</tr>
<tr>
<td>India</td>
<td>Banerjee et al. (2009)</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Pitt et al. (2006)</td>
</tr>
<tr>
<td>South Africa</td>
<td>Karlan and Zinman (2009)</td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>Ngo (2008)</td>
</tr>
<tr>
<td>South Africa</td>
<td>Kim et al. (2007)</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
</tr>
</tbody>
</table>

Source: Data taken from Jos Vaessen et al., 3ie Systematic Review Report
We are more likely to conclude programmes have effects using meta-analysis

Source: McCulloch and Lombardini, *J Dev Effectiveness*
Stakeholder engagement

- A structured way to determine and collaborate with potential users
- Advisory group crucial to engagement strategy from outset
- Highly effective approach: e.g. reviews on sanitation/hygiene behaviour change and agriculture certification schemes
The stakeholder analysis requires you to do the following.

**Exercise 1**: List all stakeholders you can think of.

**Exercise 2**: Who among the listed is influential and who is not?

**Exercise 3**: Among the influential, what is the interest-access nexus?

**Exercise 4**: Fill out box 1 based on exercise 2 and 3.

---

**Rationale for the 3ie stakeholder engagement and communication plan (SECP)**

3ie was set up in 2008 with a global mandate to improve the evidence base for what works, how, why and at what cost in international development policies and programmes in low- and middle-income countries (L&MICs). Part of that mission is to focus on funding studies and reviews that are policy-relevant and useful to decision-makers. 3ie does this by funding high-quality impact evaluations and full systematic reviews. 3ie supports high-quality, methodologically sound systematic reviews in recognition of the limits of single-study results and the value of synthesising evidence rigorously.
User-friendly products

• Policy-friendly products (1: 3: 25)
  • Plain language summary
  • Systematic review brief
  • Systematic review summary report

• Blogs and infographics
Exercise (15 mins): Plain Language Summaries/Policy briefs

Read the PLS/brief in your hands (5 mins)

Discuss in groups what are the key takeaways from the PLS (5 mins)

Discuss the relevance of this evidence to policy and practice (5 mins)