clear
Centers for Learning on Evaluation and Results
ANGLOPHONE AFRICA

Guideline on evaluation synthesis

What this guide is for	To provide guidance in the implementation of evaluation syntheses to answer specific policy questions
Who the guide is for	The guide can be used by government M&E units commissioning syntheses/conducting syntheses internally; independent evaluators commissioned to carry out syntheses for government and parliamentary researchers interested in assessing the performance of specific public sector interventions.
How to use it	The guide can be used for various purposes including when countries are doing evaluation synthesis as part of the VNR process. When using it as part of VNR process, the guide should be used in conjunction with the UNICEF/CLEAR AA guideline on Embedding evaluations in VNRs.

1. Introduction

The practice of evaluating public sector programmes and policies experienced significant growth in recent years. Evaluation is an important source of evidence for policymakers and implementers to base decisions about policy, programme and service delivery adaptation. However, the use of evaluation evidence in policy and practice is complicated for several reasons. One of the reasons is that no single evaluation can unequivocally answer all questions regarding a policy or programme. This is particularly true in public policies and programmes that tend to be large, cover large sections of the population and are implemented by a range of stakeholders with varying capacities and resources. For example government response to inadequate housing condition can be through capital subsidies, urban regeneration programmes, and other land management instruments. Such interventions are likely to be implemented by several government and non-government agencies. In such a policy it is likely that several evaluations will be needed to answer questions about the performance, costs, sustainability or impact of the policy intervention. This is when evaluation synthesis can be a powerful tool to bring together findings from different evaluations to answer policy questions.

2. What is an evaluation synthesis?

There are different definitions of evaluation synthesis. Evaluation synthesis has been defined as "formal technical procedure for combining the results from several empirical studies. It is a systematic procedure for organising findings from several disparate evaluations studied"[1].
Evaluation synthesis has also been defined as the process of bringing together existing studies, assessing their relevance and reliability, and drawing together their data to answer specific questions. These definitions highlight four key features of evaluation synthesis:

- It aims to bring together what is known about a particular programme/policy or service delivery issue from findings of existing research/evaluations;
- It is a scientific process, a systematic research process guided by established research principles;
- It is (or should be) a transparent process, decisions taken in different stages of doing synthesis must be recorded and possible for others to verify;
- It answers a specific policy/service delivery question;

Why evaluation synthesis is important:

- Synthesising existing research and evaluations is a very useful and cost-effective way to answer policy questions using existing evaluations;
- By answering specific policy question and applying evaluation criteria, it can give new insights from disparate primary research/evaluations;
- By pulling the findings of individual evaluations that are often limited by focus and size, evaluation synthesis increases the strength of individual evaluations and empirical research, and presents these in one systematically produced evaluation report.
- It can reveal gaps in knowledge on a topic. This can better point policy makers to what areas there is not enough known and therefore they need to act with caution if they are considering adaptation. As a result, doing a synthesis can help inform an evidence agenda in a policy area;
- A synthesis can reveal conflicting evidence and the reasons for conflict. This makes it easier for policymakers who will not be able to read all available research/evaluations reports on a given topic or who might not know how to reconcile contradictory evidence.

3. Different types of syntheses

There are several types of syntheses, though they follow similar methods they vary in terms of intensity and methodological rigour. The table below shows some of the most used:

Table 1: Different types of syntheses

Types of synthesis	Nature	When might you use It	The time required to complete
Quick scoping review	A quick review of research undertaken on a (constrained) topic. This could be systematic, but because it is quick, it is unlikely to be comprehensive or detailed.	Constrained topic Need answers urgently	1-2 weeks to 2 months
Evidence Mapping/evaluation maps	Systematically and transparently assess and map existing evidence on a particular	When it is unclear what evidence exists on a given topic	2-6 months (can be quicker

[1] Government Accounting Office (1992). Programme evaluation: the evaluation synthesis. https://www.gao.gov/assets/80/76108.pdf

	topic. It is based on systematic review methods but does not do synthesis.	Useful as a starting point (can build on for synthesis later)	depending on the scope)
Rapid evidence assessment	Systematic search but a quick overview of existing research/evaluation on a topic. Synthesis of evidence provided by these research/evaluations to answer the rapid evidence assessment question.	When the need is to get an overview of existing research/evaluation on a given topic to inform both policy decisions and evaluation agenda	2-6 months (quicker than systematic review)
Full systematic review	A systematic search ad review of existing research on a specific topic and synthesis of the evidence provided by these research/evaluation to answer the review question	When there is a focused or specific question Only possible where there is capacity and information specialists as it requires specialised methods for searching research and evaluations	8-12 months minimum
Realist review	A broad review methodology that can deal with complexity, impact of context and varied outcomes	A complex programme/policy implemented with varying beneficiaries and in a different context; different study designs	8-10 months depending on the scope
Review of reviews/ umbrella reviews	Same as the above methods but only includes reviews of other reviews	In a topic where there have been a number of reviews done	Often quicker than other types of full systematic review

Source: DPME [2]

Evaluation synthesis and other research syntheses?

Methodologically and procedurally evaluation synthesis borrows from other research syntheses and depending on the purpose it could follow the same strict procedure as in systematic reviews. However, there are some differences:

- Unlike research, evaluation entails a judgement of interventions according to their implementation processes, results and impacts. Therefore, evaluation synthesis still applies some criteria; this could be DAC OEC criteria but could be any criteria determined to be relevant for the evaluation.
- Like other evaluations forms, evaluation synthesis must have an evaluand. There must be a subject of evaluation, i.e. a policy, programme, or some government intervention;
- Evaluation synthesis is structured to answer a specific policy question, not just synthesise or aggregate what is known on a given topic. Evaluation synthesis will often seek to explain not just give an effect size.

4. When might an evaluation synthesis be appropriate?

Evaluation synthesis can be extremely helpful in these contexts:

- In a sector where there is a significant body of evidence and where there are contradictions and conflicts in the evidence or contradictory views on a policy or programme;
- When the government is contemplating significant policy change. In this context government cannot rely on a single evaluation to determine areas where changes are needed;

[2] DPME (2014). Guideline on evaluation synthesis

- When reporting on performance against SDGs, where outcomes on a goal/target are determined by the different policy and programme interventions;
- When several evaluations have been done on a given policy.

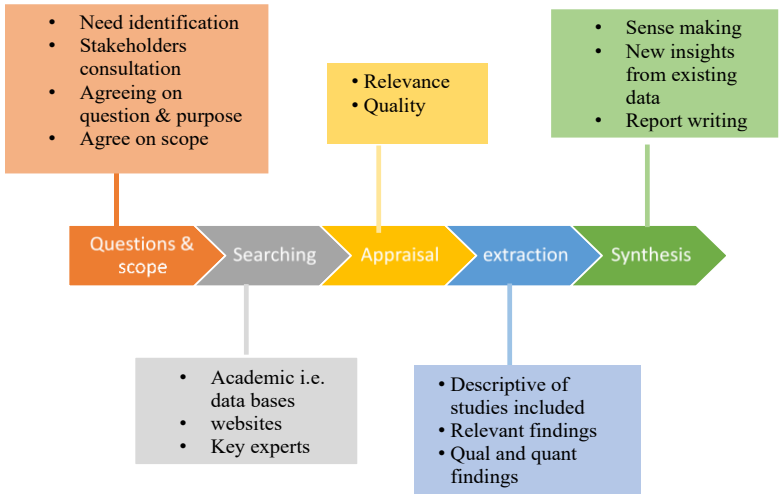
When evaluation synthesis might not be the right instrument:

- Policy areas with a thin evidence base;
- When there is no clear evaluand;
- When there is no specific policy question.

5. What is involved in evaluation synthesis?

Steps involved in evaluation synthesis are articulated in varying ways depending on the nature of the synthesis. However, all syntheses have at least five steps: Planning, searching, appraisal, extraction and synthesis. Figure 1 shows summary of these steps. Each of the step is explained further:

Figure 1: steps involved in evaluation synthesis



5.1 Planning:

5.1.1 Questions and scope clarification

The planning stage involves defining the question and scope of the synthesis. For policy-relevant evaluation synthesis, the development of the question, purpose and scope of the synthesis is an important step. This process should be done in consultation with stakeholders, both within and outside government. The process of developing and agreeing on a question can take a few meetings/workshops with stakeholders. A programme/policy theory of change and results framework can be useful in helping to define the synthesis question. For example, synthesis can ask questions about process/implementation; achievement of results (outcomes) or achievement of programme/policy impacts.

For example

- Impact: How has the provision of fully subsidised housing by government contributed to addressing poverty and building assets for the poor?
- Outcomes: To what extent has the provision of parenting programmes reduced harsh parenting and use of physical punishment by parents and care givers in low income communities in South Africa?

The DAC criteria can also be a useful guide for developing evaluation synthesis questions and sub-questions. Evaluation synthesis can be asked on any of the six criteria. However most primary research/evaluations tend to not consistently report costs related data making efficiency focused evaluation synthesis a challenge. The table two below shows how the DAC criteria can be used to

construct synthesis questions and type of synthesis that can be done depending on the level of rigour required as discussed in section three above.

Table 2: Using DAC criteria to develop synthesis questions

DAC criteria	Focus of the synthesis		Some examples
Policy relevance and appropriateness	Brings together findings from different research/evaluations measuring the extent to which the intervention objectives and design (ToC) respond to needs of beneficiaries, appropriate for the country context and an appropriate response to the problem it seeks to address etc.	• Realist review • Scoping review • Rapid evidence assessment	What international evidence is there of the cumulative effect of state housing investment in the residential property market, and its impact on poverty?
Coherence	Answers questions related to fit of the intervention with other interventions in the county. This could also answer the questions of how well different interventions in the same policy area fit together, i.e. whether they coherently respond to the problem.	• Realist review • Scoping review • Rapid evidence assessment • Evidence map can be a helpful start when what is being implemented by different partners is not well known	In a context of high levels of structural unemployment and high levels of inequality, what delivery path to housing and human settlements are best able to provide a welfare safety net, create a functioning residential property market and leverage people out of poverty?
Effectiveness	Synthesizes findings on the extent to which the intervention/policy is producing effects.	• Systematic review methods applied evaluatively • Realist reviews	Does the provision of parenting programmes reduce harsh parenting and the use of physical punishment in South Africa?
Impact	This question focuses on the difference an intervention is making, focusing on both expected and unintended high-level effects	Systematic review methods applied evaluatively	How has the provision of fully subsidised housing by the government contributed to addressing poverty and building assets for the poor?
Sustainability	Sustainability questions focus on the net benefits of an intervention, the likelihood that an intervention can be sustained. More recently evaluators have also urged the inclusion of Anthropocene concerns and transformation in measures of sustainability.	• Realist review • Rapid evidence assessment	In a context of high levels of structural unemployment and high levels of poverty what is a sustainable path to deliver housing to create a functioning residential property market and leverage people out of poverty?

It is possible for one synthesis to answer more than one question. However, it is important to avoid having too many questions, particularly if the questions address different criteria. Having too many questions can make the synthesis process too complex and reduce precision. Ideally there should be one overarching question, and if sub-questions, they should contribute to answering the main evaluation question and related to one criterion.

The same logic applies to synthesis as a primary programme/policy evaluation in how to ensure ownership of the process and likely application of the result. During the preparation stage, it can be beneficial to establish a steering committee or some reference team involving the main stakeholders,

e.g. representatives from policy units, implementing bodies, relevant civil society organisations, researchers[3].

5.1.2 Scope

Once the policy question has been agreed the scope of the synthesis can be determined. The most common used tool to scope a synthesis is what is called PICOS - Population, Intervention, Comparison, Outcomes and Studies. A PICOS defines what would be included and excluded in the synthesis. In addition to the PICOS, other variables can be added to the inclusion criteria depending on the question and methods being used. Table 3 below shows an example of an inclusion criteria with additional variables. The example is drawn from a Realist Evaluation Synthesis of the South African housing programme that was carried out in 2014 by the Department of Planning Monitoring and Evaluation and the Department of Human Settlements[4]. The question for the synthesis is repeated below:

Synthesis question: How has the provision of fully subsidised housing by the government contributed to addressing poverty and building assets for the poor?

Table 3: Inclusion criteria

Concepts	Operationalisation/definition
Population	Which population group is the synthesis interested in? Defines the type of programme participants/population group in which outcomes and effects are assessed: This could be defined by: • Geographical location-whole country or certain areas in the country? Interested in what is coming from other countries? • Population group-entire population or specific subgroups of interest? *For example:* This synthesis is focused on the South African Government housing programme. The focus is on households defined as poor. The definition is per the Housing policy and refers to South African citizens who earn less than R3500 per household. Specific interest is in areas defined as urban.
Phenomenon	Which phenomena is the synthesis concerned with? Consider how different legislations and policies have named the phenomena of interest. Are there different terms used in the sector to refer to the same social issue of interest in the synthesis? Those can be listed here. *For example:* The phenomena of interest in the evaluation is government housing interventions. This is referred to as • State-provided housing • Subsidised housing • Government housing • RDP housing • Capital subsidy • Free housing

[3] I Goldman and M Pabari. 2020. Lessons for using evidence in policy and practice. https://www.taylorfrancis.com/books/e/9781003007043/chapters/10.4324/9781003007043-13

[4] M M Amisi, M Lochner, and S Cloete Jan, "The Appropriateness of a Realist Review for Evaluating the South African Housing Subsidy Programme," South African Journal of Science 114, no. 11/12 (11/26 2018), accessed 2020/09/30, https://dx.doi.org/10.17159/sajs.2018/4472.

Concepts	Operationalisation/definition
Intervention	What kinds of interventions will be included in the synthesis? What intervention should the primary research/evaluations have evaluated? This could be: • Specific interventions • Government policy/programme or even non-government delivered programmes? *For example:* This synthesis will include research and evaluations that measured the effect of the following interventions in South Africa: State subsidised housing; fully subsidised housing, individual-level capital subsidies, project financing support, Mortgage guarantee fund, building standards, the housing Act and Housing code.
Comparison group	Most formal systematic reviews where there is a synthesis of quantitative data require specification of the comparison group. • What counterfactual should the primary evaluations have used? • Should they have used counterfactual? For qualitative syntheses, it is not necessary/important to use a comparison group *For example:* The comparison includes households living in inadequate housing, this will be households living in informal settlements and households with no access to services.
Outcomes	Which outcomes are of interest? Which outcomes should the primary evaluations have measured? Consider: • Use what is on policy/government plans to define outcomes • Synthesis cannot measure government interventions on outcomes it was not intended to achieve *For example* To be included in this synthesis, evaluations need to have collected data or have assessed the impact of housing interventions on either the primary outcomes or one of the secondary outcomes of interest in this synthesis: • Growth in property value of subsidy and old housing stock • Subsidy houses are incorporated in the residential property market • The poor's share of the value of the property market
Study design	Are there specific designs that the synthesis is particularly interested in? Are there study designs that will be excluded? Note that in most complex social interventions' evaluation designs vary a lot, so be careful about excluding too many designs or being too strict. An important issue is to ensure the quality and rigour of the primary evaluations not solely the designs. Most synthesis, except synthesis of synthesis/review of reviews, will exclude systematic reviews or other synthesis because it can amount to double counting. Literature reviews are also often excluded because they are not evaluations but summaries of literature on a topic. *For example:* In this synthesis, interest is on research or evaluations that measure the outcomes of government interventions on market participation and residential property market value. The synthesis will include all research and evaluations where there is a process of collecting data and analysis to answer a specified question. Research/evaluation will not be excluded on the basis of design if they meet rigour standard consistent with their design. The synthesis will exclude other systematic reviews as this can amount to double counting and literature reviews where there is no systematic research process.
Search process	Indicate how wide the search should be. For formal syntheses such as systematic reviews and where there would be meta-analysis, the search has to be systematic to ensure that all possible studies or evaluations are included.

Concepts	Operationalisation/definition
	For synthesis including qualitative research/evaluations, this process is more flexible but still has to be systematic. For less formal synthesis, i.e. when there are evaluations already done by the ministry and a synthesis is required, the search could be quite restricted. Indicate how far back should the search go? (period)
Type of publications	Will the synthesis exclude primary studies based on publication types? i.e. books, government reports, dissertations *For example:* In the housing example there was no exclusion by publication. A statement that captured this: In this synthesis, all relevant publications will be included. No publication will be excluded by type.

Source: Adapted from Amisi et al[5]

5.2 Searching

All synthesis involves some form of searching for relevant research/evaluations, the difference lies in how the search will be done and the coverage of the search. Broadly there are two types of searching; academic and non-academic. Academic searching of databases is done by information specialists with experience in using a string of search terms. The Cochrane Collaboration has an information retrieval support network for academic searching processes[6]. Non-academic search or grey literature search includes:

- Searching websites of relevant institutions for any relevant evaluations/research. During website search evaluators use concepts/terms in the inclusion criteria (phenomena/intervention) developed in step one to search for relevant information;
- Contacting sector experts and asking them for any work they have done or they are aware of on the topic;
- Backward and forward-searching of reference lists of some seminal work in the sector;
- Physically going to the offices of key officials in ministries or implementing agencies to check their libraries for any relevant evaluations;
- Checking with evaluation/research units within government, where evaluations tend to be unpublished;
- Checking international evaluation databases i.e. 3ie, United Nations Evaluation Group, UNICEF, African Evaluation Database etc.

Some resources:

- https://www.unicef.org/evaluation/reports
- https://www.3ieimpact.org/evidence-hub/impact-evaluation-repository
- https://www.campbellcollaboration.org/better-evidence.html
- http://uneval.org/evaluation/reports
- https://evrd.afdb.org/

[5] MM Amisi et al., Mapping the evidence: South African interventions to prevent violence against women and children, ISS, 2019, https://issafrica.org/crimehub/analysis/research/evidence-map-south-african-interventions-to-prevent-violence-against-women-and-children

[6] https://methods.cochrane.org/irmg/welcome

All accessed evaluations need to be stored in a way that the search process, what was retrieved and decisions about what is included and excluded can be recorded transparently.

In less formal syntheses where there is no formal search process, the same process of recording research/evaluation included and the reasons they were considered sufficient for the question asked and the purpose of the synthesis is still needed.

In most cases, a spreadsheet with all the accessed evaluations is kept. The principle is to be systematic and transparent.

5.3 Screening and appraisal

Screening and appraisal is a two-step process. In the first stage, screening involves screening abstracts or executive summaries of studies/evaluations to see if they are relevant for the synthesis question asked. The inclusion criteria developed in the first stage is used to determine what is relevant and what is not. This makes sure that evaluators do not spend time reading reports of evaluations that are not relevant to the question being asked.

The second aspect of screening involves a process of judging the quality of primary evaluation reports. There are contestations about when this should be done and how useful it is. However, it is important that synthesis is drawn from studies of good enough quality. In most policy evaluation syntheses, it is possible that primary studies will vary significantly in size and type. It is important that they meet quality standards for their design. This process should ideally be done before the process of extracting findings from primary studies.

Quality appraisal differs for qualitative and quantitative studies. The table below from Hannes K, 2011 illustrates the differences:

Table 4: Quality appraisal in qualitative and quantitative studies

Aspect	**Qualitative Term**	**Quantitative Term**
Truth value	Credibility	Internal Validity
Applicability	Transferability	External validity or generalisability
Consistency	Dependability	Reliability
Neutrality	Confirmability	Objectivity

There are also different tools for quality appraisal:

Table 5: different appraisal tools

Source	**What it covers**
Popay, Rogers & Williams (1998)	The primary question relates to the appropriateness of the methods used. This is followed by a detailed assessment of methodological soundness
Critical Appraisals skills programme (1998)	10 questions relating to rigour, credibility, and relevance

Quality framework (2003)	18 questions relating to key areas: Findings: design: sample: data collection: Analysis: reporting: reflexivity and neutrality: ethics and auditability
Prompts for appraising qualitative research (2004)	A generic set of prompts relating to aspects of reporting and aspects of study design and execution
Long and Godfrey (2004)	A tool to explore descriptive and evaluative elements of a study. 34 questions across 4 key areas: phenomenon studied and context; ethics: data collection: analysis and potential researcher bias; policy and practise implications
Waish and Downe (2006)	Set of prompts relating to 8 key areas: scope and purpose: design: sampling strategy: analysis: Interpretation: reflexivity: ethical dimensions: relevance and transferability

Source: York.ac.uk[7]

5.4 Extraction

Once the selection of studies that are considered relevant and of acceptable quality has been made, extraction of data can be done. Whether synthesizing qualitative and quantitative primary research/evaluation extraction of data needs to be done. Extraction is the process of "collecting" data from the primary studies. Remember, in synthesis primary research/evaluation reports are the sample. To facilitate comparable extraction of data from different research/evaluations evaluators need to develop an extraction tool relevant to the synthesis. An extraction tool is like a data collection tool. Each evaluation is read in full and relevant data extracted. This process should be managed by a seasoned evaluator but it can be done by a junior evaluator with enough understanding of the methods or the policy being evaluated.

Data extraction for qualitative and quantitative primary research/evaluations differ. However, there is standard data that should be extracted in all synthesis. This can include:

- Descriptive data about the primary study
- Intervention/s evaluated
- Outcomes measured
- Evaluation/research questions asked
- Sample size
- Types of methods used
- When the evaluation/research was done
- Where was the research/evaluation done?
- Who authored the evaluation?

What is extracted for different studies:

Table 6: Extraction in qualitative and quantitative data

Qualitative	**Quantitative** [8]

[7] https://www.york.ac.uk/crd/SysRev/!SSL!/WebHelp/6_4_ASSESSMENT_OF_QUALITATIVE_RESEARCH.htm
[8] Additional resources for extraction https://guides.library.cornell.edu/evidence-synthesis/data-extraction

Can be challenging because it is not standardised	Fairly standardised
Quotes from respondent	Effect sizes
Concepts	Mean

5.5 Analysis and synthesis

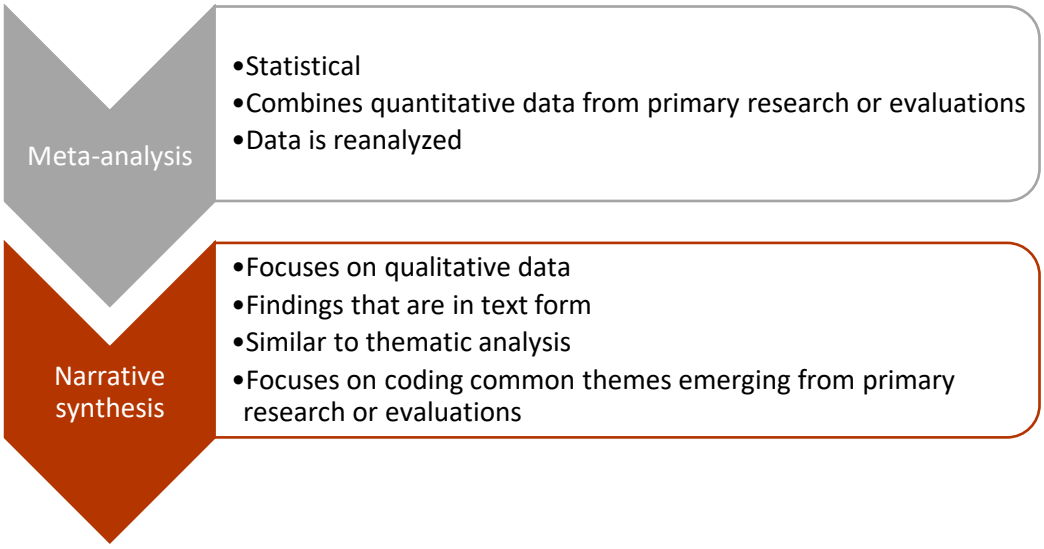
During this stage, data that have been extracted are reanalysed in view of the question being asked. Like with other evaluations, an analysis plan is needed to make sure that analysis is done in a rigorous manner. A theory of change can be a useful framework for analysing and making sense of findings from different evaluations/studies. An important thing to bear in mind is that synthesis is not a summary. The table below differentiates summary from a synthesis.

Table 7: The difference between summary and synthesis

Summary	**Synthesis**
Basic reading technique	Advanced research process
Pulls together information to highlight important points	In addition to highlighting important findings, synthesis adds evaluator's own conclusions
Restates/re articulates findings from different studies	Combines and contrasts findings from different research/evaluations
Shows what original authors wrote	In addition to reflecting original findings, in synthesis evaluators also make new findings/discoveries
Addresses one set of information at a time, and each study/source remains distinct	Involves combining parts and elements from different research/evaluations. The evaluator combines findings from different studies. The primary studies do not always remain distinct
Presents overview	Focuses on main ideas, details, findings and also new meanings
Focuses on overall findings from primary studies	Will examine the data in primary studies

The analysis and synthesis differ for qualitative and quantitative data as shown in the figure two below.

Figure 2: Meta-analysis and narrative synthesis



For qualitative data, synthesis process follows the same principles of qualitative data analysis in primary evaluations. Evaluators code new themes emerging from the data that has been extracted and reanalyse. Evaluators can use qualitative data management software for analysis such as NVIVO, Atlast Ti, etc. The process involves:

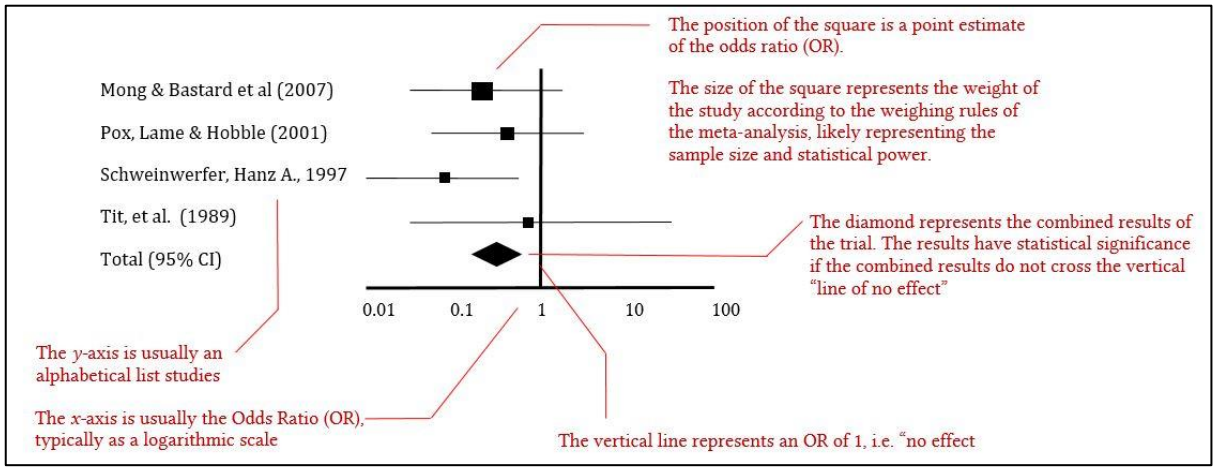
- Coding descriptive themes from primary studies-closest to what is in primary studies;
- Developing analytical themes-new themes emerging from reading across studies;
- Interpreting the data and generating new insight/meanings using the analytical plan and responding to the synthesis question.

For quantitative data, synthesis process usually involves meta-analysis. A meta-analysis is "a statistical procedure that integrates the results of several independent studies considered to be combinable." [9]. A forest plot is generated from the meta-analysis showing estimate of the size of the effect of combined studies. Individual primary research/evaluations are weighted to account for the difference in precision, for example, research/evaluations with smaller sample sizes have less weight in the meta-analysis. Where there is a definite intervention whose effect has been quantitatively measured in different evaluations, meta-analysis allows evaluators to arrive at the best possible understanding of the effect of intervention from existing studies. Figure three below shows an example of a forest plot.

Some synthesis around complex policy interventions will do both qualitative and quantitative analysis. In that case, they will apply both the qualitative extraction and analysis methods for the qualitative data and quantitative procedures for quantitative data.

Figure 3: Example of a forest plot

[9] Egger et al, BMJ 1997



Source: Deranged Physiology [10]

Once data has been analysed and new insights generated it can be reported answering the specific evaluation questions the synthesis pose. A generic report structure for an evaluation synthesis is attached in annexure b. Each report will differ depending on the synthesis but also on the purpose it was done. However, the annex gives a general outline to consider what could be included in the report.

6. Possible steps involved if outsourcing

Most syntheses done in the public sector are outsourced to external evaluators. Table 8 below shows the steps that would be involved if the evaluation synthesis is outsourced.

Table 8: possible steps involves in commissioning synthesis

Step in the process	What is involved	Possible output
Step 1: Initiation	Consult stakeholders, agree on the need for synthesis and establish project management mechanisms	• Agreement with key stakeholders • Established reference team/steering committee
Step 2: Questions and scope	Hold workshops with stakeholders to develop questions, consider what are key policy questions that policy actors and other stakeholders have	Agreement on questions, purpose and scope (inclusion criteria)
Step 3: Developing ToRs (include the inclusion criteria)	Use questions and scope to develop ToRs that expand what the role of external evaluators would be, what expertise they need to have in their teams, how long the project would take and the project costs. In rapid synthesis, initiation, questions and ToRs might be developed concurrently.	Approved ToRs
Step 4: Appointing the team	What is involved in this step is determined by rules that govern	Team appointed. Ideally, the team should have an evaluator

[10] https://derangedphysiology.com/main/sites/default/files/sites/default/files/CICM%20Primary%20Exam/A-%20Research%20methods/Anatomy%20of%20the%20forest%20plot.JPG

	procurement in an organisation. However, this should include a call for proposal, assessment of the proposal and ideally, it should also involve some interviews with shortlisted candidates. It is important to have more than one individual responsible for assessing the quality of the proposals.	and programme/policy/sector expert. For more formal synthesis an information specialist might be needed
Step 5: Project inception	Once a team has been appointed, an inception meeting can take place. The appointed team should meet with the reference team to present how they are going to approach the synthesis. Note that it is possible for the appointed team to make additional suggestions on the inclusion criteria or the overall approach.	Project officially initiated
Step 6: Searching	The breadth of the search would have been determined in earlier stages. A possible set of search terms are developed depending on the nature of the synthesis. What is involved in searching is detailed in section 5.2. If a synthesis is outsourced, the involvement of government staff in searching can help build capacity but also improve the quality of grey literature search	Relevant studies/evaluation identified
Step 7: Screening and appraisal	It is possible that the search process even if narrow, will find irrelevant material. Synthesis requires a process of excluding irrelevant studies and checking quality. This process should be pragmatic, guided by question and purpose of the synthesis.	Final list of relevant studies/evaluation identified
Step 8: data extraction	Data extraction tools developed and data extracted. It is important that policy actors are involved in the process of developing the extraction tool. Even if it is only to offer guidance on what could be important to extract.	Data extracted
Step 9: Analysis	What is involved in analysis is determined by the type of synthesis. It is important that the project manager and some policy people get involved in this step. Analysis should focus on explaining the findings not only summarising the findings. Emerging findings and analysis can also be presented to the steering committee.	Extracted data analysed

Report writing	The report should have policy-relevant findings and clear policy implications, not a retelling of primary studies.	Evaluation report

7. Planning an evaluation synthesis

When considering an evaluation synthesis there are some important things to remember:

- Evaluation synthesis applies research synthesis methods, however, it does so evaluatively. Using a theory of change and/or evaluation criteria helps evaluation synthesis to combine findings from different individual research and evaluation reports in an evaluative way. Evaluation synthesis should answer the "why" question not just the aggregate effect size of interventions.
- For policy-relevant synthesis to inform policymakers and programme implementers, it is important to Involve stakeholders earlier in the definition of questions and inclusion criteria. The question needs to be developed collaboratively with those who are likely to use the information. This ensures that the right questions are asked and the right parameters set for the synthesis.
- The evaluation synthesis process has to be transparent and pragmatic. Even when outsourced the appointed evaluation team has to keep records which project managers in government and other evaluators should have access to if needed. Also, though this guide outlines steps for the synthesis and all are important in the process, those doing synthesis need to apply the steps as determined by the nature of the synthesis, resources available and time pressure. Choices can be made about what is crucial for the quality of the synthesis and what is not. The important thing is to document decisions made so as to be transparent;
- Syntheses do not need a large number of studies/evaluations. It is possible to do a good quality synthesis with a few research/evaluations if that is what exists on a given question. The important thing about synthesis is not so much the number of primary research/evaluation included-though it is important to make sure that all research/evaluations that exists and qualify to be in the synthesis are included- but generating new insights by combing findings of the different studies.
- In most instances doing a good quality synthesis will require a team including if possible, an information specialist (especially if the intention is to do a formal synthesis such as systematic review). A team is also needed for evaluation synthesis of complex social intervention; even rapid syntheses are best done by a team of at least three evaluators.

8. Limitations

All forms of evaluation or research processes have limitations. The following are limitations to keep in mind when considering using synthesis:

Quality of the evaluation synthesis is dependent on quality and reporting of primary research/evaluations. Researchers/evaluators of primary research and evaluation do not always report all important information that might be needed for synthesis. This is particularly a challenge when synthesizing qualitative primary research/evaluations. Evaluators tend to present qualitative findings in different ways. Others use verbatim quotes from sources while others tend to summarise

data from different sources to make an argument. What is considered data in qualitative evaluation is also not standard. This can be a challenge when trying to synthesize because data can be presented in different forms in different primary research/evaluations. Similar challenge is experienced with syntheses answering implementation questions, few primary research/evaluations report implementation findings in details. This can make synthesis a challenge.

Formal syntheses, like systematic reviews, can be quite restrictive. They follow defined steps requiring a protocol to be approved and once approved, for researchers/evaluators to minimise deviations from the protocol. This approach works well in clinical research where there is minimal disparity in methods used and how results are reported. However outside of clinical research context, such restrictive approach is not feasible. Therefore, there is limitation to where formal synthesis methods can be applied.

The quality of evaluation synthesis also relies on insights and interpretive capability of evaluators. This is an issue for synthesis because evaluators rely on what other evaluators have reported and do not have opportunity to interact with programme/policy implementers and beneficiaries, or observe where policy/programme has been implemented to gain additional insights from different sources of data. To strengthen interpretation of existing research/evaluation, it is important to have more than one evaluator working on the evaluation synthesis. In addition, it is important to have views and experiences of policy actors and programme implementers included in the analysis/sense making process. This can be done for example by holding workshop for evaluators to interact with the evaluation steering committee on key evaluation milestones i.e finalising inclusion criteria, analysis and report writing.

Annexure a-example of extraction tool

Human Settlements Synthesis

Data Extraction Tool (version 2.0)

CHECKLIST

Document details	**Author:**		**Date:**	November 2013
Document title:				
Document code:		**Document type:**	6 -	

Document Types:

1. *Design & Implementation Evaluation*
2. *Implementation Evaluation*
3. *Implementation and Impact Evaluation*
4. *Impact Evaluation*
5. *Rapid Assessment / Appraisal*
6. *Performance & Expenditure Review*
7. *Research Report*
8. *Other*

Assessment Key: ✓= Reported ; ✗ = Not reported; n/a = Not available

Checklist	**Assessment**	**Comments**
1. Clearly stated aims & objectives		•
2. Study design adequately described		•
3. Appropriate research methods		•
4. Appropriate use of instruments (reliability and validity)		•
5. Adequate description of source of data or sample, inclusion / exclusion criteria, response rates etc.		•
6. Appropriate analyses (statistical or qualitative)		•
7. Results clear and adequately reported		•
8. Discussion of results reported in light of study, question & relevant literature		•
9. Limitations of research & design discussed		•
10. Implications of research discussed		•

11. Any recommendations		•

THEMATIC REVIEW

Theme	Question Level 1	Question Level 2	What findings are presented?	What evidence is provided?	Quotations	Page ref
A. Analysis of the programme logics (incl. context, mechanisms and outcomes).	A.1 What was to be implemented?	A.1.1 What was the context for the programme?				
		A.1.2 What was the problem statement the programme aimed to address?				
		A.1.3 What were the envisaged programme inputs?				
		A.1.4 What were the envisaged programme outputs?				
		A.1.5 What were the envisaged programme outcomes?				
		A.1.6 What was the envisaged programme impact?				
	A.2 What delivery mechanism was envisaged??					
	A.3 What institutional arrangements were envisaged?					
	A.4 What funding structure and sources was envisaged?					
	A.5 What was the envisaged role of planning in the programme?					

Theme	Question Level 1	Question Level 2	What findings are presented?	What evidence is provided?	Quotations	Page ref
	A.6 What were the key programme assumptions?					
	A.7 Which components of the overall theory of change are impacted on by the programme?					
	A.8 Does the proposed overall theory of change capture the essential elements of the programme, and if not, should it be refined?					
	A.9 What conclusions are drawn on the programme logic and are these justified by the evidence?					
	A.10 What recommendations are made on the Programme logic / design and do these follow logically from the conclusions?					

Theme	Question Level 1	Question Level 2	What findings are presented?	What evidence is provided?	Quotations	Page ref
B. Analysis of the actual programme implementation (the way in which the programme was implemented as opposed to the programme logic). Must cover context, mechanisms and outcomes	B.1 What was actually implemented?	B.1.1 Within what context was the programme implemented?				
		B.1.2 What key changes in the implementation against the original programme logic were noted?				
		B.1.3 Why did these changes occur?				
		B.1.4 What outputs were achieved (for whom, in what circumstances, in what respects and why?)				
		B.1.5 What outcomes were achieved (for whom, in what circumstances, in what respects and why?)				
	B.2 What institutional arrangements were implemented /utilised?					
	B.3 What funding structure and sources were utilised?					
	B.4 What was the role of planning in the programme?					
	B.5 What key assumptions were not					

Theme	Question Level 1	Question Level 2	What findings are presented?	What evidence is provided?	Quotations	Page ref
	evident / what other critical assumptions are noted?					
	B.6 What conclusions are drawn and are these justified by the evidence?					
	B.7 What recommendations are made and do these follow logically from the conclusions?					
C. The role of the Programme in the overall theory of change	C.1 What is the evidence about the underlying assumptions of the overall theory of change					
	C.2 Does the proposed overall theory of change capture the essential elements of the programme as implemented, and if not, should it be refined?					
	C.3 On the basis of the conclusions and recommendations how should the overall theory of change be revised?					

Annexure b-generic report structure

Possible report structure

1. Introduction
2. Background to the synthesis
3. Understanding the intervention
    a. Background to the intervention (policy/programme)
    b. Theory of change
4. Methods followed
    a. Evaluation questions
    b. Inclusion criteria
    c. Search process
    d. Screening and appraisal
    e. Extraction process
    f. Team roles
    g. Analytical framework
5. Findings
    a. Could use the evaluation criteria to organise findings or key questions asked
6. Discussion
    a. Present new insights from the synthesis
    b. Could use the theory of change to structure discussion
7. Conclusions