

# Conducting a Literature Review

A step-by-step guide

---



EVANS SCHOOL OF PUBLIC POLICY & GOVERNANCE

UNIVERSITY *of* WASHINGTON

Evans School Policy Analysis and Research (EPAR)

# Goals of a Literature Review

A literature review can be conducted to:

- **Assess** an evidence base;
- **Identify** research gaps; and/or
- **Synthesize and document** the base of existing knowledge so that the contribution of new work is evident.

*For example, in 2015, EPAR conducted a literature review ([EPAR Technical Report #293](#)) to answer the research question: “What is the evidence in the scholarly peer-reviewed literature of the effect of morbidity on economic growth?” The goal of this review was to assess the evidence base on the effects of morbidity and economic growth, and summarize the pathways presented in the literature.*

# EPAR's Approach

The goal in following a rigorous review process is to have some level of certainty that you have **identified the relevant literature**, and that you are **collecting from the literature the information necessary to rigorously evaluate the state of current knowledge** relating to your research question(s).

To increase our confidence in the results of literature reviews, EPAR emphasizes three key components of the literature review process:

- building the sample of studies to review
- developing a review framework and systematically extracting information from the sample of studies using a coding spreadsheet
- using the coding spreadsheet to facilitate analysis of the evidence base

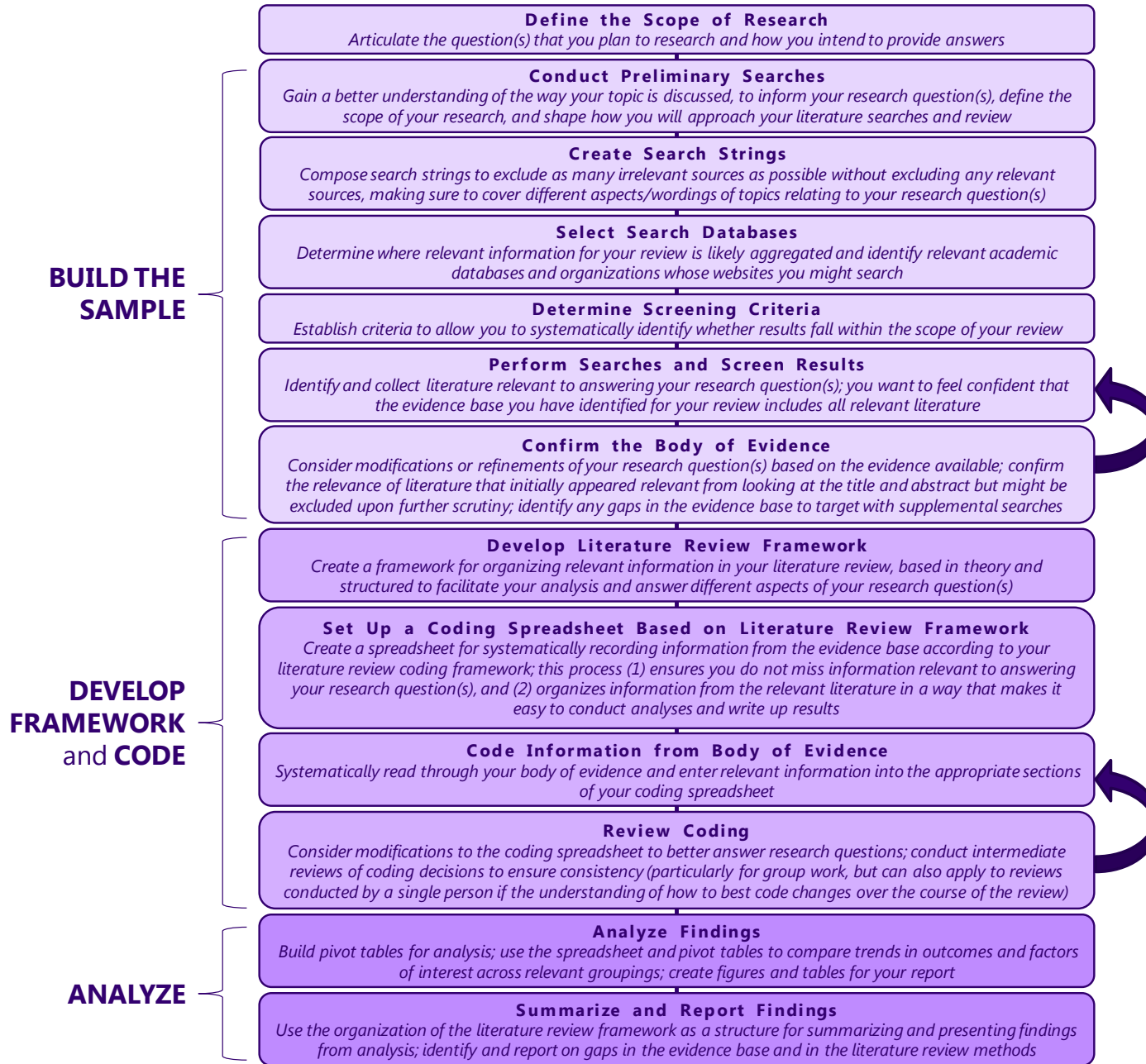
# Results

---

The intended results of a rigorous literature review are:

- **Transparency** around the scope of your research and your review methods;
- **Confidence** (among both the researchers and your audience) in your clearly defined body of evidence and the integrity of any findings you pull from it;
- **Improved teamwork** - with everyone following the same “system” in identifying, coding, and analyzing evidence;
- A **spreadsheet of evidence** from the literature coded according to a well-organized review framework, covering key aspects of the theory relevant to your research question(s); and
- A document presenting findings from your **analysis** (and relevant tables/graphics) that answers your research question(s) and identifies gaps in the evidence base (or in your review methods).

# Overview of the Literature Review Process



# Define Scope of Research (1/3)

**The first step in conducting a literature review is to articulate the question(s) that you plan to research and how you intend to provide answers**

- Spending time on this step will reduce work later stemming from changes to the questions, scope, or framework.
- You should be able to clearly articulate the contribution of your literature review to answering your research questions.

## **1. Formulate your general research question(s). To start, consider:**

- a. What do you want to be able to answer from a review of the literature?
- b. Are there any specific sub-questions you want to be able to answer?

*Example general research questions:*

- What is the evidence in the scholarly peer-reviewed literature of the effect of morbidity on economic growth? ([EPAR Technical Report #293](#))
- What is the evidence of the long-term impacts and cost-effectiveness of cash transfer programs in low- and lower middle-income countries? ([EPAR Technical Report #359](#))

# Define Scope of Research (2/3)

## 2. Get specific by articulating the following to define the scope of your research :

- a. Types of literature to review (e.g., grey (nonacademic) or unpublished literature)
- b. Years of literature to review (e.g., whether to only include recent literature)
- c. Study methods (e.g., only studies with experimental or quasi-experimental designs)
- d. Geographies of interest
- e. Interventions/policies of interest
- f. Outcomes/indicators of interest
- g. Populations of interest (e.g. rural vs. urban, farmers vs. non-farmers, etc.)
- h. Other considerations

*Example definition of research scope:*

[EPAR Technical Report #359](#) only included:

- peer-reviewed journal articles, grey literature, and unpublished literature
- only studies from 2010 or later
- only review studies and meta-analyses, or studies with experimental or quasi-experimental design
- literature on cash transfer programs and universal basic income programs

It focused on long-term impacts (sustained impacts of short-term outcomes or outcomes that are linked to cash transfers via “downstream pathways”) in relation to: general and reproductive health, nutrition, labor market participation, poverty, gender and intra-household decision-making, financial inclusion.

# Define Scope of Research (3/3)

*Example research questions and scope ([EPAR Technical Report #293](#)):*

**Research questions:** We plan to conduct a review of the literature linking morbidity to growth. Specifically, the research questions are as follows:

1. What is included in standard measures of morbidity? How are they calculated? Does cognition factor in?
2. What is the evidence that assesses the linkages between morbidity and growth at the individual/household level (e.g. cognitive impairment to less education to lower household income)?
3. What is the evidence that assesses the linkages between morbidity and growth at the macro/country level?
4. How robust is the evidence on these linkages overall and what is the quality of this evidence? What are the unanswered, or unsatisfactorily answered, questions?

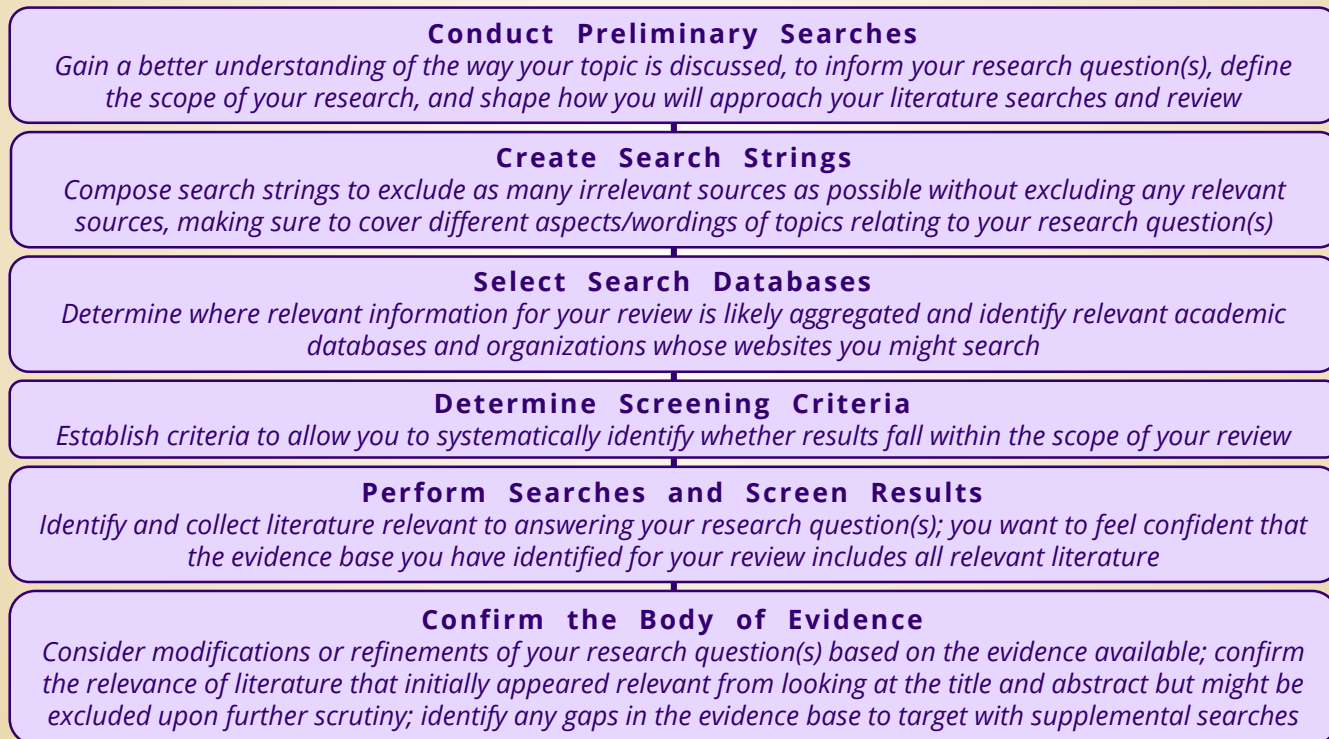
**Scope of research:** This will be an academic literature review and assessment of existing studies on morbidity and economic growth.

- Only include studies published in peer-reviewed journals
- Morbidity refers to the overall burden of disease, including both measures of the incidence or prevalence of a specific disease as well as broader measures of more general disease-related health conditions



# Build the Sample

EPAR follows six steps in the study search and screening process, refining our research question(s) as we go:



# Conduct Preliminary Searches (1/2)

The goal of preliminary searches is to gain a better understanding of the way your topic is discussed, inform and refine your research question(s), further define the scope of your research, and shape how you will approach your literature searches and review

1. **Conduct a few exploratory searches in Google and Google Scholar, using general terms related to your research question.**
  - a. Identify relevant background documents
  - b. Review any recent, highly-cited articles related to the topic (as available) and relevant theoretical pieces
2. **Read through the top results for a couple different searches, paying particular attention to the language used to define and describe your topic.**
  - a. Take notes on relevant findings to inform your research
  - b. What key terms are used in relation to your research topic?
  - c. Are there particular issues you had not considered?
3. **Go back and modify your research question and scope as useful.**

## Conduct Preliminary Searches (2/2)

While reading, consider how you might adapt your forthcoming work to the way the research is structured. This includes thinking about:

- Theory: Note explicit or implicit theories that link activities or interventions to outcomes whenever possible. Break down this theory into its component assumptions and hypotheses, as these inform the information you seek to draw from the literature.
  - For example, in reviewing the long-term impacts of cash transfers, it may be relevant to consider whether impacts of the transfers vary by program characteristics such as the amount or duration of the transfers.
- Language used: Pay attention to terms and definitions to inform your search strings. For example:
  - Measures of economic growth can include GDP, GNI, wages, income, and productivity.
  - Measures of morbidity include incidence, prevalence, life expectancy, and specific disease counts or indexes .

*Example questions targeted in preliminary searches for [EPAR Technical Report #293](#):*

What are the various measures of morbidity (such as DALYs) and of growth (such as GDP)?  
What are the hypothesized pathways between morbidity and growth, such as the effect of malnutrition on wages and productivity?

# Create Search Strings (1/4)

The goal when creating search strings is to exclude as many irrelevant sources as possible without excluding relevant sources, while covering different aspects/wordings of topics relating to your research question(s).

1. **Develop search string keywords based on your research question(s) and preliminary searches. Think carefully about your terms, making sure they relate specifically to your research questions.**
2. **Try out a few different versions of your search string and scan the first ~10 results, to see which strings produce the greatest proportion of relevant results. Try including different versions of your key terms.**
3. **Based on your initial searches, choose the search string that you will use with each database (your primary initial search string).**
  - a. If your research question(s) may be best answered by separate searches, develop separate strings targeting each aspect of your question specifically.
  - b. You may choose to conduct targeted supplemental searches if you find gaps in your sample of studies after your first round of searches and screening.

Example search string for [EPAR Technical Report #293](#): ("*Healthy life expectancy*" OR HALE\*) AND ("*economic growth*" OR GDP OR GNI OR wage\* OR income\* OR productiv\*)

Other keywords used included: morbidity, health, disease, Tuberculosis, Rotavirus, Malaria, Polio, HIV, disease index, life expectancy, self-reported health, Healthy Life Expectancy (HALE), Disability Adjusted Life Years (DALY), Quality Adjusted Life Year (QALY), Years of Life Lost (YLL), comorbidity

## Create Search Strings (2/4)

- In a code-based database, you can:
  - Use the Boolean operators “AND” “OR” “AND NOT” to increase search accuracy.
  - You may want to run a search string replacing “AND” (more restricted results) with “OR” (less restricted results) to see if any new relevant sources come up.
  - Use an asterisk at the end of a word to capture different variations of the same word—“evaluat\*” returns results with “evaluate,” “evaluation,” and “evaluating.”
  - If relevant, use proximity operators such as “W/n”, which restricts to n words between two words. This is helpful for exploring the connection between two distinct concepts, e.g., “mobile money” and “regulation.” Rules of thumb:
    - Find terms in the same phrase: W/3, W/4, or W/5
    - Find terms in the same sentence: W/15
    - Find terms in the same paragraph: W/50

Example from [EPAR Technical Report #339](#):

We searched for information on private company research on orphan crops, including taro. We included *NOT “Taro Pharmaceuticals”* in the search string for taro to exclude irrelevant information about a popular pharmaceutical company that had nothing to do with the crop.

## Create Search Strings (3/4)

- **One carefully chosen, all-encompassing search string reduces the number of times you may have to screen the same search results, as opposed to conducting multiple searches with different terms.**
  - Example: Use (“multilateral aid” AND “bilateral aid”) AND (evaluation OR analysis OR effectiveness OR evidence OR impact) rather than separate search strings covering the different elements within the parentheses.
  - You will still need to check for duplicate results from running your searches in different databases. To identify duplicate search results after conducting searches with multiple strings or across different databases, EPAR puts all of the results in a single spreadsheet, and uses conditional formatting to “Highlight Duplicates” among titles.
- **Multiple search strings may be more effective for targeted searches that focus on different aspects of your research scope.**
  - Example: If you are searching for evidence for a defined set of countries, you might conduct separate searches adding a country name to the end of your initial search string, to ensure you do not miss relevant results for those countries.
- **More “general” search strings may return a greater variety of relevant results and save time from conducting multiple searches, but also bring up a greater proportion of non-relevant results. Test out your search strings and consider this trade-off.**

## Create Search Strings (4/4)

- **You may need different search strings for Google/Google Scholar.**
  - These engines generally return more results than the academic databases, complicating the initial process of identifying relevant sources.
  - They also limit the length of your string to 32 words in Google or 256 characters in Google Scholar, so limit your Boolean search terms.
- **Document your search process. You will need to describe it in the “Methods” of your final product and be prepared to justify your choice of search string(s)/term(s).**

*Example search strings for [EPAR Technical Report #293](#):*

Search Scopus with the following search strings:

- (HIV\*) AND ("economic growth" OR GDP OR GNI OR wage\* OR income\* OR productiv\*)
- ("Healthy life expectancy" OR HALE\*) AND ("economic growth" OR GDP OR GNI OR wage\* OR income\* OR productiv\*)

Search Google Scholar with the following search strings:

- HIV ("economic growth" OR GDP OR GNI OR wage OR income OR productiv)
- ("Healthy life expectancy" OR HALE\*) ("economic growth" OR GDP OR GNI OR wage OR income OR productiv)

Note that Google Scholar automatically inserts “AND” between every word/phrase unless OR is specifically there, and it gives words at the beginning of your string more weight/higher priority.

# Select Search Databases (1/4)

The goal in selecting databases is to identify academic databases and organizations' websites where relevant information for your review is likely aggregated.

## 1. Decide what type of literature you need to include:

- a. *Academic Literature*: Published studies have likely been peer-reviewed and may therefore be more reliable sources. Some databases include unpublished studies that may be more recent, but have not yet undergone journal peer review.
  - i. The UW Library System draws from multiple academic databases.
  - ii. Researchers without access to academic databases will be restricted to studies that are published as open access.
- b. *Grey Literature*: You may want your review to include grey literature if there is significant non-academic research on your topic. If so, identify relevant organizations/ institutions conducting research related to your questions, and search their websites.
  - i. EPAR frequently searches websites of the following organizations: Abdul Latif Jameel Poverty Action Lab (J-PAL), Consultative Group to Assist the Poor (CGAP), Center for Global Development (CGDEV), Food and Agriculture Organization (FAO), World Bank, International Food Policy Research Institute (IFPRI), and any Consultative Group for International Agricultural Research (CGIAR) or UN institutions relevant to the topic.
  - ii. Grey literature databases include Open-grey, ProQuest Dissertation and Theses Database, and the Networked Digital Library of Theses and Dissertations.
  - iii. Google searches may also provide relevant results from the grey literature, which can help to address gaps in published evidence.



# Select Search Databases (2/4)

## 2. Select the databases that are relevant to your topic. As general guidelines:

- a. Undertake the primary search using one or more academic databases using [Boolean operators](#) (e.g., Scopus, PubMed, EconLit, PAIS, Cochrane Library, etc.).
  - i. EPAR literature reviews usually include searches in Scopus, a multi-disciplinary database of published academic studies. Scopus allows users to export search results to a spreadsheet, facilitating screening and documenting the process.
  - ii. Your choice of academic databases should be informed by the relevance of their disciplinary focus to your research question(s). For example, PubMed focuses on topics in health.
- b. Use Google Scholar as a check on academic search results, using the same search terms to look for any well-cited papers missed by more systematic databases.
  - i. All EPAR literature reviews include Google Scholar searches as it includes a wide variety of publishing and unpublished academic studies. Results not published in peer-reviewed journals may not be included in academic databases.
  - ii. Google Scholar is especially useful for finding new research (i.e., the last 1-2 yrs)
- c. Search selected organizations' websites to identify relevant grey literature.
- d. Following all literature searches, conduct a general Google search of your search string and review the first ~50 results as a final check to see whether the previous searches have missed any key results.
  - i. This step may be particularly useful in identifying recent research relevant to your question(s) that may not yet be included in other search databases.
  - ii. Google supports two basic Boolean commands: (1) AND, (2) OR.

## Select Search Databases (3/4)

- **Database selection should be informed by the discipline of your research.**
  - Depending on the research question(s), different discipline-specific databases may have information on different aspects related to the scope of the review. Searching multiple databases will limit the risk of missing relevant literature.
  - EPAR often searches Scopus as the primary database and always conducts searches in Google Scholar.
- **Databases have different strengths and weaknesses. For example:**
  - Some databases allow you to export and download a spreadsheet of research results (e.g., Scopus does but Google Scholar does not). This is helpful for systematically tracking your searches and screening, and allows you to go back to look at your results again later if needed. For other databases, you need to record sources manually. When possible, EPAR exports search results for review.
  - Some allow you to search title, author and abstract, but others only permit searching of the title, or some other field. Narrowing the search to certain parts of the document may help increase the likelihood that your search results are relevant to your review.
  - Some databases include literature behind a paywall (not published “open access”). It is up to you to decide whether the material is relevant enough to warrant paying for it.
  - Google Scholar only allows viewing of the first 100 pages of results.

## Select Search Databases (4/4)

- If available, you might choose to consult an academic or public librarian for advice on conducting your searches.
- UW Political Science and Public Affairs Librarian [Emily Keller](#) can help identify potentially relevant places to search.
  - Her webpage includes links to library guides for “Political Science and Public Policy” and “International Studies”.
  - Her “[Lit Review Toolkit](#)” includes links to specialized lit review resources and grey literature searches, including a Google blog search and a custom Google search engine for NGOs (under “Discipline-specific tools”).
- **[Databases for EPAR Literature Search](#)**
  - While this list is EPAR-specific, it will be relevant to other individuals/groups conducting international development research. It also provides a useful example of what a database list might look like more generally.
- ***Again, always conduct a final set of searches in Google to make sure you didn't miss any key results!***

Example from [EPAR Technical Report #293](#):

We conducted supplementary searches using Google Scholar to identify any well-cited literature that may have been missed through the Scopus keywords. These searches yielded an additional 106 studies that seemed relevant from an initial screening of the title and abstract, which were narrowed down to 43 relevant empirical studies upon further screening.

## Example Search String and Database Selection

Example from [EPAR Technical Report #293](#):

- We only reviewed **academic peer-reviewed literature**. We used the Scopus academic database, supplemented by searches in Google Scholar. We chose Scopus because it is a multi-disciplinary academic database with useful tools for filtering search results and for exporting search results to a spreadsheet for documentation and screening.
- We used **Boolean search strings** (with operators AND, OR, NOT, quotation marks "" that surround exact phrases to be searched, and parentheses () that combine terms and operators), based on each of the different measures of morbidity and keywords including economic growth: e.g. ("life expectancy" OR "years lived with disability") AND (GDP OR income OR "economic growth")
- We initially put terms like "RCT" and "randomized control trial" and "experimental" in our search strings to try to focus on studies that might allow for causality to be assessed. So few studies emerged, however, that we included more **general terms** like "empirical" and expanded the acceptable methodologies to "quasi-experimental" in our search strings.
- Initial Scopus searches using **broad keywords** such as "health" and "disease" alongside economic growth indicators returned 7,331 results for "health" and economic growth, and 2,231 results for "disease" and economic growth.
- A **narrower search** using specific morbidity search terms returned 3,948 items of published literature across multiple searches. For example, one search focusing on DALYs as a measure of morbidity used the following search string: ("Disability Adjusted Life Years" OR DALY\*) AND ("economic growth" OR GDP OR GNI OR wage\* OR income\* OR productiv\*)

# Determine Screening Criteria (1/2)

The goal in determining screening criteria is to allow you to systematically identify whether results fall within the scope of your review

1. **Determine screening criteria based on your research question and scope and logic for excluding irrelevant sources.**
  - a. “Relevance” should be defined clearly in writing at the outset and could be captured by several criteria, usually based on your review framework.
2. **Outline your screening criteria as a list of Yes/No questions that you can quickly go through for each search result**
3. **The following are helpful general screening criteria to apply to search results, depending on your research scope:**
  - a. *Geography.* e.g., only include results from Africa
  - b. *Date.* e.g., exclude results from before 2010 (with the goal of prioritizing more recent studies, which may cite or discuss older studies)
  - c. *Study type.* e.g., exclude results that are not published in peer-reviewed journals, or that are not experimental (e.g., qualitative studies)
  - d. *Empirical studies.* e.g., exclude results that only discuss theory
  - e. *Availability.* e.g., source is available online for free in full text.
    - i. If you have access to an academic library, you might be able to request an article scan.
  - f. *Language.* e.g., source is written in English (this criterion is usually implicit).

## Determine Screening Criteria (2/2)

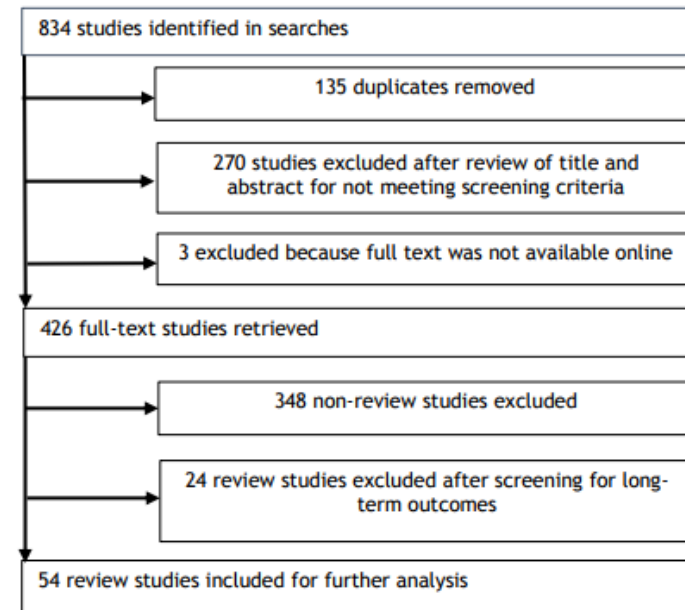
Example screening criteria from [EPAR Technical Report #359](#):

- *Type of study*: Is the study a literature review or meta-analysis?
- *Empirical evidence*: Does it include empirical evidence on the impacts of cash transfer programs?
- *Outcomes/indicators discussed*: Does it include evidence on general or reproductive health, nutrition, labor market participation, poverty, gender and intra-household dynamics, or financial inclusion?
- *Nature of outcomes/indicators*: Does it report on long-term impacts of cash transfer programs on one of these outcome areas?
- *Geography*: Does it report evidence from cash transfer programs in a low- or middle-income country?
- *Date*: Was the study published from 2010-?

The figure to the right summarizes the search and Screening process for this project.

Only 54 studies for which the answer to all of the above questions was yes were included in the literature review.

Figure 1. Summary of literature search and screening process



# Perform Searches and Screen Results (1/5)

The goal during your searches and screening is to identify and collect literature relevant to answering your research question(s). You want to feel confident that the sample of studies you have identified for your review includes all relevant literature.

## 1. Enter search strings into each database you selected.

- a. For Google Scholar, you may want to conduct one search sorted by relevance and another sorted by relevance but limited to the last couple of years (to ensure you don't miss new research).

## 2. Determine how many and which results to screen. Make sure the search database presents your results sorted by relevance, to help ensure that if you are not able to screen all results, you have screen those most likely to be relevant.

- a. For searches that yield fewer than 100 results, screen the title and abstracts of all results using your screening criteria.
- b. For searches that yield more than 100 results, screen at least the first 50. If a good proportion of them are still relevant, continue screening the next 20. Repeat this process until you see very few relevant documents in a given set of results.
- c. Note how many pages/results you reviewed for searches where you did not review every result. Google and Google Scholar stop showing results after 100 pages. If you are continuing to see relevant results, consider conducting additional searches with more specific terms.

## Perform Searches and Screen Results (2/5)

3. **Read the titles and abstracts of your results and screen them against your set of screening criteria. Exclude results that are clearly and indisputably not relevant, i.e., that do not meet your screening criteria (see example of a screening spreadsheet in *Resources*).**
  - a. If possible, export or enter your search results to a spreadsheet, and add columns for each of your screening criteria where you can code whether a given results meets each criterion.
4. **Download and save all of the relevant sources.**
  - a. If searches yield a large number of sources that appear relevant, you may want to screen more than the title and abstract to confirm their relevance before downloading.
5. **For transparency about your search methods, track the following information about each database and search string in a spreadsheet (see example of Results Tracking table in *Resources*):**
  - a. Database searched
  - b. Search string used (make sure to document filters you add to your search string in a given database)
  - c. Date of the search and name of person who performed the search (if dividing searches among multiple people)
  - d. Number of search results
  - e. Number of search results screened (in case searches turn up hundreds or more results)
  - f. Number of new relevant search results
  - g. Number of duplicate relevant search results (you will not be able to note duplicates until you have aggregated all of your search results)



## Perform Searches and Screen Results (3/5)

- Document limitations to your searches, including noting titles of sources not available for download free of charge.
- For UW students, if a published source is not available online, you should search for it in the UW library system to see if a scan might be available.
- After your initial searches, you may find that you lack studies with evidence for certain aspects of your research question(s). You may choose to conduct supplemental searches with new search strings that target gaps in your sample of studies, search particular organizations/additional databases for relevant grey literature, or search the bibliographies/references of the studies you initially identified for additional publications.
  - Document these supplemental searches as you would the initial searches, along with your rationale for conducting them.
- When downloading relevant results, consider a standardized nomenclature.
  - For example, EPAR saves sources in PDF format as “AuthorLastName1 AuthorLastName2 YEAR,” for sources with up to three authors, and “AuthorLastName1 et al YEAR” for sources with more than three authors, to make it easier to track documents.

# Perform Searches and Screen Results (4/5)

Example Search Results Tracking (abbreviated) for [EPAR Technical Report #293](#):

Table A2.1: Scopus Keyword and search results

Keywords searched	Total Number of Search Results	First-cut				Second-cut	
		Relevant Studies	Relevant and Empirical Studies	Relevant, Empirical and Individual/H H/ Firm level Studies	Relevant, Empirical and Economy level Studies	Relevant, Empirical and Individual/H H/Firm level Studies - Morbidity Focus Areas	Relevant, Empirical and Economy level Studies - Morbidity Focus Areas
morbidity*	326	104	87	47	44	23	34
Tuberculosis	99	16	15	8	7	4	3
rotavirus	29	13	12	11	1	1	
malaria	109	27	19	9	14	7	11
polio	5	0	0	0	0		
HIV*	309	79	71	32	47	11	29
"Disability Adjusted Life Years" OR DALY*	84	38	37	32	5	4	5
"Quality adjusted life year" OR QALY*	119	69	69	66	3	12	2
comorbidity* OR comorbidity* OR comorbidity*	85	28	27	23	4	12	4
"weight-for-age" OR "stature-for-age" OR height OR (weight AND height)	226	33	32	28	5	25	
"mental health" OR "Patient health questionnaire" OR "Generalized Anxiety Disorder"	291	38	37	31	7	17	6
<b>TOTAL</b>	<b>13492</b>	<b>796</b>	<b>673</b>	<b>467</b>	<b>229</b>	<b>207</b>	<b>134</b>

Note: The first-cut refers to the initial level of coding done using the title, keywords and abstract. The second cut refers to the second level of coding done using the text of the studies.

# Perform Searches and Screen Results (5/5)

Example Screening Spreadsheet (abbreviated) for [EPAR Technical Report #293](#):

Microeconomic Studies of Morbidity & Growth (Scopus)						Level of Analysis			Empirical?	
Keyword	Authors	Title	Year	Source title	Abstract	Individual/HH level study (0/1)	Firm level study (0/1)	Economy level study (0/1)	Primarily Theory (0/1)	Primarily Empirical (0/1)
BMI	Bockerman P., Jha R., Gaiha R.	Stature and life-time labor market outcomes: Accounting for body mass index, participation, duration of work and	2013	Labour Economics	We use twin	1				1
BMI	Kan K., Lee M.	Lose weight for a raise only if overweight: Marginal i	2012	Journal of Applied Ec	Some studies	1				1
BMI	Han E., Norton	Direct and indirect effects of body weight on adult w	2011	Economics and Hum	Previous estim	1				1
BMI	Brown H.	Marriage, BMI, and Wages: A double selection appro	2011	Scottish Journal of P	Obesity rates	1				1
BMI	Bozoyan C., Wc	Fat, muscles, and wages	2011	Economics and Hum	Recent studie	1				1
BMI	Kortt M., Leigh	Does size matter in Australia?	2010	Economic Record	We estimate	1				1
BMI	Han E., Norton	Weight and wages: Fat versus lean paychecks	2009	Health Economics	Past empirica	1				1
BMI	Johansson E.,	Obesity and labour market success in Finland: The di	2009	Economics and Hum	This paper ex	1				1
BMI	Cawley J., Han	Obesity and labor market outcomes among legal imm	2009	Economics and Hum	This paper st	1				1
BMI	Greve J.	Obesity and labor market outcomes in Denmark	2008	Economics and Hum	This paper an	1				1
BMI	Gates D.M., Su	Obesity and presenteeism: The impact of body mass	2008	Journal of Occupatio	OBJECTIVE: To	1				1
BMI	Sullivan P.W.,	Cardiovascular risk factor clusters and employment i	2007	Value in Health	Objective: The	1				1
Calorie intake	Price G.N.	The allometry of metabolism and stature: Worker fat	2013	Economics and Hum	If the positive	1				1
Calorie intake	Mushtaq K., Ga	An examination of calorie demand relationship in Pa	2007	Pakistan Journal of N	This paper ha	1				1
Calorie intake	Glick P., Sahn U	Health and productivity in a heterogeneous urban la	1998	Applied Economics	The effects of	1				1
cognition	Vogl T.S.	Height, skills, and labor market outcomes in Mexico	2014	Journal of Developm	Taller worker	1				1
cognition	Ramos J., Cobl	The Impact of Cognitive and Noncognitive Skills on Pe	2013	Developing Econom	Professional	1				1
cognition	Lindqvist E.	Height and leadership	2012	Review of Economics	This paper st	1				1
cognition	Macke A., LePo	Social, societal, and economic burden of mal de deb	2012	Journal of Neurology	Mal de debar	1				1
cognition	Sakellariou C.	Central government versus private sector wages and	2012	Applied Economics	The importan	1				1
cognition	Aslam M., Bari	Returns to schooling, ability and cognitive skills in Pa	2012	Education Economics	This study inv	1				1
cognition	Hall M., Farkas	Adolescent cognitive skills, attitudinal/behavioral tr	2011	Social Forces	We use pane	1				1
cognition	Lindqvist E., Ve	The labor market returns to cognitive and noncogniti	2011	American Economic J	We use data	1				1
cognition	Hartog J., Van f	If You Are So Smart, Why Aren't You an Entrepreneur?	2010	Journal of Economics	How valuable	1				1
cognition	Jones G., Schn	IQ in the production function: Evidence from immigra	2010	Economic Inquiry	We show that	1		1		1
cognition	Heineck G., An	The returns to cognitive abilities and personality trai	2010	Labour Economics	We provide th	1				1
cognition	Anger S., Heine	Cognitive abilities and earnings - first evidence for G	2010	Applied Economics L	We provide fi	1				1
cognition	Field E., Roble	Iodine deficiency and schooling attainment in Tanza	2009	American Economic J	Cognitive dan	1		1		1
cognition	McGurk S.R., M	Cognitive training and supported employment for pei	2005	Schizophrenia Bulet	This study exa	1				1
cognition	Goodman A., S	Early education and children's outcomes: How long d	2005	Fiscal Studies	We evaluate	1				1
cognition	Tyler J.H.	Basic skills and the earnings of dropouts	2004	Economics of Educati	This paper te	1				1
cognition	Mitra A.	Mathematics skill and male-female wages	2002	Journal of Socio-Ecor	Using data fro	1				1
cognition	Cawley J., Heck	Three observations on wages and measured cognitiv	2001	Labour Economics	This paper su	1			1	1

\*Include studies that use empirical data for the study of household and firm-level links between morbidity and economic growth.

# Document the Screening Process (1/2)

**Systematically documenting your search and screening process helps to increase transparency around research methods and confidence in the findings you present.**

However, rigorous documentation may not be appropriate for all literature reviews.

Consider the intended use of your review when determining how rigorous your documentation should be.

- 1. When feasible (if using Scopus or other databases that allow you to export search results to a spreadsheet, or if the quantity of search results is small), record all results of your searches in a spreadsheet. This enables you to go back and include sources if your research scope changes, or explain your inclusion/exclusion decisions to others.**
- 2. If you cannot export your results, only enter documents into your spreadsheet if they meet your screening criteria from an initial review of title and abstract.**
- 3. Your spreadsheet should include basic information about the results (i.e., citation information), the abstract (if possible), a hyperlink, and one column for each screening criterion.**
  - a. This saves time in searching for citations/links later, and allows you to code your initial descriptive information and screening questions as you review documents.
  - b. For each source, code Yes, No, or “not specified” for each screening criterion. If it is not clear from the title and abstract whether a source meets your screening criteria, you might code a “?” in the appropriate cell.
  - c. If a document is clearly not relevant by any criteria, you do not need to code the remaining criteria.

## Document the Screening Process (2/2)

4. You may also wish to code some initial descriptive information about your search results at this stage, in case you want to further refine your results, for example by geography or by study methodology.
    - a. This would allow you to prioritize a subset of your relevant search results for review, if your sample of relevant results becomes too big.
  5. **At the end of your search process, you should have a spreadsheet with information on each search result from your various searches. You should also have columns coded with information on each screening criteria for your search results. This spreadsheet will serve as the basis for identifying the sample of studies for your review.**
- **Make sure your spreadsheet includes information on the source (author, title, year, publication), the search process used to ID the source (database, search string, date), and the screening criteria.**
  - **In some cases it will be appropriate to code the reasons for excluding sources that are not clearly irrelevant (e.g., “relevant but not empirical”).**
    - The level of detail in coding excluded papers depends on the aim of the review— in a review of empirical papers on yam diseases, if you want to be able to say “of the 186 papers published on the topic of yam diseases since 2010, 27 papers used empirical data” then you need a record of both empirical studies (the focus of the review) and non-empirical studies (not the focus, but relevant background). A screening spreadsheet allows you to track this information.

# Confirm the Sample of Studies for Review (1/2)

Before proceeding to the review of your studies, you must first confirm the sample of studies for your review. The goals in confirming the sample of studies are to:

- consider modifications or refinements of your research question(s) and scope based on the studies available;
- confirm the relevance of literature that initially appeared relevant from looking at the title and abstract but might be excluded upon further scrutiny; and
- identify gaps in the sample of studies to target with supplemental searches;

**1. Consider whether you want to apply different or additional screening criteria to your studies.**

- a. This is particularly important if you have identified more relevant studies than you will be able to review. For example, you may choose to further screen for type of study, to note if studies are quasi-experimental or experimental.

**2. For all search results that initially appeared relevant from looking at the title/abstract, confirm they are relevant by briefly scanning the full text and conducting a second screening, including any additional screening criteria.**

**3. Note changes to whether a given study meets the screening criteria in the screening spreadsheet by changing the coding for the relevant criteria.**

- a. Keep results that are no longer relevant in the spreadsheet so you can confirm you screened them; code them as “No” for meeting screening criteria.

## Confirm the Sample of Studies for Review (2/2)

- Determine whether there are any gaps aspects of your research question(s) or scope that are not addressed by the results you have identified. Consider conducting supplemental searches to target these gaps.**

*EPAR Technical Report #359:* We conducted supplementary searches using Google Scholar to identify additional literature on universal basic income using the terms “universal basic income” and UBI, as only two studies in our initial sample discussed this type of cash transfer program.

- If you coded additional descriptive information about your search results during the screening phase, consider whether you will prioritize your review according to any of that info. This process of narrowing your sample for review is only necessary if a large number of results are relevant based on the initial screening criteria.**

- So, if you have 150 results, you might restrict your review to only quasi-experimental and experimental results, to have a more manageable and higher-quality sample.
- Common prioritization criteria include year of publication (i.e., filtering out old papers); number of citations; study methodology (i.e. experimental, quasi-experimental, etc.); or focus on subpopulations of interest (i.e., by geography, gender, income level, age).
- Be sure if you narrow to a subpopulation that you don't discard information from the broader group that helps construct a comparative. If you find studies reporting outcomes for youth, for example, and there are no studies reporting outcomes for adults included in your evidence, the findings from are specific to youth and not more broadly generalizable.

*EPAR Technical Report #293:* The research question states that we are analyzing morbidity's “impact on” growth, rather than morbidity's “association with” growth. We therefore prioritize evidence from experimental studies (e.g., randomized controlled trials) and large-N analyses.

# Check Point

---

**At this point, after completing your searches and screening them for relevance, you should have the following:**

- A complete search results and screening spreadsheet listing the documents that met or did not meet the screening criteria, with coding to indicate what sample of studies you will be reviewing
- A document or separate spreadsheet tab summarizing all searches and their results
- A folder with PDFs for your sample of studies

**If you have all three, you are ready to begin developing your review framework and coding the literature gathered.**



# Develop Review Framework and Code

The next part of a literature review involves developing a review framework to organize the information you want to retrieve from your sample of studies, and systematically reviewing and coding the relevant information for analysis.

## Develop Literature Review Framework

*Create a framework for organizing relevant information in your literature review, based in theory and structured to facilitate your analysis and answer different aspects of your research question(s)*

## Set Up a Coding Spreadsheet Based on Literature Review Framework

*Create a spreadsheet for systematically recording information from the evidence base according to your literature review coding framework; this process (1) ensures you do not miss information relevant to answering your research question(s), and (2) organizes information from the relevant literature in a way that makes it easy to conduct analyses and write up results*

## Code Information from Body of Evidence

*Systematically read through your body of evidence and enter relevant information into the appropriate sections of your coding spreadsheet*

## Review Coding

*Consider modifications to the coding spreadsheet to better answer research questions; conduct intermediate reviews of coding decisions to ensure consistency (particularly for group work, but can also apply to reviews conducted by a single person if the understanding of how to best code changes over the course of the review)*



# Develop Literature Review Framework (1/5)

The goal of developing a literature review framework is to outline and organize the key information that is relevant to your literature review, based in theory and structured to facilitate your analysis and answer different aspects of your research question(s). Developing a review framework makes the research question, goals, and causal pathways explicit. This is particularly important if the literature review aims to answer “why” certain outcomes are observed or to explain differences in outcomes across contexts or studies.

## What is a literature review framework?

- An outline of the types of information you want to be able to pull out of the literature you review
- An organizational tool to facilitate review and analysis

Usually, an initial review framework is developed prior to the literature search and screening, based on theoretical expectations and preliminary searches. The framework can be refined during the review as we identify new, relevant information that would be valuable to include in analysis.

*It can be helpful to organize your review framework into a bulleted outline that includes all of the information you will seek to draw out of the literature. You may not find evidence on everything in your review framework, but that does not mean you should not look for it. Remember that finding no evidence on a particular point is itself a finding!*

# Develop Literature Review Framework (2/5)

1. **Using your research question/scope and the information you noted during your preliminary searches, develop a framework to organize your review.**
  - a. Break your question down into its component questions and hypotheses.
  - b. Draw on existing theories that link activities or interventions to outcomes.
  - c. Think critically about what you want to be able to say in your review, and how you want to be able to analyze your information:
    - i. Is it important to analyze by geography? By particular intervention characteristics? By study quality? By category of outcome (e.g., general health, reproductive health, and nutritional health outcomes)?
    - ii. Be specific in what information will be useful to retrieve from the literature, and include it in your framework.
  - d. Organize your review framework into relevant categories.
    - i. For example: study characteristics, intervention characteristics, population characteristics, and findings by outcome area.
2. **(Optional) This is a good time to check-in with clients or others (your audience, relevant faculty, or outside experts) to get feedback and ensure your proposed review framework is appropriate:**
  - a. Have you covered key hypotheses and assumptions to test?
  - b. Have you included key characteristics of interest to answer the different aspects of your research question?
  - c. Is your review framework organized in a way that will facilitate analysis?

# Develop Literature Review Framework (3/5)

Selection from example review framework  
(abbreviated): [EPAR Technical Report #359](#)

- Review framework organized to collect information from each document.
- First record document information.
- Then record information on geographies of the cash transfer programs included in the reviews.
- Next record information on the methods of the reviews.
- Finally start collecting information on the findings of the reviews.

**The framework is organized to allow comparison and analysis across reviews, geographies, and outcomes.**

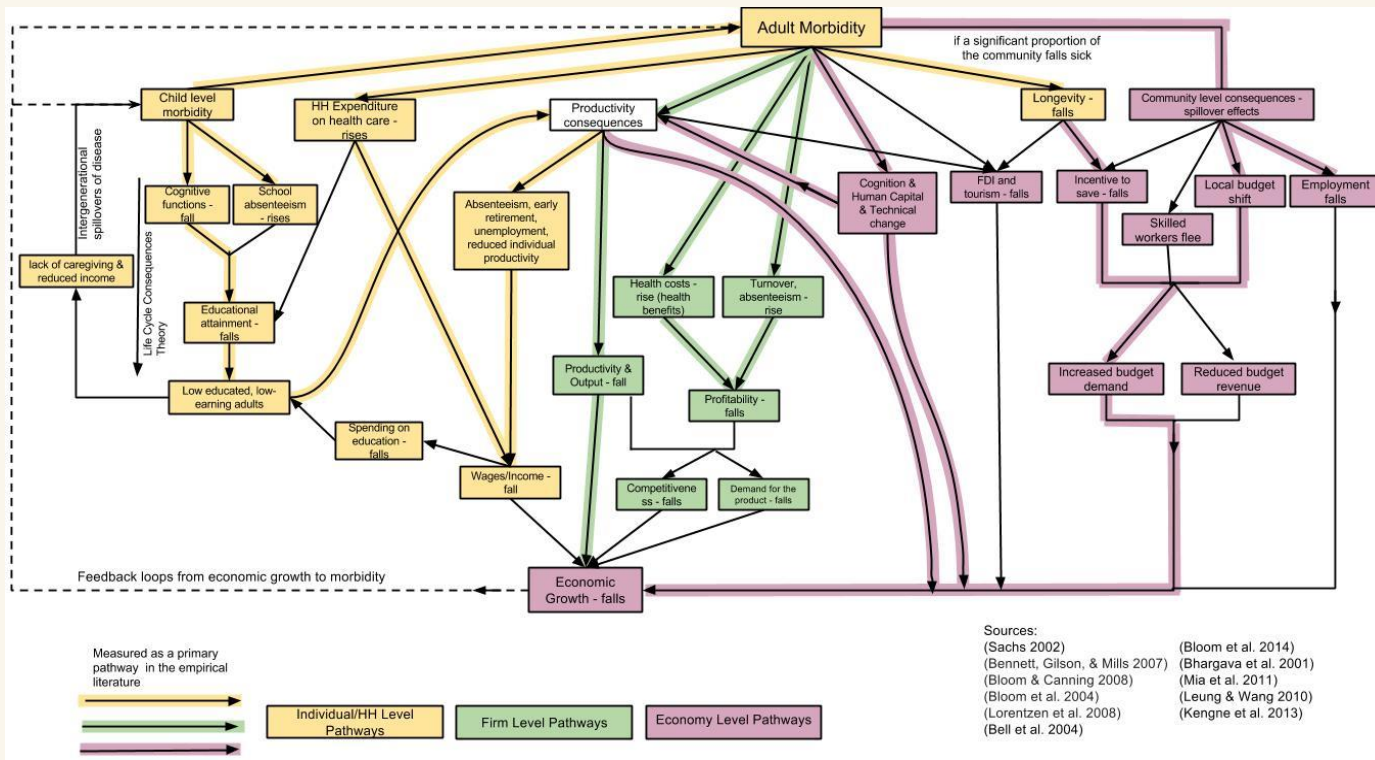
## Appendix A. Review Coding Framework

- **Document Information**
  - Title
  - Abstract
  - Database
  - Author(s) (Last name, Initial)
  - Published (Y/N)?
  - Publication Journal
  - Full APA Citation
  - Link
  - Year
- **Geography**
  - Country
    - If systematic review and multiple, number of countries
    - List countries
  - Region
    - Sub-Saharan Africa? (Y/N)
    - South Asia? (Y/N)
    - Southeast and East Asia? (Y/N)
    - Latin America? (Y/N)
    - North America? (Y/N)
    - Middle East/North Africa (MENA)? (Y/N)
- **Review Information and Inclusion Criteria**
  - Brief summary of study's key takeaways
  - Study findings
  - Describe factors affecting validity (external and internal) of findings
  - Systematic review? (Y/N)
  - Included studies start (year)
  - Included studies end (year)
  - Databases searched
  - Last search date
  - Describe additional inclusion criteria (e.g. specific geographic range, study methodology, program type [e.g. UCT CCT], etc.)
  - Describe additional sampling methods (e.g. snowball sampling, expert recommendations)
  - Number of studies identified by searches
  - Number of studies included and reviewed about CTs
  - Reports on quality of evidence? (Y/N)
    - Describe
- **Outcomes**
  - Health? (Y/N)
    - General findings of impact of CTs on health (positive, negative, mixed, no evidence)
    - Outcome 1
      - Impact (positive, negative, mixed, not significant)
    - Outcome 2
      - Impact (positive, negative, mixed, not significant)
    - Outcome 3
      - Impact (positive, negative, mixed, not significant)
    - Additional outcomes
      - Impact (positive, negative, mixed, not significant)
    - Describe findings for outcomes (specify countries if possible)
    - List outcomes which are **definitive** = impact in perpetuity (e.g. mortality, stunting, contracts HIV/AIDS)
    - List outcomes which are **sustained** = impacts measured after the end of the CT (e.g. continued breastfeeding of new children)
    - List **second-order** outcomes = impacts resulting from direct CT outcomes (e.g. higher wages resulting from increased school enrollment)

# Develop Literature Review Framework (4/5)

Example of theoretical grounding for review framework from [EPAR Technical Report #293](#):

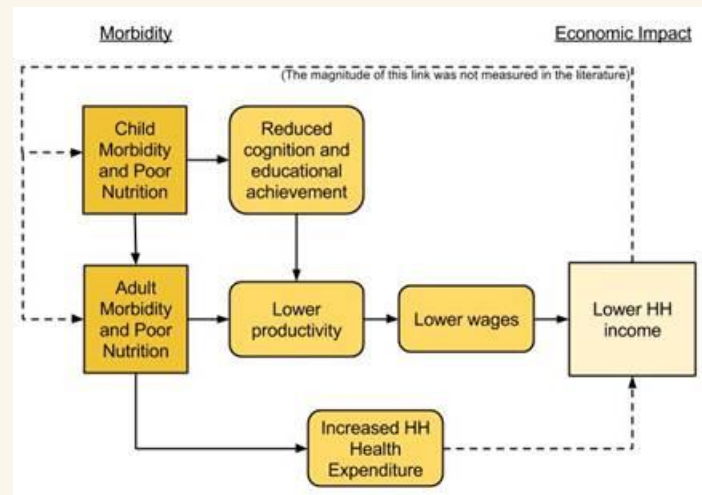
The causal pathways from morbidity to economic growth are fundamentally through disease or disability reducing productivity (lost income, or compromised physical or cognitive inputs that decrease firm outputs) or expenditures diverted from investment to health care. These pathways are theorized via three different avenues (individual/HH-level, firm-level, and economy-level) and pictured below. See the full report for more detail about each pathway.



# Develop Literature Review Framework (5/5)

Example of theoretical grounding for review framework from [EPAR Technical Report #293](#):

For example, child morbidity may result in greater school absenteeism and reduced educational attainment, resulting in lower adult productivity with implications for reduced household income. This pathway might further draw a connection from lower adult productivity to reduced economic output and lower economic growth, at the economy-level.



The review framework for our review of morbidity to economic growth includes both the measures of morbidity and economic growth used in our sample of studies, as well as the various pathways identified. We included each measure and pathway in our review framework, to be able to code yes or no for whether each study discussed that measure or pathway.

# Set Up a Coding Spreadsheet Based on Literature Review Framework (1/4)

After we have identified our sample of studies to review, we create a spreadsheet for systematically recording information from the studies according to the review framework. The goals of creating a coding spreadsheet are to (1) ensure you do not miss information relevant to answering the research question(s), and (2) organize information from the sample of studies in a way that makes it easy to conduct analyses and write up results.

1. **Begin by creating a coding spreadsheet. Using the columns from your search and screening spreadsheet, save a new version with a new title.**
  - a. Delete duplicates and documents that do not meet screening criteria – only your final sample of studies should be retained.
  
2. **Add the following columns (if not already created in screening spreadsheet):**
  - a. General information about the documents – e.g., citation, year, whether or not it has been peer-reviewed
  - b. Screening criteria
  - c. Information about the methodology of the study (if appropriate) – e.g., country/countries, years studied, hypothesis, independent and dependent variables, control group, type of experiment, etc.
    - i. For empirical studies, you may want to note the sample size, representation, sampling and study methodology, unit of analysis, etc. Also note any assumptions that the authors made that could change a reader’s interpretation of the results (e.g., “we assumed that differential rates of ARV adoption did not affect deaths from earlier HIV/AIDS cases”).

## Set Up a Coding Spreadsheet Based on Literature Review Framework (2/4)

- d. Add the components from your literature review framework.
  - i. Group sub-topics by theme, e.g. study descriptors, characteristics of an intervention/policy, outcome measures, etc.
  - ii. For each element of your framework, add a column to your spreadsheet asking a question related to that element. For example, “does the document include information about X? (Y/N)”, “What type of impact does the intervention have on Y outcome? (positive, negative, mixed, not significant, not specified)”.
    - Wording questions as Yes/No or categorical options will make it easier for you to code and analyze your information.
    - When questions are phrased to solicit categorical responses, specify the categories beforehand so that coding is consistent.
  - iii. For each coding question, include an adjacent “describe” column where you can include a qualitative description for the coding decision.

### *Example from [EPAR Technical Report #293](#):*

The coding spreadsheet columns separate distinctions in the measures of morbidity and economic growth used in the study. So, each measure (life-expectancy, stunting, etc.) has a different column, grouped under “morbidity measures” and is coded according to indicator:

- **Measures of morbidity** include direct measures (e.g., life expectancy, self-reported health), disability adjusted health metrics, nutritional measures, and cognition
- **Measures of economic impact** includes GDP or GNP measures, household income or average income, FDI, wages, productivity, absenteeism, employment, savings, and education/human capital

Other columns distinguish the level of pathway discussed in the study (individual/household, firm, economy-wide) and the pathway developed in the review framework (e.g. wages, absenteeism).



## Set Up a Coding Spreadsheet Based on Literature Review Framework (3/4)

- **The qualitative descriptions should be succinct but include any information necessary for interpreting and justifying the coding decision.**
  - Think about what information may be useful in writing the technical report (e.g., if you are including a quote, specify the page number).
- **Color code categories of questions for ease of reference (e.g., you may have a set of questions describing components of interventions, a set describing study populations, and a set describing study outcomes).**
- **Think about how you will use the spreadsheet for analysis (e.g., whether you want to create pivot tables) and structure it accordingly – for example, structure a column with categorical variables and then a subsequent “describe” column to include any relevant description or qualitative information.**
- **You may not know upfront what coding structure will work best – make it as good as possible at the outset, but be open to revising it.**
- **Overall, you want to be able to state confidently that you collected all the relevant studies and reviewed them to collect relevant information in a consistent way.**

# Set Up a Coding Spreadsheet Based on Literature Review Framework (4/4)

Example coding spreadsheet (abbreviated) for [EPAR Technical Report #359](#):

<b>Basic Information</b>				
<b>Author(s) (Last name, Initial)</b>	<b>Health (Y/N)</b>	<b>General finding of impact of CTs on health (positive, negative, mixed, no evidence)</b>	<b>Outcome 1</b>	<b>Impact (positive, negative, mixed, not significant)</b>
Samuels, F. & Stavropoulou, M.	Y	Positive	Stress	Positive
Narayanan, S.	Y	Positive	Health visits	Positive
Pantelic	Y	Positive	Immunization	Positive
Gibbs, A., Jacobson, J., & Wilson, A	Y	Mixed	Domestic violence	Positive
Neri, M.	N			
Ambia, J., & Mandala, J.	N			
Dellar, R. C., Dlamini, S., & Karim, C	N			
Wiysonge, C. S., Ngcobo, N. J., Jeer	Y	Positive	Immunization	Positive
Nores, M., & Barnett, W. S.	N			
Sanchez-Ancochea, D., Mattei, L.	Y	Not significant	Immunization	Not significant
Pettifor, A. et al.	N			
de Groot, R., Palermo, T., Handa, S.	Y	Positive	Health visits	Positive
Heise, L., Lutz, B., Ranganathan, M.	N			
Forget, E. L., Peden, A. D. & Strobel	Y	Positive	Hospitalization rates	Positive
Bastagli, F. et al.	Y	Positive	Health visits	Positive
Vincent, K., & Cull, T.	N			
Ma, Z., Bauchet, J., Steele, D. et al.	Y	Positive	Health visits	Positive
Khan, M. E., Hazra, A., Kant, A., & A	N			
Cecchini, S. & Madariaga, A.	Y	Mixed	Health visits	Positive
Segura-Perez, S., Grajeda, R., & Per	Y	Positive	Children's health	Positive
Fernald, L. C. H., Gertler, P. J., & Hi	Y	Positive	Infant mortality	Positive
Bastagli, F.	N			
Kabeer, N., & Waddington, H.	N			
Murray, S. F., Hunter, B. M., Bisht, F	Y	positive	Infant mortality	Positive
Soares, F. et al.	Y	Mixed	Immunization	mixed
Bassani, D. G., Arora, P., Gaffey, M.	Y	Not significant	Immunization	Not significant

# Code Information from Body of Evidence (1/6)

The goal during coding is to systematically read through your sample of studies and enter relevant information into the appropriate sections of your coding spreadsheet, such that you have confidence that you have captured all the relevant evidence in the same way across your sample.

1. Read each study in your sample.
2. As you read, apply the coding framework:
  - a. Consider each column and how the source does (or does not) address it.
    - i. When you first code, flag areas where the coding framework doesn't fit.
    - ii. After coding a few studies, consider whether you should adjust the coding framework to better fit the data. If you are working in a group, make sure to discuss and align around these decisions.
  - b. If a Yes/No coding question is not addressed in a given document, code "not specified" or leave the cell blank as opposed to coding "No".
    - i. This allows more accuracy in descriptions and analysis of the studies. If 10 of 50 documents report positive impacts on women's empowerment, it does not mean that the other 40 report non-positive impacts. Ten may report negative or not significant impacts, with the other 30 not reporting on this outcome area.

## Code Information from Body of Evidence (2/6)

3. **As you code, record your rationale for decisions in the “describe” columns** (see [EPAR Technical Report #351](#) Results Coding for an example – columns BD-BE):
  - a. For subjective measures, include evidence that informs your coding decision directly in the spreadsheet. Listing rationale in an extra column ensures consistency when you have more than one coder.
  - b. If the justifications for coding include distinguishable subgroups, you might add specific coding questions about them. For example, a question about whether digital products are combined with other digital financial services (DFS) may be divided into two questions: DFS digital services and DFS non-digital services.
  - c. “Describe” columns provide useful qualitative information that can be used to supplement tables and figures in the analysis and in writing up findings from the literature review. Recording page numbers for direct quotes can help in incorporating information from the describe columns.

*Example coding (column heading followed by what is entered in the cell for that study/row):*

- “Number of countries included in the study”: 1
- “Country (specify country name, or “multiple” if more than one)”: Tanzania
  - An associated “Describe” column could include the names of multiple countries when there are more than one, or information about the included country, such as “Rural areas in Northern Zone and Lake Zone only”.
- “Sub-Saharan Africa? Y/N”: Y
- “Impact on nutrition (positive, negative, mixed, not significant)”: positive
  - An associated “Describe” column would include more detail, such as “.3% increase in school absences associated with 1% decrease in measures of wasting”

## Code Information from Body of Evidence (3/6)

### 4. If you are working with a group of coders, calibrate with your group:

- a. Once you have each coded several sources (~5), convene to discuss the process and ensure everyone is coding consistently.
- b. Make sure your coding follows the same format for each question, to allow for analysis with pivot tables (i.e., entering the same words – Yes/No/Mixed – into the cell or the same numbers – 0s and 1s).
- c. Make sure everyone is following the same reasoning for coding, and is evaluating the sources according to the same standards.
- d. Determine whether or not the coding questions you selected are useful and sufficient: are there questions to add, or questions you want to reframe?
- e. Make sure the coding options are consistent: what options might you add to Yes/No? In some cases, unclear, mixed, or somewhat are all viable options.

*Examples:* We usually code one study per row (as in the review morbidity and economic growth), but if we want to compare across countries, programs, or products, for example, we may aggregate information from multiple studies discussing a given country, program, or product into a single row. For example, the coding spreadsheet for EPAR's review of land tenure technologies ([EPAR Technical Report #357](#)) aggregates information from multiple studies into rows where the unit of analysis was a specific technology, and the coding spreadsheet for EPAR's review of digital financial services consumer protection regulations in developing countries ([EPAR Technical Report #324](#)) aggregates information from multiple studies into rows where the unit of analysis was an individual country. The coding spreadsheets for each of these projects are available on the respective project webpages.

# Code Information from Body of Evidence (4/6)

Example of coding at the country level (abbreviated) for [EPAR Technical Report #324](#):

Basic Information			Basic Digital Financial Services Information												
Country	Number of Documents Reviewed	References (Author/Organization, Title, Year)	General Information				Financial Inclusion Indicators								
			Does the country's regulation enable a bank led or non-bank led model?	Describe	Number of Mobile Money (MM) Service Operators	Describe	Background Summarizing State of CEP Regulation	Access to Financial Institution Account	% of adults with at least one MFS account	% of active MFS accounts	Institution Name	Founding (Year)	Monitoring Responsibility (Finance, Telecommunications, Competition, Consumer Protection)	For what aspects of DFS CEP (risk, fraud, data privacy, protection of funds)	Describe
Bangladesh	18	Bangladesh Bank, Draft Regulatory Guidelines for Mobile Financial Services, 2015 Bangladesh Bank, Guidelines for	Bank-Led	"As part of its aim to promote financial	3+ (GSMA, The Mobile	3+ (GSMA, The Mobile Eco	29.1 percent of adults have an	2.7 percent of adults have	23% "The 2014 InterMedia survey	Bangladesh E	1971	finance	.	"...The Central Bank has the power to	
Brazil	24	Center for Financial Inclusion, Client Protection in Brazil, n.d. Central Bank of Brazil, Law 12865: The Brazilian Payments System,	Non-bank led	"Non-banks in Brazil were initially not	3+ (GSMA, The Mobile	AAccording to the 2015 Brookings report, the	68.1 percent	0.9 percent	.	Secretaria Nacional do Consumidor	2012	Consumer Protecction	.	"In Brazil the National Consumer Protection	
Colombia	17	CGAP, Emerging Risks in Branchless Banking: Key Findings from Colombia Case Study, 2014	Non-bank led	restricted.	3+ (GSMA, The Mobile Eco	The Mobile Eco	39 percent	2.2 percent	.	Ministry of Finance and Public Credit of	1923	.	"The Ministerio de Hacienda y Crédito		
Congo, Dem. Rep	8	Altai Consulting, Consumer Financial	Non-bank led	"Financial i	3+ (GSMA, The Mobile Eco	The Mobile Eco	18 percent	9.2 percent	.	Banque Cent	1964	Finance	.	"Among other	
Ecuador	22	Almazán & Frydrych, Mobile financial services in Latin America and the Caribbean: State of play, commercial models, and regulatory	Other	"Most mobil	1	1 - All money is issued by the State	In Ecuador, the Central Bank acts as the e-money	46.2 percent	Not specific	.	La Junta de Política Y Regulación Monetaria y	2014	Finance	.	Ecuador's 201
Egypt	9	African Development Bank, Financial Inclusion in Africa, 2013 Arab Republic of Egypt,	Bank-Led	"CBE's regulations on mobile payments	3+ (GSMA, The Mobile Economy, 2015)	The Mobile Economy, 2015)	14.1 percent	1.1 percent	.	National Télé.	.	Telecommunica	Competition	"In terms of licensing, NTRA is responsible	
Ghana	10	Bank of Ghana, Agent Guidelines, 2015 Bank of Ghana, Guidelines for E-Money Issuers in Ghana, 2015	Non-bank led	"The new a	3+ (GSMA, The Mobile Economy, 2015)	The Mobile Economy, 2015)	40.5 percent	13.0 percent	Second, the	Bank of Ghar	1957	Finance	Fraud, data pr	The Bank of G	

We collected information from multiple sources to analyze digital financial services regulations across selected low- and middle-income countries. In each row, we coded information from multiple sources as they applied to the coding question, and noted the source of the info in the "describe" columns. The sources are also recorded in the "references" column.

## Code Information from Body of Evidence (5/6)

- Due to length or time, you may not be able to read every single source in its entirety. Use “Ctrl+F” to search for key words (may differ by the column in the spreadsheet that you are considering) and read the surrounding sections.
  - Pay particular attention to the abstract, methods, and discussion sections. The introduction and background may provide context as well.
- **Look for different ways to think about the questions you’ve asked. A variable that you previously interpreted as Yes/No may have grey area. Consider breaking one ambiguous coding question into several specific questions.**
  - E.g., you may originally include a Y/N column for the question, “Are there caps on cash-in cash-out?” As you code, you realize that caps might be on transaction amounts *or* account balances. Break this out into two separate Y/N questions and include a description of the cap in the “Describe” column.
- **If you find relevant information that does not neatly fit into your coding spreadsheet, think about changing an existing coding question or adding a new coding question. The coding spreadsheet is a tool to organize the relevant evidence, and should not hinder you in accurately collecting that evidence.**
- **If working in group, communicate frequently! Discuss changes to the coding spreadsheet with your fellow coders to ensure you make the same changes to all versions of the spreadsheet.**

# Code Information from Body of Evidence (6/6)

Example of coding (abbreviated) for [EPAR Technical Report #293](#):

Keyword	Authors	Title	MEASURES OF MORBIDITY		MEASURES OF ECONOMIC IMPACT					Where? (Geography/Country)
			Direct Measures		GDP or GNP measures	Household income or average income	Wages	Productivity, absenteeism, employment, etc	Other (specify)	
			Disease-specific (what disease?)	Disease-specific (/1)						
comorbidity	Libby A.M., Ghosh	Economic differences in direct and indirect costs bet	epilepsy	1				1	1	United States
comorbidity	Jung D., Bhatta	Association of hearing loss with decreased employ	hearing loss	1				1		United States
comorbidity	Kruse M., Saett	Particulate emissions: Health effects and labour ma	pollution-related	1					1	Denmark
comorbidity	Fu A.Z., Chen L.	Absenteeism and short-term disability associated w	breast cancer	1					1	United States
comorbidity	Sullivan P.W., C	Influence of rheumatoid arthritis on employment, fu	arthritis	1				1	1	United States
comorbidity	Fu A.Z., Qu Y.,	Health care and productivity costs associated with d	diabetes	1					1	United States
comorbidity	Jofre-Bonet M.	Poor mental health and smoking: Interactive impact	on wages					1		United States
comorbidity	Bell B., Chalkli	Burden of dysthymia and comorbid illness in adults	dysthymia	1				1		Canada
comorbidity	Birbaum H.G.	Direct and indirect costs of asthma to an employer	asthma	1					1	United States
Disability Adj	Rao P.S., Darlo	Disability adjusted working life years (DAWLs) of le	leprosy	1				1	1	India
Disability Adj	Sharieff W., Ho	Economic gains of a home fortification program: Eval	diarrhea	1				1		Pakistan
HIV	Marinescu I.	HIV, wages, and the skill premium	HIV/AIDS	1				1		sub-Saharan Africa
HIV	Chicoine L.	AIDS mortality and its effect on the labor market: Evi	HIV/AIDS	1				1	1	South Africa
HIV	Olang'O C.O., N	Children as caregivers of older relatives living with H	HIV/AIDS	1						Kenya
HIV	Pennap G.R.I.,	A review of the impact of HIV/AIDS on education, the	HIV/AIDS	1					1 firm costs	multiple countries in Africa
HIV	Collins D.L., Le	The financial impact of HIV/AIDS on poor household	HIV/AIDS	1			1		HH costs	South Africa
HIV	Dorward A.R., M	Labor market and wage impacts of HIV/AIDS in rural	HIV/AIDS	1					1	Malawi
HIV	Rosen S., Vince	The cost of HIV/AIDS to businesses in southern Africa	HIV/AIDS	1					1 firm costs	multiple countries in Africa
HIV	Rosen S., Simo	AIDS Is Your Business	HIV/AIDS	1					firm costs	South Africa, Botswana
HIV	Fraser F.K., Gra	The impact of HIV/AIDS on small and medium enterp	HIV/AIDS	1					1 firm costs	South Africa

We first captured basic document characteristics such as the author(s), title, abstract, geography, research design (experimental and quasi-experimental studies, meta-analyses and systematic reviews), etc. The spreadsheet further included information for each study on measures of morbidity and measures of economic impact, as outlined in the review framework, with yes/no entries coded as 1/0 to facilitate analysis with pivot tables and adjacent cells to provide qualitative descriptions (not shown in the above figure). Finally, for each study, we coded information on the specific pathways analyzed as connecting morbidity to economic growth. The [full coding spreadsheet](#) is available on our website.



# Review Coding (1/1)

The goals of reviewing your coding are to consider modifications to the coding spreadsheet to better answer research question(s), and to ensure consistency in coding across studies (particularly for group work, but can also apply to reviews conducted by a single person if the understanding of how to best code changes over the course of the review).

- 1. Conduct a quality check of your completed coding spreadsheet—to ensure that each question was coded consistently, to standardize the responses, and, if useful, to break out particular questions (especially categorical questions) or descriptions into new columns.**
  - a. EPAR typically reviews coding column by column to consider whether a given category has been coded consistently over time and/or across different coders.
  - b. If you are working in a group, divide the work by assigning certain sections of the coding spreadsheet to each coder for review.
- 2. Discuss possible changes to your coding spreadsheet, either as a group or with advisers.**
- 3. Return to the previous step and re-code any necessary information.**

# Analyze the Evidence

A coding spreadsheet organizes the information from a literature review into tabular form (rows, columns, and cells). This method of organizing the evidence from a sample of studies facilitates the final step of a literature review—analysis using pivot tables, and the creation of tables and figures to illustrate key findings.

## **Analyze Findings**

*Build pivot tables for analysis; use the spreadsheet and pivot tables to compare trends in outcomes and factors of interest across relevant groupings; create figures and tables for your report*

## **Summarize and Report Findings**

*Use the organization of the literature review framework as a structure for summarizing and presenting findings from analysis; identify and report on gaps in the evidence base and in the literature review methods*



**EVANS SCHOOL OF PUBLIC POLICY & GOVERNANCE**

UNIVERSITY of WASHINGTON

Evans School Policy Analysis and Research (EPAR)

# Analyze the Evidence (1/5)

The goal of analyzing the evidence is to summarize key findings responding to all of your research questions, to present tables and figures to illustrate those findings, and to identify gaps in the evidence.

1. **Identify the logical sections to analyze. This will depend on the structure of your review framework, and should always relate back to the research question(s).**
  - a. Consider what columns in your spreadsheet you might want to use to sub-divide your sample for analysis, and what columns you want to compare across sub-groups.
2. **Use spreadsheet filtering and pivot tables to examine results in your coding spreadsheet and begin developing a story.**
  - a. Depending on your research question and data, this could include filtering for the count of documents that mention a certain topic, calculating proportions, or comparing findings between documents.
  - b. Filter relevant columns in the spreadsheet by the information you are interested in (use “Sort & Filter” within Editing on the Home tab).
  - c. Construct pivot tables to help analyze how different columns and rows in your spreadsheet relate to one another (select your full spreadsheet of data and click “Create Pivot Table” under the Insert tab, making sure to place the Pivot Table on a new worksheet within the same workbook).
    - i. If you have more than one row of column headings, you will need to consolidate the headings on a single row to be able to create a pivot chart.

## Analyze the Evidence (2/5)

- **If you set-up your literature review coding spreadsheet to categorize data in a consistent way (i.e., 0s and 1s or uniform text options, such as Yes, No, Sometimes), it will be easy to use your spreadsheet to create pivot tables.**
- **To generate and use pivot tables in Excel:**
  - Select all of your coded data (including headings) and choose “PivotTable” from the Insert menu.
  - Click inside the pivot table that is generated (usually on a new sheet) and the “PivotTable Fields” dialogue box will appear allowing you to choose your which column labels, row labels, cell values, and filters will appear in the pivot table.
  - You can create multiple pivot tales depending on how you want the data summarized and displayed.
  - Before creating a pivot table from your raw coding data, make sure you have just one row of headings/column labels selected along with your data.
  - Additional resources: [GCF Intro to Pivot Tables](#) and [Microsoft Pivot Table Tutorial](#)
- **Once you have created a pivot table, you can simply click on a cell in the table and then select the PivotChart option (also under the “Insert” top menu option in Excel) and choose from bar or line graphs, pie charts, etc. Once you select your graphic and hit “OK”, it will appear on the same spreadsheet tab as the pivot table and can be copied and moved to your report (like the bar chart below).**
- **Remember to “Refresh” any existing pivot table if you update any data in the coding spreadsheet.**

# Analyze the Evidence (3/5)

Example of coding spreadsheet and pivot tables (abbreviated) from [EPAR Technical Report #359](#):

Basic Information				
Author(s) (Last name, Initial)	Health (Y/N)	General finding of impact of CTs on health (positive, negative, mixed, no evidence)	Outcome 1	Impact (positive, negative, mixed, not significant)
Samuels, F. & Stavropoulou, M.	Y	Positive	Stress	Positive
Narayanan, S.	Y	Positive	Health visits	Positive
Pantelic	Y	Positive	Immunization	Positive
Gibbs, A., Jacobson, J., & Wilson, A	Y	Mixed	Domestic violence	Positive
Neri, M.	N			
Ambia, J., & Mandala, J.	N			



A	B	C	D	E	F	G	H
<b>Article Summary</b>		<b>Geographic Summary</b>			<b>Outcomes</b>		
<b>Coded documents by published status</b>		<b>Included region by study</b>			<b>Categories included in reviews</b>		
Relevant Y/N	Y		Relevant Y/N	Y	Relevant Y/N	Y	
<b>Row Labels</b>	<b>Count of Doc. #</b>	<b>Region</b>	<b>Number of reviews</b>		<b>Outcome category</b>	<b>Number of reviews</b>	
N	8	Sub-Saharan Africa	36		Health	27	
Y	46	South Asia	18		Reproductive health	24	
<b>Grand Total</b>	<b>54</b>	Southeast and East Asia (SEA)	10		Nutrition	20	
		Latin America	37		Financial inclusion	2	
		North America	5		Labor Market	12	
<b>Coded documents by year published</b>		Middle East/North Africa (MENA)	4		Poverty	17	
Relevant Y/N	Y				Gender and Intra-household decision mak	10	
<b>Row Labels</b>	<b>Count of Doc. #</b>				<b>Reviews reporting on the implementation of CTs</b>		
2010	3				Relevant Y/N	Y	
2011	7						
2012	6				<b>Implementation of CTs</b>	<b>Number of reviews</b>	
2013	7				Compares different types of CTs	20	
2014	3				Compares CTs to a different intervention	13	
2015	8				Compares delivery method of CTs	3	
2016	11				Mentions digital delivery?	2	
2017	9				Discusses scalability	8	
<b>Grand Total</b>	<b>54</b>						



# Analyze the Evidence (4/5)

Example of pivot table and pivot chart from [EPAR Technical Report #293](#):

In our review of the evidence on morbidity and economic growth, we coded what countries were analyzed in each Study, and whether the study presented quasi-experimental evidence, non-experimental evidence, or was a review. We used that coding information to create the pivot table and chart to the right.

	Non-experimental	Quasi-experimental	Review	Grand Total
Multiple countries	10	1	5	39
United States	6	1		7
China	5			5
India	3	1		4
South Africa	4			4
Tanzania	3			3
(blank)	1		2	3
Germany	2			2
Korea	2			2
Malawi	2			2
Australia	1			1
Botswana	1			1
Brazil	1			1
Canada		1		1
Europe			1	1
Europe	1			1
France	1			1
Georgia	1			1
Ghana, Ivory Coast	1			1
Honduras	1			1
Hong Kong	1			1
Ireland	1			1
Kenya	1			1
KwaZulu-Natal	1			1
Mexico	1			1
Mozambique	1			1
New Mexico	1			1
Nigeria	1			1
Phillipines	1			1
Russia	1			1
rwanda	1			1
Singapore	1			1
South Asia (India, Pakistan)	1			1
South Korea	1			1
sub-Saharan Africa	1			1
Texas	1			1
Thailand	1			1
Uganda	1			1
UK	1			1
Vietnam			1	1
Zimbabwe	1			1
<b>Grand Total</b>	<b>89</b>	<b>4</b>	<b>9</b>	<b>102</b>



# Analyze the Evidence (5/5)

- **Coding spreadsheets, in addition to being able to summarize and graph the data in a spreadsheet package, can be read into more specialized statistical packages like Stata or multi-purpose packages like R for analysis.**
  - Sometimes importing these data or spreadsheets into other programs requires you to save your spreadsheet as a .csv or “comma-separated values” file.
  - R in particular can be used to create a variety of interesting graphics using the data from a coding spreadsheet.
- **Spreadsheets can also be imported into visualization software programs such as Tableau or Power BI to create dynamic visualizations of your data and coding.**
  - This is particularly useful for data analysis, but can also be useful for visualizing results of a literature review if you include many yes/no or categorical questions in your coding framework.
  - [Tableau Desktop](#) offers a free one-year subscription for educational purposes, though you need to register and request a personal product key for the software. After creating a Tableau visualization using the Tableau Desktop software, you can upload your visualization to [Tableau Public](#) (free to use after registering) to share with others or embed into a web page. Tableau offers a variety of useful training videos [here](#).
  - [Power BI Desktop](#) can be downloaded for free, and similarly allows users to create an interactive visualization dashboard and publish it to the web. The [Power BI website](#) provides a variety of videos, samples, and in-depth documentation to support users in learning about the software.

# Summarize and Report Findings (1/2)

The goals in summarizing and reporting the findings of your literature review are to (1) use the organization of the literature review framework as a structure for summarizing and presenting findings from analysis, and (2) identify and report on gaps in the sample of studies and in the literature review methods

- 1. Using the tables and figures you just created, begin organizing relevant results about the trends and patterns arising from the data. Consider:**
  - Do you go beyond basic descriptions of the studies to truly analyze? For example, do you look at trends in the evidence – by geography, study population, type of intervention, outcome area, etc.?
  - Are there any gaps in the studies that have been revealed by analysis? Are there additional searches you can undertake to fill those gaps?
  - Are there any unexpected or surprising trends? How do you explain them?
- 2. As you begin outlining findings for your final output (typically a report or slide deck), think about the following questions:**
  - Do you answer all the research questions?
  - Are the findings organized in the most logical order?
  - Do you present useful analysis in each of your sections, and is this analysis supported by the collected evidence?
  - Where might a table or figure help illustrate or summarize the findings?



## Summarize and Report Findings (2/2)

- **A complete literature review output (often a report) will include**
  - an introduction with background and a discussion of the theoretical framework and grounding for the review (which informed the literature review framework),
  - a transparent overview of the search, screening, and coding methods,
  - tables and figures summarizing and presenting the findings, accompanied with relevant text providing further interpretation and analysis, and
  - a discussion of the relevance of the findings to the research question(s) and of any research gaps.
- **Your analysis of the literature should be a cohesive story that gives an objective summary of your findings. This story should follow directly from your literature review framework and coding categories.**
- **Run an outline of your findings by a professor, your client, or a friend. They may be able to suggest different ways to analyze the data, logical ways to organize content, or gaps in the discussion. At this point in your review, it can be hard to step back enough to see these opportunities; outside feedback can help.**

*Example categories of key findings from [EPAR Technical Report #293](#):*

- Measures of Morbidity
- Measures of Economic Growth
- Links between Morbidity and Economic Growth
- Findings: Literature on Morbidity and Growth
- Pathways from Morbidity to Growth: Individual/Household and Firm Level
- Pathways from Morbidity to Growth: Economy Level
- Conclusions and Research Gaps

# Evans School Policy Analysis & Research Group (EPAR)

Professor C. Leigh Anderson, Principal Investigator

Professor Travis Reynolds, co-Principal Investigator

**EPAR uses an innovative student-faculty team model to provide rigorous, applied research and analysis to international development stakeholders. Established in 2008, the EPAR model has since been emulated by other UW schools and programs to further enrich the international development community and enhance student learning.**

*Please direct comments or questions about this research to Principal Investigators C. Leigh Anderson and Travis Reynolds at [eparinfo@uw.edu](mailto:eparinfo@uw.edu).*