Promoting sustainable energy development through access, renewables and efficient technologies: An evidence gap map and systematic review

Constanza Gonzalez Parrao
International Initiative for Impact Evaluation (3ie)

Cem Yavuz
3ie

Miriam Berretta
3ie

Zafeer Ravat
3ie

Dina Rodrigues
3ie

Benjamin Sovacool
Boston University and University of Sussex

Alison Bethel
University of Exeter

Solomon Asfaw
Sustainable Energy for All (SEforALL)

Frederick Gaved
SEforALL

Birte Snilstveit
3ie

EGM and SR Protocol

June 2023

About 3ie

The International Initiative for Impact Evaluation (3ie) develops evidence on how to effectively transform the lives of the poor in low- and middle-income countries. Established in 2008, we offer comprehensive support and a diversity of approaches to achieve development goals by producing, synthesizing and promoting the uptake of impact evaluation evidence. We work closely with governments, foundations, NGOs, development institutions and research organizations to address their decision-making needs. With offices in Washington DC, New Delhi and London and a global network of leading researchers, we offer deep expertise across our extensive menu of evaluation services.

3ie evidence gap maps

3ie evidence gap maps (EGMs) are thematic collections of impact evaluations and systematic reviews that measure the effects of international development policies and programs. The maps provide a visual display of completed and ongoing systematic reviews and impact evaluations in a sector or sub-sector, structured around a framework of interventions and outcomes.

The EGM protocol provides supporting documentation for implementing the map, including background information for its scope and the methods that will be applied to systematically search and screen the evidence base, extract and analyze data, and develop the EGM report.

About this evidence gap map protocol

This report presents the protocol for a systematic search to identify and map the evidence base of impact evaluations and systematic reviews of interventions that aim to promote sustainable energy in low- and middle-income countries. The EGM will be co-funded and co-produced by 3ie and Sustainable Energy for All (SEforALL). The content of this report is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. Please direct any comments or queries to the corresponding author, Constanza Gonzalez Parrao cgonzalez@3ieimpact.org.

Suggested citation: Gonzalez Parrao, C, Yavuz, C, Berretta, M, Ravat, Z, Rodrigues, D, Sovacool, B, Bethel, A, Asfaw, S, Gaved, F, Snilstveit, B. 2023 EGM Protocol: Promoting sustainable energy development through access, renewables and efficient technologies: An evidence gap map and systematic review. New Delhi: International Initiative for Impact Evaluation (3ie).

Contents

List of figures and tables

1. Background

1.1. Challenges in the energy sector

Halfway through the implementation of the 2030 Sustainable Development Goals Agenda, progress on SDG 7 is mixed. SDG 7 aims to ensure sustainable energy for all through access to energy, the adoption of renewables and an increase in energy efficiency. By working towards these three pillars, not only will energy targets be met, but advancements in climate and human welfare goals can also be achieved, which highlights the far-reaching consequences of ensuring sustainable energy for all (International Council for Science, 2017).

Substantial headway has been made in access to electricity, with the share of the population served in L&MICs increasing approximately 10% in the past decade, currently standing at 89% access in L&MICs (World Bank, n.d.). However, progress has been concentrated on populous countries in South Asia, mostly India and Bangladesh. In low-income countries, most of the population (59%) is still unserved, with the figure increasing to an overwhelming 72% if considering the world's rural poor (World Bank, n.d.). Furthermore, energy access does not equate to the use of green energy sources. The International Energy Agency (IEA) reports that of 50 energy system components, which are key to green energy transitions globally, only three are on target to achieve the 2050 goals: solar photovoltaics, electric vehicles and lighting (IEA, 2023). Much of Africa's electricity, for example, is sourced through non-renewable means despite the fact the continent has an abundance of natural resources, such as 60% of the world's solar resources (Climate Analytics, 2022).

There are further energy disparities when looking within regions. In Africa, there is a large divide between urban and rural areas in terms of access to electricity (UNDESA, 2023a). Nigeria, Ethiopia and the Democratic Republic of Congo have the highest concentration of the world's unserved poor, highlighting the need for financing environments and policy packages that consider the challenges of providing grid-electrification in fragile and conflict-affected states (IEA et al., 2023). In Latin America and the Caribbean, though electrification is near universal, other energy inequalities exist, such as gender inequalities in terms of access to clean energy sources, particularly for cooking (UNDESA, 2023b). Indeed, the IEA estimates that mitigating the electricity deficit requires an annual electrification growth rate of 1% from 2021 onwards, substantially higher than the 0.7% average growth in the past decade. If the current policy mix is

to be maintained, 660 million people in sub-Saharan Africa and 70 million in South and Southeast Asia will remain unserved by 2030 (IEA, 2022).

Covid-19 and Russia's invasion of Ukraine have increased pressure on global energy supply chains, particularly on natural gas. Compounded with the climate crisis and threats to food security, both of which are disproportionately felt by poor households in low-income countries, current energy challenges go well beyond basic electrification and include the transition to clean energy sources and improved energy efficiency. The IEA (2022) projects that by 2050, demand for cooling solutions (particularly air-conditioning) in emerging economies will be the equivalent of introducing another European Union to today's global electricity demand. Moreover, the IEA (2022) estimates that due to price increases and other factors, about 75 million people who recently gained access to electricity are likely to lose the ability to pay for it, and 100 million people who have gained access to clean cooking may revert to the use of traditional biomass.

Addressing limited dissemination of clean and efficient energy technologies in L&MICs can contribute to human and economic development and is critical to mitigate climate change. In the absence of substantial progress in clean energy transitions, meeting global demand will require further investment in oil and gas - or poor households will face further shortages and price volatility. Such scenarios would critically compromise the 1.5 C global average temperature rise established in the Paris Agreement.

1.2. Global policy responses

Over the past couple of decades, the international community has identified the necessity of providing sustainable energy for all with a particular focus on developing frameworks to guide global efforts. The most significant framework has been SDG 7, whose goals are aligned with several other major frameworks. Both the United Nations Framework Convention on Climate Change (UNFCCC) and the Paris Agreement also highlight the role of energy in mitigating climate change (Gao et al. 2017).

Frameworks such as these have prompted substantial funding strategies to tackle challenges in the energy sector, which can be used to implement a variety of policies and activities. A prominent example is the USD100 billion pledge made by developed countries to support mitigation and adaptation efforts to tackle climate change (UN Climate Change. n.d.). Furthermore, although there has been a recent decline in funding, USD10.8 billion of

international financial flows in 2021 have been provided for clean energy initiatives in developing countries (IEA et al., 2023). Directly related to the creation of international frameworks, the UNFCCC established the Green Climate Fund (GCF), the largest of its kind. GCF channeled over USD10 billion (over USD40 billion with co-financing) on climate change initiatives during its first programming stage. Currently transitioning into its second stage, GCF is planning an ambitious replenishment based on, among other pillars, supplying low-carbon energy for all (GCF, n.d.).

However, energy investments are far from matching projected demand from the global south. The large investment target to meet SDG 7 has been disproportionately focused on high- and middle-income countries (UNDP, 2019), and annual investment in clean energy in L&MICs (excluding China) has remained stagnant since 2015 (IEA, 2022). Yet, the IEA projects more than USD100 trillion being invested into energy infrastructure by 2050 (IEA, 2021),

1.3. Why it is important to do this EGM?

The large investments made towards the energy sector coupled with the dramatic effects of climate change, which have forced alterations to the way energy is transmitted, have led to the need to better understand what sustainable energy interventions work, and what evidence on effectiveness is available. An EGM that provides easy access to the available rigorous evidence on the effectiveness of sustainable energy interventions can help guide evidence-informed policymaking and highlight the existing evidence gaps that may be filled through future research investments.

3ie's Development Evidence Portal (DEP) currently houses over 12,000 impact evaluations (IEs) and over 1,000 systematic reviews (SRs). A preliminary scoping on this platform shows that there is a potentially large evidence base on sustainable energy interventions of up to 468 IEs and 13 SRs. Through additional scoping, we have identified various synthesis efforts within the energy sector, particularly on energy efficiency (Tsang et al. 2012; Kivimaa & Martiskainena 2017; Rasmussen 2017, Nehler 2018, Solnørdal & Foss 2018; Gonzalez-Caceres et al. 2019; Kamal et al. 2019; Khanna et al. 2021; Berretta et al. 2021a; Berretta et al. 2021b), electrification (Mathur et al. 2015; Bayer et al. 2020; Irwin et al. 2020; Hamburger et al. 2019; Moore et al. 2020), and renewable energy (Akella et al. 2009; Owusu & Asumadu-Sarkodie, 2016). However, none of these syntheses provide a clear overview on the state of the evidence across the three SDG 7 pillars of electrification, renewables and energy efficiency.

There is also reason to believe that some of these synthesis efforts are not of high-confidence or do not include the most up-to-date evidence. Of the potentially relevant SRs on the DEP, only two have been assessed by 3ie as high-confidence and one as medium-confidence (the appraisal tool used is presented in Appendix F). While one high confidence review is recent (Berretta et al. 2021a), it focuses solely on energy efficiency. The remaining high-confidence and medium-confidence reviews (Bensch et al. 2016; Thakur et al. 2018), can be considered outdated by now. With this map, we aim to fill this gap by implementing an ambitious framework that covers evidence on the effects of interventions targeting any or all of these three pillars. We also aim to facilitate the use of existing systematic reviews by appraising their confidence and summarizing their main results so that researchers and practitioners can understand what works, in which contexts and why.

1.4. Study objectives and questions

EGMs are tools to help policymakers and researchers working in a sector or thematic area make evidence-informed decisions. They make existing evidence more accessible and ease the prioritization of future research by mapping existing studies in a field on a framework of interventions and outcomes. Studies are mapped onto a framework of interventions and outcomes, providing a visual display of the volume of evidence for combinations of interventions and outcomes, the type of evidence, an indication of research gaps and, for systematic reviews, a confidence rating reflecting the study quality.

The results will be displayed on 3ie's online platform, which provides a graphical and interactive display of the evidence in a matrix framework. There will also be filters which users can apply to sort the evidence in the EGM according to different dimensions, including study design, country and population. To complement the map, we will conduct descriptive analyses to address the key research questions.

The specific objectives of this EGM are twofold:

- Identify and describe the characteristics of IEs and SRs evaluating the effects of sustainable energy interventions on environmental and welfare outcomes in L&MICs.
- Identify potential primary evidence and synthesis gaps.

To meet these objectives, we will address the research questions shown in Table 1. The reasoning for our conceptual scope is discussed in detail in Section 2.

Table 1. EGM research questions

#	Research Question	Type
RQ1	What is the extent and characteristics of empirical evidence, IEs and SRs, on the effects of sustainable energy interventions on intermediate, environmental and welfare outcomes in L&MICs?	Coverage
RQ2	What are the major primary and synthesis evidence gaps in the literature?	Gaps
RQ3	What intervention/outcome areas could be prioritized for primary research and/or evidence synthesis?	Research

2. Scope

The interventions and outcomes framework for this EGM encompasses a set of activities and goals that align with the three pillars of SDG 7. Interventions are separated into four groups: legal and regulatory framework and policies; financial incentives and market enabling activities; electrification and energy infrastructure; and information and capacity development. The outcomes of interest are separated into three groups: intermediate/behavior change outcomes; energy and environmental outcomes; and socio-economic and community welfare outcomes. To ensure the framework remains entirely focused on sustainable energy, general environmental rules and regulations, which may or may not affect the energy sector, are excluded.

2.1 Conceptual framework

In order to develop the framework for the EGM, we used SDG 7 as a starting point, which contains the following three main targets:

- 7.1: By 2030, ensure **universal access** to affordable, reliable and modern energy services.
- 7.2: By 2030, increase substantially the share of **renewable energy** in the global energy mix.
- 7.3: By 2030, double the global rate of improvement in **energy efficiency**.

These three targets (access, renewables and efficiency) are used to define how we understand sustainable energy and provide a seemingly intuitive basis for categorizing interventions in a framework ; yet, due to the overlap in mechanisms and activities which can achieve these goals, using these as the basis for interventions is a challenge.

Given that access, renewables and efficiency are targets, it is expected that programs or policies may attempt to affect one or more of these goals at once. For instance, the International Council for Science (2017) highlighted that the distribution of energy through renewable sources (e.g. solar lamps) may increase the share of renewables used, while also increasing access to electricity. We quickly identified this challenge when testing this framework approach. Projects such as the Asian Development Bank's support to rural electrification in Bhutan (ADB, 2010)

and the elementary village electrification systems in Nepal (Zahnd & Kimber, 2009) aimed to increase access to electricity through renewable sources.

Considering this, and drawing from previous 3ie projects (Moore et al., 2020; Berretta et al., 2021b), we have developed a framework that includes interventions aiming to overcome the most common barriers to energy access, the use of renewables and energy efficiency in L&MICs. Given this broad scope, it is difficult to conceptualize the framework based on any one theory. Instead, we first provide an overview of how the included interventions can overcome the common barriers to achieving sustainable energy, and then discuss how we anticipate the included interventions can affect the outcomes of interest for this project.

Legal and regulatory frameworks and policies can overcome barriers to the development of sustainable energy by encouraging or forcing governments, firms or households to meet codes and targets in the energy sector. While countries have aimed to diversify their energy mix, to date, many L&MICs do not always have strong policies in place to support the transition to sustainable energy (Bhattarai et al. 2022). Organizations' behaviors are pressed to change with the implementation of mandatory targets, particularly when previous priorities had been placed on production (Hasanbeigi et al., 2010), or where decisions had been limited by bounded rationality (Iwaro & Mashwa, 2010). Energy management reforms may also benefit situations where the provision of energy projects is hindered by limited institutional support and tension (Mawhood & Gross, 2014).

Financial incentives and market enabling activities can overcome the financial barriers limiting achievement towards SDG 7. Financial limitations have hindered electrification both on the supply and demand side. For those who supply energy, there has been few financial resources made available to overcome transaction costs (Bhattacharyya 2013; Bos et al. 2018; Palit & Kumar, 2022); while for households, the up-front cost of grid connectivity is often cited as a common hurdle (IEG, 2008; Jimenez, 2017; Bonan et al., 2017; Bos et al., 2018; Palit & Kumar, 2022). The inability to afford new technologies has also been identified as a major barrier for both the increase in renewables share (Gribkova & Milshina, 2022; Singh & Ru, 2022) and the use of efficient technologies (Painuly et al., 2003).

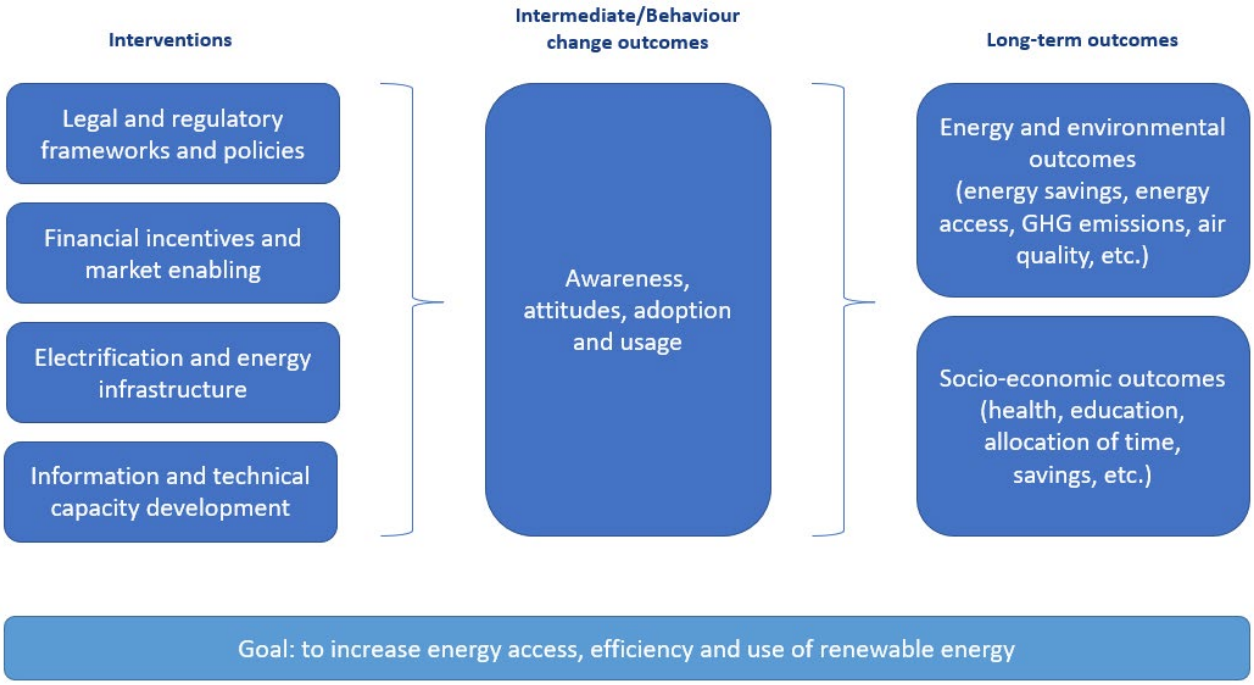
Electrification and energy infrastructure encompass activities which directly provide electricity, through renewable and non-renewable sources, as well as the provision of energy efficient upgrades and the maintenance of electrification systems. These activities can directly overcome

barriers related to access, which can be particularly prevalent for rural communities (Liao et al., 2021), while the upgrade of efficient infrastructure can help prevent non-technical losses, a common issue in L&MICs (Fowlie & Meeks, 2021).

Finally, the provision of information and capacity strengthening can overcome information barriers on both the supply and demand side. Limited technical capacity to implement electrification has been cited as a challenge particularly common to rural settings (Jimenez 2017; Almeshqab & Ustin, 2019), in relation to renewable technologies (Numata et al., 2020) and to the benefits of efficient buildings (Iwaro & Mashwa, 2010). On the demand side, there may be limited public awareness of the benefits of renewable technologies (Qazi et al., 2019) and constrained opportunities for households to learn about and improve their energy consumption habits (Dianshu et al., 2010). The inverse has been shown to be true as well, that is, many of the most successful, effective energy access and poverty reduction projects at the national level spent more money on capacity building than they did on installing renewable energy or equipment (see, for example, Sovacool 2013; Sovacool & Drupady 2012).

As shown in Figure 1, we anticipate that the four sets of interventions included in the EGM will work towards achieving SDG 7 through two steps. In the first step, interventions may change the behaviors of governments, firms, communities and individuals, leading to the adoption and use of energy efficient and renewable technologies. In the second step, these behavioral adjustments can lead to long-term and more general outcomes related to environmental and socio-economic and welfare measures.

Figure 1. Conceptual framework



3. Research Methods

This section presents the overall approach of the EGM and provides an overview of the search, screening, data extraction, analysis and reporting processes that will be implemented to address the EGM objectives. The last section describes the general approach we will take to conduct further synthesis of the evidence, as informed by the EGM results.

We will follow the standards and methods for EGMs developed by 3ie (Snilstveit et al., 2016; Snilstveit et al., 2017). EGMs are developed using systematic methods to identify and describe all completed and ongoing impact evaluations and systematic reviews relevant to research objectives. We outline the methods employed in more detail below.

3.1 Developing the framework

The scope and framework for this EGM were developed collaboratively by 3ie and SEforALL, with additional consultation with Benjamin Sovacool, the subject expert for this project. Further discussions and feedback from an advisory group of experts helped to ensure the framework was relevant to practitioners and researchers in the wider energy sector. Further details on the advisory group are available in Appendix A.

3.2 Criteria for including or excluding studies

3.2.1 Context, location, or setting

We will include studies that evaluate interventions in L&MICs as defined by the World Bank income group classification (see Appendix B). Studies will be classified based on the year the intervention began. Multi-country studies will be included if there is at least one effect estimate reported for an L&MIC.

Within these countries, we will not exclude studies based on their setting. This means we will include interventions implemented in rural and urban areas at the local, regional, national or international level, as well as those targeting firms, communities and households.

3.2.2 Population

We will include all populations with no limitations based on individual or community demographic, social or health characteristics.

3.2.3 Interventions

For the purpose of this project, an "intervention" is defined as a decision or set of activities deliberately undertaken by a specific entity—such as a government body, NGO, private firm or consortium—intended to influence events or outcomes, which another such entity could emulate. Studies where the independent variable is not an activity but a state of affairs will be excluded. For example, a households' access to electricity. This may be influenced by policies/programs but when the variation in the independent variable has arisen as a result of uncoordinated activity (i.e., some people have electricity and some do not) we argue there is no intervention. This is true even when a counterfactual study design is used to identify the causal impact of the independent variable, rather than mere associations.

Relevant interventions are grouped into four levels: energy-related policies and regulations, financial incentives to access energy technologies, construction and upgrade of energy infrastructure, and information and capacity development. The full list of included categories is detailed in Table 2 and to be included in the map, a study must have evaluated one of these interventions.

To ensure the EGM scope remains manageable, we will not include studies evaluating the effect of general environmental policies or regulations with no explicit linkage to the energy sector (e.g., Yu & Morotomi, 2022; Guo et al., 2022). Likewise, we will also exclude carbon markets and emission trading schemes unless they are specifically set up on the energy sector (e.g., Dong et al., 2022). Renewable energy credits (RECs) and tax credits, which are both specific to energy, are included within the framework. To ensure interventions are coded consistently with the literature, we will make use of extant research to create distinctions between concepts. For instance, how to differentiate between on- and off-grid systems (Sovacool, 2014).

.

Table 2. List of included interventions

Domain	Category	Definition	Example activities
Legal and regulatory frameworks and policies	Regulatory frameworks for energy infrastructure	Regulatory frameworks for the construction of energy infrastructure. Infrastructure may focus on providing electricity or may be related to energy efficiency codes.	• Energy efficiency standards for buildings • Standards reforms for the design of micro-grid systems • Improved standards for off-grid components and designs
	Energy targets and enforcement mechanisms	Energy targets may be set by governments or private firms and require actors to ensure they meet standards on the quantity or type of energy used. Enforcement mechanisms related to targets being met may include rewards in the form of environmental certifications or may include audits which ensure actors are meeting regulations. Where rewards for compliance are financial (e.g. accessing subsidies or better prices) these interventions will be included within the Financial Incentives group.	• Environmental compliance audits • Energy intensity targets • Renewable portfolio standards (RPS) / renewable electricity standards
	Energy management reforms	Reforms to the bodies which provide energy to populations for consumption and productive use.	• Energy service contracts • Privatization of the energy sector management
	Financial regulations for investment	Government financial regulations to incentivize investment (mainly private) in energy access, renewables or efficiency. This could be through tax deductions or tradeable credits specifically set up for the energy sector.	• Feed-in tariffs • Green taxes • Renewable Energy Credits (RECs)
	Other energy regulations and policies	Other regulations or policies related to energy but not concerned with infrastructure, targets, management or investment. Regulations must have a specific focus on the energy sector, and general environmental regulations which do not have a specific link to energy will be excluded.	• Daylight savings time • Harmful subsidy cuts (e.g. fossil fuel subsidies) • Energy planning regulations

Domain	Category	Definition	Example activities
Financial incentives and market enabling activities	Credits and loans	Provision of credit or loans to households or small and medium enterprises (SMEs) to mitigate liquidity constraints that may deter last-mile connectivity or the adoption of improved energy technologies, including the refitting of domestic appliances. Credit and loan schemes may also be offered to energy suppliers, for example, to enable the expansion of grid infrastructure. This also includes leasing schemes on energy-efficient technologies.	• Micro-credits to access energy technologies • Leasing of improved productive inputs and machinery
	Subsidies and other transfers	Subsidies, transfers, and other monetary incentives (e.g., vouchers or pay-as-you-go schemes) for households, SMEs, and energy suppliers that reduce the cost of adopting energy technologies and behaviors, as well as the setup, expansion, and maintenance of energy grids.	• Subsidies for utility companies to expand last-mile connectivity • Household transfers to adopt efficient cooling solutions or renewable energies
	Insurance and other risk guarantee instruments	Insurance and other risk guarantee instruments which lower the possible risk a household, community or enterprise might face when upgrading systems and infrastructure to a more renewable or efficient energy source.	• Insurance for SMEs + for communities
	Utility revenue collection	Interventions that introduce innovations and improve the ability of energy suppliers to recover energy bills and liabilities.	• Pre-paid electricity meters to avoid non-technical losses • Mobile money payments • Enforcement of energy contracts or disconnection notices
	Energy pricing	Demand-side interventions that reduce the price of electricity and other energy sources.	• Time-of-use pricing tariffs
	Push and pull finance	Push and pull financial instruments that align profitability with energy sustainability objectives and whose disbursement tends to be conditional on project performance indicators evaluated against predefined targets.	• Results-based financing • Impact investing • Advanced market commitments

Domain	Category	Definition	Example activities
Electrification and energy infrastructure	On-grid systems	Activities to construct new systems or expand existing on-grid electrification connections. This also includes when technologies/energy are received for free or purchased by households.	• Construction/expansion of on-grid systems • Rural electrification programs • Network densification measures • Local manufacturing of (renewable) energy technologies
	Off-grid systems	Activities to construct new systems or expand existing off-grid electrification connections. This also includes when technologies/energy are received for free or purchased by households.	• Construction/expansion of off-grid systems • Set up of home energy systems (e.g. solar lamps)
	Sustainable upgrades	Activities to manage and maintain energy systems aimed to lower the cost of producing and/or delivering energy; reduce service interruptions (including through the storage of energy); and improve the quality of energy supply. This category covers technical assistance only; interventions also providing information or capacity development will be included within the Information and capacity development group.	• Energy-efficient housing • Provision of improved cookstoves • Electrification replacement with renewable energy systems
	Management and maintenance of systems	Activities to manage and maintain energy systems aimed to lower the cost of producing and/or delivering energy; reduce service interruptions (including through the storage of energy); and improve the quality of energy supply. This category covers technical assistance only; interventions also providing information or capacity development will be included within the Information and capacity development group.	• Use of improved equipment to increase generation and improve efficiency • Measures to reduce transmission and distribution losses • Advance notifications on service interruptions and service restoration times • Energy storage technologies

Domain	Category	Definition	Example activities
Information and capacity development	Information dissemination	Information interventions to nudge individuals to change their behaviors and practices to (i) reduce their energy consumption by adopting energy efficiency technologies or (ii) encourage the adoption of renewable energy to substitute fossil fuel energy sources.	• Communication/advocacy campaigns through flyers or face-to-face discussions • Energy efficiency appliances labels • Energy ratings or certificates
	Monitoring and displaying energy consumption	Interventions that provide households or firms with their energy consumption records to induce them to use less energy.	• Home energy reports comparing own consumption with mean consumption in the area • Smart meters delivered to households
	Energy audits	Interventions where auditors visit and provide tailored information to households and firms to raise awareness of cost-effective energy efficiency or renewables technologies that would reduce their energy consumption and/or increase the adoption of renewables.	• Energy audits
	Technical assistance and capacity strengthening	Interventions that provide access to specific technical assistance to implement and scale up technologies or practices. The assistance is usually aimed to build people's or communities' (including local authorities) capacities to understand, implement, and use energy technologies, services and practices.	• Formal or informal training programs • Delivering workshops associated with the introduction of a new technology • One-stop shops that provide information on potential opportunities and funding access for investments
	Advocacy and diplomacy	Energy advocacy initiatives and diplomatic efforts, which convene stakeholders to foster collaboration and cooperation in the energy sector. These interventions are outside of market enabling activities.	• Support or coordination of international or in-country efforts for energy transition

3.2.4 Outcomes

The list of behavioral, environmental and welfare outcomes of interest is detailed in Table 3. Studies need to report on at least one of these outcomes to be included in the map. Outcomes were selected based on previous electrification and energy efficiency frameworks, as well as the outcomes presented within studies identified during preliminary scoping.

Table 3. List of included outcomes

Domain	**Category**	**Description**
Intermediate/ Behavior change	Knowledge, awareness and attitudes	Measures of knowledge, awareness and attitudes of individuals towards energy, including energy access, efficiency and the use of renewables. This excludes indicators or survey findings related to the adoption, usage or uptake of behaviors or technology.
	Behavior adoption (excluding technology usage)	Measures of adoption of behaviors related to energy, for example, energy conservation behavior. This excludes behaviors related to the uptake or usage of technology.
	Technology uptake	Measures of the uptake of an energy technology, including the uptake of a new technology or the replacement/upgrade with a more efficient or renewable technology. This excludes measures related to how the technology is used.
	Usage of the technology	Measures of the usage of an energy technology after it has been received or bought. This includes usage measures of alternative technologies, for example, those the interventions aimed to replace.
Energy and environmental outcomes	Energy net savings or consumption	Measures of net energy or demand savings as the portion of gross savings that is attributable to the programme. This involves separating out impacts that are a result of other influences, such as consumer self-motivation.
	Energy security	Measures of energy security as the uninterrupted availability of energy sources at an affordable price. This includes measures of energy poverty and individual indicators of energy access, quality and reliability.

Domain	**Category**	**Description**
	GHG emissions	Measures of carbon related emissions (CO2) and non-carbon related emissions e.g. methane (CH4), nitrous oxide (N2O) and fluorinated gases.
	Water consumption	Measures of (net) water consumption.
	Air quality/pollution	Measures of air pollution or greenhouse gases that would have been emitted if more energy had been consumed in the absence of the intervention. These emissions, usually of SOX, NOX, black carbon or particulate matter, can be from combustion of fuels at an electrical power plant or from combustion of heating fuels, such as natural gas and fuel oil, at a project site.
	Energy innovations	Measures of research and development or innovations in the energy sector, for example, green energy patents.
Socio-economic and community welfare	Income, savings and expenditures	Measures of income, revenue, savings and expenditures.
	Assets value	Measures of assets value including, for example, building stock or land value.
	Employment	Measures of employment including, for example, hours worked or new jobs created.
	Health status, comfort and wellbeing	Measures of health status, comfort or wellbeing of individuals and communities.
	Education	Measures of education including learning outcomes (e.g. test scores), access to education (e.g. school drop-out) or hours dedicated to education.
	Women's empowerment	Measures of women's empowerment including indices and individual indicators of, for example, decision making. Related outcomes disaggregated by sex will be included in the respective categories (e.g. indicators of women's time allocation will be included in the Allocation of time category).
	Allocation of time	Measures of time allocated to activities excluding employment and education.

Domain	**Category**	**Description**
	Crime and security	Measures of crime and security at the individual (e.g. crime perception) or community levels (e.g. crime rates).
	Public service delivery	Measures of public service delivery (e.g. health services utilization).

3.2.5 Study designs

We will include studies that look at the effect of an intervention aimed at promoting sustainable energy in L&MICs. Specifically, we will include studies that adopt methods estimating effects that can be attributed to an intervention, as compared to what would have happened in the absence of the intervention. We will include studies that implement at least one of the following study designs that are widely used to evaluate intervention effectiveness (Aloe et al. 2017; Reeves et al. 2017):

1. Randomised controlled trials (RCTs)
2. Natural experiments with clearly defined intervention and comparison groups
3. Regression discontinuity designs (RDD)
4. Instrumental variables (IV)
5. Endogenous treatment-effects models, endogenous switching regression, and other methods synonymous to the Heckman two step model
6. Difference-in-differences (DID), two-way fixed-effects (TWFE), and two-way Mundlak regressions (TWM)
7. Interrupted time series (ITS) models
8. Weighting and matching approaches which control for observable confounding
9. Synthetic control methods
10. Systematic effectiveness reviews

The specific criteria required for inclusion is presented in Appendix C, drawing on commonly accepted standards for impact evaluations (Gertler et al. 2016) and systematic reviews (Waddington et al., 2012).

3.2.6 Other inclusion and exclusion criteria

- **Language:** We will include studies published in any language, although the search terms used will be in English only.
- **Publication date:** We will include studies published in 2000 or after. From 3ie's experience developing other EGMs, less than one per cent of impact evaluations and systematic reviews of interventions implemented L&MICs predate the year 2000, hence the likelihood of missing eligible studies is small. This will ensure the map remains manageable as well as relevant to the current technologies and programs being used in the energy sector.
- **Publication status:** We will include ongoing and completed impact evaluations and systematic reviews, both peer-reviewed studies and 'grey' literature. For on-going studies, we will include prospective study records, protocols and trial registrations.

3.3 Search strategy

We will adopt a systematic search strategy following guidelines for systematic literature searching (Kugley et al. 2017). The search strategy has been developed in collaboration with the information specialist Alison Bethel. This has been designed to address potential publication bias issues by systematically searching academic bibliographic databases and implementing additional searches for grey literature in specialist organizational websites, websites of bilateral and multilateral agencies and repositories of research in international development. To minimize the likelihood of missing studies, we will conduct forward and backward citation tracking of included studies. In addition, and where possible, the review team will contact key experts and organizations through the advisory group to identify additional studies that meet the inclusion criteria.

The full list of literature sources to be used and an example of the search strings for one database is presented in Appendix D. The precise strings and logic (e.g., index terms and truncation operators) will be adapted for each database and platform.

3.4 Screening protocol

The selection of studies will be managed using EPPI-Reviewer software (Thomas et al. 2022) and will be completed by implementing the following steps:

- **Import study records**: All output files (e.g. RIS or txt files) of the search strategy will be imported into EPPI.
- **Removal of duplicate studies**: An automated process within EPPI will be used to remove known duplicate files.
- **Training of screeners:** A team of consultants will be trained on the protocol by the core project team, with a focus on understanding the subject matter and the screening process. All coders will screen the same set of studies until an 85 per cent level of consistency is achieved in terms of the decision to include or exclude a study (i.e. did screeners make the decision that was consistent with the core team?).
- **Title and abstract screening:** The title and abstract (T/A) of all imported and de-duplicated studies will be screened by one coder, who will give a judgment of include, exclude, or unsure. Items marked as unsure will be screened by a second screener (an approach that has been demonstrated to produce comparable results to double screening at significantly lower cost; Shemilt et al. 2016). Several exclude codes will be available to provide information on the exclusion reason in each case. Screening codes will be applied in a hierarchical order so that consistent comparisons can be made about why studies were excluded and at what stage in the screening process. Periodic meetings will be held by members of the core team to address studies flagged for a second opinion and make any refinements to the screening approach. The output of this process will be a set of potentially relevant studies that will be put forward for full text screening.
  We will use the machine-learning features of EPPI to accelerate the T/A screening process. We will first run a classifier model based on 3ie's DEP data to identify studies more likely to be included (i.e. impact evaluations or systematic reviews conducted in L&MICs). If the number of records is still deemed high, we may run a second classifier model based on the screening decision of a random subset of 100 records across likelihood ranges from the first model (i.e. energy-relevant studies). We will combine the information from these two models and screen all records with an inclusion probability score of at least 0.30 in any of them. As a precautionary measure, we will screen a random sample of records with lower probability scores to determine if any should be

included for full-text screening. If more than one percent of this sample is found to be includable, we will proceed to screen all records.

- **Full text screening:** We will retrieve the full text for each study that meets the inclusion criteria at T/A. Two reviewers will independently examine each full text in detail against the protocol. A third core team member will resolve any disagreements. A code will be applied to each study reflecting if it is included or why it is excluded. The output of this stage will be a set of studies deemed suitable to be included in the EGM.
- **Checks for linked publications:** The project team will group publications that evaluate the same intervention and study population (i.e., publications that report on the same study). This typically occurs in cases where a group of authors publishes more than one paper in relation to one particular study on a specific population, for example a working paper before a journal article. The latest publication will be classified as the main record for that group of linked studies. Descriptive information will only be extracted once for each group of linked publications, drawing on all linked publications so that extraction is as comprehensive as possible.

Each step in this process will be documented in detail and graphically presented in a flow chart in the final report to facilitate replication of the project's approach.

3.5 Data extraction and critical appraisal

We will systematically extract data from all included studies into 3ie's DEP admin panel based on the categories presented in Appendix E. Many of these categories will then become filters in the online interactive map. The data will cover the following areas:

- **Study and publication information:** We will extract the general characteristics of the study including author(s), publication date and status, study location, intervention type, outcomes reported, definition of outcome measures, population of interest, program and evaluation funders, time periods for delivery and analysis.
- **Topical cross-cutting issues:** We will extract data on a number of cross-cutting issues, including gender, inclusion and equity dimensions, and energy-related categories (e.g., SDG 7 pillar (i.e., access, renewables, efficiency), energy source, technology used).
- **Critical appraisal:** All included systematic reviews will be critically appraised following the practices adopted by 3ie's systematic review database protocol, which draws on

Lewin and colleagues (2009). This appraisal assesses systematic reviews according to criteria relating to the search, screening, data extraction, and synthesis activities conducted, and covers the most common areas where biases are introduced. Each systematic review will be rated as low, medium, or high confidence drawing on guidance provided in Snilstveit and colleagues (2017). The critical assessment tool we will use for this process is presented in Appendix F. We will not appraise impact evaluations, as this is beyond the scope of EGMs.

The following processes will be implemented to collect data from included studies:

- **Develop and refine data extraction tools and codebooks:** The draft tools developed for this project will be reviewed and potentially refined considering any feedback received by the EGM advisory group and insights from project implementation.
- **Data extraction training and pilot:** Coders assigned to each data extraction task will undergo theory- and practice-based training in using the tools provided and the DEP admin panel. Each coding group will code a 'training set' of studies and assessments of inter-rater reliability will be calculated. Additional group training will be completed as required prior to the main-stage extraction.
- **Main-stage extraction:** In the case of descriptive and equity-based information, studies will be coded by two coders independently. In the case of critical appraisal assessments, studies will first be single coded and then reviewed by a systematic review methods expert. Meetings will be held periodically with coders on the project to provide support and resolve queries.
- **Quality checks:** As the data extraction process develops, the project team will run regular data quality checks. In practice, a member of the core team will check the consistency of data extracted in duplicate by consultants and measures of consistency will be calculated and used to inform the checking process.

3.6 Analysis and reporting

3.6.1 Analysis of the evidence

We will conduct a range of descriptive analyses to provide an overview of the size and extent of the evidence, based on the following dimensions:

- Publication year
- Geography
- Interventions
- Outcomes
- Study design
- Results of the critical appraisal of systematic reviews
- Equity and cross cutting themes considerations (e.g. SDG 7 pillars, intervention implementation levels, gender considerations)
- Research funding and implementing agencies

3.6.2 Dealing with multicomponent interventions

Each study will be coded according to the list of included interventions. If a study evaluated multiple intervention (non-control) arms or activities, the study will be categorized as multicomponent to prevent biasing the extent of evidence available. We anticipate identifying some intervention packages and will create new categories as appropriate, to reflect the most common combinations of intervention components.

3.6.3 EGM reporting

Where appropriate, we will consider running cross-tabs to provide a more nuanced overview of the evidence identified. We will produce the following analytical outputs:

- **Interactive EGM:** An interactive evidence gap map that visually presents the current evidence base categorized by coverage of the pre-determined interventions-outcomes framework, confidence rating and study completion status. Filters will be incorporated into the map to enable a more targeted use – for example, by restricting the studies to a specific unit of analysis (see other filter examples in Appendix E). The map will be stored on the 3ie website and shared as a public good.
- **Presentation:** A presentation will provide an overview of the emerging findings of the EGM. This will be presented by the project team and will provide an opportunity for SEforALL to comment on the findings and to collaboratively discuss opportunities for additional analyses, presentation of results and implications. It will be designed such that it can be used by SEforALL and other stakeholders for their internal learning purposes.

- **EGM technical report:** The EGM technical report will include an overview of the conceptual framework, methods, and key results of the EGM. It will provide a high level of analytical detail and will be supported by technical annexes. This report will conclude by directly addressing the EGM research questions by providing a discussion of the main evidence clusters and gaps in the literature (following the intervention and outcome framework) and the implications of these findings for policymakers and researchers in terms of priorities for future programming and research. The technical report will be published by 3ie and shared as a public good.
- **EGM executive summary:** This short report will provide a high-level summary of the results and will primarily focus on addressing the research questions using non-technical language.

3.7 Systematic review of the evidence

Once the map of the evidence available is completed, the project team will look closer into this evidence through a systematic review. The specific synthesis topic will be informed by the findings of the EGM and will be decided in conjunction with SEforALL. Depending on the topic, we may aim to address questions on effectiveness, heterogeneity of effects, cost-effectiveness, influencing factors to achieve impact, or a combination of these questions. We will not run another search for literature but depending on the research question, the team may decide to complement the search conducted for the EGM to identify additional documentation (e.g., process evaluations of relevant programs).

For this review, we will follow the Methodological Expectations of Campbell Collaboration Intervention Reviews (MECCIR) conduct and reporting Standards (2019a, 2019b) and the process will be based on recognized guidelines for systematic reviews of effectiveness in international development (Waddington et al., 2012). Further details on the data extraction processes, studies appraisal, data treatment and analyses approaches for this synthesis are presented in Appendix G.

4.4. Additional information about the EGM

4.1 Timeline

The EGM is expected to be completed by September and the technical report by October 2023. The systematic review, along with its report, is expected to be completed by December 2023. All final outputs will be published on 3ie's Development Evidence Portal.

4.2 Engagement and communication plan

It is important that the results of the EGM are shared with SEforALL and its internal audiences, and more broadly with the development sector working on energy. We will complete the following activities to engage with key stakeholders to attempt to ensure the results of the project accurately reflect their policy and research needs:

- **Develop an EGM advisory group:** The project team, in collaboration with SEforALL, will engage with key stakeholders with academic and/or practitioner energy expertise. We will set up an advisory group which will provide pro-bono support to the project at several key stages of the project, including developing the protocol, reviewing the search results produced, reviewing and interpreting emerging findings and developing and optimizing the analytical outputs produced to aid evidence uptake and use.
- **Develop a Stakeholder Engagement and Communication Plan (SECP):** The aim of this plan is to ensure that findings from the EGM are effectively disseminated to the appropriate audiences, in an engaging and accessible format. The plan includes a provisional analysis of key stakeholder groups, focusing on their relevant interests, the extent to which 3ie and/or SEforALL have access to them, and an assessment of the outputs with most value-added to aid evidence uptake and use. The SECP is a 'live document' and will be refined as additional information needs or dissemination opportunities are identified by the project team, advisory group or SEforALL.
- Further outreach: This stage will involve two activities. First, in order to ensure the map is as comprehensive as possible we will publish a blog calling for additional studies for both the EGM and SR. This will allow relevant experts and practitioners to suggest studies

which may be unpublished or ongoing. Secondly, based on the SECP, we will seek to disseminate both the EGM and SR findings at relevant events.

4.3 Roles and responsibilities

- Scope and framework development: Constanza Gonzalez Parrao, Cem Yavuz, Miriam Berretta, Dina Rodrigues, Benjamin Sovacool, Solomon Asfaw, Frederick Gaved, Birte Snilstveit
- EGM methodology: Constanza Gonzalez Parrao, Cem Yavuz
- Search strategy: Alison Bethel, Constanza Gonzalez Parrao, Cem Yavuz
- Screening and data extraction: Cem Yavuz, Zafeer Ravat, Dina Rodrigues
- Analysis and reporting: Constanza Gonzalez Parrao, Cem Yavuz, Miriam Berretta, Zafeer Ravat, Dina Rodrigues, Birte Snilstveit

4.4 Declaration of interest

The authors declare no conflicts of interest. While 3ie and SEforALL have collaboratively developed the EGM scope and framework, SEforALL will not influence the implementation of the EGM or its analysis. 3ie and SEforALL will collaborate again at the reporting and dissemination stages of the project to help ensure the outputs are useful and relevant to the energy sector.

5.5. References

ADB. 2010. “Asian Development Bank’s Assistance for Rural Electrification in Bhutan - Does Electrification Improve the Quality of Rural Life?” IES: BHU 2010-27. Independent Evaluation Department. Asian Development Bank. https://www.adb.org/documents/asian-development-banks-assistance-rural-electrification-bhutan-does-electrification.

Akella, A.K., R.P. Saini, and M.P. Sharma. 2009. “Social, Economical and Environmental Impacts of Renewable Energy Systems.” Renewable Energy 34 (2): 390–96. https://doi.org/10.1016/j.renene.2008.05.002.

Almeshqab, Fatema, and Taha Selim Ustun. 2019. “Lessons Learned from Rural Electrification Initiatives in Developing Countries: Insights for Technical, Social, Financial and Public Policy Aspects.” *Renewable and Sustainable Energy Reviews* 102 (March): 35–53. https://doi.org/10.1016/j.rser.2018.11.035.

Aloe, Ariel M., Betsy Jane Becker, Maren Duvendack, Jeffrey C. Valentine, Ian Shemilt, and Hugh Waddington. 2017. “Quasi-Experimental Study Designs Series—Paper 9: Collecting Data from Quasi-Experimental Studies.” *Journal of Clinical Epidemiology* 89 (September): 77–83. https://doi.org/10.1016/j.jclinepi.2017.02.013.

Bayer, Patrick, Ryan Kennedy, Joonseok Yang, and Johannes Urpelainen. 2020. “The Need for Impact Evaluation in Electricity Access Research.” Energy Policy 137 (February): 111099. https://doi.org/10.1016/j.enpol.2019.111099

Bensch, Gunther, Maximiliane Sievert, Jörg Langbein and Nadine Kneppel. 2016.”Effects and mechanisms of market-based reforms on access to electricity in developing countries.” 3ie Systematic Review 31. London: International Initiative for Impact Evaluation. https://www.3ieimpact.org/sites/default/files/2017-11/sr31-market-research.pdf

Berretta, Miriam, Joshua Furgeson, Yue (Nicole) Wu, Collins Zamawe, Ian Hamilton, and John Eyers. 2021a. “Residential Energy Efficiency Interventions: A Meta-analysis of Effectiveness Studies.” *Campbell Systematic Reviews* 17 (4): e1206. https://doi.org/10.1002/cl2.1206.

Berretta, Miriam, Collins Zamawe, Paul J Ferraro, Neal Haddaway, Jan Minx, Birte Snilstveit, and John Eyers. 2021b. “Mapping Energy Efficiency Interventions.” 3ie Evidence Gap Map Report 17. New Delhi: International Initiative for Impact Evaluation (3ie). https://doi.org/10.23846/EGM017.

Bhattacharyya, Subhes C. 2013. “Financing Energy Access and Off-Grid Electrification: A Review of Status, Options and Challenges.” *Renewable and Sustainable Energy Reviews* 20 (April): 462–72. https://doi.org/10.1016/j.rser.2012.12.008.

Bhattarai, Utsav, Tek Maraseni, and Armando Apan. 2022. “Assay of Renewable Energy Transition: A Systematic Literature Review.” *Science of The Total Environment* 833 (August): 155159. https://doi.org/10.1016/j.scitotenv.2022.155159.

Bonan, Jacopo, Stefano Pareglio, and Massimo Tavoni. 2017. “Access to Modern Energy: A Review of Barriers, Drivers and Impacts.” *Environment and Development Economics* 22 (5): 491–516. https://doi.org/10.1017/S1355770X17000201.

Borenstein, Michael, Larry V. Hedges, Julian P. T. Higgins, and Hannah R. Rothstein. 2009. *Introduction to Meta-Analysis*. 1st ed. Wiley. https://doi.org/10.1002/9780470743386.

Bos, Kristine, Duncan Chaplin, and Arif Mamun. 2018. “Benefits and Challenges of Expanding Grid Electricity in Africa: A Review of Rigorous Evidence on Household Impacts in Developing Countries.” *Energy for Sustainable Development* 44 (June): 64–77. https://doi.org/10.1016/j.esd.2018.02.007.

CASP. n.d. “Critical Appraisal Skills Programme Checklists.” https://casp-uk.net/casp-tools-checklists/.

Climate Analytics. 2022. ‘Renewable Energy Transition in Sub-Saharan Africa’. https://climateanalytics.org/publications/2022/renewable-energy-transition-in-sub-saharan-africa.

De La Rue, Lisa, Joshua R. Polanin, Dorothy L. Espelage, and Terri D. Pigott. 2013. “PROTOCOL: School-based Interventions to Reduce Dating and Sexual Violence: A Systematic Review.” *Campbell Systematic Reviews* 9 (1): 1–43. https://doi.org/10.1002/CL2.106.

Dhaliwal, I, E Duflo, R Glennerster, and C Tulloch. 2013. “Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education.” In *Education Policy in Developing Countries*, edited by Paul Glewwe. University of Chicago Press. https://doi.org/10.7208/chicago/9780226078854.001.0001.

Dianshu, Feng, Benjamin K. Sovacool, and Khuong Minh Vu. 2010. “The Barriers to Energy Efficiency in China: Assessing Household Electricity Savings and Consumer Behavior in Liaoning Province.” *Energy Policy* 38 (2): 1202–9. https://doi.org/10.1016/j.enpol.2009.11.012.

Dong, Zhaoyingzi, Shaojian Wang, Weiwen Zhang, and Huijun Shen. 2022. ‘The Dynamic Effect of Environmental Regulation on Firms’ Energy Consumption Behavior-Evidence from China’s Industrial Firms’. Renewable and Sustainable Energy Reviews 156 (March): 111966. https://doi.org/10.1016/j.rser.2021.111966.

Egger, M., G. D. Smith, M. Schneider, and C. Minder. 1997. “Bias in Meta-Analysis Detected by a Simple, Graphical Test.” *BMJ* 315 (7109): 629–34. https://doi.org/10.1136/bmj.315.7109.629.

Ellis, Paul D. 2010. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO9780511761676.

Fisher, Zachary, and Elizabeth Tipton. 2015. “Robumeta: An R-Package for Robust Variance Estimation in Meta-Analysis.” arXiv. http://arxiv.org/abs/1503.02220.

Fowlie, Meredith, and Robyn Meeks. 2021. “The Economics of Energy Efficiency in Developing Countries.” *Review of Environmental Economics and Policy* 15 (2): 238–60. https://doi.org/10.1086/715606.

Gao, Yun, Xiang Gao, and Xiaohua Zhang. 2017. “The 2 °C Global Temperature Target and the Evolution of the Long-Term Goal of Addressing Climate Change—From the United Nations Framework Convention on Climate Change to the Paris Agreement.” *Engineering* 3 (2): 272–78. https://doi.org/10.1016/J.ENG.2017.01.022.

GCF. n.d. “Resource Mobilisation.” Green Climate Fund. Accessed June 14, 2023. https://www.greenclimate.fund/about/resource-mobilisation.

Gertler, Paul, Sebastian Martinez, Patrick Premand, Laura Rawlings, and Christel Vermeersch. 2016. *Impact Evaluation in Practice*. Second Edition. Washington, DC: Inter-American Development Bank and World Bank. http://hdl.handle.net/10986/25030.

Goldacre, Ben, Henry Drysdale, Anna Powell-Smith, Aaron Dale, Ioan Milosevic, Eirion Slade, Philip Hartley, Cicely Marston, Kamal Mahtani, and Carl Heneghan. 2016. “The COMPare Trials Project.” COMPare. 2016. http://compare-trials.org.

Gonzalez-Caceres, A, M Rabani, and P A Wegertseder Martínez. 2019. “A Systematic Review of Retrofitting Tools for Residential Buildings.” *IOP Conference Series: Earth and Environmental Science* 294 (1): 012035. https://doi.org/10.1088/1755-1315/294/1/012035.

Gribkova, Darya, and Yulia Milshina. 2022. “Energy Transition as a Response to Energy Challenges in Post-Pandemic Reality.” *Energies* 15 (3): 812. https://doi.org/10.3390/en15030812.

Guo, Bingnan, Yu Wang, Yu Feng, Chunyan Liang, Li Tang, Xiafei Yao, and Feng Hu. 2022. “The Effects of Environmental Tax Reform on Urban Air Pollution: A Quasi-Natural Experiment Based on the Environmental Protection Tax Law.” *Frontiers in Public Health* 10 (August): 967524. https://doi.org/10.3389/fpubh.2022.967524.

Hamburger, David, Joel Jaeger, Patrick Bayer, Ryan Kennedy, Joonseok Yang, and Johannes Urpelainen. 2019. “Shades of Darkness or Light? A Systematic Review of Geographic Bias in Impact Evaluations of Electricity Access.” Energy Research & Social Science 58 (December): 101236. https://doi.org/10.1016/j.erss.2019.101236.

Hasanbeigi, Ali, Christoph Menke, and Peter du Pont. 2010. “Barriers to Energy Efficiency Improvement and Decision-Making Behavior in Thai Industry.” *Energy Efficiency* 3 (1): 33–52. https://doi.org/10.1007/s12053-009-9056-8.

Hedges, Larry V., Elizabeth Tipton, and Matthew C. Johnson. 2010. “Robust Variance Estimation in Meta-Regression with Dependent Effect Size Estimates.” *Research Synthesis Methods* 1 (1): 39–65. https://doi.org/10.1002/jrsm.5.

Hedges, Larry V. 2009. “Effect Sizes in Nested Designs.” In *The Handbook of Research Synthesis and Meta-Analysis*, edited by Harris M. Cooper, Larry V. Hedges, and Jeff C. Valentine, 2nd ed, 337–55. New York: Russell Sage Foundation.

Higgins, JPT, J Thomas, J Chandler, M Cumpston, T Li, MJ Page, and VA Welch, eds. 2020. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.1 (Updated September 2020). Cochrane. www.training.cochrane.org/handbook.

Higgins, Julian PT, Jonathan AC Sterne, Jelena Savovic, Matthew J. Page, Asbjørn Hróbjartsson, Isabelle Boutron, Barney Reeves, and Sandra Eldridge. 2016. "A Revised Tool for Assessing Risk of Bias in Randomized Trials." *Cochrane Database of Systematic Reviews* 10 (Suppl 1): 29–31.

Hombrados, G J, and Hugh Waddington. 2012. "Internal Validity in Social Experiments and Quasi-Experiments: An Assessment Tool for Reviewers." Working Paper. 3ie.

IEA. 2021. 'Net Zero by 2050: A Roadmap for the Global Energy Sector'. International Energy Agency. https://www.iea.org/reports/net-zero-by-2050.

———. 2022. 'World Energy Outlook 2022'. International Energy Agency. https://www.iea.org/reports/world-energy-outlook-2022.

———. 2023. 'Tracking Clean Energy Progress 2023'. Paris: International Energy Agency. https://www.iea.org/reports/tracking-clean-energy-progress-2023.

IEA, IRENA, UNSD, World Bank, and WHO. 2023. 'Tracking SDG 7: The Energy Progress Report'. Tracking SDG 7: The Energy Progress Report 2023. Washington, DC.: World Bank. https://trackingsdg7.esmap.org/data/files/download-documents/sdg7-report2023-full_report.pdf.

IEG. 2008. "The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits." Independent Evaluation Group. Washington, DC.: World Bank. http://hdl.handle.net/10986/6519.

International Council for Science. 2017. "A Guide to SDG Interactions: From Science to Implementation." https://council.science/wp-content/uploads/2017/05/SDGs-Guide-to-Interactions.pdf.

Irwin, Bridget R., Klesta Hoxha, and Karen A. Grépin. 2020. "Conceptualising the Effect of Access to Electricity on Health in Low- and Middle-Income Countries: A Systematic Review." Global Public Health 15 (3): 452–73. https://doi.org/10.1080/17441692.2019.1695873.

Iwaro, Joseph, and Abraham Mwasha. 2010. "A Review of Building Energy Regulation and Policy for Energy Conservation in Developing Countries." *Energy Policy* 38 (12): 7744–55. https://doi.org/10.1016/j.enpol.2010.08.027.

Jimenez, Raul. 2017. "Barriers to Electrification in Latin America: Income, Location, and Economic Development." *Energy Strategy Reviews* 15 (March): 9–18. https://doi.org/10.1016/j.esr.2016.11.001.

Kamal, Athar, Sami G. Al-Ghamdi, and Muammer Koc. 2019. "Revaluing the Costs and Benefits of Energy Efficiency: A Systematic Review." *Energy Research & Social Science* 54 (August): 68–84. https://doi.org/10.1016/j.erss.2019.03.012.

Keef, Stephen P., and Leigh A. Roberts. 2004. “The Meta-Analysis of Partial Effect Sizes.” *British Journal of Mathematical and Statistical Psychology* 57 (1): 97–129. https://doi.org/10.1348/000711004849303.

Khanna, Tarun M., Giovanni Baiocchi, Max Callaghan, Felix Creutzig, Horia Guias, Neal R. Haddaway, Lion Hirth, et al. 2021. “A Multi-Country Meta-Analysis on the Role of Behavioural Change in Reducing Energy Consumption and CO2 Emissions in Residential Buildings.” *Nature Energy* 6 (9): 925–32. https://doi.org/10.1038/s41560-021-00866-x.

Kivimaa, Paula, and Mari Martiskainen. 2018. “Innovation, Low Energy Buildings and Intermediaries in Europe: Systematic Case Study Review.” *Energy Efficiency* 11 (1): 31–51. https://doi.org/10.1007/s12053-017-9547-y.

Kugley, Shannon, Anne Wade, James Thomas, Quenby Mahood, Anne-Marie Klint Jørgensen, Karianne Hammerstrøm, and Nila Sathe. 2017. “Searching for Studies: A Guide to Information Retrieval for Campbell Systematic Reviews.” *Campbell Systematic Reviews* 13 (1): 1–73. https://doi.org/10.4073/cmg.2016.1.

Lewin, Simon, Andrew D Oxman, John N Lavis, and Atle Fretheim. 2009. “SUPPORT Tools for Evidence-Informed Health Policymaking (STP) 8: Deciding How Much Confidence to Place in a Systematic Review.” *Health Research Policy and Systems* 7 (S1): S8. https://doi.org/10.1186/1478-4505-7-S1-S8.

Liao, Chuan, James T. Erbaugh, Allison C. Kelly, and Arun Agrawal. 2021. “Clean Energy Transitions and Human Well-Being Outcomes in Lower and Middle Income Countries: A Systematic Review.” *Renewable and Sustainable Energy Reviews* 145 (July): 111063. https://doi.org/10.1016/j.rser.2021.111063.

Macdonald, Geraldine, Julian PT Higgins, Paul Ramchandani, Jeffrey C. Valentine, Latricia P. Bronger, Paul Klein, Roland O’Daniel, et al. 2012. “Cognitive-Behavioural Interventions for Children Who Have Been Sexually Abused: A Systematic Review.” *Campbell Systematic Reviews* 8 (1): 1–111. https://doi.org/10.4073/csr.2012.14.

Mathur, Kavita, Sandy Oliver, and Janice Tripney. 2015. “PROTOCOL: Access to Electricity for Improving Health, Education and Welfare in Low- and Middle-Income Countries: A Systematic Review.” Campbell Systematic Reviews 11 (1): 1–55. https://doi.org/10.1002/CL2.140.

Mawhood, Rebecca, and Robert Gross. 2014. “Institutional Barriers to a ‘Perfect’ Policy: A Case Study of the Senegalese Rural Electrification Plan.” *Energy Policy* 73 (October): 480–90. https://doi.org/10.1016/j.enpol.2014.05.047.

Miles, Matthew B, and A. Michael Huberman. 1994. *An Expanded Source Book: Qualitative Data Analysis, Second Edition*. London: SAGE Publications.

Moore, Nick, Douglas Glandon, Janice Tripney, Tomasz Kozakiewicz, Shannon Shisler, John Eyres, Rami Zalfou, et al. 2020. “Effects of Access to Electricity Interventions on Socio-Economic Outcomes in Low- and Middle- Income Countries.” 3ie Systematic Review 45. London: International Initiative for Impact Evaluation (3ie). https://doi.org/10.23846/SR00045.

Nehler, Therese. 2018. “A Systematic Literature Review of Methods for Improved Utilisation of the Non-Energy Benefits of Industrial Energy Efficiency.” *Energies* 11 (12): 3241. https://doi.org/10.3390/en11123241.

Numata, Masako, Masahiro Sugiyama, and Gento Mogi. 2020. “Barrier Analysis for the Deployment of Renewable-Based Mini-Grids in Myanmar Using the Analytic Hierarchy Process (AHP).” *Energies* 13 (6): 1400. https://doi.org/10.3390/en13061400.

Oliver, S, K Dickson, M Bangpan, and M Newman. 2017. “Getting Started with a Review.” In *An Introduction to Systematic Reviews*, Second Edition, 71–92. SAGE Publications.

Owusu, Phebe Asantewaa, and Samuel Asumadu-Sarkodie. 2016. “A Review of Renewable Energy Sources, Sustainability Issues and Climate Change Mitigation.” Edited by Shashi Dubey. Cogent Engineering 3 (1): 1167990. https://doi.org/10.1080/23311916.2016.1167990.

Painuly, J.P, H Park, M.-K Lee, and J Noh. 2003. “Promoting Energy Efficiency Financing and ESCOs in Developing Countries: Mechanisms and Barriers.” *Journal of Cleaner Production* 11 (6): 659–65. https://doi.org/10.1016/S0959-6526(02)00111-7.

Palit, Debajit, and Atul Kumar. 2022. “Drivers and Barriers to Rural Electrification in India – A Multi-stakeholder Analysis.” *Renewable and Sustainable Energy Reviews* 166 (September): 112663. https://doi.org/10.1016/j.rser.2022.112663.

Peters, Jaime L., Alex J. Sutton, David R. Jones, Keith R. Abrams, and Lesley Rushton. 2008. “Contour-Enhanced Meta-Analysis Funnel Plots Help Distinguish Publication Bias from Other Causes of Asymmetry.” *Journal of Clinical Epidemiology* 61 (10): 991–96. https://doi.org/10.1016/j.jclinepi.2007.11.010.

Polanin, Joshua R., Emily E. Tanner-Smith, and Emily A. Hennessy. 2016. “Estimating the Difference Between Published and Unpublished Effect Sizes: A Meta-Review.” *Review of Educational Research* 86 (1): 207–36. https://doi.org/10.3102/0034654315582067.

Qazi, Atika, Fayaz Hussain, Nasrudin ABD. Rahim, Glenn Hardaker, Daniyal Alghazzawi, Khaled Shaban, and Khalid Haruna. 2019. “Towards Sustainable Energy: A Systematic Review of Renewable Energy Sources, Technologies, and Public Opinions.” *IEEE Access* 7: 63837–51. https://doi.org/10.1109/ACCESS.2019.2906402.

R Core Team. 2020. “R: A Language and Environment for Statistical Computing.” Viena, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Rasmussen, Josefine. 2017. “The Additional Benefits of Energy Efficiency Investments—a Systematic Literature Review and a Framework for Categorisation.” *Energy Efficiency* 10 (6): 1401–18. https://doi.org/10.1007/s12053-017-9528-1.

Reeves, Barnaby C., George A. Wells, and Hugh Waddington. 2017. “Quasi-Experimental Study Designs Series—Paper 5: A Checklist for Classifying Studies Evaluating the Effects on Health Interventions—a Taxonomy without Labels.” *Journal of Clinical Epidemiology* 89 (September): 30–42. https://doi.org/10.1016/j.jclinepi.2017.02.016.

Sánchez-Meca, Julio, Fulgencio Marín-Martínez, and Salvador Chacón-Moscoso. 2003. “Effect-Size Indices for Dichotomized Outcomes in Meta-Analysis.” *Psychological Methods* 8 (4): 448–67. https://doi.org/10.1037/1082-989X.8.4.448.

Shemilt, Ian, Nada Khan, Sophie Park, and James Thomas. 2016. “Use of Cost-Effectiveness Analysis to Compare the Efficiency of Study Identification Methods in Systematic Reviews.” *Systematic Reviews* 5 (1): 140. https://doi.org/10.1186/s13643-016-0315-4.

Shemilt, I, M Mugford, S Byford, M Drummond, E Eisenstein, M Knapp, J Mallender, et al. 2008. “The Campbell Collaboration Economics Methods Policy Brief.” Campbell Collaboration. https://www.campbellcollaboration.org/images/pdf/plain-language/Economic_Methods_Policy_Brief.pdf.

Singh, Sanjeet, and Jayaram Ru. 2022. “Accessibility, Affordability, and Efficiency of Clean Energy: A Review and Research Agenda.” *Environmental Science and Pollution Research* 29 (13): 18333–47. https://doi.org/10.1007/s11356-022-18565-9.

Snilstveit, B., R. Bhatia, K. Rankin, and B. Leach. 2017. “3ie Evidence Gap Maps: A Starting Point for Strategic Evidence Production and Use.” 3ie Working Paper 28. New Delhi: International Initiative for Impact Evaluation (3ie). https://www.3ieimpact.org/sites/default/files/2019-01/wp28-egm.pdf.

Snilstveit, Birte, Martina Vojtkova, Ami Bhavsar, Jennifer Stevenson, and Marie Gaarder. 2016. “Evidence & Gap Maps: A Tool for Promoting Evidence Informed Policy and Strategic Research Agendas.” *Journal of Clinical Epidemiology* 79 (November): 120–29. https://doi.org/10.1016/j.jclinepi.2016.05.015.

Snilstveit, Birte. 2012. “Systematic Reviews: From ‘Bare Bones’ Reviews to Policy Relevance.” *Journal of Development Effectiveness* 4 (3): 388–408. https://doi.org/10.1080/19439342.2012.709875.

Solnørdal, Mette, and Lene Foss. 2018. “Closing the Energy Efficiency Gap—A Systematic Review of Empirical Articles on Drivers to Energy Efficiency in Manufacturing Firms.” *Energies* 11 (3): 518. https://doi.org/10.3390/en11030518.

Sovacool, Benjamin K., Jonn Axsen, and Steve Sorrell. 2018. “Promoting Novelty, Rigor, and Style in Energy Social Science: Towards Codes of Practice for Appropriate Methods and Research Design.” *Energy Research & Social Science* 45 (November): 12–42. https://doi.org/10.1016/j.erss.2018.07.007.

Sovacool, Benjamin K., ed. 2014. Energy, Poverty and Development: Critical Concepts in Development Studies. New York, NY: Routledge.

Sovacool, Benjamin K. 2013. “A Qualitative Factor Analysis of Renewable Energy and Sustainable Energy for All (SE4ALL) in the Asia-Pacific.” Energy Policy 59 (August): 393–403. https://doi.org/10.1016/j.enpol.2013.03.051.

Sovacool, Benjamin K., and Ira Martina Drupady. 2012. Energy Access, Poverty, and Development: The Governance of Small-Scale Renewable Energy in Developing Asia. 1st ed. London: Routledge. https://doi.org/10.4324/9781315579535

Sterne, Jonathan AC, Miguel A Hernán, Barnaby C Reeves, Jelena Savović, Nancy D Berkman, Meera Viswanathan, David Henry, et al. 2016. “ROBINS-I: A Tool for Assessing Risk of Bias in Non-Randomised Studies of Interventions.” *BMJ*, October, i4919. https://doi.org/10.1136/bmj.i4919.

Tabachnick, Barbara G., and Linda S. Fidell. 2001. *Using Multivariate Statistics*. 4th ed. Boston, MA: Allyn and Bacon.

Thakur, Megha, Paulien A W Nuyts, Esther A Boudewijns, Javier Flores Kim, Timor Faber, Giridhara R Babu, Onno C P Van Schayck, and Jasper V Been. 2018. ‘Impact of Improved Cookstoves on Women’s and Child Health in Low and Middle Income Countries: A Systematic Review and Meta-Analysis’. Thorax 73 (11): 1026–40. https://doi.org/10.1136/thoraxjnl-2017-210952.

The Campbell Collaboration. 2019. “Campbell Systematic Reviews: Policies and Guidelines (Version 1.4).” Campbell Policies and Guidelines Series No. 1. DOI: 10.4073/cpg.2016.1.

The Methods Group of the Campbell Collaboration. 2019a. “Methodological Expectations of Campbell Collaboration Intervention Reviews: Conduct Standards.” Campbell Policies and Guidelines Series No. 3. The Campbell Collaboration. DOI: 10.4073/cpg.2016.3.

———. 2019b. “Methodological Expectations of Campbell Collaboration Intervention Reviews: Reporting Standards.” Campbell Policies and Guidelines Series No. 4. The Campbell Collaboration. DOI: 10.4073/cpg.2016.4.

Thomas, James, S Graziosi, J Brunton, Z Ghouze, P O’Driscoll, M Bond, and A Koryakina. 2022. “EPPI-Reviewer: Advanced Software for Systematic Reviews, Maps and Evidence Synthesis.” EPPI Centre: UCL Social Research Institute, University College London.

Tsang, Flavia, Peter Burge, Stephanie Diepeveen, Benoit Guerin, Samuel Drabble, and Evan Bloom. 2012. “What Works in Changing Energy Using Behaviours in the Home? A Rapid Evidence Assessment.” Final Report for the UK Department for Energy and Climate Change. RAND. https://www.gov.uk/government/publications/what-works-in-changing-energyusing-behaviours-in-the-home-a-rapid-evidence-assessment.

UN Climate Change. n.d. “COP26 Outcomes: Finance for Climate Adaptation.” Accessed June 14, 2023. https://unfccc.int/process-and-meetings/the-paris-agreement/the-glasgow-climate-pact/cop26-outcomes-finance-for-climate-adaptation.

UNDESA. 2023a. ‘Advancing SDG 7 in Africa’. United Nations Department of Economic and Social Affairs. https://sdgs.un.org/sites/default/files/2023-06/2023%20Advancing%20SDG7%20in%20the%20Africa-061523.pdf.

———. 2023b. ‘Advancing SDG 7 in Latin America and the Caribbean’. United Nations Department of Economic and Social Affairs. https://sdgs.un.org/sites/default/files/2023-06/2023%20Advancing%20SDG7%20in%20the%20Latin%20America%20and%20the%20Caribbean-061623_0.pdf.

UNDP. 2019. “Financing SDG 7.” Policy Brief 3. United Nations Development Programme. https://sustainabledevelopment.un.org/content/documents/24090pb3_cover.pdf.

Valentine, Jeffrey C., Ariel M. Aloe, and Timothy S. Lau. 2015. “Life After NHST: How to Describe Your Data Without ‘p-Ing’ Everywhere.” *Basic and Applied Social Psychology* 37 (5): 260–73. https://doi.org/10.1080/01973533.2015.1060240.

Viechtbauer, Wolfgang. 2010. “Conducting Meta-Analyses in *R* with the Metafor Package.” *Journal of Statistical Software* 36 (3). https://doi.org/10.18637/jss.v036.i03.

Waddington, Hugh, Howard White, Birte Snilstveit, Jorge Garcia Hombrados, Martina Vojtkova, Philip Davies, Ami Bhavsar, et al. 2012. “How to Do a Good Systematic Review of Effects in International Development: A Tool Kit.” *Journal of Development Effectiveness* 4 (3): 359–87. https://doi.org/10.1080/19439342.2012.711765.

Wilson, David B., David Weisburd, and David McClure. 2011. “Use of DNA Testing in Police Investigative Work for Increasing Offender Identification, Arrest, Conviction and Case Clearance.” *Campbell Systematic Reviews* 7 (1): 1–53. https://doi.org/10.4073/csr.2011.7.

World Bank. n.d. “World Development Indicators.” Data Bank. Accessed June 15, 2023. https://databank.worldbank.org/source/world-development-indicators.

Yu, Chunling, and Toru Morotomi. 2022. “The Effect of the Revision and Implementation for Environmental Protection Law on Ambient Air Quality in China.” *Journal of Environmental Management* 306 (March): 114437. https://doi.org/10.1016/j.jenvman.2022.114437.

Zahnd, Alex, and Haddix McKay Kimber. 2009. “Benefits from a Renewable Energy Village Electrification System.” *Renewable Energy* 34 (2): 362–68. https://doi.org/10.1016/j.renene.2008.05.011.

6.6. Appendices

6.1 Appendix A: EGM advisory group

EGM advisory groups are a requirement for all 3ie EGMs to help ensure the project is policy-relevant and useful in informing decision-making. The following advisory group members have provided feedback to refine the interventions-outcomes framework:

- Carol Mungo, Research Fellow, Energy and Climate Change Programme, Stockholm Environment Institute (SEI) Africa
- Jan Minx, Head of Research Group on Applied Sustainability Sciences, Mercator Research Institute on Global Commons and Climate Change (MCC); and Visiting Professor for Climate Change and Public Policy, University of Leeds
- Neelima Jain, Deputy Director and Senior Fellow, Chair in U.S.-India Policy Studies, Center for Strategic and International Studies (CSIS)
- Steven Hunt, Senior Energy & Innovation Advisor, UK's Foreign, Commonwealth and Development Office (FCDO)

6.2 Appendix B: List of low- and middle-income countries

The following list is based on the World Bank's country and lending groups: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

LOW- AND MIDDLE-INCOME COUNTRIES (L&MICs)			
Afghanistan	Eritrea	Marshall Islands	Syrian Arab Republic
Albania	Ethiopia	Mauritania	Tajikistan
Algeria	Fiji	Mexico	Tanzania
Angola	Gabon	Micronesia, Federal States	Thailand
Armenia	Gambia, The	Moldova	Timor-Leste
Azerbaijan	Georgia	Mongolia	Togo
Bangladesh	Ghana	Montenegro	Tonga
Belarus	Grenada	Morocco	Tunisia
Belize	Guatemala	Mozambique	Turkey
Benin	Guinea	Myanmar	Turkmenistan
Bhutan	Guinea-Bissau	Namibia	Tuvalu
Bolivia	Guyana	Nauru	Uganda
Bosnia and Herzegovina	Haiti	Nepal	Ukraine
Botswana	Honduras	Nicaragua	Uzbekistan
Brazil	India	Niger	Vanuatu
Bulgaria	Indonesia	Nigeria	Vietnam
Burkina Faso	Iran, Islamic Republic	Pakistan	West Bank and Gaza
Burundi	Iraq	Papua New Guinea	Yemen, Republic
Cambodia	Jamaica	Paraguay	Zambia
Cameroon	Jordan	Peru	Zimbabwe
Cape (Cabo) Verde	Kazakhstan	Philippines	
Central African Republic	Kenya	Rwanda	
Chad	Kiribati	Samoa	
China	Korea, Democratic Republic	São Tomé and Principe	
Colombia	Kosovo	Senegal	
Comoros	Kyrgyz, Republic	Serbia	
Congo, Democratic Republic	Lao PDR	Sierra Leone	
Congo, Republic	Lebanon	Solomon Islands	
Costa Rica	Lesotho	Somalia	

Côte d'Ivoire (Ivory Coast)	Liberia	South Africa	
Cuba	Libya	South Sudan	
Djibouti	Macedonia, FYR	Sri Lanka	
Dominica	Madagascar	St. Lucia	
Dominican Republic	Malawi	St. Vincent and the Grenadines	
Ecuador	Malaysia	Sudan	
Egypt, Arab Republic	Maldives	Suriname	
El Salvador	Mali	Swaziland	

FORMER LOW- AND MIDDLE-INCOME COUNTRIES
Czechoslovakia
Gibraltar (High income: 2009-2010)
Mayotte (High income: 1990)
Netherlands Antilles (High income: 1994-2009)
Serbia and Montenegro
USSR
Yugoslavia

TRANSITIONAL COUNTRIES		
Name	**L&MIC period**	**High-income country period**
American Samoa	1990-present	1987-1989
Antigua and Barbuda	1987-2001; 2003-2004; 2009-2011	2002; 2005-2008; 2012-present
Argentina	1987-2013; 2015-2016; 2018-present	2014; 2017
Aruba	1991-1993	1987-1990; 1994-present
Bahrain	1990-2000	1987-1989; 2001-present
Barbados	1987-1988; 1990-1999; 2001; 2003-2005	1989; 2000; 2002; 2006-present
Chile	1987-2011	2012-present
Croatia	1992-2007; 2016	2008-2015; 2017-present
Cyprus	1987	1988-present
Czech Republic	1992-2005	2006-present
Equatorial Guinea	1987-2006; 2015-present	2007-2014
Estonia	1991-2005	2006-present
Guam	1990-1994	1987-1989; 1995-present
Greece	1987-1995	1996-present
Hungary	1987-2006; 2012-2013	2007-2011; 2014-present
Isle of Man	1990-2001	1987-1989; 2002-present
Latvia	1991-2008; 2010-2011	2009; 2012-present
Lithuania	1991-2011	2012-present
Macao (SAR)	1987-1993	1994-present
Malta	1987-1988; 1990-1997; 1999; 2001	1989; 1998; 2000; 2002-present

Mauritius	1987-2018; 2020-present	2019
Nauru	2016-2018	2015; 2019-present
New Caledonia	1987-1994	1995-present
Northern Mariana Islands	1992-1994; 2002-2006	1995-2001; 2007-present
Oman	1987-2006	2007-present
Palau	1987-2015; 2021-present	2016-2020
Panama	1987-2016; 2020	2017-2019; 2021-present
Poland	1987-2008	2009-present
Portugal	1987-1993	1994-present
Puerto Rico	1987-1988; 1990-2001;	1989; 2002-present
Republic of Korea (aka South Korea)	1987-1994; 1998-2000	1995-1997; 2001-present
Romania	1987-2018; 2020	2019; 2021
Russia	1991-2011; 2015-present	2012-2014
Seychelles	1987-2013	2014-present
Slovak Republic	1992-2006	2007-present
Slovenia	1992-1996	1997-present
Saudi Arabia	1990-2003	1987-1989; 2004-present
St. Kitts and Nevis	1987-2010	2011-present
Trinidad and Tobago	1987-2005	2006-present
Uruguay	1987-2011	2012-present
Venezuela	1987-2013; 2015-present	2014

6.3 Appendix C: Study designs

Systematic review: A systematic review is a synthesis of the research evidence on a particular topic, such as the effectiveness of water supply and sanitation, obtained through an exhaustive systematic literature search for all relevant studies using widely accepted scientific strategies to minimize error associated with appraising the design and results of studies. A systematic review uses methods of internal and external quality assurance that make it more similar to a primary study (e.g. double-coding of data, calculation of effect sizes from data reported, synthesis of finding). Studies will be excluded if they do not evaluate the effectiveness of an intervention delivered in a rea-world setting, i.e., experiments conducted in tightly-controlled settings, like those of a laboratory, will be excluded.

Impact evaluation: An impact evaluation is a study that uses rigorous methods to provide a quantitative estimate of the impact of an intervention. This is accomplished by constructing a counterfactual, which provides evidence about what would have happened in the absence of the intervention. In an impact evaluation, the outcomes of those who receive the intervention are compared with those of a comparison group that does not receive the intervention. The comparison group may be a specific population in the study area that does not receive the treatment (as in an RCT), or may be constructed by researchers (as in propensity score matching or interrupted time series). For an impact evaluation to be valid, there must be a sound statistical basis for claiming that the comparison group represents what would have happened to the treatment group had they not received the intervention.

.

Effectiveness studies: Studies will be excluded if they do not evaluate the effectiveness of an intervention delivered in a real-world setting (i.e., experiments conducted in tightly controlled settings, like those of a laboratory) will be excluded. Screening questions used to help determine whether a study qualifies as an effectiveness study will include (answering yes signals the study may have been conducted in a lab setting):

- Is the study primarily designed to determine to what extent a specific technique, technology, treatment, procedure or service works under ideal conditions rather than attempt to answer a question relevant to the roll- out of a large programme?

- Is the intervention being carried out by the researchers themselves (e.g., by applying fertilizer in test plots to measure effects on plant growth), or by the people who would carry it out at scale (e.g., farmers applying fertilizer to their crops)?
- Does the study evaluate an intervention that is, or could easily be implemented as, a social policy or programme, or is it “basic science” research on biophysical mechanisms?
- Are participants exposed to an artificial scenario created by the researchers to simulate a policy or programme, rather than being exposed to such a programme in the real world (i.e., is the study an artefactual/framed field experiment)?

We will exclude studies that test the efficacy of energy technologies. Often utilizing quantitative energy models (Sovacool et al. 2018), efficacy studies test the effects of technologies under simulated environments, abstract from real-world settings.

Particularly, we will include studies that implement at least one of the following study designs:

A) Prospective studies that allocate participants to treatment and control groups using random assignment or quasi-experimental methods:

1. Randomised controlled trials (RCTs), with assignment at individual, household, community, or other cluster level, and quasi-RCTs using prospective methods of assignment (such as alternation).
2. Natural experiments with clearly defined intervention and comparison groups, which exploit natural randomness in implementation assignment by decision makers (e.g., public lottery) or random errors in implementation.

B) Non-randomized designs with selection on unobservables:

1. Regression discontinuity designs (RDD) or fuzzy-RDD
2. Instrumental variables (IV)
3. Endogenous treatment-effects models, endogenous switching regression, and other methods synonymous to the Heckman two step model.
4. Difference-in-differences (DID), two-way fixed-effects (TWFE), and two-way Mundlak regressions (TWM):
    a) DiD models will include an interaction term between a time and intervention variable in a regression model. They may also regress an intervention variable on an outcome

variable measuring the changes in outcomes over time or present a t-test comparing changes in outcomes over time between the intervention and control group.

b) TWFE regressions must include time fixed-effects and unit fixed-effects at the level of the intervention (or lower). For example, if the intervention varies at a village level, it must include either village fixed-effects or fixed-effects of a smaller unit, such as households.

5. Interrupted time series (ITS) models, with or without a contemporaneous comparison group. The ITS regression model must adjust for autocorrelation, or it can use autoregressive integrated moving-average (ARIMA) models. An ITS model should include pre-intervention outcome data for a minimum of three time periods.
6. Weighting and matching approaches which control for observable confounding, including non-parametric approaches (e.g., statistical matching, covariate matching, coarsened-exact matching, propensity score matching) and parametric approaches (e.g., propensity-weighted multiple regression analysis).
7. Synthetic control methods.

Natural experiments where the assignment to intervention and control groups was not part of a planned experiment could use different includable designs (e.g., RCT, RDD, ITS). These cases will be categorized as RCT, RDD, ITS, etc.

C) Systematic effectiveness reviews will be included if they describe the search, data collection and synthesis methods according to the 3ie database of systematic reviews protocols (Snilstveit et al., 2016). Any evidence reviews, such as literature reviews, that do not adopt these methods will be excluded. We will exclude systematic reviews that are not effectiveness reviews (i.e., those which do not aim to synthesize the evidence of the effects of a relevant intervention on priority outcomes of interest), such as systematic reviews of driving factors of programs. We will include SRs if they focus exclusively on L&MICs or if they report a separate effect for L&MICs (e.g., through subgroup analysis or meta-regression with LMIC vs. HIC as a moderator).

6.4 Appendix D: Search strategy

Academic/bibliographic databases:

Academic databases
Applied Social Sciences Index and Abstracts (ASSIA)
CAB Abstracts
EBSCO Discovery (RePEc, World Bank eLibrary, Oxfam Policy and Practice)
Econlit (via EBSCOhost)
EBSCO: Gender Studies, GreenFILE, Africa-wide
Environment Complete
Epistemonikos
Global Health
IBSS
Medline
Latin American and Caribbean Health Science (LILACS)
ProQuest Dissertations & Theses Global
SciELO (Scientific Electronic Library Online)
Scopus
Web of Science: Science Citation Index (1990)
Web of Science: Social Science Citation Index (1990)
Web of Science: Arts and Humanities Citation Index (1975-)
Web of Science: Emerging Sources Citation Index (2015-)

Grey literature and supplementary sources, including specialist organizational websites, websites of bilateral and multilateral agencies and repositories of research in international development:

Specialist organizations and other international development organizations	
Abdul Latif Jameel Poverty Action Lab (J-Pal)	https://www.povertyactionlab.org/evaluations
African Development Bank (AfDB)	https://www.afdb.org/en/documents/evaluation-reports
Asian Development Bank (ADB)	https://www.adb.org/publications
British Library for Development Studies	https://guides.lib.sussex.ac.uk/c.php?g=655545&p=4613793
Campbell Collaboration Evidence Portal	https://www.campbellcollaboration.org/better-evidence.html
Collaboration for Environmental Evidence Database of Evidence Reviews (CEEDER)	https://environmentalevidence.org/ceeder-search/
Center for Global Development	https://www.cgdev.org/section/publications
Centre for Effective Global Action (CEGA)	https://cega.berkeley.edu/our-research

Cochrane Database of Systematic Reviews	https://www.cochranelibrary.com/
Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ)	https://mia.giz.de/esearcha/browse.tt.html
Energy Policy Institute at the University of Chicago (EPIC)	https://epic.uchicago.edu/research/
EU Publications	https://op.europa.eu/en/web/general-publications/publications
European Commission	https://ec.europa.eu/info/publications_en
Foreign, Commonwealth and Development Office (FCDO)	www.gov.uk/research-for-development-outputs
German Institute for Development Evaluation (DEval) DEval Rigorous Impact Evaluation	https://www.deval.org/en/evaluations/our-evaluations https://rie.deval.org
Innovations for Poverty Action (IPA)	https://www.poverty-action.org/search-studies
Inter-American Development Bank (IDB)	https://publications.iadb.org/en/publications
International Initiative for Impact Evaluation (3ie) Development Evidence Portal	https://developmentevidence.3ieimpact.org
International Monetary Fund	https://www.elibrary.imf.org
National Bureau of Economic Research (NBER) – Working Papers	https://www.nber.org
Overseas Development Institute	https://odi.org/en/publications
OECD iLibrary	https://www.oecd-ilibrary.org/search/advancedsearch
Oxfam	https://www.oxfam.org.uk/
Registry for International Development Impact Evaluations	https://ridie.3ieimpact.org
Social Science Research Network	https://www.ssrn.com/index.cfm/en
Swedish International Development Cooperation Agency (SIDA)	https://www.sida.se/en/publications
UN DAC resource center	https://resourcecenter.undac.org
United Nations Evaluation Group	http://www.uneval.org/evaluation/reports
USAID Development Experience Clearinghouse (DEC)	https://dec.usaid.gov/dec/content/evaluations.aspx

Website/platforms considered but not included as part of the search strategy		
Power for all - PEAK	http://www.powerforall.org/peak/	Does not include IEs
GOGLA	https://www.gogla.org/reports-publications/	Does not include IEs
International Energy Agency (IEA)	https://www.iea.org/analysis?type=report	Does not include IEs
Energy Efficient End-Use Equipment (4E), IEA	https://www.iea-4e.org/publications/	Does not include IEs
Global Energy Alliance for People and Planet (GEAPP)	https://www.energyalliance.org/reports/	Does not include IEs

Example of search strings:

Database: CAB Abstracts <1973 to 2023 Week 25>

Search Strategy:

1 (afghanistan or albania or algeria or american samoa or angola or "antigua and barbuda" or antigua or barbuda or argentina or armenia or armenian or aruba or azerbaijan or bahrain or bangladesh or barbados or republic of belarus or belarus or byelarus or belorussia or byelorussian or belize or british honduras or benin or dahomey or bhutan or bolivia or "bosnia and herzegovina" or bosnia or herzegovina or botswana or bechuanaland or brazil or brasil or bulgaria or burkina faso or burkina fasso or upper volta or burundi or urundi or cabo verde or cape verde or cambodia or kampuchea or khmer republic or cameroon or cameron or cameroun or central african republic or ubangi shari or chad or chile or china or colombia or comoros or comoro islands or iles comores or mayotte or democratic republic of the congo or democratic republic congo or congo or zaire or costa rica or "cote d'ivoire" or "cote d' ivoire" or cote divoire or cote d ivoire or ivory coast or croatia or cuba or cyprus or czech republic or czechoslovakia or djibouti or french somaliland or dominica or dominican republic or ecuador or egypt or united arab republic or el salvador or equatorial guinea or spanish guinea or eritrea or estonia or eswatini or swaziland or ethiopia or fiji or gabon or gabonese republic or gambia or "georgia (republic)" or georgian or ghana or gold coast or gibraltar or greece or grenada or guam or guatemala or guinea or guinea bissau or guyana or british guiana or haiti or hispaniola or honduras or hungary or india or indonesia or timor or iran or iraq or isle of man or jamaica or jordan or kazakhstan or kazakh or kenya or "democratic people's republic of korea" or republic of korea or north korea or south korea or korea or kosovo or kyrgyzstan or kirghizia or kirgizstan or kyrgyz republic or kirghiz or laos or lao pdr or "lao people's democratic republic" or latvia or lebanon or lebanese republic or lesotho or basutoland or liberia or libya or libyan arab jamahiriya or lithuania or macau or macao or republic of north macedonia or macedonia or madagascar or malagasy republic or malawi or nyasaland or malaysia or malay federation or malaya federation or maldives or indian ocean islands or indian ocean or mali or malta or micronesia or federated states of micronesia or kiribati or marshall islands or nauru or northern mariana islands or palau or tuvalu or mauritania or mauritius or mexico or moldova or moldovian or mongolia or montenegro or morocco or ifni or mozambique or portuguese east africa or myanmar or burma or namibia or nepal or netherlands antilles or nicaragua or niger or nigeria or oman or muscat or pakistan or panama or papua new guinea or new guinea or paraguay or peru or philippines or philipines or phillipines or phillippines or poland or "polish people's republic" or portugal or portuguese republic or puerto rico or romania or russia or russian federation or ussr or soviet union or union of soviet socialist republics or rwanda or ruanda or samoa or pacific islands or polynesia or samoan islands or navigator island or navigator islands or "sao tome and principe" or saudi arabia or senegal or serbia or seychelles or sierra leone or slovakia or slovak republic or slovenia or melanesia or solomon island or solomon islands or norfolk island or norfolk islands or somalia or south africa or south sudan or sri lanka or ceylon or "saint kitts and nevis" or "st. kitts and nevis" or saint lucia or "st. lucia" or "saint vincent and the grenadines" or saint vincent or "st. vincent" or

grenadines or sudan or suriname or surinam or dutch guiana or netherlands guiana or syria or syrian arab republic or tajikistan or tadjikistan or tadzhikistan or tadzhik or tanzania or tanganyika or thailand or siam or timor leste or east timor or togo or togolese republic or tonga or "trinidad and tobago" or trinidad or tobago or tunisia or turkey or turkmenistan or turkmen or uganda or ukraine or uruguay or uzbekistan or uzbek or vanuatu or new hebrides or venezuela or vietnam or viet nam or middle east or west bank or gaza or palestine or yemen or yugoslavia or zambia or zimbabwe or northern rhodesia or global south or africa south of the sahara or sub-saharan africa or subsaharan africa or africa, central or central africa or africa, northern or north africa or northern africa or magreb or maghrib or sahara or africa, southern or southern africa or africa, eastern or east africa or eastern africa or africa, western or west africa or western africa or west indies or indian ocean islands or caribbean or central america or latin america or "south and central america" or south america or asia, central or central asia or asia, northern or north asia or northern asia or asia, southeastern or southeastern asia or south eastern asia or southeast asia or south east asia or asia, western or western asia or europe, eastern or east europe or eastern europe or developing country or developing countries or developing nation? or developing population? or developing world or less developed countr* or less developed nation? or less developed population? or less developed world or lesser developed countr* or lesser developed nation? or lesser developed population? or lesser developed world or under developed countr* or under developed nation? or under developed population? or under developed world or underdeveloped countr* or underdeveloped nation? or underdeveloped population? or underdeveloped world or middle income countr* or middle income nation? or middle income population? or low income countr* or low income nation? or low income population? or lower income countr* or lower income nation? or lower income population? or underserved countr* or underserved nation? or underserved population? or underserved world or under served countr* or under served nation? or under served population? or under served world or deprived countr* or deprived nation? or deprived population? or deprived world or poor countr* or poor nation? or poor population? or poor world or poorer countr* or poorer nation? or poorer population? or poorer world or developing econom* or less developed econom* or lesser developed econom* or under developed econom* or underdeveloped econom* or middle income econom* or low income econom* or lower income econom* or low gdp or low gnp or low gross domestic or low gross national or lower gdp or lower gnp or lower gross domestic or lower gross national or lmic or lmics or third world or lami countr* or transitional countr* or emerging economies or emerging nation?).ti,ab,hw. (3684931)

2 (afghan or afghans or afghani or albanian? or algerian? or american samoan? or angolan? or antiguan? or barbudan? or argentine? or argentinian? or argentinean? or armenian? or aruban? or azerbaijani? or bahraini? or bangladeshi? or bangalees or bajan? or belarusian? or byelorussian? or belizean? or beninese? or bhutanese or bolivian? or bosnian? or botswana or batswana or brazilian? or brasilian? or bulgarian? or burkinabe or burkinese or burundian? or cape verdean? or cabo verdean? or cambodian? or khmer or cameroonian? or central african? or chadian? or chilean? or chinese or colombian? or comorian? or congolese or costa rican? or ivorian? or croatian? or cuban? or cypriot? or

czech? or djiboutian? or dominican? or ecuadorian? or egyptian? or salvadoran? or equatorial guinean? or equatoguinean? or eritrean? or estonian? or swazi? or swati? or ethiopian? or fijian or gabonese or gabonaise or gambian? or georgian? or ghanaian? or gibraltarian? or greek? or grenadian? or guamanian? or guatemalan? or guinean? or bissau guinean? or guyanese or haitian? or honduran? or hungarian? or indian? or indonesian? or iranian? or iraqian? or iraqi? or manx or jamaican? or jordanian? or kazakhstani? or kenyan? or kirabati or kirabatian? or north korean? or korean? or kosovar? or kosovan? or kyrgyz* or lao or laotian? or latvian? or lebanese or lesothan? or lesothonian? or mosotho or basotho or liberian? or libyan? or lithuanian? or macanese or macedonian? or malagasy or madagascan? or malawian? or malaysian? or maldivian? or malian? or maltese or marshallese? or mauritanian? or mauritian? or mexican? or micronesian? or moldovan? or mongolian? or mongol or montenegrin? or moroccan? or mozambican? or burmese or myanma or namibian? or nauruan? or nepali or nepalese or netherlands antillean? or nicaraguan? or nigerien? or nigerian? or northern mariana islander? or mariana? or omani? or pakistani? or palauan? or panamanian? or papua new guinean? or paraguayan? or peruvian? or philippine? or philipine? or phillipine? or phillippine? or filipino? or filipina? or polish or pole or poles or portuguese or puerto rican? or romanian? or russian? or soviet people or soviet population or rwandan? or rwandese or ruandan? or ruandese or samoan? or sao tomean? or santomean? or saudi arabian? or saudi? or senegalese or serbian? or montenegrin? or seychellois or seychelloise? or sierra leonean? or slovak? or slovene? or solomon islander? or somali? or south african? or south sudanese or sri lankan? or ceylonese or kittitian? or nevisian? or saint lucian? or vincentian? or sudanese or surinamese? or syrian? or tajik? or tajikistani? or tanzanian? or tanganyikan? or thai or timorese? or togolese or tongan? or trinidadian? or tobagonian? or tunisian? or turk? or turkish or turkmen? or tuvaluan? or ugandan? or ukrainian? or uruguayan? or uzbek? or vanuatu* or venezuelan? or vietnamese or yemeni? or yemenite? or yemenese or yugoslav? or yugoslavian? or zambian? or zimbabwean? or african? or asian? or pacific islander? or latin american? or central american? or south american? or caribbean? or west indian? or iberoamerican? or middle eastern?).ti,ab,hw. (1437850)

3 1 or 2 (3958566)

4 ((match* adj2 (propensity or coarsened or covariate or neighbo?r)) or "propensity score" or ("difference* in difference*" or "difference-in-difference*" or "differences-in-difference*" or "double difference*") or (quasi-experiment$2 or "quasi experiment$2") or (estimator and evaluat*) or ("instrumental variable*" or (IV adj2 (estimation or approach))) or (Heckman adj3 (model* or approach*)) or ((two-stage or "two stage" or three-stage or "three stage" or four-stage or "four stage") adj3 (control* or function* or regression* or "least squares")) or 2SLS or "regression discontinuity" or "synthetic control?" or "time series" or counterfactual or "segment* regression" or (non adj2 participant*) or ((control or comparison) adj2 (group* or condition* or area* or village* or household* or intervention)) or (panel$1 adj2 (data or household* or model*)) or ((exploit* or "tak* advantage") adj3 (variation* or variety or exogen* or heterogen*)) or (econometric adj2 (model* or adjust*)) or (select* adj2 (bias* or self))).ti,ab,hw.

(281868)

5 ((experiment$4 adj2 (design or study or research or evaluat* or evidence or vary or varies or variation)) or ((random or randomi#ed or randomly) adj2 (trial or assign* or treatment or control* or allocat* or experiment$2 or evaluat* or vary or varies or variation or choose or chose*))).ti,ab,hw. (216489)

6 ((impact? or effect*) adj2 (evaluat* or assess or assessing or assessment or analyze or analyse or analyzing or analysing or analysis or analytical or estimate or estimating or estimation or cause or causal)).ti,ab,hw. (295222)

7 ("program* evaluation" or "project evaluation" or "evaluation research" or "natural experiment*" or "program* effectiveness" or "outcome assessment" or "evaluation study" or "field experiment").ti,ab,hw. (77660)

8 4 or 5 or 6 or 7 (779554)

9 ((Systematic* or synthes*) adj3 (research or evaluation* or overview or finding* or thematic* or report or descriptive or explanatory or narrative or meta* or review* or data or literature or studies or evidence or map or mapping or quantitative or study or studies or paper or impact or impacts or effect* or compar*)).ti,ab,hw. (99191)

10 ("Meta regression" or "meta synth*" or "meta-synth*" or "meta analy*" or "metaanaly*" or "meta-analy*" or "metanaly*" or "Metaregression" or "Meta-regression" or "Methodologic* overview" or "pool* analys*" or "pool* data" or "Quantitative* overview" or "research integration").ti,ab,hw. (42184)

11 ((effectiveness or effects or systemat* or synth* or integrat* or gap or methodologic* or quantitative or evidence or literature or rapid or scoping) adj3 (review or map)).ti,ab,hw. (92711)

12 9 or 10 or 11 (170119)

13 3 and 8 (294697)

14 3 and 12 (52379)

15 13 or 14 (339527)

16 (Bioenergy or biofuel* or Biomass* or coal or electricity or electrificat* or energy or fuel or gas or geothermal* or hydro or hydroelectric* or hydropower or nonrenewable* or oil or power or renewable* or solar or wave or water or wind).tw. (2700994)

17 exp renewable energy/ or exp biofuels/ or exp water power/ or exp wave power/ or exp wind power/ (99027)

18 16 or 17 (2702338)

19 (audit* or access* or adequa* or affordab* or alternative* or availab* or capacity or clean or connect* or conservation or consumer* or consump* or cook* or cool* or coverage or credit* or delivery or demand* or design* or development* or dissemination or distribut* or efficien* or end-user* or Enforcement* or expansion or Financ* or framework* or generat* or grid or grids or heat* or improv* or incentiv* or information* or infrastructur* or insurance or Invest* or light* or loan* or maintenance or market* or micro or monitor or network* or outage* or performance or planning or policies or policy or power or privat* or production or pre-paid or program* or project* or provision* or quality or reform* or regulat* or reliab* or

renewable or report* or resource* or sector* or service* or source* or standard* or storage or subsid* or supply* or supplies or supplier* or sustainab* or Target* or tariff* or technolog* or training or transmission or system* or usage or utility or workshop*).tw. (10185979)

20 "use".tw. (1461408)

21 (system* adj1 maintenanc*).tw. (615)

22 (system* adj1 management).tw. (31102)

23 19 or 20 or 21 or 22 (10254515)

24 15 and 17 (2116)

25 15 and 16 and 23 (93660)

26 24 or 25 (93695)

6.5 Appendix E: Data extraction template

Code	Subcode
Study Information	Study ID
	Coder name
	Title
	Foreign Title
	Short title
	Language
Author Information	Author Name
	Author Affiliation Institution
	Author Affiliation Country
Publication Information	Publication Type
	DOI
	Study status
	Abstract
	Keywords
	Journal name
	Other journal name
	Journal volume
	Journal issue
	Pages
	Year of Publication
	URL
	Publisher location
	Open access
Sector Information	Sector name
	Sub-sector name
	DAC rank
	Primary DAC Code
	Secondary DAC Code
	CRS-Voluntary (tertiary) Code
	SDGs
	World Bank (WB) first theme
	WB first sub-theme
	WB second theme
	WB second sub-theme
	WB third theme
	WB third sub-theme
	Other topics
	Equity focus
	Equity dimension
	Equity description
Geographic Information	First year of intervention
	Continent name
	Country name
	Additional country
	Country income level

Code	Subcode
	Region name
	State/province name
	District name
	City/town name
	Location name
Target population and cost data	Age
	Sex
	Setting
	Sexual orientation
	Specific population group
	Cost data
	Type of cost data
Methodological information	Evaluation Design
	Evaluation Method
	Mixed Method
	Additional quantitative Methods
	Unit of Observation
Program, Funding and Implementation Information	Project Name
	Implementation Agency Category
	Implementation Agency Name
	Program Funding Agency Category
	Program Funding Agency Name
	Research Funding Agency Category
	Research Funding Agency Name
Intervention Information	Treatment group/Arm 1
	Treatment group/Arm 1 Description
	Intervention group/Arm 2
	Treatment group/Arm 2 Description
	(Create additional options as necessary)
Outcome Information	Outcome
	Outcome description
	(Create additional options as necessary)
Energy-relevant categories	SDG Pillar: access, renewables, efficiency
	Energy Source (aggregated): non-renewable, renewable, both, not specified
	Energy Source (disaggregated): coal, oil, gas, solar, wind, geothermal, hydroelectric, biomass, not specified
	Energy Use: cooking, lighting, heating, cooling, productive
	Technology: improved cookstoves, engines, solar lighting, heat pumps, efficient cooling systems, insulation, energy efficient windows, hydropower plants, wind plants, biodigesters, etc.

6.6 Appendix F: Critical appraisal tool

Question	Criteria
Section A: Methods used to identify, include, and critically appraise studies	
A.1 Were the criteria used for deciding which studies to include in the review reported? Did the authors specify: ▪ Types of studies ▪ Participants/ settings/ population ▪ Intervention(s) ▪ Outcome(s)	Yes; partially; no; can't tell Coding guide - check the answers above YES: All four should be yes NO: All four should be no PARTIALLY: Any other
A.2 Was the search for evidence reasonably comprehensive? Were the following done: ▪ Language bias avoided (no restriction of inclusion based on language) ▪ No restriction of inclusion based on publication status ▪ Relevant databases searched (Minimum criteria: All reviews should search at least one source of grey literature such as Google; for health: Medline/ PubMed + Cochrane Library; for social sciences IDEAS + at least one database of general social science literature and one subject specific database) ▪ Reference lists in included articles checked ▪ Authors/experts contacted	Yes; partially; no; can't tell Coding guide - check the answers above: YES: All five should be yes PARTIALLY: Relevant databases and reference lists are both reported NO: Any other

Question	Criteria
A.3 Does the review cover an appropriate time period? Is the search period comprehensive enough that relevant literature is unlikely to be omitted?	Yes; can't tell (only use if no information about time period for search); no; unsure Coding guide: YES: Generally, this means searching the literature at least back to 1990 NO: Generally, if the search does not go back to 1990 CAN'T TELL: No information about time period for search Note: With reference to the above – there may be important reasons for adopting different dates for the search, e.g. depending on the intervention. If you think there are limitations with the timeframe adopted for the search which have not been noted and justified by the authors, you should code this item as a NO and specify your reason for doing so in the comment box below. Older reviews should not be downgraded, but the fact that the search was conducted some time ago should be noted in the quality assessment. Always report the time period for the search in the comment box.
A.4 Was bias in the selection of articles avoided? Did the authors specify: ▪ Independent screening of full text by at least 2 reviewers ▪ List of included studies provided ▪ List of excluded studies provided	Yes; partially; no Coding guide: YES: All three should be yes, although reviews published in journals are unlikely to have a list of excluded studies (due to limits on word count) and the review should not be penalized for this. PARTIALLY: Independent screening and list of included studies provided are both reported NO: All other. If list of included studies provided, but the authors do not report whether or not the screening has been done by 2 reviewers review is downgraded to NO.

Question	Criteria
A.5 Did the authors use appropriate criteria to assess the quality and risk of bias in analyzing the studies that are included? ▪ The criteria used for assessing the quality/ risk of bias were reported ▪ A table or summary of the assessment of each included study for each criterion was reported ▪ Sensible criteria were used that focus on the quality/ risk of bias (and not other qualities of the studies, such as precision or applicability/external validity). "Sensible" is defined as a recognized quality appraisal tool/ checklist, or similar tool which assesses bias in included studies. Please see footnotes for details of the main types of bias such a tool should assess.	Yes; partially; no Coding guide: YES: All three should be yes PARTIALLY: The first and third criteria should be reported. If the authors report the criteria for assessing risk of bias and report a summary of this assessment for each criterion, but the criteria may be only partially sensible (e.g. do not address all possible risks of bias, but do address some), we downgrade to PARTIALLY. NO: Any other
A.6 Overall – how much confidence do you have in the methods used to identify, include, and critically appraise studies? Summary assessment score A relates to the 5 questions above. High confidence applicable when the answers to the questions in section A are all assessed as 'yes' Low confidence applicable when any of the following are assessed as 'NO' above: not reporting explicit selection criteria (A1), not conducting reasonably comprehensive search (A2), not avoiding bias in selection of articles (A4), not assessing the risk of bias in included studies (A5) Medium confidence applicable for any other – i.e. section A3 is assessed as 'NO' or can't tell and remaining sections are assessed as 'partially' or 'can't tell'	**Low confidence** (limitations are important enough that the results of the review are not reliable) **Medium confidence** (limitations are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found) **High confidence** (only minor limitations)

Question	Criteria
Section B: Methods used to analyze the findings	
B.1 Were the characteristics and results of the included studies reliably reported? Was there: □ Independent data extraction by at least 2 reviewers □ A table or summary of the characteristics of the participants, interventions, and outcomes for the included studies □ A table or summary of the results of all the included studies	Yes; no; partially; not applicable (e.g. no included studies) Coding guide: YES: All three should be yes PARTIALLY: Criteria one and three are yes, but some information is lacking on second criteria. No: None of these are reported. If the review does not report whether data was independently extracted by 2 reviewers (possibly a reporting error), we downgrade to NO. NOT APPLICABLE: if no studies/no data
B.2 Are the methods used by the review authors to analyze the findings of the included studies clear, including methods for calculating effect sizes if applicable?	Yes; partially; no; not applicable Coding guide: YES: Methods used clearly reported. If it is clear that the authors use narrative synthesis, they don't need to say this explicitly. PARTIALLY: Some reporting on methods but lack of clarity NO: Nothing reported on methods NOT APPLICABLE: if no studies/no data
B.3 Did the review describe the extent of heterogeneity? Did the review ensure that included studies were similar enough that it made sense to combine them, sensibly divide the included studies into homogeneous groups, or sensibly conclude that it did not make sense to combine or group the included studies? Did the review discuss the extent to which there were important differences in the results of the included studies? If a meta-analysis was done, was the I^2, chi square test for heterogeneity or other appropriate statistic reported? If no statistical test was reported, is a qualitative justification made for the use of random effects?	Yes; partially; no; not applicable Coding guide: YES: First two should be yes, and third category should be yes if applicable should be yes PARTIALLY: The first category is yes NO: Any other NOT APPLICABLE: if no studies/no data

Question	Criteria
B.4 Were the findings of the relevant studies combined (or not combined) appropriately relative to the primary question the review addresses and the available data? How was the data analysis done? ▪ Descriptive only ▪ Vote counting based on direction of effect ▪ Vote counting based on statistical significance ▪ Description of range of effect sizes ▪ Meta-analysis ▪ Meta-regression ▪ Other: specify ▪ Not applicable (e.g. no studies or no data) How were the studies weighed in the analysis? ▪ Equal weights (this is what is done when vote counting is used) ▪ By quality or study design (this is rarely done) ▪ Inverse variance (this is what is typically done in a meta-analysis) ▪ Number of participants (sample size) ▪ Other: specify ▪ Not clear ▪ Not applicable (e.g. no studies or no data) Did the review address unit of analysis errors? ▪ Yes - took clustering into account in the analysis (e.g. used intra-cluster correlation coefficient) ▪ No, but acknowledged problem of unit of analysis errors ▪ No mention of issue ▪ Not applicable - no clustered trials or studies included	Yes; partially; no; not applicable (e.g. no studies or no data); can't tell. Coding guide: YES: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors addressed (if appropriate). PARTIALLY: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors not addressed (and should have been). NO: If narrative OR vote counting (where quantitative analyses would have been possible) OR inappropriate reporting of table, graph, or meta-analyses. NOT APPLICABLE: if no studies/no data CAN'T TELL: if unsure (note reasons in comments below)

Question	Criteria
B.5 Does the review report evidence appropriately? The review makes clear which evidence is subject to low risk of bias in assessing causality (attribution of outcomes to intervention), and which is likely to be biased, and does so appropriately Where studies of differing risk of bias are included, results are reported and analyzed separately by risk of bias status	Yes; partially; no; not applicable Coding guide: YES: Both criteria should be fulfilled (where applicable) NO: Criteria not fulfilled PARTIALLY: Only one criterion fulfilled, or when there is limited reporting of quality appraisal (the latter applies only when inclusion criteria for study design are appropriate) NOT APPLICABLE: No included studies Note on reporting evidence and risk of bias: For reviews of effects of 'large n' interventions, experimental and quasi-experimental designs should be included (if available). For reviews of effects of 'small n' interventions, designs appropriate to attribute changes to the intervention should be included (e.g. pre-post with assessment of confounders)
B.6 Did the review examine the extent to which specific factors might explain differences in the results of the included studies? Were factors that the review authors considered as likely explanatory factors clearly described? Was a sensible method used to explore the extent to which key factors explained heterogeneity? ▪ Descriptive/textual ▪ Graphical ▪ Meta-analysis by sub-groups ▪ Meta-regression ▪ Other	Yes; partially; no; not applicable Coding guide: YES: Explanatory factors clearly described and appropriate methods used to explore heterogeneity PARTIALLY: Explanatory factors described but for meta-analyses, sub-group analysis or meta-regression not reported (when they should have been) NO: No description or analysis of likely explanatory factors NOT APPLICABLE: e.g. too few studies, no important differences in the results of the included studies, or the included studies were so dissimilar that it would not make sense to explore the heterogeneity of the results

Question	Criteria
B.7 Overall - how much confidence do you have in the methods used to analyze the findings relative to the primary question addressed in the review? Summary assessment score B relates to the 5 questions in this section, regarding the analysis. High confidence applicable when all the answers to the questions in section B are assessed as 'yes.' Low confidence applicable when any of the following are assessed as 'NO' above: critical characteristics of the included studies not reported (B1), not describing the extent of heterogeneity (B3), combining results inappropriately (B4), reporting evidence inappropriately (B5). Medium confidence applicable for any other: i.e. the "Partial" option is used for any of the 6 preceding questions or questions and/or B.2 and/ or B.6 are assessed as 'no'.	**Low confidence** (limitations are important enough that the results of the review are not reliable) **Medium confidence** (limitations are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found) **High confidence** (only minor limitations)
Section C: Overall assessment of the reliability of the review	
C.1 Are there any other aspects of the review not mentioned before which led you to question the results?	▪ Additional methodological concerns – only one person reviewing ▪ Robustness ▪ Interpretation ▪ Conflicts of interest (of the review authors or for included studies) ▪ Other ▪ No other quality issues identified
C.2 Are there any mitigating factors which should be considered in determining the review's reliability?	▪ Limitations acknowledged ▪ No strong policy conclusions drawn (including in abstract/ summary) ▪ Any other factors
C.3 Based on the above assessments of the methods how would you rate the reliability of the review? **Low confidence in conclusions about effects**: **Medium confidence in conclusions about effects**: The systematic review has the following limitations... **High confidence in conclusions about effects**: If applicable: The review has the following minor limitations...	

Question	Criteria
Coding guide: **High confidence in conclusions about effects**: high confidence noted overall for sections A and B, unless moderated by answer to C1. **Medium confidence in conclusions about effects**: medium confidence noted overall for sections A or B, unless moderated by answer to C1 or C2. **Low confidence in conclusions about effects**: low confidence noted overall for sections A or B, unless moderated by answer to C1 or C2. Limitations should be summarized above, based on what was noted in Sections A, B and C.	

6.7 Appendix G: Systematic review methods

The particular approach to the systematic review will depend on its scope and the research questions we will aim to address. The following sections describe the general guidelines for extracting, processing and analyzing data to address questions on effectiveness, cost-effectiveness, and barriers and facilitating factors for effective interventions. We may use all or sections of these guidelines once we transition to the synthesis stage of the project.

6.7.1 Criteria for determination of independent findings

Complex data structures are a common occurrence in meta-analyses of impact evaluations. There are several scenarios through which these complex structures with dependent effect sizes might occur. For instance, there could be several publications that stem from one study, or several studies based on the same data set. Some studies might have multiple treatment arms that are all compared to a single control group. Other studies may report outcome measurements from several time points or use multiple outcome measures to assess related outcome constructs. All such cases yield a set of statistically dependent effect size estimates (Borenstein et al., 2009).

The research team will assess the extent to which relationships exist across the studies included in the review. We will make every attempt to avoid double counting of identical evidence by linking papers prior to data analysis. Where we have several publications reporting on the exact same effect, we will use effect sizes from the most recent publication. We will utilize information provided in studies to support these assessments, such as samples sizes, programme characteristics and key implementing and/or funding partners.

We will extract effects reported across different outcomes or subgroups within a study, and where information is collected on the same programme for different outcomes at the same or different periods of time, we will extract information on the full range of outcomes over time. Where studies report effects from multiple model specifications, we will use the author's preferred model specification. If this is not stated or is unclear, we will use the specification with the most controls. Where studies report multiple outcome sub-groups for the same outcome construct, we may calculate a "synthetic effect size" (Borenstein et al., 2009, chapter 24). Where studies report multiple outcomes or evidence according to sub-groups of participants, we will

record and report data on relevant sub-groups separately. Further information on criteria for determining independent effect sizes is presented below.

We will deal with dependent effect sizes in one of two ways, either through the use of robust variance estimation (RVE: Fisher and Tipton, 2015; Hedges et al., 2010), or through data processing and selection techniques. RVE using a small sample adjustment will be the preferred analytic method when feasible. The RVE approach allows us to use all available data in the effect size estimates, even data that is statistically dependent. However, these analyses must have >4 degrees of freedom in order to make valid inferences. In cases where analyses do not meet this criteria, data processing and selection techniques will be used to deal with dependent effect sizes.

If RVE analyses are not feasible for a meta-analysis of any given intervention or outcome group, we will utilize several criteria to select one effect estimate per study. Where we have several publications reporting on the same study, we will use effect sizes from the most recent publication. For studies with outcome measures at different time points, we will follow De La Rue and colleagues (2013) and synthesize outcomes measured immediately after the intervention (defined as 1-6 months) and at follow-up (longer than six months) separately. If multiple time points exist within these time periods, we will use the most recent measure. We anticipate many of the interventions we include in the review will be ongoing programmes and the follow-up will, therefore, reflect duration in a program rather than time since intervention. When such studies report outcome measures at different time points, we will identify the most common follow-up period and include the follow up measures that match this most closely in the meta-analysis. When studies include multiple outcome measures to assess related outcome constructs, we will follow Macdonald and colleagues (2012) and select the outcome that appears to most accurately reflect the construct of interest without reference to the results. If studies include multiple treatment arms with only one control group and the treatments represent separate treatment constructs, we will calculate the effect size for treatment A versus control and treatment B versus control and include in separate meta-analyses according to the treatment construct. If treatments A and B represent variations of the same treatment construct, we will calculate the weighted mean and standard deviation for treatment A and B before calculating the effect size for the merged group versus control group, following the procedures outlined in Borenstein and colleagues (2009, chapter 25). Where different studies report on the same programme but use different samples (e.g. from different regions) we will include both

estimates, treating them as independent samples, provided effect sizes are measured relative to separate control or comparison groups.

6.7.2 Data extraction and management

Because the main characteristics of the intervention and evaluation will have been extracted for the EGM, we will focus on extracting descriptive, methodological and quantitative data from each included study using a standardized data extraction forms:

- Information on intervention design, including how the intervention incorporates participation, inclusion, transparency and accountability characteristics, participant adherence, contextual factors, and programme mechanisms.
- Methodological information on analysis method and type of comparison group (if relevant).
- Quantitative data for outcome measures, including outcome descriptive information, sample size in each intervention group, outcomes means and standard deviations, and test statistics (e.g. t-test, F-test, p-values, 95% confidence intervals).

We will extract these data using Excel. Descriptive data will be single coded by one reviewer and checked by a second reviewer. Two independent reviewers will double code quantitative data for outcomes analysis, and any disagreement will be resolved through discussion with a third reviewer (who must be a core team member).

Once all effect sizes are calculated and converted to a standardized mean difference (as described in detail below), we will examine the data for outliers. We will define outliers as any effect sizes ± 3.29 standard deviations from the mean, following the guidance of Tabachnick and Fidell (2001). Sensitivity to outliers will be examined as discussed in the section on sensitivity analysis below.

6.7.3 Assessment of risk of bias in included studies

We will assess the risk of bias in the included studies by drawing on the signaling questions in the 3ie risk of bias tool, which covers both internal validity and statistical conclusion validity of experimental and quasi-experimental impact evaluation designs (Hombrados and Waddington, 2012). It includes the bias domains and extensions to Cochrane's ROBINS-I tool and RoB2.0 (Higgins et al., 2016; Sterne et al., 2016). The risk of bias assessment helps us to determine the

extent to which the findings in each study are reliable. Two reviewers will undertake the risk of bias assessment independently. If there are disagreements, we will resolve them through discussion and the involvement of a third reviewer, as necessary. We will assess the risk of bias at the paper level, noting any potential differences in methods and risk of bias by different outcomes.

We will assess risk of bias based on the following criteria, coding each paper as “Yes”, “Probably Yes”, “Probably No”, “No” and “No Information” according to how they address each domain:

- Factors relating to baseline confounding and biases arising from differential selection into and out of the study (e.g. assignment mechanism).
- Factors relating to bias due to missing outcome data (e.g. assessment of attrition).
- Factors relating to biases due to deviations from intended interventions (e.g. performance bias and survey effects) and motivation bias (Hawthorne effects).
- Factors relating to biases in outcomes measurement (e.g. social desirability or courtesy bias, recall bias).
- Factors relating to biases in reporting of analysis.

We will report the results of the assessment for each of the assessed criteria for each study. In addition, we will use the results of the risk of bias assessments to produce an overall rating for each study as either “High risk of bias”, “Some concerns” or “Low risk of bias”, drawing on the decision rules in RoB2.0 (Higgins et al., 2016), rating studies as follows:

- “High risk of bias”: if any of the bias domains were assessed as “No” or “Probably No”.
- “Some concerns”: if one or several domains were assessed as “No Information” and none were “No” or “Probably No”.
- “Low risk of bias”: if all of the bias domains were assessed as “Yes” or “Probably Yes”.

In addition, we will attempt to explore whether there are systematic differences in outcome effects between primary studies with different risk of bias. If meta-analysis is feasible, we will conduct sensitivity analysis to assess the robustness of the results to the risk of bias in included studies.

6.7.4 Measures of treatment effect

An effect size expresses the magnitude (or strength) and direction of the relationship of interest (Valentine et al., 2015; Borenstein et al., 2009). We will extract data from each individual study to calculate standardized effect sizes for cross-study comparison wherever possible. For continuous outcomes comparing group means in a treatment and control group, we will calculate the standardized mean difference (SMDs), or Cohen's d, its variance and standard error using formulae provided in Borenstein and colleagues (2009). A SMD is a difference in means between the treatment and control groups divided by the pooled standard deviation of the outcome measure. Cohen's d can be biased in cases where sample sizes are small. Therefore, in all cases we will simply adjust d using Hedges' method, adjusting Cohen's d to Hedges' g using the following formula (Ellis, 2010):

$$g \cong d(1 - \frac{3}{4(n_1 + n_2) - 9})$$

We will choose the appropriate formulae for effect size calculations in reference to, and dependent upon, the data provided in included studies. For example, for studies reporting means (X) and pooled standard deviation (SD) for treatment (T) and control or comparison (C) at follow up only:

$$d = \frac{x_{Tp+1} - x_{Cp+1}}{SD}$$

If the study does not report the pooled standard deviation, it is possible to calculate it using the following formula:

$$SD_{p+1} = \sqrt{\frac{(n_{Tp+1} - 1)SD_{Tp+1}^2 + (n_{Cp+1} - 1)SD_{Cp+1}^2}{n_{Tp+1} + n_{Cp+1} - 2}}$$

Where the intervention is expected to change the standard deviation of the outcome variable, we will use the standard deviation of the control group only.

For studies reporting means ($\underline{X}$) and standard deviations (SD) for treatment and control or comparison groups at baseline (p) and follow up (p+1):

$$d = \frac{\Delta\underline{X}_{p+1} - \Delta\underline{X}_{p}}{SD_{p+1}}$$

For studies reporting mean differences ($\Delta\underline{X}$) between treatment and control and standard deviation (SD) at follow up (p+1):

$$d = \frac{\Delta\underline{X}_{p+1}}{SD_{p+1}} = \frac{\underline{X}_{Tp+1} - \underline{X}_{Cp+1}}{SD_{p+1}}$$

For studies reporting mean differences between treatment and control, standard error (SE) and sample size (n):

$$d = \frac{\Delta\underline{X}_{p+1}}{SE\sqrt{n}}$$

As primary studies have become increasingly complex, it has become commonplace for authors to extract partial effect sizes (e.g. a regression coefficient adjusted for covariates) in the context of meta-analysis. For studies reporting regression results, we will follow the approach suggested by Keef and Roberts (2004) using the regression coefficient and the pooled standard deviation of the outcome. Where the pooled standard deviation of the outcome is unavailable, we will use regression coefficients and standard errors or t-statistics to do the following, where sample size information is available in each group:

$$d = t\sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$$

where n denotes the sample size of treatment group and control. We will use the following where only the total sample size information (N) is available, as suggested in Polanin and colleagues, 2016):

$$d = \frac{2t}{\sqrt{N}} \qquad Var_d = \frac{4}{N} + \frac{d^2}{4N}$$

We will calculate the t-statistic (t) by dividing the coefficient by the standard error. If the authors only report confidence intervals and no standard error, we will calculate the standard error from the confidence intervals. If the study does not report the standard error, but report t, we will extract and use this as reported by the authors. In cases in which significance levels are reported rather than t or SE (b), then t will be imputed as follows:

Prob > 0.1: t = 0.5

0.1 ≥ Prob > 0.05: t = 1.8

0.05 ≥ Prob > 0.01: t = 2.4

0.01 ≥ Prob: t = 2.8

Where outcomes are reported in proportions of individuals, we will calculate the Cox-transformed log odds ratio effect size (Sánchez-Meca et al., 2003):

$$d = \frac{ln\ (OR)}{1.65}$$

where OR is the odds ratio calculated from the two-by-two frequency table.

Where outcomes are reported based on proportions of events or days, we will use the standardized proportion difference effect size:

$$d = \frac{p_T\ -\ p_C}{SD(p)}$$

Where p_t is the proportion in the treatment group and p_c the proportion in the comparison group, and the denominator is given by:

$$SD(p) = \sqrt{p\ (1-p)}$$

where p is the weighted average of p_c and p_t:

$$p = \frac{n_T\ p_T\ +\ n_C\ p_C}{n_T\ +\ n_C}$$

An independent reviewer will evaluate a random selection of 10 per cent of effect sizes to ensure that the correct formulae were employed in effect size calculations. In all cases after synthesis, we will convert pooled effect sizes to commonly used metrics such as percentage changes and mean differences in outcome metrics typically used (e.g. weight in kg) whenever feasible.

6.7.5 Unit of analysis issues

Unit of analysis errors can arise when the unit of allocation of a treatment is different to the unit of analysis of effect size estimate, and this is not accounted for in the analysis (e.g. by clustering standard errors at the level of allocation). We will assess studies for unit of analysis errors (The Campbell Collaboration, 2019), and where they exist, we will correct for them by adjusting the standard errors according to the following formula (Higgins et al., 2020; Waddington et al., 2012; Hedges, 2009):

$$SE(d)' = SE(d) * \sqrt{1 + (m-1)c}$$

where m is the average number of observations per cluster and c is the intra-cluster correlation coefficient. Where included studies use robust Huber-White standard errors to correct for clustering, we will calculate the standard error of d by dividing d by the t-statistic on the coefficient of interest.

6.7.6 Dealing with missing data

In cases of relevant missing or incomplete data in studies identified for inclusion, we will make every effort to contact study authors to obtain the required information. If we are unable to obtain the necessary data, we will report the characteristics of the study but state that it could not be included in the meta-analysis or reporting of effect sizes due to missing data.

6.7.7 Assessment of heterogeneity

We will assess heterogeneity by calculating the Q-statistic, I2, and Tau2 to provide an estimate of the amount of variability in the distribution of the true effect sizes (Borenstein et al., 2009). We will complement this with an assessment of heterogeneity of effect sizes graphically using forest plots.

6.7.8 Assessment of reporting biases

In order to reduce the possibility of publication bias, we will search for and include unpublished studies in the review. We will also test for the presence of publication bias through the use of contour-enhanced funnel graphs (Peters et al., 2008) and statistical tests (Egger et al., 1997) for

outcomes for which we identify at least 10 studies. Capitalizing on recent shifts towards pre-registration of studies and their associated pre-analysis plans, we will also examine whether studies that were pre-registered (e.g. on platforms such as ClinicalTrials.gov, the Open Science Foundation, the American Economic Association's trial registry, or the Registry for International Development Impact Evaluations (RIDIE)) report on all of the outcomes that were proposed in their pre-analysis plans. This additional analysis of outcome reporting bias may draw on methodologies used in previous work, such as the COMPare Trials Project (Goldacre et al., 2016).

6.7.9 Data synthesis

We will conduct meta-analyses of studies that we assess to be sufficiently similar. While the inclusion criteria for the review will follow the results of the EGM, within intervention categories, we may include studies that report on diverse interventions and outcomes. It is therefore difficult to predict how meta-analysis will be used in the review prospectively. However, minimum criteria will be to only combine studies using meta-analysis when we identify two or more effect sizes using a similar outcome construct and where the comparison group state is judged to be similar across the two, similar to the approach taken by Wilson and colleagues (2011). We provisionally suggest that we combine studies in the same analysis when they evaluate the same intervention type, or the same outcome type. Moderator analyses can take into account multiple interventions as moderator variables, allowing us to also examine the impact of different intervention types by outcome. Where there are too few studies, or included studies are considered too heterogeneous in terms of interventions or outcomes, we will present a discussion of individual effect sizes along the causal chain. As heterogeneity exists in theory due to the variety of interventions and contexts included, we will use inverse-variance weighted, random effects meta-analytic models (Higgins et al., 2020).

We will use the metafor package (Viechtbauer, 2010) and/or the robumeta package (Fisher and Tipton, 2015) in R software to conduct the meta-analyses (R Core Team, 2020).

The qualitative information we will extract from studies may subsequently be coded quantitatively to be used in moderator analysis. It may also be used to classify intervention mechanisms in synthesis or in the further development of intervention causal chains. These characteristics may include: intervention objectives (to change processes, behaviors or both); whether interventions are strategic (complex, adaptable strategy to realize change) or tactical

(tool-based); the source of intervention (local, NGO, government or researcher-led); the scale of the intervention (pilot experiment versus adoption of formal policy/law); extent to which members of both targeted groups are engaged (equally or primarily one group); and initial power differences between the groups targeted.

6.7.10 Subgroup analysis and investigation of heterogeneity

Whenever feasible, we will conduct moderator analyses to investigate sources of heterogeneity. Following the PROGRESS-PLUS approach (Oliver et al., 2017), we will assess moderators falling into three broad categories of extrinsic, methodological and substantive characteristics to address inequity aspects within the aquaculture context. Examples of these categories include:

- Extrinsic characteristics: funder of the study (e.g. NGO vs private sector vs government investments), publication type, publication date.
- Methodological characteristics: study design, risk of bias, study quality characteristics, evaluation period, length of follow-up.
- Substantive characteristics: participant characteristics (gender, age, socio-economic status, education, land ownership), context (geographical setting, market access), intervention type, intervention features, type of implementing agency.

We will use random effects meta-regression to investigate the association between moderator variables and heterogeneity of treatment effects (Borenstein et al., 2009) and sub-group analyses to investigate heterogeneity by treatment sub-groups (e.g. men and women, poor and non-poor, and so on). If the latter strategies are not possible (that is, if we do not have sufficient number of studies or data), we will discuss and explore the factors which may be driving heterogeneity of results narratively by conducting cross-case comparisons (Miles and Huberman, 1994).

6.7.11 Sensitivity analysis

We will conduct sensitivity analysis to assess whether the results of the meta-analysis are sensitive to the removal of any single study. We will do this by removing studies from the meta-analysis one-by-one and assessing changes in results. We will also assess sensitivity of results to inclusion of high risk of bias studies by removing these studies from the meta-analysis and

comparing results to the main meta-analysis results. Finally, we will assess sensitivity to outliers by comparing results with and without outliers included.

6.7.12 Treatment of qualitative research

We will use qualitative research to supplement the findings of the interventions covered by included studies. We will not seek out all qualitative studies relating to energy interventions in low- and middle-income countries, but under the “effectiveness+” framework (Snilstveit, 2012), we will look for qualitative studies to provide additional information about the context and implementation of relevant interventions. This may include feasibility studies, stakeholder analyses, formative evaluations, process evaluations, project reports, among other documents.

Two coders will independently appraise these studies and documents based on an adapted version of the Critical Appraisal Skills Programme checklist (CASP, n.d.). We will assess the quality of qualitative and descriptive quantitative studies by appraising the adequacy of reporting, data collection, presentation, analysis and conclusions drawn. In turn, the assessment of process evaluations will focus on sampling and methods of data collection. Finally, project documents provide information about the design or resources available for a project. As these documents present factual information about interventions, we will not formally appraise the quality of such documents but will assess the relevance of the documents against the interventions included in the review.

6.7.13 Treatment of cost data

We will use cost data reported in the set of included studies or in additional studies identified through the second search of references. Following Shemilt and colleagues (2008), relevant studies may include full economic evaluations (e.g. cost-benefit, cost-effectiveness, or cost-utility analyses), partial economic evaluations (e.g. cost analyses, cost-comparison studies, cost-outcome descriptions), or any other documentation reporting cost data of included interventions.

Full and partial economic evaluation studies will be appraised in terms of the cost and/or effectiveness components reported and used in the analyses. In turn, general descriptions of cost information of included interventions will be synthesized narratively. If there is relevant data on the costs and effects of an intervention reported separately, we will extract data on the

resources, unit and/or total costs with the aim of examining both components. In these cases, we will focus on comparable outcomes if possible. We will also note when included studies found statistically non-significant effects, however, we will not include non-significant impacts in the cost-effectiveness analysis (Dhaliwal et al., 2013). If this impact is precisely measured, then there is little relevance in examining non-effective interventions; whereas if the impact is measured with less precision, there will be uncertainty around the real effectiveness of the intervention, which would affect the analysis around its cost.