Inception report (protocol): Mapping and synthesizing evidence on root causes and drivers of irregular migration

Sanghwa Lee
International Initiative for Impact Evaluation (3ie)

María Daniela Anda León
3ie

Shannon Shisler
3ie

Carolyn Huang
3ie

Anilkrishna B. Thota
3ie

Evidence Synthesis Inception Report (Protocol)

October 2025

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 are thematic collections of information about impact evaluations or 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.

About this EGM update, expansion and evidence assessment inception report

The inception report provides all the supporting documentation for the map, including the thematic background information, and details of the methods that will be applied to systematically search and screen the evidence base, extract, and analyze data, and develop the EGM report. It also details the methods that will be used to synthesize the effectiveness of the intervention categories with sufficient evidence and produce a synthesis report. This work was developed by 3ie with funding from the Ministry of Foreign Affairs of the Kingdom of the Netherlands, through the Policy and Operations Evaluation Department (IOB).

The content of this report is the sole responsibility of the authors and does not represent the opinions of the IOB, 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, slee@3ieimpact.org.

Suggested citation: Lee, Sanghwa, María Daniela Anda León, Shannon Shisler, Carolyn Huang, and Anilkrishna B. Thota. 2025. *Inception report (protocol): Mapping and synthesizing evidence on root causes and drivers of irregular migration*. London: International Initiative for Impact Evaluation (3ie).

Acknowledgements

We are grateful to IOB for their support and feedback, particularly to Caspar Lobbrecht, Zeineb Romdhane, and Johannes Claes. We would also like to thank the members of our advisory group who reviewed and provided valuable feedback on this inception report, particularly Jon Kurtz and Florian Trauner. Finally, we acknowledge 3ie's internal reviewers, Constanza Gonzalez Parrao and Pierre Marion, for their feedback.

Introduction

Migration dynamics are shaped by interrelated social, political, geographic, economic, and environmental factors that cannot be explained in linear terms, and irregular migration reflects this complexity. Despite policy attention and a growing number of interventions, including external migration management policies, evidence on the effectiveness of these interventions remains fragmented and uneven.

An evidence gap map (EGM) provides a systematic and transparent way to map what is known and to highlight where critical knowledge gaps remain, which can be a valuable tool for policymakers, practitioners and researchers in the sector. In addition, evidence synthesis based on the studies captured in the EGM can reveal what existing research shows about intervention effectiveness, supporting more evidence-informed policy decisions.

This work aims to inform the ongoing evaluation of Dutch policy on migration cooperation and partnerships, conducted by the independent Policy and Operations Evaluation Department (IOB) of the Ministry of Foreign Affairs. In this protocol, we establish the scope and methods to be used in (1) updating and expanding 3ie's 2023 EGM of impact evaluations and systematic reviews that assess the effects of interventions addressing root causes and other drivers of irregular migration (hereafter, the original EGM),[1] and (2) synthesizing the effects of selected intervention categories to produce evidence summaries and a synthesis report. Since the original EGM was published in 2023, we expect that the evidence base has continued to expand, which makes updating and expanding the EGM both necessary and timely.

As in the original EGM, we conceptualize multidimensional drivers of irregular migration. For the EGM update, we will expand the search strategy using the original framework to include relevant studies published from the date of the previous search (April 2023) to August 2025. This will cover four intervention domains addressing a) economic drivers; b) environmental and climate-related stressors; c) drivers related to conflict, violence, and insecurity; and d) the lack of legal pathways and information (see Section 2 and Appendix B for more details). In line with the original EGM, we focus on cross-border irregular migration, with internal migration and internal displacement (i.e., movements within a single country) outside the scope of this work. We will include quantitative impact evaluations and systematic reviews from low- and middle-income countries (L&MICs), as well as studies evaluating interventions from the fourth domain that are implemented in high-income countries (HICs) but target migrants or potential migrants from L&MICs.

In addition to updating the existing map, this study introduces a new intervention domain, e) **external migration management interventions**, to enhance our understanding of policy efforts in Dutch and European Union (EU) migration priorities. These include interventions such as migration partnerships, deals, externalization of border control, deterrence policies and remote-control measures aimed at managing or reducing irregular migration through actions taken outside EU borders. For this intervention group, we will include only those policies that target or involve cooperation with Afghanistan, Pakistan, Türkiye, and L&MICs in Sub-Saharan Africa and the Middle East and North Africa. The difference in geographic

[1] The online map is available at: https://developmentevidence.3ieimpact.org/egm/addressing-root-causes-and-drivers-of-irregular-migration-an-evidence-gap-map.

scope is intended to align with IOB's evaluation needs. Also, for this new domain only, we will include qualitative evaluation studies that meet our study design criteria. Including qualitative studies will expand the knowledge base available for synthesis on the interventions evaluated in IOB's evaluation project. Both the update and the expansion will continue to use the original outcome framework, which covers outcomes related to human mobility as well as intermediate outcomes such as migration aspirations and intentions.

In addition to updating and expanding the original EGM, this study will conduct a focused evidence synthesis. This will entail producing a set of concise evidence summaries, each covering a selected intervention category. Each summary will be structured to include: (1) a brief overview of the theoretical relationship between the intervention and expected outcomes; (2) main findings from the identified studies; and (3) overarching conclusions and policy-relevant implications. We do not assess the merits of 'root causes' theories, but map interventions targeting specific drivers and, through synthesis, examine their links to migration outcomes and the validity of such theories across domains. These summaries will inform a final synthesis report, which will integrate findings across interventions to answer the study's overarching research questions and support evidence-informed decision-making in the sector.

This protocol outlines the background and reasons behind this map update, expansion, and evidence synthesis (Section 1), the scope and conceptual framework guiding this work (Section 2), the map eligibility criteria, including the interventions and outcomes of interest (Section 3), and the rigorous methods we will follow to update and expand this map (Section 4) and to synthesize the effects of selected intervention categories (Section 5). The protocol will be made publicly available to support research transparency and reproducibility, with any deviations from this protocol noted in the final synthesis report.

Contents

List of figures and table

Acronyms

AG	Advisory Group
ALMPs	Active Labor Market Policies
AMIF	Asylum, Migration and Integration Fund
CAMELOT	Critical Appraisal for Methodological Limitations of Qualitative Research Tool
DEP	3ie Development Evidence Portal
EGM	Evidence Gap Map
EU	European Union
FAO	Food and Agriculture Organization of the United Nations
HICs	High-Income Countries
ICC	Intra-Cluster Correlation Coefficient
ICMPD	International Centre for Migration Policy Development
ILO	International Labour Organization
IOB	Policy and Operations Evaluation Department (Ministry of Foreign Affairs of the Kingdom of the Netherlands)
IOM	International Organization for Migration
IE	Impact Evaluation
L&MICs	Low- and Middle-Income Countries
MENA	Middle East and North Africa
MFF	Multiannual Financial Framework
NDICI	Neighbourhood, Development and International Cooperation Instrument
ODA	Official Development Assistance
PICOS	Population, Intervention, Comparator, Outcomes, and Study Design
RVE	Robust Variance Estimation
SECP	Stakeholder Engagement and Communication Plan
SR	Systematic Review
SSA	Sub-Saharan Africa
UNHCR	United Nations High Commissioner for Refugees

1. Background

1.1 The problem, condition or issue

Migration dynamics are shaped by interrelated social, political, geographic, economic, and environmental factors that cannot be explained in linear terms (de Haas et al. 2019; Kuhnt 2019; Czaika and Reinprecht 2022), and irregular migration[2] reflects this complexity. International migration is guided by a rights-based approach that respects individuals' decisions to migrate, and views migration management as integral to sustainable development (UN 2018; UNSD 2022), as highlighted in 3ie's 2023 EGM of interventions addressing root causes and other drivers of irregular migration (hereafter, the original EGM; Berretta et al. 2023a). The "root causes" approach to migration management is often cited, yet this framing has been controversial because it can oversimplify complex mobility dynamics while supporting restrictive policies (Czaika and Reinprecht 2022; de Haas 2023; Gent 2002; Knoll and Sheriff 2017; UNHCR 2022a; Vutha et al. 2011; Yayboke and Gallego 2019).

As highlighted in the original EGM, a precise global prevalence of irregular migration remains unknown (Schewel and Debray 2024; Slootjes and Sohst 2024). This is also the case in South–South migration corridors, which now account for a growing share of international migration but remain poorly captured in official statistics, especially with respect to irregular movements (Crawley and Teye 2024). Partial estimates suggest that between 13.6 million and 14.8 million irregular migrants lived in 12 European countries[3] and the United States between 2016 and 2023, with the United States accounting for the largest share[4] (Kierans and Vargas-Silva 2024). However, these data are incomplete and context-dependent, requiring careful interpretation to avoid misleading conclusions in policy and public discourse (Hendow et al. 2024). Despite ongoing efforts to apply innovative methods to measure irregular migration, such as using social media data and machine learning to estimate migrant stocks (Rodríguez-Sánchez and Tjaden 2025) and using big data to predict migration patterns to the EU (Nicaise and Bircan 2024), these approaches also face issues, including data reliability (Nicaise and Bircan 2024; Rodríguez-Sánchez and Tjaden 2025).

While comprehensive global figures on migration-related financing remain unavailable, OECD has compiled tracking data of migration-related ODA (OECD 2025). The trend fluctuated from 2019 to 2023, reaching US$1.21 billion in 2023, a slight drop from 2022 (OECD 2025). However, the data is constrained by incomplete reporting, limited sector coverage, and the exclusion of activities not primarily focused on the development of recipient countries (OECD n.d.). As of 2024, IOM's 2025 operational budget was projected at US$2.19 billion in total (IOM 2024a). This budget includes components that address irregular migration, such as return and reintegration assistance, border and identity management, and counter-trafficking (IOM 2024a). The UNHCR secured a global budget of around US$11 billion annually for the period 2022-2025, up from US$9 billion in 2020-2021, to support

[2] Defined by the IOM (2019) as movement occurring outside legal or regulatory frameworks, involving unauthorized entry, transit, or residence, which places individuals at heightened risk of exploitation, violence, and death.

[3] Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Poland, Spain and the UK.

[4] Between 11.08 million and 11.62 million.

people forced to flee conflict and persecution, or those denied nationality (UNHCR 2025). The funds aim to create safe and supportive conditions, protect individual rights, strengthen community capacities, promote gender equity, and identify long-term solutions (UNHCR 2025).[5]

At regional and national levels, various actors have also continued to merge migration policies and funding efforts with development and security objectives, while migration remains a politically contested and sensitive issue in both donor and partner countries. For example, the EU has explicitly linked migration to external cooperation instruments in the 2021–2027 Multiannual Financial Framework (MFF), notably through €79.5 billion allocated to the Neighbourhood, Development and International Cooperation Instrument (NDICI) (European Commission n.d.). The EU also funded the Asylum, Migration and Integration Fund (AMIF) with €9.88 billion until 2027 to support asylum systems, legal migration, return and reintegration efforts, and solidarity among Member States (EU 2025).

In the Netherlands, the government commissioned the COMPASS program with IOM and 14 partner countries to support evidence-based migration policy by enhancing data use and decision-making, investing €100 million in a second phase set for 2024–2027 (Government of the Netherlands 2024; IOM 2024b). The Netherlands has also prioritized action against migrant smuggling through its 2023–2026 national policy framework and a pending bill to increase penalties and jurisdiction, while advancing a new action plan to strengthen the national and international fight against human trafficking (Government of the Netherlands 2024). Within this broader policy landscape, the independent Policy and Operations Evaluation Department (IOB) of the Dutch Ministry of Foreign Affairs has also launched an evaluation project that assesses Dutch migration partnerships.

Despite sustained policy attention and a growing number of interventions, including those from destination countries, evidence on the effectiveness of these interventions remains fragmented and uneven (Czaika and de Haas 2013; Czaika and Reinprecht 2022; Anda León et al. 2023). Previous efforts, including the original EGM (Berretta et al. 2023a) and Anda León et al.'s (2023) systematic review, have begun to address this gap, but the lack of consolidated and up-to-date evidence continues to hinder informed policy decision-making and evaluations.

1.2 Study objectives and questions

This study aims to update the 2023 evidence gap map (EGM) on interventions addressing irregular migration (Berretta et al. 2023a), and expand its framework to include external migration management. This study will systematically identify and describe impact evaluations and systematic reviews published between April 2023 and August 2025.

We will host the EGM on 3ie's online platform, using an interactive matrix framework that visually maps studies by intervention and outcome. Users can filter results by study design, geography, population group, and other metadata. The EGM will be accompanied by descriptive analyses that address key research questions on evidence coverage, gaps, and synthesis opportunities.

[5] It should be noted that these figures were based on original projections, and subsequent shifts in the funding landscape (e.g., reductions following U.S. policy changes) may have altered them. However, the updated numbers were not available at the time of writing this report.

Based on this mapping, we will also develop a set of concise evidence summaries and a synthesis report for selected intervention categories with a sufficient evidence base, defined as those supported by multiple comparable impact evaluations and/or high-confidence systematic reviews. An evidence summary is a concise report that provides a brief overview of the theoretical relationship between an intervention and its outcomes, synthesizes the main findings from identified studies, presents overarching conclusions and recommendations based on the available evidence, and includes a complete reference list. The synthesis component will present key messages on intervention effectiveness and policy relevance, integrating them into short evidence summaries and a final synthesis report aligned with the broader IOB evaluation on migration partnerships.

The objectives of this study are:

- **To update** the original EGM by identifying and mapping new impact evaluations and systematic reviews on interventions addressing the root causes and drivers of irregular migration in L&MICs.
- **To expand** the original EGM to include a new intervention group, external migration management, that targets or involves cooperation with Afghanistan, Pakistan, Türkiye, and L&MICs in Sub-Saharan Africa (SSA) and the Middle East and North Africa (MENA).
- **To describe** the characteristics of identified studies (e.g., region, design, population, intervention type).
- **To identify** primary evidence gaps and synthesis gaps in the literature.
- **To synthesize** findings from selected intervention categories into evidence summaries and an overarching synthesis report.

To meet these objectives, we will address the research questions shown in Table 1 below.

Table 1: Research Components Aligned with Research Questions

No.	Research Question	Research Components
RQ1	What is the scope of the latest available evidence and quality of systematic reviews on the effectiveness of interventions aimed at reducing irregular migration, and where key evidence gaps are that warrant further research?	Updated interactive EGM EGM report
RQ2	What does existing research reveal about the effectiveness of development cooperation interventions in addressing the root causes of irregular migration in L&MICs?	Evidence summaries and synthesis report
RQ3	What does existing research reveal about the effectiveness of external migration management policies in addressing irregular migration?	Expanded EGM Evidence summaries and synthesis report

Note: We will consult with IOB on the number and focus of intervention categories for the evidence summaries once the map update and expansion are sufficiently advanced.

1.3 Why is it important to do this work?

As described in Section 1.1, despite policy attention and a growing number of interventions, including external migration management policies, evidence on the effectiveness of interventions remains fragmented and uneven.

An EGM provides a systematic and transparent way to map what is known and to highlight where critical knowledge gaps remain, which can be a valuable tool for policymakers, practitioners and researchers in the sector. The original migration EGM (Berretta et al. 2023a) identified 89 impact evaluations and 7 systematic reviews, with limited evidence on migration outcomes. Given the steady growth of the evidence base, with about half of the studies published between 2019 and 2023, we expect to identify new studies published since the original EGM search. Updating the EGM ensures that emerging evidence remains systematically mapped and accessible for evaluation and policy use.

In addition, evidence synthesis based on the studies captured in the EGM can reveal what existing research shows about intervention effectiveness, supporting more informed policy and partnership decisions. Our evidence synthesis findings will not only inform the ongoing IOB evaluation project also contribute to research on the effectiveness of interventions addressing irregular migration. To our knowledge, there is still a scarcity of reliable evidence syntheses on the effectiveness of migration partnerships, or on the wider set of interventions covered in this EGM. The original EGM identified only seven systematic reviews, all rated as low confidence due to methodological limitations. 3ie conducted a systematic review of active labor market policies, and found no significant effects on migration intentions, decisions, or mobility outcomes, though findings were constrained by small sample sizes and high risks of bias across most studies (Anda León et al. 2023). Other recent research syntheses examined climate change and environmental pressures as key drivers of migration but did not address the effectiveness of interventions to mitigate such drivers (e.g., Gupta et al. 2025; Larotta Silva 2025; Mukherjee and Fransen 2024).

2. Scope

We reference the same conceptual framework as Berretta et al. (2023a), but also expand it to include a new intervention group on external migration management (see Section 3 for more details on eligibility criteria).

The scope covers interventions that affect migration decisions along a spectrum from voluntary choice to necessity or involuntary movement, consistent with international definitions. For consistency with the original map, we will use the following definitions to frame the scope of this work:

- **Root causes** refer to “the social and political conditions that induce departures, especially poverty, repression, and violent conflict” (Carling and Talleraas 2016), as well as environmental and economic shocks and stressors.
- **Drivers** encompass a broader set of factors that influence irregular migration, including but not limited to root causes. This distinction is maintained to align with both policy terminology and to achieve analytical clarity.
- **Irregular migration** is defined as the “movement of persons that takes place outside the laws, regulations, or international agreements governing the entry into or exit from the State of origin, transit, or destination” (IOM 2019).[6]

[6] This definition emphasizes migrants who enter a country irregularly. However, it is important to note that other definitions also include migrants who become irregular after arrival, for example, by overstaying or letting documents expire, or by taking up employment that is not permitted under their residence status. In our EGM, we exclude internal migration and internal displacement (i.e., migration

2.1 Scope selection process

This update references the intervention and outcome framework and domains previously developed in the original EGM, which we developed through reviewing policy documents, academic literature, and consultations with an Advisory Group (AG) of experts in the field (Berretta et al. 2023a). The domains covered the most relevant and salient interventions at that time.

In this EGM update, we add a new intervention group, *external migration management*, to reflect evolving EU policy priorities (European Commission 2024), including current policies and partnerships enacted by the Government of the Netherlands. We developed the expansion in consultation with IOB and reviewed it with an AG established for this research project. The AG is composed of migration experts who will engage at key stages, including the EGM report, evidence summaries, and synthesis report, to enhance the credibility and policy relevance of our work and ensure diverse perspectives are incorporated throughout the process (see Appendix A for the full list of AG members).

2.2 Conceptual framework

The original EGM (Berretta et al. 2023a) adopted Carling's (2002) theory of migration aspirations and abilities. In the current protocol, we extend this by drawing on Carling and Schewel's (2018) two-step approach, which revisits and refines the 2002 theory, alongside structural theories linking conditions in origin countries to migration outcomes. The two-step approach extends the original focus on involuntary immobility by incorporating the capabilities approach and more complex interactions between aspiration and ability (Carling and Schewel 2018). The original EGM covered interventions that influence migration aspirations, shaped by macro conditions in origin countries (economic insecurity, violence, and climate-related shocks), as well as global factors, also known as migration intervening factors (IOM 2021). Migration intervening factors include the availability or lack of legal pathways and information, which can facilitate or impede migration. Many eligible intervention categories in the original EGM framework target micro-level factors, particularly at the household or individual income level, through which these macro-level drivers influence migration decisions and behavior.

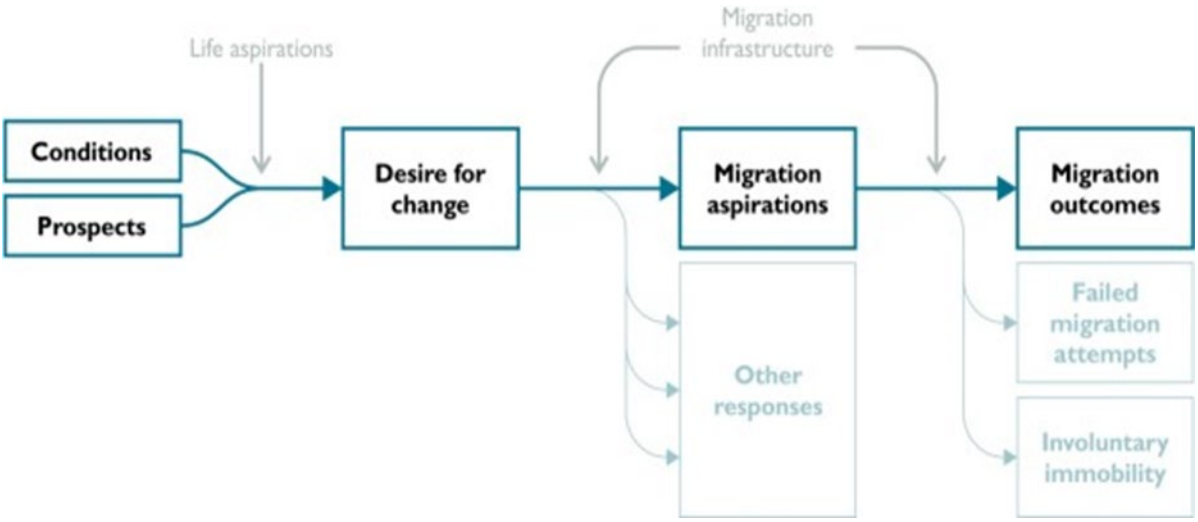
Figure 1 presents Carling and Talleraas's (2016)[7] simplified conceptual model of individual migration decision-making, illustrating how conditions in origin countries and migration infrastructure might influence decisions to migrate. In this model, external migration management (the new domain to be added to the EGM update) is a facet of the migration infrastructure. The interventions in this domain are hypothesized to influence migration aspirations by increasing perceived risks and costs of irregular migration and shaping the perception of outcomes. In the stepwise process, where migration aspirations may or may not result in actual migration depending on the individual's ability to migrate, external migration management is assumed to reduce irregularity by providing more accessible legal pathways. When staying and moving can be considered complementary expressions of

within a single country).

[7] This visual representation of the conceptual model provides a simplified snapshot of individual migration decision-making, but it does not fully capture what Carling and Schewel's (2018) two-step approach, including the dynamic interplay between aspiration and ability.

agency, shaped by conditions that support the realization of aspirations (de Haas 2021), interventions that strengthen individuals' capabilities to pursue those aspirations can reduce irregular migration.

Figure 1: Theory of change by Carling and Talleraas (2016)



The "root causes" theories on irregular migration have been debated by both academic and policy debates. We do not examine their merits, but instead present below interventions that aim to address each driver and the underlying theories as to how they relate to irregular migration. During the evidence synthesis, we will aim to identify findings on the links between interventions addressing specific presumed root causes and migration outcomes and, in the cross-domain synthesis, to assess the validity of root-cause theories where possible.

Economic drivers (e.g., poverty, unemployment, and lack of decent work) can influence life aspirations and desires for change and make migration an economically viable option for achieving those aspirations (Carling 2002; de Haas 2010). Root-cause approaches often assume that sustained economic growth and expanded opportunity can reduce aspirations to migrate (Clemens 2014, 2020), while some research indicates that the relationship between economic development and migration does not necessarily follow a simple, linear "migration hump" (Bencek and Schneiderheinze 2020; Berthiaume et al. 2021; Casentini, Hammond and Bakewell 2024). In the absence of accessible legal pathways, economic interventions such as cash transfers, skills programs, and credit access may reduce migration aspirations but may also increase individuals' ability to migrate, as they reduce financial barriers; thereby inadvertently leading to greater irregular migration (McKenzie 2017). Casentini, Hammond and Bakewell (2024) propose that understanding other forms of inequality as well as access to social networks and safe routes can help disentangle the relationship between income inequality and migration.

Environmental and climate-related stressors disproportionately affect people's lives and food security (FAO 2021). In contexts of high vulnerability, climate shocks contribute to humanitarian emergencies and can trigger stepwise patterns of mobility, starting with local displacement and potentially leading to international migration (Almulhim et al. 2024). Climate change can also shape the conditions migrants encounter in their destinations (Huang 2023). However, the way in which climate change affects mobility is not

deterministic, as migration decisions are shaped by a mix of economic, social, and political factors, not just environmental and climate stressors (Boas, Gautama and Olayiwola 2024; Mixed Migration Centre 2022). Resilience-building interventions, such as insurance, infrastructure, and safety nets, may reduce the pressure to migrate (Mueller et al. 2019).

Conflict, violence, and insecurity are one of the major drivers of migration (Clemens 2021). Weak justice systems and related institutions can heighten risks of violence, while broader governance failures, such as limited access to basic services, low trust in authorities, corruption, and political instability, can increase incentives for irregular migration (Beine et al. 2021). Community safety initiatives and violence prevention programs could improve perceptions of conditions, self-agency, and empowerment, thereby lowering migration intentions.

The lack of legal pathways and information drives irregular migration by removing safe, authorized options to move and forcing reliance on unauthorized routes (Casentini, Hammond, and Bakewell 2024). Legal migration pathways, especially for work, can reduce irregular migration by offering viable alternatives, though their effectiveness depends on awareness of and access to such pathways (McKenzie and Gibson 2010; Clemens and Postel 2017). Triandafyllidou, Bartolini, and Guidi (2019) emphasizes that ensuring viable regular migration alternatives is essential to reducing reliance on irregular entry. Migration information campaigns can influence decision-making, but their effects vary across contexts and delivery modalities (e.g., top-down versus peer-to-peer messaging; Dunsch, Tjaden, and Quiviger 2019) and often remain limited (Tjaden and Dunsch 2021; Caso and Carling 2024; Trauner et al. 2024).

External migration management interventions, such as partnerships, deterrence, remote control, and externalization, are often based on the assumption that they can shape the accessibility and opportunity to migrate irregularly (Frelick, Kysel and Podkul 2016; Xanthopoulou 2024). While restrictive measures may raise perceived risks and alter migration routes, their effectiveness depends on offering credible alternatives and addressing underlying drivers (de Haas 2007; Cummings et al. 2015; Triandafyllidou, Bartolini, and Guidi 2019). Tafani and Riccaboni (2025), for instance, suggest that the EU–Turkey statement narrowly targets one corridor, simply displacing risks to longer and riskier route, calling for a more integrated policy design. The effectiveness of external migration management also hinges on how it is mediated by partner countries, whose cooperation can reinforce, adapt, or undermine destination-country policies depending on their interests and capacities.

The decision to migrate arises from the interplay of aspirations and opportunities, shaped by structural and contextual conditions (Carling and Schewel 2018). Among those who aspire to move, only those with sufficient resources, networks, and enabling circumstances are able to realize these aspirations (Mckenzie and Rapoport 2007; Beegle, Weerdt and Dercon 2011; Carling and Schewel 2018). Framing migration through the aspiration–capability nexus highlights its multidimensional nature and shows why irregular migration remains challenging to address: interventions must consider both the factors shaping aspirations and the unequal capacities that enable mobility.

3. Eligibility criteria

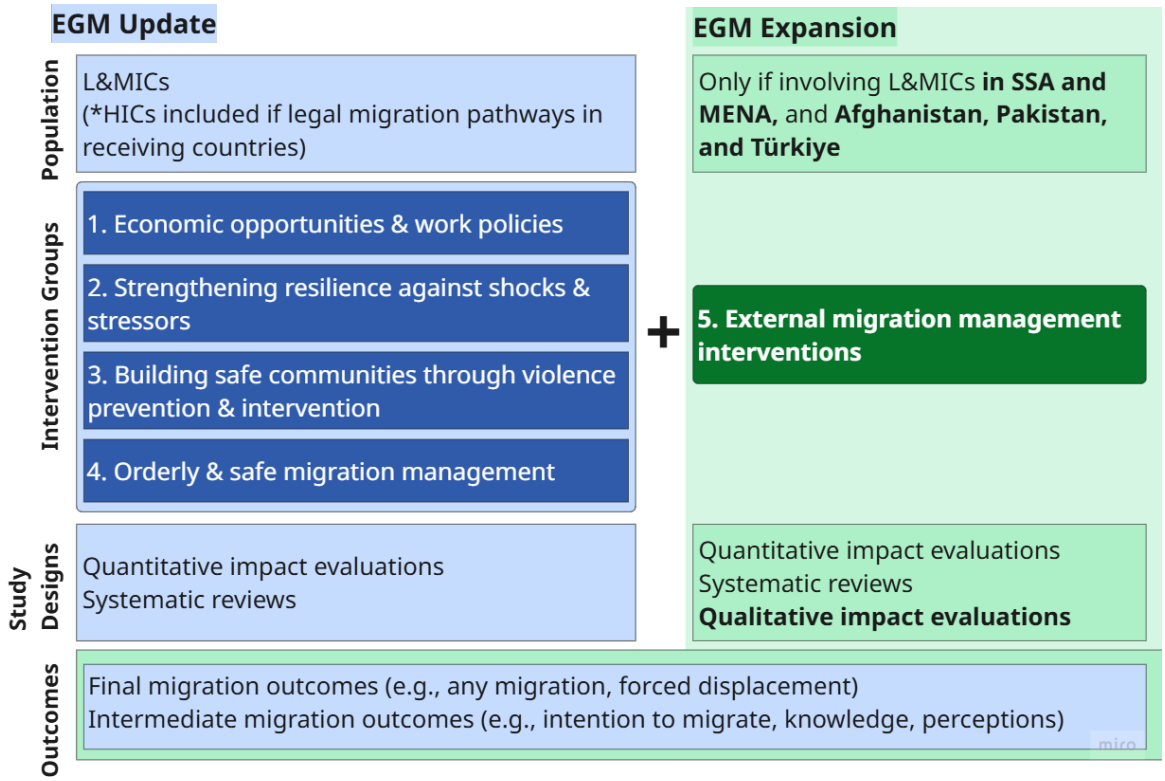
The EGM will remain a single integrated product covering all L&MICs, with the expansion adding a new intervention domain that applies more specific population criteria and includes additional study designs (see Figure 2).

For the EGM update, the eligibility criteria will follow those in the original protocol (Berretta et al. 2023a), with the key change being the publication date range. The updated timeframe extends from 1990 to August 2025 (the original timeframe ended in April 2023) to include studies published after the original searches. See more details in Appendix B.

For the expansion of the EGM, which involves incorporating a new intervention domain that captures external migration management interventions, focuses specifically on interventions targeting or involving cooperation with L&MICs in the SSA and MENA regions, as well as Afghanistan, Pakistan, and Türkiye. For this new intervention group, besides quantitative impact evaluations and systematic reviews, we will include qualitative evaluations without a comparison group, provided they meet minimum quality standards listed in Appendix B. Finally, for the expansion, we will focus on studies published from 2015 onwards, as the literature in this area is more recent and reflects the novelty of interventions in the emerging domain of external migration management policies.

We will include studies in any language and conduct academic searches in English. Appendix B provides more details on the full EGM update and expansion eligibility criteria.

Figure 2: Summary of PICOS criteria for the EGM update and expansion



Note: The EGM update refers to updating the original EGM (Berretta et al. 2023a). The EGM expansion refers to adding the new intervention group, external migration management, to the original EGM. Abbreviations: HICs = high-income countries; L&MICs = low- and middle-income countries; MENA = Middle East and North Africa; SSA = Sub-Saharan Africa.

4. Methods

4.1 EGM update and expansion

4.1.1 Overall methodological approach

We will follow the standards and methods for EGMs developed by 3ie (Snilstveit et al. 2016, 2017), as used in the original 3ie Migration EGM (Berretta et al. 2023a). These standards involve systematic methods to search, identify, and describe all completed and ongoing impact evaluations and systematic reviews relevant to the research objectives. We describe the methods in detail below.

4.1.2 Search strategy

The search strategy for this EGM update will follow the approach used in the original 3ie Migration EGM (Berretta et al. 2023a), which applied systematic searching principles based on established guidelines (Kugley et al. 2017). We will search for new studies across academic bibliographic databases, applying an updated search period (between April 2023 and August 30, 2025) to cover newly identified publications since the last search. We will tailor search strings and logics to each intervention domain and database, drawing on prior EGMs where relevant (e.g., Doherty et al. 2020; Berretta et al. 2023b). We will skip the grey literature search for the update portion of the EGM due to the low yield and high manual effort involved.

We will also design a search strategy for the new domain as part of expanding the intervention framework. We will conduct searches in both academic bibliographic databases and grey literature sources, covering specialist organizational websites, bilateral and multilateral agencies, as well as repositories focused on international development (see Appendix C for more details). For the grey literature search, we will conduct searches using both English and French keywords and explore AI-based web scraping tools to support searches in large organizational databases, where feasible. When time permits, the search will also incorporate citation tracking and outreach to domain experts through the AG to identify additional studies not yet indexed in public repositories.

4.1.3 Screening, data extraction, and critical appraisal of studies

We will adopt the same systematic processes for study screening, data extraction, and critical appraisal of systematic reviews as outlined in the protocol of the original 3ie Migration EGM (Berretta et al. 2023a; see Appendix D for details). In summary, we will use EPPI-Reviewer software (Thomas et al. 2023) for the search results management, including duplicate removal, and selection of studies. For screening, we will implement a structured training and consistency-check process; we will conduct title/abstract screening independently by two reviewers until they achieve a minimum inter-rater reliability of 85%, using predefined eligibility criteria; and we will conduct full text screening independently in duplicate with any disagreements resolved by a third review team member. For title and abstract screening, we will apply EPPI-Reviewer's machine learning classifiers (EPPI-Centre 2022) to help prioritize the screening workload.

For data extraction, we will systematically extract information from all included studies using a predefined data extraction tool (see Appendix E for the tool). This will cover core study and publication details (e.g., authors, location, outcomes, population, funders) as well as cross-cutting issues such as equity and population targeting (e.g., Indigenous populations). One

coder will extract EGM data, with independent quality assurance on a randomly selected subset performed by a second review team member.

We will critically appraise all included systematic reviews using an adapted version of the 3ie systematic review database protocol, based on Lewin et al. (2009) and Snilstveit et al. (2017). We will assign confidence ratings (low, medium, or high) based on the extent to which the review used gold standard methods to minimize the introduction of bias while conducting and reporting the review. For full details of these procedures, see Appendix F.

4.1.4 Analysis and reporting

We will conduct descriptive analyses of the included studies by publication year and type, geography, participant characteristics, interventions, outcomes, study design, equity, and other cross-cutting themes, with cross-tabulations where appropriate. The outputs will include an updated and expanded interactive evidence gap map, publicly available on the 3ie website, with filters such as region, country, study design, and equity dimensions. We will also produce an updated and expanded EGM technical report that details the methods, key findings, and policy implications.

4.2 Evidence summaries

To synthesize the evidence, we will generate evidence summaries, which will be 3-4 pages each and include: (1) a brief overview of the theoretical relationship between the interventions and the outcomes; (2) main findings from the identified studies; (3) overarching conclusions and recommendations based on all studies; and (4) the reference list of studies included in the summary.

4.2.1 Selection of intervention categories

The selection of intervention categories for the evidence summaries will begin once the updated and expanded EGM is sufficiently advanced. The volume of available evidence, the absence of existing synthesis, and policy relevance will guide the selection process in consultation with IOB. We will prioritize up to seven intervention categories for evidence summaries.

4.2.2 Quantitative data extraction and management

For quantitative impact evaluation studies included in the selected intervention categories, we will extract the following types of data using a standardized data extraction form (see Appendix G for the full provisional tool):

- **Descriptive information**, such as study authors, publication date, publication status, country, intervention and outcome types, and intervention design.
- **Methodological details**, including study design, analytical methods, and type of comparison group, where applicable.
- **Quantitative outcome data**, including descriptions of outcome measures, sample sizes for intervention and comparison groups, outcome means and standard deviations (SDs), and relevant test statistics (e.g., t-tests, F-tests, p-values, and 95% confidence intervals).
- **Cost evidence,** including whether any cost evidence is reported, and types of cost evidence, if available.

Some of these data will have already been coded as part of the EGM update and expansion. For newly extracted fields, a trained reviewer will single-code descriptive and methodological data, and a second reviewer will check the data for accuracy. We will perform duplicate independent extraction of any quantitative data that will be used to calculate effect sizes, per Cochrane guidance (Higgins et al. 2024). We will resolve any discrepancies through discussion with a third reviewer.

4.2.3 Risk of bias assessment for quantitative impact evaluations

We will assess the risk of bias of the studies included in the selected categories for evidence summaries using an abridged version of 3ie's tools. It includes bias domains and extensions from Cochrane's ROBINS-I and RoB 2.0 tools (Higgins et al. 2016; Sterne et al. 2016). This assessment addresses both the internal validity and statistical conclusion validity of experimental and quasi-experimental impact evaluation designs (Waddington et al. 2012). Two reviewers will independently conduct the risk of bias assessment, with disagreements resolved by a third reviewer as needed.

Within included studies, we will assess the risk of bias for each estimate based on the following domains by answering whether the estimate is free from each bias, with a response set of "Yes," "Probably Yes," "Probably No," "No," and "No Information":

- 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 the results of the analysis.

We will report the assessment results for each risk of bias domain. In addition, we will assign an overall risk of bias rating to each study: "High risk of bias", "Some concerns" or "Low risk of bias", drawing on the decision rules in RoB2.0 (Sterne et al. 2019):

- "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 the bias domains were assessed as "Yes" or "Probably Yes".

We will describe the reliability of the findings of included studies in our analysis and explore whether differences in estimated effects are associated with varying levels of risk of bias across studies. We will conduct a sensitivity analysis of the risk of bias assessment on the pooled effect, as described in Section 4.2.11.

4.2.4 Measures of treatment effect

An effect size (or treatment effect) represents the direction and magnitude of the difference in outcomes between groups, such as the difference between an intervention and a comparison group (Borenstein et al. 2021; Valentine et al. 2015). However, effect sizes reported in empirical studies are often dependent on the specific outcome scale or unit used, making direct comparisons across studies difficult. To enable cross-study comparisons, we will extract data from each study and calculate standardized effect sizes. We will select the

appropriate formulas for these calculations based on the type of outcome and the data available in each study. Details on the formulas used are provided in Appendix H.

When multiple outcome measures fall within the same outcome category, we will convert estimates to the most commonly used standardized metric to enhance comparability. To do this, we will apply the standard transformations to convert between different measures of standardized effects (Polanin and Snilstveit 2016; Borenstein et al. 2021).

4.2.5 Criteria for the determination of independent findings

It is essential that our analysis accurately captures and accounts for co-dependencies among study estimates. Standard meta-analytic methods assume that effect size estimates are independent; however, failing to recognize when estimates are derived from the same intervention can lead to distorted or exaggerated conclusions.

Dependent effect sizes can arise in several scenarios. For example, when multiple publications stem from a single study, when multiple studies use the same dataset, or when studies include multiple treatment arms compared to a shared control group. Dependencies also occur when outcomes are measured at different time points or when multiple related outcome measures are reported. These situations result in statistically dependent effect size estimates (Borenstein et al. 2021).

We will assess the degree of relatedness across included studies. To avoid double-counting evidence, we will identify and link related publications prior to analysis. Supporting information, such as sample sizes, intervention characteristics, and implementing or funding partners, will be used to guide this process. When multiple publications report the same effect, we will select one primary source for data extraction, supplementing with information from related reports as needed. Priority will be given to journal articles; in cases involving multiple reports or working papers, the most recent publication will be used.

We will extract effects across different interventions, outcomes, and subgroups within each study. In our experience, the structure of our data is nearly always dependent, thus we anticipate the use of robust variance estimation (RVE) methods (Fisher and Tipton 2015; Hedges et al. 2010) to incorporate all relevant data while accounting for statistical dependence, provided the minimum degrees of freedom for valid inference are met. If we are underpowered to perform the analysis using robust variance techniques, we will apply data processing and selection strategies to address dependent effect sizes. In such cases, we will apply predefined criteria to select a single effect estimate per outcome per study (see Appendix I).

4.2.6 Unit of analysis issues

Unit of analysis errors can occur when the unit at which the intervention is implemented differs from the unit of analysis for the effect size estimate, and this discrepancy is not accounted for in the analysis (e.g., by clustering standard errors (SEs) at the level of allocation). We will assess the included studies for the prevalence of these issues and, where they are identified, adjust the reported SEs using the following formula (Hedges 2009; Higgins et al. 2022):

$$(d)' = (d) \cdot 1 + (m-1)c$$

Where *d* is the effect size, *m* is the average number of observations per cluster, and *c* is the intra-cluster correlation coefficient (ICC). If the included studies use robust Huber-White SEs to account for clustering, we will calculate the SE of *d* by dividing by the t-statistic for the coefficient of interest.

If the ICC is not presented in the manuscript, we will search the literature for an appropriate ICC value. If no such value is available, we will assume an ICC of 0.05 (Waddington et al. 2014).

4.2.7 Dealing with missing data

In cases of missing or incomplete data, we will make every effort to contact study authors to obtain the required information. If the necessary data cannot be obtained, we will report the study's characteristics and note that it was excluded in any meta-analysis or effect size reporting due to missing data.

Following the recommendations of Mullan et al. (2009) for data collection in systematic reviews, we will document the number of studies for which authors were contacted, the specific information requested, the methods used to elicit responses, and the outcomes of those efforts. Where applicable, we will also report the impact of the information obtained on the results, including any effects observed in sensitivity analyses.

4.2.8 Data synthesis

We will use meta-analysis to combine studies when there are at least two effect sizes with comparable outcome constructs and similar comparison group conditions, following the approach outlined by Higgins et al. (2024). Provisionally, studies will be grouped for meta-analysis if they assess either the same type of intervention or the same type of outcome.

We will use inverse-variance weighted, random-effects models to account for heterogeneity across interventions and contexts (Higgins et al. 2020). We will conduct all meta-analyses using R software (R Core Team 2022), specifically the *metafor* package (Viechtbauer and Cheung 2010). If data are dependent, we will use the correlated and hierarchical effects model (CHE). If meta-analysis is not feasible due to high heterogeneity or insufficient data, we will synthesize findings narratively.

For cost data, we will narratively summarize it in a broad, descriptive way, which involves identifying whether each study reports cost information (yes/no) and noting the type of cost data included (e.g., unit costs, total program costs). We will briefly summarize key cost descriptions without standardizing or adjusting figures. Where possible, we will highlight broad patterns or gaps. For example, whether cost data are consistently reported.

4.2.9 Assessment of reporting biases

To reduce the possibility of publication bias, we will identify and include both published and unpublished studies in the review. We will use two approaches to explore potential publication bias. First, we will compare unpublished versus published studies by dummy coding publication status as a moderator. Second, we will use selection modeling to explicitly model the probability of reporting based on the *p*-value (e.g., statistically significant positive intervention effects are generally more likely to be reported than nonsignificant effects) using the *weightr* package in R (Coburn 2019). Despite the advances in analytic methods to assess publication bias, the results of these approaches should be viewed as sensitivity analyses, rather than conclusive bias-corrected estimates (Carter et al. 2019).

4.2.10 Subgroup analysis and investigation of heterogeneity

In our analysis, we intend to examine and discuss the distribution of estimated effects across intervention and outcome types. Following the PROGRESS-PLUS approach (Lipsey 2009), we will assess moderators falling into three broad categories: extrinsic, methodological, and substantive characteristics. Examples include:

- Extrinsic characteristics, such as the study funder (e.g., NGO, private sector, and government investments), publication type, and publication date.
- Methodological characteristics, such as study design, risk of bias, length of follow-up, and types of outcome measures.
- Substantive characteristics, such as participant characteristics (e.g., gender, age, socioeconomic status, and education), context (e.g., geographical setting), intervention type, intervention features, and type of intervention implementers.

Where meta-analysis is feasible, we will statistically assess heterogeneity by calculating the Q statistic, I2, and τ2 to estimate the amount of variability in the distribution of effect sizes and we will report the prediction interval for the average effect (Borenstein et al. 2021). We will complement this assessment with a graphical analysis (e.g., using forest plots or MARC plots). Whenever feasible, we will conduct moderator analyses using random-effects meta-regression to investigate potential sources of heterogeneity.

If these strategies are not possible (e.g., due to insufficient studies or data), we will explore potential drivers of heterogeneity of results narratively by conducting cross-case comparisons (Miles and Huberman 1994).

4.2.11 Sensitivity analysis

We will conduct a sensitivity analysis to determine whether the results of the meta-analysis are sensitive to the removal of individual studies. This will involve excluding each study one at a time and assessing how the results change, assessing sensitivity to potential outliers. Specifically, we will use studentized residuals to identify studies that may have unusually large effect estimates (Viechtbauer and Cheung 2010). Studies with a studentized residual exceeding the $100 \times (1 - 0.05/(2 \times k))$th percentile of a standard normal distribution will be considered as potential outliers. Additionally, when relevant, we will evaluate the impact of including studies with a high risk of bias by removing these studies from the analysis and comparing the findings with the main meta-analysis results.

4.2.12 Treatment of qualitative evidence: for intervention categories of external migration management

If we conduct evidence summaries for any intervention category on external migration management, we will appraise and extract data from qualitative evaluations included in the selected categories. For the quality appraisal of qualitative evidence, we will use the *Critical Appraisal for Methodological Limitations of Qualitative Research Tool (CAMELOT)* (Munthe-Kaas et al. 2024). CAMELOT is designed specifically for use in qualitative evidence syntheses and supports the assessment of how the design and conduct of studies may influence confidence in review findings, through appraisal across 12 domains (see Appendix J for details and an assessment template).

In CAMELOT, reviewers extract and note concerns for each of the 12 domains and then assess the appropriateness of fit between relevant domains. Based on this holistic assessment, rather than assigning summary scores (e.g., high, medium, low quality), an

overall level of concern for methodological limitations is assigned:

- **No or very minor concerns:** The study meets the expectations for methodological quality in this domain. Any issues identified are negligible and unlikely to compromise the credibility of the findings.
- **Minor concerns:** There are some limitations in how the study addresses this domain, but they are relatively small and unlikely to meaningfully affect the trustworthiness of the relevant finding.
- **Moderate concerns:** The domain shows more noticeable weaknesses that may limit confidence in the study's contribution to the synthesis. The methodological limitations raise questions about how well the findings reflect the phenomena of interest.
- **Serious concerns:** There are major methodological flaws in this domain that significantly reduce confidence in the study's findings. These limitations could undermine the credibility or relevance of the data used in the synthesis.

One reviewer will conduct initial assessments, and a second will cross-check these assessments whenever possible. The reviewers will resolve discrepancies through discussion. If agreement cannot be reached, a third reviewer will be consulted for reconciliation. We will report appraisal ratings of the qualitative evaluations transparently.

For all of the qualitative impact evaluations under the selected categories, we will extract qualitative evaluation data thematically to generate evidence on the theoretical relationships between the intervention(s) and outcome(s), main findings, and overarching conclusions and recommendations. If quantitative impact evaluation data are also available, whether from meta-analysis or narrative synthesis, in the same intervention category, we will use the coding from qualitative studies to support interpretive synthesis and triangulate these findings with the quantitative results. Thematic coding of qualitative studies will include intervention characteristics (e.g., type, delivery modality, duration), contextual and implementation factors (e.g., geographic and environmental settings, demographic factors), and sustainability features (e.g., long-term institutional or financial arrangements, systems to maintain interventions).

4.3 Synthesis report

The synthesis report is a concise integrative product that outlines the methodology and brings together findings from the updated and expanded EGM as well as evidence summaries to answer the research questions. Findings across summaries will be triangulated to enhance interpretability, highlight complementary evidence, and surface areas of convergence and divergence across study types and contexts. The final synthesis report will be no more than 15 pages, excluding annexes, and will be written in accessible language to support its use among policymakers and practitioners. Feedback from IOB and the AG will be incorporated before finalization.

4.4 Timeline

The approximate date for publication per deliverable is:

- **Interactive EGM and Technical Report**: November 2025
- **Evidence Summaries and Synthesis Report**: February 2026

We will publish this inception report (protocol) and the final EGM report on 3ie's website and make them publicly available.

4.5 Engagement and communication plan

We will share the results of the EGM, evidence summaries, and synthesis report with IOB, its internal audiences, the AG members, and, more broadly, stakeholders in the migration sector. We will engage with key stakeholders and ensure the results of the project accurately reflect their policy and research needs.

4.5.1 Develop an advisory group

In collaboration with IOB, we have engaged with key stakeholders with academic and practitioner expertise in the field of migration. The AG will provide support to the project at several key stages. These stages include updating the project inception report (protocol), reviewing the search results, reviewing and interpreting emerging findings, and developing and optimizing the analytical deliverables produced to support evidence uptake and use.

4.5.2 Develop a Stakeholder Engagement and Communication Plan

We will draft a stake-holder engagement and communication plan (SECP).[8] The aim of this plan is to ensure that findings from the EGM and evidence synthesis are effectively disseminated to the appropriate audiences in an engaging and accessible format. This plan includes a provisional analysis of key stakeholder groups, focusing on their relevant interests and the extent to which 3ie and/or IOB have access to them, as well as an assessment of what the most value-added EGM project outputs might be to aid evidence uptake and use. The SECP is a 'living document'. We will refine it as additional information needs or dissemination opportunities are identified by the project team, AG members and IOB.

[8] A template example is available at https://www.3ieimpact.org/sites/default/files/2018-12/3ie_SECP%20template.pdf.

Appendix A: Advisory Group

3ie formed an advisory group (AG) of experts from policy, practice, and academia. The first kick-off meeting was held in August 2025. AG members participate in a personal capacity and provide independent advice. Their contributions do not represent, and should not be attributed to, their institutional affiliations or their member states. Membership includes:

1. **Kurtz, Jon**
   Senior Director for Research & Learning, Mercy Corps

2. **Le Coz, Camille Le Coz**
   Director, MPI Europe (Migration Policy Institute Europe)

3. **Manke, Marina**
   Senior Global Advisor on Humam Mobility, engaging in independent capacity and not representing IOM or its member states.

4. **Pécoud, Antoine**
   Professor of Sociology, Université de University of Sorbonne Paris Nord

5. **Slootjes, Jasmijn**
   Deputy Director Europe, MPI Europe

6. **Trauner, Florian**
   Dean, Brussels School of Governance, Vrije Universiteit Brussel (VUB)

7. **Van Liempt, Ilse**
   Associate Professor, University Utrecht

8. **Vezzoli, Simona**
   Postdoctoral Researcher and Leader of the PACES Project at the International Institute of Social Studies (ISS), Erasmus University Rotterdam

9. **Zardo, Federica**
   Senior Researcher, Department for Migration and Globalisation, Danube University Krems

Terms of reference for an advisory group

For the project mapping and synthesizing evidence on root causes and drivers of irregular migration

Background

The Ministry of Foreign Affairs of the Kingdom of the Netherlands has commissioned the International Initiative for Impact Evaluation (3ie) to: 1) update and expand the 2023 evidence gap map (EGM) on interventions addressing the root causes and drivers of irregular migration, and 2) synthesize findings from relevant evaluations to answer priority policy questions. This research aims to inform Dutch migration partnerships and development cooperation by identifying where robust evidence exists and where gaps remain. Key outputs include an updated and expanded EGM, up to seven evidence summaries, and a synthesis report.

The IOB and 3ie are seeking Advisory Group members to provide inputs for key deliverables (i.e., inception report, EGM report, evidence summaries, and synthesis report) and to ensure that outputs are policy relevant and meet the needs in the sector. Advisory groups are comprised of sectoral experts with knowledge of the policy and evidence landscape.

Terms and responsibilities

Advisory Group members will receive acknowledgement in the written research outputs. The total time commitment is not likely to exceed 2 days over the project's implementation period, 4 July 2025 to 20 March, 2026.

Advisory Group members may be asked to participate in the following ways:

- Advise on key decisions regarding the new intervention group (external migration management) of the evidence gap map, including refining the intervention categories.
- Suggest relevant background literature and studies for inclusion.
- Participate in virtual meetings for the duration of the project (e.g., draft inception report; draft final reports).
- Provide written comments on draft inception report and final reports.
- Help the team draw policy implications from the evidence gap map and evidence summaries. This may involve participating in a brainstorming/focus group session to review lessons and implications.

Appendix B: Population, Intervention, Comparator, Outcomes, and Study Design (PICOS) for EGM Update

Population

Updating the original EGM (Berretta et al. 2023)
We will include interventions implemented in low- and middle-income countries (L&MICs). An exception is when interventions fall under the intervention category *legal pathways* of the *orderly and safe migration management* domain (see Table B1 below); in this case, we will include interventions implemented in high-income countries (HICs). This is because legal pathway interventions often occur in neighboring and/or developed destination countries (e.g., seasonal work or humanitarian visas) (Freier and Holloway 2018). We will define country income level by using the most recent World Bank income status classification (see Appendix K). We will classify the income level of a country according to the starting year of the study's intervention implementation. If the study does not explicitly specify the intervention's first year of implementation, we will classify the country income level based on the study publication year.

The same applies to multi-country studies. Except for the interventions falling under the 'legal pathways' intervention domain, we will include a multi-country study if it measures at least one estimate of effectiveness for a population based in L&MICs. In this case, the study must provide results for HICs and L&MICs separately.

For the systematic reviews, when they include a mixture of evidence from both HICs and L&MICs, we will include them if they present disaggregated evidence for L&MICs, or if more than 50 percent of the evidence of non-disaggregated results is from L&MICs. Where there are no disaggregated results for L&MICs and more than 50 per cent of the evidence for

consolidated findings in a systematic review comes from high-income countries, or where it is impossible to ascertain the composition of evidence by income level, the studies are excluded.

Expanding the EGM to cover external migration management
We will include interventions involving any of L&MICs in SSA and MENA, according to the WB classification (Appendix K). We will also include Afghanistan, Pakistan, and Türkiye.

For systematic reviews that include evidence from both HICs and LMICs, we will include them if they either (i) present disaggregated evidence for L&MICs in SSA and MENA and/or Afghanistan, Pakistan, and Türkiye, or (ii) contain non-disaggregated results in which more than 50 percent of the evidence is drawn from these countries.

Intervention(s)

Updating the original EGM (Berretta et al. 2023)
In this EGM, we will exclude interventions only focusing on internal migration and internal displacement (i.e., migration within a single country) in alignment with the original EGM.

The interventions include four domains. The first three of these domains are widely acknowledged as root causes, and the last one as another driver of irregular migration:

- Limited economic and work opportunities;
- Inability of households and communities to adapt to, mitigate, or recover from covariate shocks or stressors due to diminished resilience;
- Violence or crime creating insecure and unsafe communities;[9] and
- Drivers that increase the risk of migration through irregular pathways.

In the *legal pathways* category, we will remove two components: *bilateral agreements* and *administrative capacity*-focused interventions, to avoid overlap with a new intervention domain described below. We will review and recode any relevant studies included under this category in the original EGM.

The fourth included domain addresses the reality that migration and human mobility are a fundamental development strategy for some individuals and households. However, there are factors outside of origin countries that may increase migration through irregular pathways. These include information asymmetries, lack of awareness of legal or labor rights, or the absence of migration governance structures for achieving sustainable migration. The final set of interventions covers micro, meso, and macro-level root causes and one additional driver (IOM GMDAC 2021). For instance, disaster risk financing policies are usually implemented at a macro level, while cash transfers target causes at a micro level; information campaigns on the risks of irregular migration target factors at all levels, while policies expanding legal migration pathways have effects at the macro level.

Expanding the EGM to cover external migration management
External migration management interventions, including migration partnerships, migration deals, externalization, deterrence, and remote control. These external migration

[9] The focus of this intervention category is on crime prevention and community-based initiatives to increase safety and security. Interventions that specifically address issues arising from climate change should be placed in the corresponding resilience categories.

management interventions also cover interventions at micro, meso, and macro-levels (IOM GMDAC 2021). We will exclude interventions only focusing on internal migration and internal displacement (i.e., migration within a single country) in alignment with the original EGM.

These intervention categories are not fully mutually exclusive (for example, migration deals often include externalization activities). For EGM coding, our approach will be pragmatic: if an impact evaluation measures the effect of externalization specifically (even within the context of a partnership/deal), we will code it under *externalization*; if it assesses the impact of a partnership/deal as a whole, we will code it under *migration partnership/deal*. These categories may be refined once more literature is reviewed. If refined, we will report changes in the final report.

Table B1: Interventions included in the map

Domain	Intervention	Level	Description
Updating the original EGM (Berretta et al. 2023)			
Economic opportunities and work policies	**Active labor market policies**	Macro, meso	Demand-side interventions aimed to increase individuals' access to employment and entrepreneurship opportunities. This may include skills-based interventions such as technical and vocational education training, business skills training, mentorships, internships/apprenticeships, entrepreneurship workshops; job placement centers and matching programs, employment pipelines/pathways within communities; wage subsidies; or public works schemes.
	Access to large credit markets	Macro, meso	Interventions to improve or increase access to large capital credit or loans for the purposes of establishing a business or facilitating industry growth. This does not include microcredit or indexed insurance (for microcredit or index insurance, see Strengthening Resilience domain).
	Work policies	Macro, meso, micro	Supply-side interventions that create opportunities for work aimed to be productive and deliver fair incomes, occupationally safe and secure workplaces, social protection benefits (for social protection benefits delivered by employers – for example, health insurance policies or programs provided by employers. If delivered by government – for example, unemployment assistance, see Strengthening Resilience domain), prospects for personal development and social integration, freedom to express concerns, organize, and participate in decisions that affect workers' lives or treatment.

Domain	Intervention	Level	Description
	Microcredit and microinsurance schemes	Macro, meso, micro	Provision of, or increasing availability and access to, microcredit and/or microinsurance for households, entrepreneurs, or agricultural producers.
	Human capital strengthening interventions (non-food)	Macro, meso, micro	Interventions that financially support human capital development outcomes directly (e.g., costs of schooling or health services) or indirectly by supporting non-food basic needs. This includes the following: - Cash transfers: Giving cash to assist in meeting needs of recipients. The intervention could target eligible populations or be universal. Examples include unconditional, labelled (no conditions attached, but explicitly label the purpose of cash transfer), and conditional cash transfer. Retirement or senior citizen pensions, giving cash universally and unconditionally to citizens, is included in this category. - Health insurance or interventions that increase access to health services. *If an intervention is 'Cash for Work (employing participants for public work and giving them cash)', it applies to the 'employment assistance' intervention category of the Resilience domain. * If an intervention supports food/nutrition related basic needs, it applies to the 'food and nutrition interventions' category under the Resilience domain.
Strengthening Resilience against shocks and stressors[10]	**Disaster risk financing policies and index-based insurances**	Macro, meso	Public financing policies that aim to manage disaster risks. Examples include risk transfer instruments (e.g., public agricultural, index-based livestock or weather- based insurance policies), loans (e.g., public contingent credit, borrowing and concessional financing), or revenue generation/fiscal policies (e.g., co-financing incentives for in-country stakeholders).
	Early warning systems	Macro, meso	Early warning preventative responsive policies which provide information to the households and communities about potential risks and how to face them. *If interventions adopt a new technology or technical assistance, including

[10] See Appendix L for a list of covariate or macro-level shocks/stressors included at the end of this document. Each of the intervention's domain, to be included, should have been designed to prepare, manage, or recover from one or more of those shocks/stressors.

Domain	Intervention	Level	Description
			renewable energy and energy efficiency, they will belong to the 'technology-based assistance' intervention category.
	Natural resource management	Meso, micro	Community-based natural resources management program that bring together the civil society and state actors to take care of a natural resource. *If interventions adopt a new technology or technical assistance, including renewable energy and energy efficiency, they will belong to the 'technology-based assistance' intervention category.
	Technology-based assistance	Macro, meso, micro	Providing technology-based materials to improve risk reduction. Examples include new technology-based crop failure safeguards, improved seeds (flood-, salt-, or temperature-tolerant), water purification/supply, water harvesting, recycling, drip irrigation, and water storage. This category also includes renewable energy and energy efficiency-focused materials
	Infrastructure (re)construction and maintenance	Macro, meso	This includes the construction, maintenance, and reconstruction of environmental infrastructure, including the reconstruction of market infrastructures (e.g., road to markets, and agricultural facilities) for post-disaster recovery.
	In-kind social assistance	Macro, meso, micro	Direct provision of goods or services, or subsidies to increase access. (E.g., in-kind transfer for social security, provision of non-food items, commodity vouchers, agriculture recovery and restoration programs). This does not include health insurance schemes.
	Food and nutrition interventions	Macro, meso, micro	Direct provision of food-focused goods or subsidies. Examples include commodity vouchers, food stamps, nutritional supplementation, and agricultural inputs (e.g., seeds, machine transfer). *If an intervention provides new technology-based materials (e.g., drought tolerant seed transfer), it belongs to the 'technology-based assistance' intervention category. *If an intervention is 'food/voucher for ''work'' (employing participants for public work and giving them food or voucher)', it applies to the 'employment assistance' intervention category.
	Employment	Macro,	Interventions providing cash or in-kind support

Domain	Intervention	Level	Description
	assistance	meso, micro	for employment or during unemployment. Examples include public works (e.g., cash for work, food for work, and vouchers for work), employment guarantee schemes, and unemployment assistance in the context of shocks/stressors.
	Local coordination mechanisms in support of service provision	Meso, micro	Activities/mechanisms that bring uncoordinated and disparate actors together to collaborate and integrate provision of resilience strengthening services for all or eligible populations. Examples: hotlines and referral systems (e.g., Link and Referral programs) that link vulnerable and/or refugee populations to different social protection providers to qualifying services they may not have been aware of (e.g., humanitarian assistance to social protection, social protection to other social protection providers, or specific policies/programs to others); policies or coordinating groups that bring together Ministries working on different issues affecting the same populations (e.g., Labor, Welfare or Social Security, Women and Children, Emergency Response). This does not include health and education services.
	Services Communication and advocacy	Macro, meso, micro	Communication, awareness-raising, dissemination, or public campaigns to increase knowledge of, access to, or uptake of social protection services. * If communication or awareness campaign relating to local opportunities, legal pathways, labor rights, etc., they will belong in the Safe and Orderly migration domain.
Build Safe Communities through Violence Prevention and Intervention	**Diversion to probation or appropriate services**	Macro, meso, micro	Arrest and pre-trial diversion programs that share the objective of diverting populations with mental health issues out of the criminal justice system and into behavioral healthcare and other more appropriate services.
	Psychosocial support and education programs	Macro, meso, micro	Targeting groups or individuals who are potentially vulnerable to engaging in crime, including in gangs, drugs, or gender-based violence or in other crime, with education interventions or school-based programming to promote alternatives to violence and crime or mental health and psychosocial support. For example, cognitive behavior therapy, anger replacement therapy and family counselling-

Domain	Intervention	Level	Description
			based initiatives.
	Preventative programs for ex-offenders	Macro, meso, micro	In-facility and out-of-facility rehabilitation. Interventions to support prisoners to integrate effectively. These may include vocational training, economic interventions such as employment training programs, life skills provision, or psycho-social support and may take place in order outside of correctional institutions.
	Social services for victims of crime and violence	Macro, meso, micro	The creation and resourcing of services and interventions that can provide crisis intervention, emergency treatment, and referrals for services (physical or mental support) to adult and child victims that have been referred by a relevant justice actor or institution. This could include court-ordered placement of children into social services or mental health support for crime victims referred by a Victims' Advice Bureau. The use and strengthening of approaches that engage the person involved in addressing the problems, specifically in relation to social care.
	Protection for at risk legal actors, political prisoners and witness protection services	Macro, meso, micro	Interventions that seek to protect either justice actors or justice seekers from harm that may be posed by their attempts to seek justice for themselves or others. Interventions to support the fair trial and safe treatment of political prisoners. Interventions to ensure that witnesses do not come to harm for their willingness to provide evidence.
	Society-led crime prevention and reporting initiatives	Meso, micro	Systems- or citizen-led interventions to support reporting and prevention of crime in their locality. Locally-led campaigns to promote anti-violence and anti-crime values, including anti-gender-based violence. Strengthening the ability of actors in no-legal services who come into contact with victims of crime and abuse to notice and report issues. For example, teachers are trained to recognize child abuse among pupils. Includes: Neighborhood watch schemes, School or community anti-crime or violence campaigns, and reporting and referral by non-legal service providers.
	Behaviour change communication	Macro, meso, micro	Communications to address harmful norms related to discrimination and violence (e.g. gender-based violence, stigmatization of health

Domain	Intervention	Level	Description
	against violence		conditions), and promote rights-affirming behaviors (e.g. willingness to report violence, treating people with respect). Activities may include: classes or workshops (e.g. on de-stigmatization of HIV), community mobilization activities (e.g. to create a concern to combat gender-based violence), and campaigns (e.g. using traditional and/or non-traditional media.
Orderly and safe migration management	**Information campaigns on legal rights, risks of irregular migration, legal alternatives, and/or working conditions**	Macro, meso, micro	Information on legal rights, workers' rights, and working conditions, such as visa recruitment processes, fees, indicators of abuse, exploitation, and/or contract violations; risks of irregular migration during the journey, return, or within destination country; or legal alternatives to irregular migration (local employment opportunities or legal pathways). Booklets, meetings, counselling, tours, mass media, posters, workshops and seminars might be used to disseminate the information.
	Legal pathways	Macro, micro	Creating/expanding legal migration pathways in receiving countries. This includes: *access to mobility channels* such as temporary, seasonal, sector, work- based visas, or long-term visas, humanitarian visas; or other incentives (travel subsidies). Could be delivered by governments or non-governmental organizations.
Expanding the EGM to cover external migration management			
External migration management	**Migration partnerships/deals**	Macro, meso	A bilateral or multilateral policy aiming to reduce irregular migration. These may encompass development assistance, return measures (forced or voluntary), trade, visa arrangements, political concessions, and security cooperation. Partnerships vary in scope and duration: some are framed as long-term cooperation to address systemic drivers, while others resemble short-term, transactional arrangements (often labelled externally as "deals"). Examples include the UK–France agreement and the EU–Turkey deal.[11] Code a study under this category if it assesses the effects of a partnership/deal as a whole. If

[11] In the EU-Turkey deal, the EU committed to resettling Syrian refugees from Turkey, easing visa restrictions for Turkish citizens, providing €6 billion in aid for Syrian migrant communities, updating the customs union, and renewing negotiations on Turkey's potential EU accession (European Council 2016).

Domain	Intervention	Level	Description
			a study evaluates a single component within the context of partnership/deal, code it under the relevant specific category instead.
	Externalization	Macro, meso	Shifting/delegating responsibility for migration control and border enforcement beyond a destination country to origin/transit countries (e.g. funding and/or capacity building border control, anti-smuggling enforcement, non-entrée cooperation, and establishing detention centers in third countries).
	Deterrence	Macro, meso, micro	Measures aimed at discouraging irregular migration through intimidation, restriction of rights, or increasing the perceived and actual risks of irregular migration. These activities include limiting due process for migrants, creating a climate of fear. The activities can occur within destination countries or be communicated externally in origin/transit countries. Examples include stricter domestic surveillance and monitoring within destination countries such as workplace raids/ID checks, detention policies, return policies, restrictions on employment opportunities, or the absence of social protections for irregular migrants in destination countries.
	Remote control	Macro, meso, micro	Pre-departure border enforcement activities through formal migration pathways. For instance, increased visa checks, pre-screening requirements, and carrier sanctions. It includes return measures as part of pre-entry mechanisms.
Multi-component			
Multi-component	**Multi-component**	Macro, meso, micro	This category captures interventions that combine two or more distinct activities/components that each fall under different categories of the framework (e.g., externalization and deterrence, or information campaigns alongside work policies). These interventions are implemented as integrated packages rather than as stand-alone measures. See Appendix D for dealing with multi-component interventions.

Outcomes of interest

For both EGM update and expansion, we include the same outcomes. The primary focus of this evidence gap map is to examine irregular migration outcomes. We will exclude intermediate development outcomes that may establish that an intervention is working (e.g., a program aiming to improve health conditions of individuals).

During a preliminary review of relevant impact evaluations from 3ie's Development Evidence Portal,[12] we found that studies examining observed migration behavior or "final outcomes" often did not test whether migration was internal versus international, or if the latter, occurring through regular versus irregular pathways. We will therefore include all migration outcomes when the type is unspecified (it is not clear if it is internal or international), as the outcome may be a proxy for international and/or irregular migration and to provide a bridge for developing state of the evaluative literature on migration programming. If studies explicitly discuss migration occurring due to reasons of forced international displacement, they will be coded as such.

The final outcomes will include indicators relevant to the micro and meso levels, such as "migration"; to the macro level, such as international migration stock and flows; and to all levels, such as reception of remittances, from internationally-based or geographically unspecified family members.

We also include some intermediate outcomes within the theory of change, which are not development outcomes, but rather directly related to migration aspirations and intentions and perceptions that are theorized as precursors (perceptions of current conditions, expectations). We include intention to migrate outcomes measured by any household members, in addition to individuals surveyed.

Table B2: Outcomes included in the map

Final outcomes- observed migration behavior		Definition
Any migration (micro)	Unspecified	The individual is no longer residing in their usual place of residence. Unspecified as to whether the study evaluates internal or international migration. Exclude if internal.
	International-unspecified	The number or rate of movement of persons (individuals/households) from their place of usual residence and across **international borders** to a country of which they are not nationals. Unspecified as to whether this movement is taking place outside the law, regulations, or international agreements governing the entry from transit or origin countries.
	International-regular	Movement of persons that occur **in compliance with** the laws of the country of

[12] https://developmentevidence.3ieimpact.org

Final outcomes- observed migration behavior		Definition
		origin, transit, or destination.
	International-irregular	Movement of persons that occurs **outside the laws, regulations, or international agreements** governing the entry into or exit from the state of origin, transit or destination.
	Forced displacement-unspecified	Movement due to persecution, conflict, violence, climate change, and human rights violations. Only code if unspecified as to whether internal or international; exclude if internal.
	Forced displacement-international	International movement due to persecution, conflict, violence, climate change, and human rights violations.
International migration flow (macro)	Unspecified	The number of international migrants arriving in a country (immigrants) or the number of international migrants departing from a country (emigrants) over the course of a specific period. Unspecified to whether it is regular or irregular.
	Regular	The number of international migrants arriving in a country (immigrants) or the number of international migrants departing from a country (emigrants) over the course of a specific period *through means that are **in compliance** with countries of origin, transit, and destination.*
	Irregular	The number of international migrants arriving in a country (immigrants) or the number of international migrants departing from a country (emigrants) over the course of a specific period *through mechanisms **outside of** the laws, regulations, and agreements governing entry/exit.*
International migration stock (macro)	Unspecified	The total number of international migrants present in a given country/area/region at a particular point in time who have ever changed their country of usual residence. Unspecified to whether it is regular or irregular.
	Regular	The total number of international migrants present in a country/area/region at a particular point in time who have changed their country of usual residence through

Final outcomes- observed migration behavior		Definition
		means *that are in compliance with countries of origin, transit, or destination.*
	Irregular	The total number of international migrants present in a country/area/region at a particular point in time who have changed their country of usual residence through mechanisms *outside of the laws, regulations, and agreements governing entry/exit.*
Intermediate outcomes		
Intention to migrate (micro)	Unspecified	Individual plans to move in the next 12 months. Unspecified as to whether it is in compliance or outside the laws, regulations, and agreements governing entry/exit of people.
	Regular	Individual plans to move in the next 12 months through means that *are in compliance with the laws, regulations, and agreements governing entry/exit in countries of origin, transit, or destination.*
	Irregular	Individual plans to move in the next 12 months through mechanisms *outside of the laws, regulations, and agreements governing entry/exit in countries of origin, transit, or destination.*
Knowledge, perceptions, attitudes, and Expectations (meso, macro)	Perception/psychosocial condition of current situation	The desire for change, feelings of inescapable stagnation, and challenges due to conditions that cannot be addressed. Only include if the study also examines another intermediate or final migration outcome; exclude if the outcome does not relate to migration.
	Expectations, awareness, knowledge, or attitudes on risks, benefits, costs, and/or consequences of movement through irregular channels	Outcomes relating to what is understood about the potential costs, benefits, and/or risks of irregular migration (e.g., physical risks or harm, expulsion, exploitation risks, labor opportunities, wages in destination countries, smuggling or recruiter fees).
	Knowledge or awareness of legal pathways, legalization processes, or asylum-seeking processes	Any knowledge about regular migration pathways (schemes, programs, processes, or other options).

Final outcomes- observed migration behavior		Definition
	Knowledge or awareness of migrant labor rights	Any knowledge of worker's rights. This may include those relating to labor or contract violations, labor exploitation, freedom from discrimination, freedom of movement.

Study Designs

For both updating the original EGM (Berretta et al. 2023) and expanding the EGM to cover external migration management, we will include studies that look at the migration impacts of interventions addressing root cause conditions and drivers that affect irregular migration. Specifically, we will include studies that adopt methods estimating the effects that can be attributed to an intervention, as compared to what would have happened in the absence of the intervention. We define the specific criteria required for inclusion below, drawing on commonly accepted standards for impact evaluations (Gertler *et al.* 2016) and systematic reviews (Waddington *et al.* 2012).

We will include both impact evaluations and systematic reviews:

- **Impact evaluation:** An impact evaluation is a study that uses a counterfactual to provide a quantitative estimate of the impact of an intervention. The counterfactual provides evidence about what would have happened in the absence of the intervention. The impact of a program is measured by comparing the outcomes of those who receive the intervention 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 a randomized control trial) 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.

  *For expanding the EGM to cover external migration management, we will also include comparative studies with or without a formal control group, and qualitative studies without a comparison group, if eligible for study design details below.

- **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. Systematic effectiveness reviews will be included if they describe the search, data collection and synthesis methods according to the 3ie database of systematic review 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 drivers of migration. If the review includes studies using multiple research designs, we will include these if at least 50 percent of studies use one impact evaluation design, as specified above.

Effectiveness studies

For both updating the original EGM (Berretta et al. 2023) and expanding the EGM to cover external migration management, 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 and therefore leads to its exclusion):

- Is the study primarily designed to determine to what extent a specific technique, technology, treatment, procedure or service works under ideal condition rather than attempt to answer a question relevant to the roll-out of a large program (i.e. lab-in-the field)?
- Is the intervention being carried out by the researchers themselves (e.g., by applying fertilizer in test plots to measure effects on plant growth), rather than 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 "basic science" research on biophysical mechanisms?

Study designs included

For both updating the original EGM (Berretta et al. 2023) and expanding the EGM to cover external migration management, we will include studies that implement at least one of the following study designs widely used to evaluate intervention effectiveness (Aloe *et al.* 2017; Reeves, Wells and Waddington 2017):

A) Prospective studies that allocate participants to treatment and control groups using random assignment or quasi-experimental methods:
   1. Randomized 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) Quasi-experimental designs where treatment arms are created without random assignment:
   1. Regression discontinuity designs (RDD), either sharp or fuzzy designs
   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)
   5. Interrupted time series (ITS) models, with or without a contemporaneous comparison group. 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

Note that 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.

For expanding the EGM to cover external migration management, we will also include qualitative evaluations that meet IOB criteria (Ministry of Foreign Affairs of the Netherlands 2024): Specifically, evaluations that use realist evaluation, process tracing, contribution analysis, contribution tracing, qualitative impact assessment protocol, general elimination methodology and qualitative comparative analysis. We will also include the outcome harvesting method, a participatory approach that identifies observed changes and then traces back to explore the program's possible contributions (Wilson-Grau, 2018; HM Treasury 2020). Qualitative studies that do not clearly identify the evaluation method used will be excluded. They should explicitly state in the title, abstract, or full text that they use one of these methodologies. Further details on each eligible qualitative methodology are provided below:

- **Realist Evaluation**
  Realist evaluations are based on the assumption that projects and programs work under certain conditions and are heavily influenced by how different stakeholders respond to them. Authors must clearly state a theory tested through an intervention, indicating how and for whom a program would work. They compare contexts, mechanisms, and outcomes within a program (not with a control). There is a strong emphasis on the social and historical context and the comparison of those who benefited from the program and those who did not (White and Phillips 2012). A realist evaluation is therefore not just designed to assess whether a development intervention worked. It addresses questions such as: *What works (or doesn't work)? For whom (and to what extent)? In which circumstances? How and why does it work?* (INTRAC 2017d).
- **Process Tracing**
  Develop a set of (competing) hypotheses linking an intervention to an outcome, including how these hypotheses could be (in)validated. Gather relevant evidence to determine which hypothesis most closely matches observed data. In its pure form, process tracing is based around a set of formal tests designed to assess causation. These are applied to all possible explanations for how a particular change might have occurred, to confirm or eliminate them. Within process tracing, these different explanations are called hypotheses (INTRAC 2017b).
- **Contribution Analysis**
  Contribution analysis is a methodology used to identify the contribution a development intervention has made to a change or set of changes. The aim is to produce a credible, evidence-based narrative based on a theory of change that a reasonable person would likely agree with, rather than to produce conclusive proof. It can be used during a development intervention, at the end, or afterwards (INTRAC 2017a).
- **Contribution Tracing**
  Contribution tracing is a participatory mixed-method (qual-quant) approach to establish the validity of contribution claims. It uses explicit criteria to guide evaluators in data collection and Bayesian updating to quantify the level of confidence in a claim. It includes a contribution 'trial' with all stakeholders to establish what will prove/disprove the claim (HM Treasury 2007).

- **Qualitative Impact Assessment Protocol (QuIP)**
  QuIP studies provide an independent reality check of a predetermined theory of change, helping stakeholders to assess, learn from, and demonstrate the social impact of their work. The QuIP gathers evidence of a project's impact through narrative causal statements collected directly from intended beneficiaries. Respondents are asked to talk about the main changes in their lives over a pre-defined recall period and to attribute these changes—often to multiple sources (Avard and Remnand 2017).
- **General Elimination Methodology (GEM)**
  Scriven's GEM (2008) builds upon his earlier Modus Operandi Method (1976) to substantiate causal claims. The methodology entails systematically identifying and then ruling out alternative causal explanations of observed results. It is based on drawing up Lists of Possible Causes (LOPCs) for an outcome and identifying each cause's "footprints" or *Modus Operandi*—the conditions or sequences of events that must be present when the cause is effective (Scriven 2008; White and Phillips 2012).
- **Qualitative Comparative Analysis (QCA)**
  QCA enables the analysis of multiple cases in complex situations and can help explain why change happens in some cases but not in others. Designed for an intermediate number of cases (typically 10–50), it is useful when there are too few cases for conventional statistical analysis (INTRAC 2017c).
- **Outcome Harvesting**
  Outcome harvesting is designed to collect evidence of change (the 'outcomes') and then work backwards to assess whether or how an organization, program, or project contributed to that change. Outcomes are defined as changes in the "behavior writ large" (e.g. actions, relationships, policies, practices) of one or more social actors influenced by an intervention (Wilson-Grau 2018; HM Treasury 2020). To strengthen the credibility of outcome harvesting, evaluators should independently verify reported outcomes and explicitly test alternative causal explanations (Sharma Waddington et al. 2023).

Additional studies that will be excluded

- Before-after studies without a comparison group or cross-sectional studies that do not attempt to control for selection bias or confounding.
- Studies that only examine willingness-to-pay for goods, services, process and business models.

Other eligibility criteria

For updating the original EGM (Berretta et al. 2023)

- **Language:** We will include studies published in any language, although the search terms used will be in English only.
- **Publication date**: For this EGM update, we will include studies published from 2023 to August 2025.[13]
- **Status of studies:** We will include ongoing and completed impact evaluations and

[13] The original 3ie EGM (Berretta et al. 2023) covered studies published between 1990 and early 2023, based on a search conducted in April 2023 that included all studies published from 1990 up to that point.

systematic reviews, both peer-reviewed studies and 'grey' literature. For on-going studies, we will include prospective study records, protocols and trial registrations. Providing an indication of the prevalence and characteristics of on-going evaluation evidence is expected to enrich the analysis of current evidence gaps and support decision making in relation to evidence generation.

For expanding the EGM to cover external migration management

- **Language and status of studies:** We will include studies published in any language, although the search terms used will be in English only for academic database and in English and French for grey literature searches.
- **Publication date**: For this EGM expansion, we will include studies published from 2015.
- **Status of studies:** Same as the update component above.

References used in this appendix

Aloe, A. M., B. J. Becker, M. Duvendack, J. C. Valentine, I. Shemilt, and H. Waddington. 2017. "Quasi-experimental study designs series—Paper 9: Collecting data from quasi-experimental studies." *Journal of Clinical Epidemiology* 89: 77–83. https://doi.org/10.1016/j.jclinepi.2017.02.013.

Avard, R., and F. Remnand. 2017. "Qualitative Impact Assessment Protocol (QuIP)." https://www.betterevaluation.org/en/plan/approach/QUIP.

Berretta, M., M. D. A. Anda León, C. Huang, and S. Shisler. 2023a. *Addressing root causes and drivers of irregular migration: An evidence gap map*. International Initiative for Impact Evaluation (3ie). https://doi.org/10.23846/EGM030.

European Council. 'EU–Turkey Statement, 18 March 2016', 2016. https://www.consilium.europa.eu/en/press/press-releases/2016/03/18/eu-turkey-statement/.

Freier, L. F., and K. Holloway. 2019. "The impact of tourist visas on intercontinental South–South migration: Ecuador's policy of 'open doors' as a quasi-experiment." *International Migration Review* 53 (4): 1171–1208. https://doi.org/10.1177/0197918318801068.

Gertler, P. J., S. Martinez, P. Premand, L. B. Rawlings, and C. M. J. Vermeersch. 2016. *Impact evaluation in practice*. 2nd ed. Washington, DC: Inter-American Development Bank and World Bank. https://doi.org/10.1596/978-1-4648-0779-4.

HM Treasury. 2020. *Supplementary Guide: Handling Complexity in Policy Evaluation (Magenta Book 2020).* London: HM Treasury. https://www.gov.uk/official-documents.

———. 2007. *The Magenta Book: Guidance Notes for Policy Evaluation and Analysis. Magenta Book Background Papers.* London: HM Treasury.

IOM GMDAC. 2021. *Migration drivers*. https://www.migrationdataportal.org/themes/migration-drivers.

INTRAC. 2017a. "Contribution Analysis." https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Contribution-analysis.pdf.

INTRAC. 2017b. “Process Tracing.” https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Process-tracing.pdf.

INTRAC. 2017c. “Qualitative Comparative Analysis.” https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Qualitative-comparative-analysis.pdf.

INTRAC. 2017d. “Realist Evaluation.” https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Realist-evaluation.pdf.

Reeves, B. C., G. A. Wells, and H. 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: 30–42. https://doi.org/10.1016/j.jclinepi.2017.02.016.

Sharma Waddington, H., H. Umezawa, and H. White. 2025. ‘What Can We Learn From Qualitative Impact Evaluations About the Effectiveness of Lobby and Advocacy? A Meta-Evaluation of Dutch Aid Programmes and Assessment Tool’. *Evaluation Review*, 11, 0193841X251314731. https://doi.org/10.1177/0193841X251314731.

Scriven, M. 1976. “Maximizing the Power of Causal Investigations: The Modus Operandi Method.” *Evaluation Studies Review Annual* 1: 101–118.

Scriven, M. 2008. “A Summative Evaluation of RCT Methodology: An Alternative Approach to Causal Research.” *Journal of Multidisciplinary Evaluation* 5 (9): 11–24.

Snilstveit, B., M. Vojtkova, A. Bhavsar, J. Stevenson, and M. Gaarder. 2016. “Evidence & gap maps: A tool for promoting evidence informed policy and strategic research agendas.” *Journal of Clinical Epidemiology* 79: 120–29. https://doi.org/10.1016/j.jclinepi.2016.05.015.

Waddington, H., H. White, B. Snilstveit, J. G. Hombrados, M. Vojtkova, P. Davies, A. 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.

White, H., and D. Phillips. 2012. “Addressing Attribution of Cause and Effect in Small N Impact Evaluations: Towards an Integrated Framework.” 3ie Working Paper 15. https://www.3ieimpact.org/evidence-hub/publications/working-papers/addressing-attribution-cause-and-effect-small-n-impact.

Wilson-Grau, R. 2018. *Outcome Harvesting: Principles, Steps, and Evaluation Applications*. Charlotte, NC: Information Age Publishing.

Appendix C: Search strategy

We will adopt a systematic search strategy following guidelines for systematic literature searching (Kugley et al. 2017). The strategy will be 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.

Two of the framework domains, strengthening resilience and building safe communities, leverage pre-existing EGMs such as *mapping evidence of what works to strengthen resilience to shocks and stressors (*Berretta et al. 2023) and *the effects of rule of law interventions on justice outcomes: an evidence gap map* (Doherty et al. 2020). We will re-screen all included studies and all studies that were excluded due to irrelevance of outcomes. We will run an updated search strategy from the date of the previous search (April 2023) to date (August 2025) and screen the results against the eligibility criteria of this EGM.

For the other two intervention domains, economic opportunities and orderly and safe migration management, we have developed new search strings, and given the nature of interventions within those domains, reported changes in outcomes are expected to occur in a number of development sectors. As such, the strategy will consider sector specific databases where appropriate, as indicated in below in this section. Finally, where possible, the review team will contact key experts and organizations through our advisory group (presented in Appendix A) to identify additional studies that meet the inclusion criteria.

The precise strings and logic (e.g., index terms and truncation operators) will be adapted for each database and platform.

For the domain on resilience to climate change, from which we have taken numerous categories from the Mapping evidence of what works to strengthen resilience to shocks and stressors (Berretta et al. 2023), the following databases will be searched:

- CAB Abstracts (EBSCO)
- CAB Global Health (OVID)
- Africa-Wide (EBSCO)
- Academic Search Complete (EBSCO)
- APA PsycInfo (OVID)
- Web of Science (SSCI)
- Econlit (EBSCO)
- Social Science Research Network (SSRN)
- World Bank (EBSCO Discovery)
- Agris (EBSCO Discovery)
- RePEc (EBSCO Discovery)
- Campbell library

For the domain on violence prevention, from which we have taken numerous categories from The effects of rule of law interventions on justice outcomes: an evidence gap map (Doherty et al. 2020) the following databases will be searched:

- Scopus
- Social Science Citations Index

- International Political Science Abstracts
- Communication and Mass Media Complete
- Research Papers in Economics (RePEc)

For the two domains on Economic opportunities and Orderly and safe migration management we will search the following databases:

- Scopus
- Social Science Citations Index
- International Political Science Abstracts
- Research Papers in Economics (RePEc)
- CAB Abstracts
- Africa-Wide
- Academic Search Complete
- Web of Science
- Econlit
- Social Science Research Network (SSRN)
- World Bank
- Campbell library

For the external migration management intervention domain (the EGM expansion), we will search the same or most of the databases listed above.

Also, for the EGM expansion, we will search for grey literature on the websites of organizations relevant to our project. We selected these organizations in consultations with IOB, based on their action and work in migration-related matters such as the International Organization of Migration (IOM), Center for Global Development, European Council on Refugees and Exiles (ECRE), among others. We will also search on other websites from referential international development and research organizations, including Amnesty International and Carnegie Endowment for International Peace. An indicative list of organizations and websites are presented below. If time permits, we will also conduct the forward and backward citation tracking of all the included studies to reduce the risk of missing relevant studies.

Table C1: Indicative list of organizations websites for grey literature search

Website name	Link
Amnesty International	https://www.amnesty.org/
Carnegie Endowment for International Peace	https://carnegieendowment.org/
Centre for Global Development	https://www.cgdev.org/
Chatham House (The Royal Institute of International Affairs)	https://www.chathamhouse.org/
Clingendael Institute	https://www.clingendael.org/
CONCORD	https://concordeurope.org/
DGAP	https://dgap.org/
ECDPM (European Centre for Development Policy Management)	https://ecdpm.org/
EPC (European Policy Center)	https://www.epc.eu/

Website name	Link
Euromesco	https://www.euromesco.net/
ECRE (European Council on Refugees and Exiles)	https://ecre.org/
Friedrich-Ebert-Stiftung	https://www.fes.de/en/
Forced Migration Review	https://www.fmreview.org/
GIGA (German Institute for Global and Area Studies)	https://www.giga-hamburg.de/en
Global Asylum Governance and the European Union's Role	https://www.asileproject.eu/
GPPi (Global Public Policy Institute)	https://gppi.net/
Heinrich Böll Stiftung	https://www.boell.de/en
Human Rights Watch	https://www.hrw.org/
Istituto Affari Internazionali (IAI)	https://www.iai.it/en
IEMed (European Institute of the Mediterranean)	https://www.iemed.org/
International Migration Institute	https://www.migrationinstitute.org/
International Refugee Rights Association	https://www.umhd.org.tr/en/
ISPI (Italian Institute for International Political Studies)	http://www.ispionline.it/
Maastricht Centre for Citizenship, Migration and Development	https://macimide.maastrichtuniversity.nl/
MPI (Migration Policy Institute)	https://www.migrationpolicy.org/
Mixed Migration Centre	https://mixedmigration.org/
Mirekoc	https://mirekoc.ku.edu.tr/
Oxfam Policy and Practice	https://policy-practice.oxfam.org/
ODI (Overseas Development Institute)	https://odi.org/
Refugee Studies Centre (RSC), University of Oxford	https://www.rsc.ox.ac.uk/
SWP (Stiftung Wissenschaft und Politik)	https://www.swp-berlin.org/
Statewatch	https://www.statewatch.org/
Tunisian Forum for Economic and Social Rights (FTDES)	http://ftdes.net/
UN Network on Migration	http://migrationnetwork.un.org/
International Organization for Migration (IOM)	https://www.iom.int/
ICMPD	https://www.icmpd.org/
UNHCR	https://www.unhcr.org/

References used in this appendix

Berretta, M., S. Lee, M. Kupfer, C. Huang, W. Ridlehoover, D. Frey, F. Ahmed, et al. 2023. *Strengthening resilience against shocks, stressors and recurring crises in low- and middle-income countries: An evidence gap map*. International Initiative for Impact Evaluation (3ie), 10 May. https://doi.org/10.23846/EGM023.

Doherty, J., A. Sonnenfeld, D. Glandon, T. Kozakiewicz, B. Snilstveit, D. Sabet, R. Zelfo, and

M. Malik. 2020. *Protocol: The effects of rule of law interventions on justice outcomes: An evidence gap map*. New Delhi: International Initiative for Impact Evaluation (3ie). https://www.3ieimpact.org/sites/default/files/2020-10/Protocol-for-Rule-ofLaw-EGM.pdf.

Kugley, S., A. Wade, J. Thomas, Q. Mahood, A.-M. K. Jørgensen, K. Hammerstrøm, and N. 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.

Appendix D: Screening, data extraction, and critical appraisal process

Screening

This subsection provides an overview of the processes we will adopt to systematically screen, critically appraise and extract data from studies identified by the search.

The selection of studies for data extraction as part of the map will be managed using EPPI-Reviewer 4 software (Thomas et al. 2023), developed by the Evidence for Policy and Practice Information and Co-ordinating Centre, 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.
- **Title and abstract screening:** We will implement a structured training and consistency-check process for screeners; we will conduct title/abstract screening independently by two reviewers until they reach a minimum inter-rater reliability of 85%, using predefined eligibility criteria. Several exclude codes will be available to provide more information on the reasons for exclusion 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 (screening codes are available below). 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.
  We will use a machine learning classifier trained on data from 3ie's Development Evidence Portal repository of impact evaluations and systematic reviews of social and economic interventions, regardless of sector. The classifier model will attempt to replicate human labelling decisions such as whether the study mentions an intervention, utilizes a relevant study design (impact evaluations or systematic reviews), and whether the intervention was conducted in a low- or middle-income country where applicable. The model is built on the studies included and excluded in the DEP, drawing on a large sample (more than 80,000 samples). The model learned from the inclusion and exclusion decision taken during the DEP screening in relation to the criteria explained above, and showed a precision (0.70) and recall (0.8) higher than other DEP models run in 2022 and 2021. This "generic" classifier model will be applied to our search results to rank studies by the likelihood of relevance to our inclusion criteria, with a percentage ranking 0-100%. We will direct resources to screening highest likelihood studies first and proceed in descending priority rank order. We will automatically exclude studies which ranks below the 20%, indicating a low possibility of being included. Before doing so, we will check a random sample of them to ensure they are irrelevant. In case there seems to be still relevant studies among them, we will screen another portion until it appears they are all not includable studies. The output of this process will be a set of screened studies that have been put forward for full text screening.
- **Full-text screening:** We will retrieve the full text manuscript for each study that meet all the T/A inclusion criteria. Two reviewers will examine each full text in detail

against the protocol. Again, we will apply a code to each study that reflects either that the study is included, or why the study is excluded. The output of this stage will be a set of studies deemed suitable to include in the EGM.

- **Checks for linked publications:** The project team will attempt to group publications that focus on the same intervention and study population (i.e., publications that report on the same study). This typically occurs in cases where an author group publishes more than one paper in relation to one particular study on a specific population. 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 approach.

Title and abstract screening codes:

- EXCLUDE – Publication year.
- EXCLUDE – No intervention.
- EXCLUDE – Lab/efficacy.
- EXCLUDE – Not a quantitative effectiveness study (skip this code for the EGM expansion component studies)
- EXCLUDE – Not a qualitative effectiveness study (only for the EGM expansion component studies)
- EXCLUDE – High-income country.
- EXCLUDE – L&MICs not in SSA/MENA, nor Afghanistan, Pakistan, or Türkiye (only for the EGM expansion component studies)
- EXCLUDE – Not a systematic review.
- EXCLUDE – Intervention not relevant.
- INCLUDE

Full text screening codes:

- EXCLUDE – Publication year.
- EXCLUDE – No intervention.
- EXCLUDE – Lab/efficacy.
- EXCLUDE – Cost analysis only.
- EXCLUDE – Not a quantitative effectiveness study (skip this code for the EGM expansion component studies)
- EXCLUDE – Not a qualitative effectiveness study (only for the EGM expansion component studies)
- EXCLUDE – High-income country.
- EXCLUDE – L&MICs not in SSA/MENA, nor Afghanistan, Pakistan, or Türkiye (only for the EGM expansion component studies)
- EXCLUDE – No valid causal inference.
- EXCLUDE – Insufficient clusters.
- EXCLUDE – Not a systematic review.
- EXCLUDE – Intervention not relevant.
- EXCLUDE – Outcomes not relevant.
- INCLUDE

Data extraction and critical appraisal

We will systematically extract data from all included studies using the data extraction tool available in Appendix E. The data will cover the following broad areas:

- **Basic study and publication information:** This coding will focus on capturing the general characteristics of the study including authors, publication date and status, study location, intervention type, outcomes reported, definition of outcome measures, population of interest, study and program funders, time periods for delivery and analysis;
- **Topical cross-cutting issues:** We will extract data on a number of cross-cutting issues, including equity, targeted population (e.g indigenous people).
- **Critical appraisal:** All included systematic reviews will be critically appraised following the practices adopted by the 3ie systematic review database protocol, which draws on Lewin et al. (2009). This appraisal assesses systematic reviews according to criteria relating to the search, screening, data extraction, and synthesis activities conducted, and covers all 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 et al. (2017). The tool used for this process is presented in Appendix F. We will not critically appraise impact evaluations, as this is typically beyond the scope of EGMs.

The following processes will be implemented to collect this information:

- **Develop and refine data extraction tools and codebooks:** The draft tools developed for this project will be reviewed and potentially refined in light of any feedback received by the EGM advisory group and lessons 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. Each coding group will all 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 one coder. 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:** Once the data extraction is near completion, the project team will check all extracted data. In practice, a member of the core team will check the consistency of data extracted and measures of consistency will be calculated and used to inform the checking process.

Dealing with multicomponent interventions

Depending on the number and nature of multi-component interventions included, the project team will adopt a consistent approach to coding these in the map. The decision will be taken based on how many similar combinations of intervention categories or domains there are, and consequently, the most appropriate option to represent those studies will be chosen. This approach may be (i) to determine the main intervention of focus in the study and

grouping the study with others that focus on that main component, (ii) grouping all multicomponent studies together in a 'package', (iii) grouping studies by packages of interventions domains (rather than interventions categories) or (iiii) a combination of those approaches. The approach adopted and the associated limitations will be clearly stated in the final report.

References used in this appendix

Lewin, S., A. D. Oxman, J. N. Lavis, and A. 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.

Shemilt, I., N. Khan, S. Park, and J. 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.

Snilstveit, B., R. Bhatia, K. Rankin, and B. Leach. 2017. *3ie evidence gap maps: A starting point for strategic evidence production and use*. New Delhi: International Initiative for Impact Evaluation (3ie). https://www.3ieimpact.org/sites/default/files/2019-01/3ie_evidence_gap_maps.pdf.

Thomas, J., S. Graziosi, J. Brunton, Z. Ghouze, P. O'Driscoll, M. Bond, and A. Koryakina. 2023. *EPPI-Reviewer: Advanced software for systematic reviews, maps and evidence synthesis*. EPPI Centre, UCL Social Research Institute, University College London. https://eppi.ioe.ac.uk/cms/About/AboutEPPIReviewer/tabid/2967/Default.aspx.

Appendix E: EGM data extraction template

Code	Subcode
Study Information	Study EPPI internal ID
	Coder name
	Title name
	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

Code	Subcode
	Additional country
	Country income level
	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 3 different treatment options in case there is more than one intervention group.
Outcome Information	Outcome
	Outcome description

Appendix F: Critical appraisal tool for SRs included in the EGM

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
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)
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

Question	Criteria
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
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
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

Question	Criteria
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... 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.	

Appendix G: Provisional data extraction form for evidence summaries

Variable group	Variable	Description
Publication Information	Study ID	The unique ID code that is assigned to each included study
	Estimate ID	The unique ID code that assigned to each individual estimate
	Study status	Select one of the following: i) Completed; ii) Protocol; iii) Ongoing
	Author Name	Authors last names [Open Answer]
	Year of Publication	Year published (publication date, not preprint or first online publication dates)
Intervention Information	Intervention code	Choose one or more intervention code(s) for each corresponding effect size from the provided intervention table.
	Country	Country of intervention
	Exposure to intervention (in months)	For how long are the observations exposed to the intervention?
	Evaluation period (in months)	The total number of months elapsed between the end of an intervention and the point at which an outcome measure is taken post intervention, or as a follow-up measurement. If less than one month, use decimals (e.g., measurement immediately after the intervention end would be coded as 0, one week would be .25, etc.)
	Intervention description	Provide detailed description of the intervention and its different components such that a reader could easily understand what happened. Include page numbers for quick reference. If two or more interventions are being evaluated, please provide descriptions for each intervention arm under separate rows.
Method information	Cost	Report any cost data provided or comments on cost effectiveness, include the authors' comments on cost data even if quantifications are not provided. Provide details of what the cost relates to or how they have been calculated if possible. Include any information identified from cited documents or linked studies.
	Evaluation Design	Select one of the options below: 1. Experimental (defined as prospective randomised assignment, where randomisation is implemented by researchers (or by decision makers in the context of an evaluation study) 2. Quasi-experimental (including natural experiments and non-randomised studies).

Variable group	Variable	Description
	Evaluation Method	If Experimental, then select: Randomised controlled trial If Quasi-experiment or natural experiment, then select: Natural experiment in which exposure to treatment is random Regression Discontinuity Design (RDD) Difference-in-Differences (DID) / Fixed effects estimation Instrumental variable (IV) estimation Endogenous treatment-effects models (including endogenous switching regression, and other methods synonymous to the Heckman two step model) Statistical matching (includes PSM or statistical weighting) Interrupted time series (ITS) Synthetic controls
	Additional Methods	Select additional method if any. If none, select not applicable. [Open Answer]
Estimate Information	Analysis type for this effect size	Free text, what type of analysis was used (Regression, 2SLS, ANCOVA, etc.)
	Estimate Type	Type of data for this effect size: 1 = Continuous - means and SDs, 2 = Continuous - mean difference and SD, 3 = Dichotomous outcome - proportions, 4 = Regression data - dichotomous outcome, 5 = Regression data - continuous outcome
	Treatment Effect	1=Intention to Treat (ITT), 2=Average Treatment Effect on the Treated (ATET), 3=Average Treatment Effect (ATE) 4 = Local Average Treatment Effect (LATE)
	Unit of analysis	What is the unit of analysis? UOA for this effect size: 1= Individual, 2= Household, 3= Group (e.g., community organisation), 4= Village, 5 = Other, 6 = Not clear
	Source	Note the page number, table number, column, and row you used to extract the data [Open Answer]
Treatment variable information	Treatment	Record the treatment variable as written in the model (e.g., the variable name the author uses, such as ("Intervention x Time") [Open Answer]
	Treatment type	Describe the types of treatment variable used: i) binary; ii) continuous; iii) categorical; iv) other
	Comparison	1=No intervention (service delivery as usual), 2=Other intervention, 3=Pipeline (waitlist) control

Variable group	Variable	Description
	Describe Comparison Group	Describe the comparison group [Open Answer]
	Subgroup	Is this analysis of a subgroup or estimating heterogeneous effects? 0=no, 1=yes
	Subgroup information	Describe the subgroup or variable interacted with the treatment variable if applicable (e.g., boys, girls). If no subgroup, select not applicable [Open Answer]
Outcome Information	Outcome description	Record the outcome for the corresponding effect size. Use this open answer field to enter, in the author's own words, a description of the outcome. Be selective and concise with the excerpts being transcribed here as to ensure accurate and precise descriptions of the outcome. To the extent possible, be sure to include numbers, units, population, and comparators. Include page numbers with every excerpt extracted.
	Outcome codes	Choose an outcome code for each corresponding effect size from the provided outcome table.
	Outcome sub-group	Choose an outcome sub-group code for each corresponding effect size from the provided outcome table.
	Outcome description	Record the outcome for the corresponding effect size. Use this open answer field to enter, in the author's own words, a description of the outcome. Be selective and concise with the excerpts being transcribed here as to ensure accurate and precise descriptions of the outcome. To the extent possible, be sure to include numbers, units, population, and comparators. Include page numbers with every excerpt extracted.
	Post-intervention or change from baseline?	0 = Post-intervention, 1 = Change from baseline
Estimate data	Mean treatment	Outcome mean for the treatment group
	SD treatment	Outcome standard deviation for treatment group
	Mean Control	Outcome mean for the comparison group
	SD Control	Outcome standard deviation for control group
	Mean difference	Overall mean difference (treatment - control)
	SE difference	Standard error of the overall mean difference
	Tstat difference	t-statistic of mean difference
	p-value difference	p-value of mean difference
	Odds ratio	Odds ratio reported in the study

Variable group	Variable	Description
	SE odds ratio	Odds ratio standard error reported in the study
	Risk ratio	Risk ratio reported in study
	SE risk ratio	Risk ratio standard error
	Coeff reg	Report the regression coefficient of the treatment effect
	SE reg	Report the associated standard error of the regression coefficient.
	Tstat reg	Report the associated t statistic of the effect size (coefficient/SE)
	CI_LB reg	Report the associated Lower bound of the 95% Confidence interval of the effect size. If CI is reported for a different confidence level, indicate that in the notes section.
	CI_UP reg	Report the associated Upper bound of the 95% Confidence interval of the effect size. If CI is reported for a different confidence level, indicate that in the notes section.
	P value exact	Exact p value if given, if not, record as written in the manuscript (e.g., p < .001, or p > .05)
	Clusters treatment	Number of clusters - treatment group
	Clusters control	Number of clusters - control group
	Clusters total	Number of clusters - total sample
	N treatment	Sample size - treatment group
	N control	Sample size - control group
	N total	Sample size - total sample
	periods (1 if cross sectional)	Record how many time-period there are in the evaluation (e.g., cross section is 1, panel data with 3 measurements is 3)
	Does the sample size need to be corrected?	Often in panel data, models will report number of observations rather than number of participants. In this column you will indicate 1="Yes" if the sample size needs to be divided by the number of periods, and 0="No" if either it is cross-sectional data, or if the authors have already divided the number of observations by the number of panel assessments and thus no correction is necessary.

Appendix H: Calculating standardized effects

An effect size expresses the magnitude (or strength) and direction of the relationship of interest (Valentine et al. 2015; Borenstein et al. 2021). We will extract data from each individual study to calculate standardised effect sizes for cross-study comparison wherever possible. For continuous outcomes comparing group means in a treatment and control group, we will calculate the standardised mean difference (SMDs), or Cohen's d, its variance and standard error using formulae provided in Borenstein et al. (2021). 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:

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

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 et al. (2016):

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

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 :

$$d = LogOddsRatio * \frac{\sqrt{3}}{\pi}$$

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 standardised 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 percent 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.

References used in this appendix

Borenstein, M., L. V. Higgins, J. P. T., Hedges, and H. R. Rothstein. 2021. *Introduction to meta-analysis.* Chichester, UK: Wiley.

Ellis, P. D. 2010. *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511761676.

Keef, S. P., and L. 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.

Polanin, J. R., E. E. Tanner-Smith, and E. 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.

Sánchez-Meca, J., F. Marín-Martínez, and S. 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.

Valentine, J. C., A. M. Aloe, and T. 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.

Appendix I: Criteria determining selection of effect estimates for data extraction

We will extract effects reported across different interventions, outcomes and subgroups within a study. We will address dependent effect sizes using data processing and selection techniques. We will utilise several criteria to select one effect estimate per outcome per study:

- Where studies report effects from multiple estimators, we will use the author's preferred specification. If no preference is indicated, we will select the estimate based on a model specification with the most controls. If data is available, we will use the specification that appears most robust to falsification tests (e.g. according to sensitivity analysis for propensity score matching or placebo tests for difference-in-difference estimators).
- Where different studies report on the same program 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.
- Where studies report evidence according to subgroups of participants, we will record and report data on relevant subgroups separately.
- For studies with outcome measures at different time points, we will synthesise short- and long-term outcomes separately, following De La Rue et al. (2013).
- When a study uses multiple outcome measures for a particular construct, we will use the measure that appears to most accurately capture the construct and without regard to the results reported. In these cases, if authors do not present an effect for the full sample, we may calculate a “synthetic effect size” using the sample-weighted average and applying appropriate formulae to recalculate variances (Borenstein et al. 2021).
- 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 them in separate meta-analyses according to the intervention type. Where multiple treatment arms represent the same treatment construct, we may calculate a “synthetic effect size”.
- Our analysis will prioritise synthesising outcomes using composite or aggregate indicators. If a study does not report a composite measure, we will use the outcome that most closely relates to the intervention type or perform outcome mapping to identify the outcome in each study that appears most frequently across studies.

References used in this appendix

Borenstein, M., L. V. Higgins, J. P. T., Hedges, and H. R. Rothstein. 2021. *Introduction to meta-analysis.* Chichester, UK: Wiley.

Appendix J: CAMELOT template for quality appraisal of qualitative evaluations

We will use the *Critical Appraisal for Methodological Limitations of Qualitative Research Tool (CAMELOT)* (Munthe-Kaas et al. 2024a). CAMELOT is designed specifically for use in qualitative evidence syntheses and supports the assessment of how study conduct and reporting may influence confidence in review findings across 12 domains:

META domains

- Research aim and question(s) – Purpose of the study and the questions being explored.
- Stakeholders – Individuals or groups with an interest in the study's findings (not the same as research participants).
- Researchers – Investigators conducting the study and their relationship to the topic, context, or participants.
- Context – The setting (local, national, or international) in which the study took place.

METHOD – Research design domains

- Research strategy – The overarching plan or approach for carrying out the study.
- Ethical considerations – How ethical principles were integrated into the study's design and conduct.
- Equity, diversity, and inclusion considerations – How equity, diversity, and inclusion were addressed in representation, participation, and research processes.
- Theory – Conceptual frameworks or systems used to explain or understand the phenomenon.

METHOD - Research conduct domains

- Participant recruitment and selection – How participants were identified, recruited, and chosen.
- Data collection – How qualitative information was gathered from participants or observations.
- Analysis and interpretation – How data was examined to identify themes, patterns, and insights.
- Presentation of findings – How results were organized, communicated, and aligned with the data.

CAMELOT does not calculate overall scores such as "high," "medium," or "low" quality. Instead, it supports reviewers in considering how well the Method and Meta domains fit together, and then summarizing methodological limitations as no or minimal concerns, minor concerns, moderate concerns, or serious concerns.

The template is as follows:

STUDY ID		
META domains	**Data extracted from primary study**	***Optional comments (notes to self, including any problems or missing information)***
Research aim & question(s)		
Stakeholders		
Researchers		
Context		
METHOD domains		
Research design domains		
	Data extracted from primary study	***Optional comments (notes to self, including any problems or missing information)***
Research strategy		
Ethical considerations		
Equity, diversity & inclusion considerations		
Theory		
Research conduct domains		
	Data extracted from primary study	***Optional comments (notes to self, including any problems or missing information)***
Participant recruitment & selection		
Data collection		
Analysis and interpretation		

Presentation of findings									
Describe if you have any concerns about the fit between the following domains	**Fit between Research design domains and Research aim & question**	**Fit between Research design domains and Stakeholders**	**Fit between Research design domains and Researchers**	**Fit between Research design domains and Context**	**Fit between Research conduct domains and Research aim & question(s)**	**Fit between Research conduct domains and Stakeholders**	**Fit between Research conduct domains and Researchers**	**Fit between Research conduct domains and Context**	**Fit between Research design domains and Research conduct domains**
Indicate concerns regarding fit using Serious, Moderate, Minor, No or minimal or Unclear:									
OVERALL ASSESSMENT OF LIMITATIONS:(No or minimal, minor, moderate, serious)									
Explanation for overall assessment									

Note: This follows the template of Munthe-Kaas et al. (2024b) (accessed on August, 2025).

References used in this appendix

Munthe-Kaas, H. M., A. Booth, I. Sommer, S. Cooper, R. Garside, K. Hannes, J. Noyes, and The CAMELOT Development Group. 2024a. "Developing CAMELOT for assessing methodological limitations of qualitative research for inclusion in qualitative evidence syntheses." *Cochrane Evidence Synthesis and Methods* 2 (6): e12058. https://doi.org/10.1002/cesm.12058.

Munthe-Kaas, H. M., A. Booth, I. Sommer, S. Cooper, R. Garside, K. Hannes, and J. Noyes. 2024b. "Appendix 4 & 5. CAMELOT primary study table and CAMELOT guidance." *Zenodo*, 4 November. https://doi.org/10.5281/ZENODO.14035136.

Appendix K: World Bank Classification of Countries

LOW-INCOME ECONOMIES ($1,135 OR LESS)

Afghanistan	Korea, Dem. People's Rep	Somalia
Burkina Faso	Liberia	South Sudan
Burundi	Madagascar	Sudan
Central African Republic	Malawi	Syrian Arab Republic
Chad	Mali	Togo
Congo, Dem. Rep	Mozambique	Uganda
Eritrea	Niger	Yemen, Rep.
Gambia, The	Rwanda	
Guinea-Bissau	Sierra Leone	
Afghanistan	Korea, Dem. People's Rep	

LOWER-MIDDLE INCOME ECONOMIES ($1,136 TO $4,495)

Angola	India	Papua New Guinea
Bangladesh	Jordan	Philippines
Benin	Kenya	São Tomé and Principe
Bhutan	Kiribati	Senegal
Bolivia	Kyrgyz Republic	Solomon Islands
Cambodia	Lao PDR	Sri Lanka
Cameroon	Lebanon	Tajikistan
Comoros	Lesotho	Tanzania
Congo, Rep.	Mauritania	Timor-Leste
Côte d'Ivoire	Micronesia, Fed. Sts.	Tunisia
Djibouti	Morocco	Uzbekistan
Egypt, Arab Rep.	Myanmar	Vanuatu
Eswatini	Namibia	Viet Nam
Ghana	Nepal	West Bank and Gaza
Guinea	Nicaragua	Zambia
Haiti	Nigeria	Zimbabwe
Honduras	Pakistan	
Angola	India	
Bangladesh	Jordan	

UPPER-MIDDLE-INCOME ECONOMIES ($4,496 TO $13,935)

Albania	Equatorial Guinea	Moldova
Algeria	Fiji	Mongolia
Argentina	Gabon	Montenegro

Armenia	Georgia	North Macedonia
Azerbaijan	Grenada	Paraguay
Belarus	Guatemala	Peru
Belize	Indonesia	**Samoa**
Bosnia and Herzegovina	Iran, Islamic Rep.	Serbia
Botswana	Iraq	South Africa
Brazil	Jamaica	St. Lucia
Cabo Verde	Kazakhstan	St. Vincent and the Grenadines
China	Kosovo	Suriname
Colombia	Libya	Thailand
Cuba	Malaysia	Tonga
Dominica	Maldives	Türkiye
Dominican Republic	Marshall Islands	Turkmenistan
Ecuador	Mauritius	Tuvalu
El Salvador	Mexico	Ukraine
Albania	Equatorial Guinea	

HIGH-INCOME ECONOMIES ($13,935 OR MORE)

American Samoa	Gibraltar	Panama
Andorra	Greece	Poland
Antigua and Barbuda	Greenland	Portugal
Aruba	Guam	Puerto Rico
Australia	Guyana	Qatar
Austria	Hong Kong SAR, China	Romania
Bahamas, The	Hungary	Russian Federation
Bahrain	Iceland	San Marino
Barbados	Ireland	Saudi Arabia
Belgium	Isle of Man	Seychelles
Bermuda	Israel	Singapore
British Virgin Islands	Italy	Sint Maarten (Dutch part)
Brunei Darussalam	Japan	Slovak Republic
Bulgaria	Korea, Rep.	Slovenia
Canada	Kuwait	Spain
Cayman Islands	Latvia	St. Kitts and Nevis
Channel Islands	Liechtenstein	St. Martin (French part)
Chile	Lithuania	Sweden
Costa Rica	Luxembourg	Switzerland
Croatia	Macao SAR, China	Taiwan, China
Curaçao	Malta	Trinidad and Tobago

Cyprus	Monaco	Turks and Caicos Islands
Czechia	Nauru	United Arab Emirates
Denmark	Netherlands	United Kingdom
Estonia	New Caledonia	United States
Faroe Islands	New Zealand	Uruguay
Finland	Northern Mariana Islands	Virgin Islands (U.S.)
France	Norway	

Note: This follows the latest update available at https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups accessed on 5 August 2025.

Appendix L: List of shocks and stressors required for interventions included in the resilience domain

Emergency	Definition
Rapid/ sudden onset emergencies:	
emerge suddenly and affect rapidly usually from a single distinct event	
Earthquake	Earthquake is defined as sudden movement of a block of the Earth's crust along a geological fault and associated ground shaking. (EM-DAT 2025) Earthquakes often trigger landslides, tidal waves and tsunamis. Powerful aftershocks frequently occur, causing further damage and increasing psychological stress (IFRC 2024).
Flood	Flood is a general term for the overflow of water from a stream channel onto normally dry land in the floodplain (riverine flooding), higher-than- normal levels along the coast and in lakes or reservoirs (coastal flooding) as well as ponding of water at or near the point where the rain fell (flash floods). (EM-DAT 2025) Flash Flood is defined as rapid inland floods due to intense rainfall A flash flood describes sudden flooding with short duration. In sloped terrain the water flows rapidly with a high destruction potential. (EM-DAT 2025)
Land Slide	Land Slide is defined as the usually rapid downward movement of a mass of rock, earth, or artificial fill on a slope. Covers all mass movements other than Mudslide (MS) and Avalanche (AV). (EM-DAT 2025) Mud slide is defined as a type of landslide, which occurs when the slope is saturated with water. This more destructive flow can pick up rocks, trees, houses and cars. As the debris moves into river and stream beds, bridges can become blocked or even collapse, making a temporary dam that can flood neighboring areas. (GLIDE) Snow Avalanche is defined as mass of snow and ice falling suddenly down a mountain slope and often taking with i t earth, rocks and rubble of every description. (EM-DAT 2025)
Tropical Cyclone	"Hurricane", "cyclone" and "typhoon" (GLIDE hazard code: TC) are different terms for the same weather phenomenon which is accompanied by torrential rain and maximum sustained wind speeds (near center) exceeding 119 kilometers per hour.
Tsunami	Tsunami is defined as a series of waves (with long wavelengths when traveling across the deep ocean) that are generated by a displacement of massive amounts of water through underwater earthquakes, volcanic eruptions, or landslides. Tsunami waves travel at very high speed across the ocean but as they begin to reach shallow water they slow down, and the wave grows steeper. (EM-DAT 2025)
Emergency	**Definition**
Volcano	Volcanic eruption with disastrous effects: eruption and

Emergency	Definition
	emission of gas and ashes, stone falls (pyroclast), flows of lava, etc.
Slow onset emergencies:	
emerge gradually over time, often based on a confluence of different events	
Cold wave	Cold Wave is defined as a period of abnormally cold weather. Typically, a cold wave lasts two or more days and may be aggravated by high winds. The exact temperature criteria for what constitutes a cold wave vary by location (EM-DAT 2025). It can cause respiratory problems, adverse effects on livelihoods and food security (ACAPS 2014).
Drought	Drought is defined as an extended period of unusually low precipitation that produces a shortage of water for people, animals and plants. Drought is different from most other hazards in that it develops slowly, sometimes even over years, and its onset is generally difficult to detect. Drought is not solely a physical phenomenon because its impacts can be exacerbated by human activities and water supply demands. Drought is therefore often defined both conceptually and operationally. Operational definitions of drought, meaning the degree of precipitation reduction that constitutes a drought, vary by locality, climate and environmental sector. (EM-DAT 2025)
Epidemic	Epidemic is defined as either an unusual increase in the number of cases of an infectious disease, which already exists in the region or population concerned, or the appearance of an infection previously absent from a region. (EM-DAT 2025)
Heat Wave	Heat Wave is defined as a prolonged period of excessively hot and sometimes also humid weather relative to normal climate patterns of a certain region. Heat waves like in Central Europe 2003. (EM-DAT 2025)
International Displacement	International displacement involves a number of people crossing international borders, being in need of assistance regardless of their status.
Emergencies with variable onset period:	
can emerge either rapidly or slowly depending on cause, triggering event and situations	
Conflict/Gangs/Terrorism/War	Any civil or political conflicts, gangs, terrorism or war as intervention contexts. war.
Emergency	**Definition**
*Famine/Starvation	A situation where households face an extreme lack of food that can bring a lack of other basic needs as well. Famine is intertwined with starvation, death, destitution, and extremely critical acute malnutrition levels. (IPC 2024)
Other:	
All disasters that do not fall into any of the other disaster types	
*Climate Change	This refers to general climate change being addressed as a stressor in a study. We will select this only if other specific

Emergency	Definition
	shocks/stressors are not mentioned.
*Economic Crisis	A sharp decline in economic performance of a country, which includes drastic increases in unemployment and business bankruptcy. (Dzingirai and Ndava 2022)
Technological Disaster	Danger originating from technological or industrial accidents, dangerous procedures, infrastructure failures or certain human activities, which may cause the loss of life or injury, property damage, social and economic disruption, or environmental degradation. Includes Explosions/Chemical explosion/Nuclear explosion/Radiation/Mine explosion; Pollution/Chemical pollution/Atmosphere pollution; Acid rain
Wildfire	Wildfire (GLIDE hazard code: WF) is defined as any uncontrolled and non-prescribed combustion or burning of plants in a natural setting such as a forest, grassland, brush land or tundra, which consumes the natural fuels and spreads based on environmental conditions (e.g., wind, topography). Wildfires can be triggered by lightning or human actions. (EM-DAT 2025)

References used in this appendix

ACAPS. 2014. “Bangladesh: Pre-Disaster Secondary Data – River Flooding (March 2014).” Geneva: ACAPS. https://reliefweb.int/report/bangladesh/bangladesh-pre-disaster-secondary-data-river-flooding-march-2014.

Dzingirai, M., and R. Ndava. 2022. “Entrepreneurship Education as a Driver for Economic Recovery in the Post-Pandemic Era.” In *Advances in Business Strategy and Competitive Advantage*, edited by S. Hai-Jew, 27–50. Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-7998-8626-6.ch002.

EM-DAT. 2025. “Emergency Events Database (EM-DAT).” Brussels: Centre for Research on the Epidemiology of Disasters. https://www.emdat.be.

Integrated Food Security Phase Classification (IPC). 2024. “IPC Famine Factsheet.” Rome: IPC Global Support Unit, Food and Agriculture Organization of the United Nations. https://www.ipcinfo.org/.

International Federation of Red Cross and Red Crescent Societies. 2024. “Earthquakes: Safety Tips for Volunteers & Personnel Working in Disaster-Affected Areas.” Geneva: IFRC. https://ifrcstaysafe.org/disasters-tips/earthquakes.

References

Almulhim, A. I., G. N. Alverio, A. Sharifi, R. Shaw, S. Huq, M. J. Mahmud, S. Ahmad, and I. R. Abubakar. “Climate-Induced Migration in the Global South: An In-Depth Analysis.” *npj Climate Action* 3, no. 1 (14 June 2024): 47. https://doi.org/10.1038/s44168-024-00133-1.

Anda León, M. D. A., Carolyn Huang, Miriam Berretta, Promise Nduku, Andile Madonsela, and Shannon Shisler. ‘Systematic Review on the Effect of Skills-Based Active Labor Market Interventions on Migration Outcomes’. International Initiative for Impact Evaluation (3ie), 2023. https://doi.org/10.23846/SR00048.

Beegle, K., J. D. Weerdt, and S. Dercon. ‘Migration and economic mobility in Tanzania: Evidence from a tracking survey’. *The Review of Economics and Statistics* 93, no. 3 (2011): 1010–33. http://www.jstor.org/stable/23016092.

Beine, M., D. Khoudour, and J. Tarvainen. 2021. *Can targeted interventions mitigate the adverse drivers of irregular migration and forced displacement?* New York: United Nations Development Programme. https://www.undp.org/sites/g/files/zskgke326/files/2021-11/UNDP-Can-Targeted-Interventions-Mitigate-the-Adverse-Drivers-of-Irregular-Migration-and-Forced-Displacement.pdf.

Berretta, M., M. D. A. Anda León, C. Huang, and S. Shisler. 2023a. *Addressing root causes and drivers of irregular migration: An evidence gap map.* New Delhi: International Initiative for Impact Evaluation (3ie). https://doi.org/10.23846/EGM030.

Berretta, M., S. Lee, M. Kupfer, C. Huang, W. Ridlehoover, D. Frey, F. Ahmed, G. Hug, E. McFee, R. Glennerster, S. Kumar, and A. R. Pebley. 2023b. *Strengthening resilience against shocks, stressors and recurring crises in low- and middle-income countries: An evidence gap map.* New Delhi: International Initiative for Impact Evaluation (3ie). https://doi.org/10.23846/EGM023.

Bencek, D., and C. Schneiderheinze. 2020. *Higher economic growth in poor countries, lower migration flows to the OECD: Revisiting the migration hump with panel data.* Kiel Working Paper. Kiel: Kiel Institute for the World Economy. https://hdl.handle.net/10419/231567.

Berthiaume, N., N. Leefmans, N. Oomes, H. A. R. Romagosa, and T. Vervliet. 2021. *A reappraisal of the migration–development nexus: Testing the robustness of the migration transition hypothesis.* Policy Research Working Paper. Washington, DC: World Bank Group. http://documents.worldbank.org/curated/en/786201611170919983.

Boas, I., A. Gautam, and A. Olayiwola. 2024. “Climate change and human mobility in the global south.” In *The Palgrave Handbook of South–South Migration and Inequality*, edited by H. Crawley and J. K. Teye, 435–53. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-39814-8_20.

Borenstein, M., L. V. Higgins, J. P. T., Hedges, and H. R. Rothstein. 2021. *Introduction to meta-analysis.* Chichester, UK: Wiley.

Carling, J. 2002. “Migration in the age of involuntary immobility: Theoretical reflections and Cape Verdean experiences.” *Journal of Ethnic and Migration Studies* 28 (1): 5–42. https://doi.org/10.1080/13691830120103912.

Carling, J., and K. Schewel. 2018. “Revisiting aspiration and ability in international migration.” *Journal of Ethnic and Migration Studies* 44 (6): 945–63. https://doi.org/10.1080/1369183X.2017.1384146.

Carling, J., and C. Talleraas. 2016. *Root causes and drivers of migration: Implications for humanitarian efforts and development cooperation.* Oslo: Peace Research Institute Oslo. https://www.prio.org/publications/10005.

Carter, E. C., F. D. Schönbrodt, W. M. Gervais, and J. Hilgard. 2019. “Correcting for bias in psychology: A comparison of meta-analytic methods.” *Advances in Methods and Practices in Psychological Science* 2 (2): 115–44. https://doi.org/10.1177/2515245919847196.

Casentini, G., L. Hammond, and O. Bakewell. 2024. “Poverty, income inequalities and migration in the global south.” In *The Palgrave Handbook of South–South Migration and Inequality*, edited by H. Crawley and J. K. Teye, 371–92. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-39814-8_17.

Caso, N., and J. Carling. 2024. “The reach and impact of migration information campaigns in 25 communities across Africa and Asia.” *Migration Policy Practice* 13 (1): 3–11. https://www.prio.org/publications/13925.

Clemens, M. A. 2014. *Does development reduce migration?* Bonn: Institute of Labor Economics (IZA). https://docs.iza.org/dp8592.pdf.

———. 2020. *The emigration life cycle: How development shapes emigration from poor countries.* Washington, DC: Center for Global Development. https://www.cgdev.org/publication/emigration-life-cycle-how-development-shapes-emigration-poor-countries.

———. 2021. “Violence, development, and migration waves: Evidence from Central American child migrant apprehensions.” *Journal of Urban Economics* 124: 103355. https://doi.org/10.1016/j.jue.2021.103355.

Clemens, M. A., and H. Postel. 2017. “Temporary work visas as US-Haiti development cooperation: A preliminary impact evaluation.” *IZA Journal of Labor & Development* 6 (1): 4. https://doi.org/10.1186/s40175-016-0070-x.

Coburn, K. M. 2019. *weightr: Estimating weight-function models for publication bias* (R package version 2.0.2). https://CRAN.R-project.org/package=weightr.

Cummings, C., J. Pacitto, D. Lauro, and M. Foresti. 2015. *Why people move: Understanding the drivers and trends of migration to Europe.* ODI Working Paper 430. London: Overseas Development Institute. https://odi.org/en/publications/why-people-move-understanding-the-drivers-and-trends-of-migration-to-europe/.

Crawley, H., and J. K. Teye, eds. 2024. *The Palgrave Handbook of South–South Migration and Inequality.* Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-39814-8.

Czaika, M., and H. de Haas. 2013. 'The Effectiveness of Immigration Policies'. *Population and Development Review* 39(3): 487–508. https://doi.org/10.1111/j.1728-4457.2013.00613.x.

Czaika, M., and C. Reinprecht. 2022. "Migration drivers: Why do people migrate?" In *Introduction to Migration Studies*, edited by P. Scholten, 49–82. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-92377-8_3.

de Haas, H. 2007. "Turning the tide? Why development will not stop migration." *Development and Change* 38 (5): 819–41. https://doi.org/10.1111/j.1467-7660.2007.00435.x.

———. 2010. "Migration and development: A theoretical perspective." *International Migration Review* 44 (1): 227–64. https://doi.org/10.1111/j.1747-7379.2009.00804.x.

———. 2021. "A theory of migration: The aspirations–capabilities framework." *Comparative Migration Studies* 9 (1): 8. https://doi.org/10.1186/s40878-020-00210-4.

———. 2023. *How migration really works: The facts about the most divisive issue in politics.* New York: Basic Books.

de Haas, H. Czaika, M. Flahaux, M. L. Mahendra, E. Natter, K. Vezzoli, S. and Villares-Varela, M. 2019. International Migration: Trends, Determinants, and Policy Effects. *Population and Development Review,* 45(4): 885–922. https://doi.org/10.1111/padr.12291.

Doherty, J., S. Hughes, L. Pritchard, and H. Jones. 2020. *Protocol: The effects of rule of law interventions on justice outcomes: An evidence gap map.* New Delhi: International Initiative for Impact Evaluation (3ie). https://www.3ieimpact.org/sites/default/files/2020-10/Protocol-for-Rule-ofLaw-EGM.pdf.

Dunsch, F., D. Tjaden, and W. Quiviger. 2019. *Migrants as Messengers: The Impact of Peer-to-Peer Communication on Potential Migrants in Senegal. Impact Evaluation Report.* Geneva: International Organization for Migration.

EPPI-Centre. 2022. *Classifiers in EPPI Reviewer.* London: UCL Social Research Institute, May 18, 2022. https://eppi.ioe.ac.uk/cms/Portals/35/Manuals/Classifiers%20in%20EPPI%20Reviewer%20180522.pdf.

European Commission. 2024. *Pact on migration and asylum.* https://home-affairs.ec.europa.eu/policies/migration-and-asylum/pact-migration-and-asylum_en.

———. 2025. *Asylum, Migration and Integration Fund (2021–2027).* Migration and Home Affairs – Funding section. https://home-affairs.ec.europa.eu/funding/asylum-migration-and-integration-funds/asylum-migration-and-integration-fund-2021-2027_en.

———. n.d. *Neighbourhood, Development and International Cooperation Instrument – Global Europe (NDICI – Global Europe).* Enlargement & Technical Assistance section. https://enlargement.ec.europa.eu/funding-technical-assistance/neighbourhood-development-and-international-cooperation-instrument-global-europe-ndici-global-europe_en.

FAO. 2021. *The impact of disasters and crises on agriculture and food security: 2021.* Rome: Food and Agriculture Organization of the United Nations. https://doi.org/10.4060/cb3673en.

Fisher, Z., and E. Tipton. 2015. "Robumeta: An R package for robust variance estimation in meta-analysis." *arXiv.* https://arxiv.org/abs/1503.02220.

Frelick, B, I. M. Kysel, and J. Podkul. 2016. 'The Impact of Externalization of Migration Controls on the Rights of Asylum Seekers and Other Migrants'. *Journal on Migration and Human Security* 4 (4): 190–220. https://www.hrw.org/sites/default/files/supporting_resources/jmhs.pdf.

Gent, S. 2002. *The root causes of migration: Criticising the approach and finding a way forward.* Brighton: Sussex Centre for Migration Research.

Gibson, J., and D. McKenzie. 2010. *The development impact of a best practice seasonal worker policy.* Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-5488.

Gupta, D., Pankaj Kumar, Naoyuki Okano, and Manish Sharma. 'Climate-Induced Migration in India and Bangladesh: A Systematic Review of Drivers, Impacts, and Adaptation Mechanisms'. Climate 13, no. 4 (21 April 2025): 81. https://doi.org/10.3390/cli13040081.

Hedges, L. V. 2009. "Effect sizes in nested designs." In *The handbook of research synthesis and meta-analysis,* 2nd ed., edited by H. Cooper, L. V. Hedges, and J. Valentine, 337–55. New York: Russell Sage Foundation.

Hedges, L. V., E. Tipton, and M. 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.

Hendow, M., N. Wichmann, A. Spehar, and M. Neske. 2024. *How fit is the available data: Understanding strengths and weaknesses in data on irregular migration in Europe.* MIrreM Working Paper No. 6. Krems: University for Continuing Education Krems (Danube University Krems). https://irregularmigration.eu/wp-content/uploads/2024/10/MIRREM_Hendow-et-al-2024-How-fit-is-the-available-data-v1.pdf.

Higgins, J. P. T., J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, and V. A. Welch. 2024. *Cochrane handbook for systematic reviews of interventions.* Version 6.5 (updated August 2024). Cochrane. www.training.cochrane.org/handbook.

Higgins, J. P. T., J. A. C. Sterne, J. Savovic, M. J. Page, A. Hróbjartsson, I. Boutron, B. Reeves, and S. Eldridge. 2016. "A revised tool for assessing risk of bias in randomized trials." *Cochrane Database of Systematic Reviews* 10 (Suppl 1).

Higgins, J. P. T., J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, and V. A. Welch, eds. 2020. *Cochrane handbook for systematic reviews of interventions.* Version 6.1 (updated September 2020). London: Cochrane. https://training.cochrane.org/handbook/archive/v6.1.

Huang, L. 2023. *Climate migration 101: An explainer.* Migration Information Source, November 16, 2023. https://www.migrationpolicy.org/article/climate-migration-101-explainer.

IOM. 2019. *Glossary on migration.* Geneva: International Organization for Migration. https://publications.iom.int/system/files/pdf/iml_34_glossary.pdf.

———. 2021. *World migration report 2022.* Geneva: International Organization for Migration. https://doi.org/10.18356/9789292680763.

———. 2024a. *Programme and budget for 2025 (C/115/7).* 115th Session of the Council. Geneva: International Organization for Migration. https://governingbodies.iom.int/system/files/en/council/115/c-115-7-programme-and-budget-for-2025.pdf.

———. 2024b. *Guiding safe migration: A new phase for the COMPASS programme (2024–2027).* Geneva: International Organization for Migration. https://www.iom.int/sites/g/files/tmzbdl486/files/documents/2024-06/compass-phase-ii-programme-brief.pdf.

Kierans, D., and C. Vargas-Silva. 2024. *The irregular migrant population of Europe.* MIrreM Working Paper. https://irregularmigration.eu/wp-content/uploads/2024/10/MIRREM-Kierans-and-Vargas-Silva-2024-Irregular-Migrant-Population-in-Europe-v1.pdf.

Knoll, A., and A. Sheriff. 2017. *Making waves: Implications of the irregular migration and refugee situation on official development assistance spending and practices in Europe.* Stockholm: Expertgruppen för biståndsanalys (EBA). https://www.oecd.org/derec/sweden/201701-ECDPM-rapport.pdf.

Kugley, S., A. Wade, J. Thomas, Q. Mahood, A.-M. K. Jørgensen, K. Hammerstrøm, and N. 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.

Kuhnt, J. 2019. *Why do people leave their homes? Is there an easy answer? A structured overview of migratory determinants.* Discussion Paper 9/2019. Bonn: German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE). https://doi.org/10.23661/dp9.2019.

Larotta Silva, S. P. ‘Understanding of Environmental Change as a Driver of Migration in Latin America through a Systematic Review of Literature’. Journal of Population Research 42, no. 3 (2025): 42. https://doi.org/10.1007/s12546-025-09391-0.

Lewin, S., A. D. Oxman, J. N. Lavis, and A. 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.

Lipsey, M. W. 2009. “Identifying interesting variables and analysis opportunities.” In *Handbook of research synthesis and meta-analysis*, 2nd ed., 147–58. New York: Russell Sage Foundation.

McKenzie, D. ‘Poverty, Inequality, and International Migration: Insights from 10 Years of Migration and Development Conferences’: *Revue d’économie Du Développement* Vol. 25, no. 3 (26 June 2018): 13–28. https://doi.org/10.3917/edd.313.0013.

Mckenzie, David, and Hillel Rapoport. ‘Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico’. *Journal of Development Economics* 84, no. 1 (2007): 1–24. https://doi.org/10.1016/j.jdeveco.2006.11.003.

Miles, M. B., and A. M. Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*. 2nd ed. Thousand Oaks, CA: Sage.

Mixed Migration Centre. ‘Climate Change, Environmental Stressors, and Mixed Migration’, 2022. http://www.mixedmigration.org/.

Mukherjee, M., and Sonja Fransen. ‘Exploring Migration Decision-Making and Agricultural Adaptation in the Context of Climate Change: A Systematic Review’. World Development 179 (2024): 106600. https://doi.org/10.1016/j.worlddev.2024.106600.

Mullan, R. J., D. N. Flynn, B. Carlberg, I. M. Tleyjeh, C. C. Kamath, M. L. LaBella, P. J. Erwin, G. H. Guyatt, and V. M. Montori. 2009. “Systematic reviewers commonly contact study authors but do so with limited rigor.” *Journal of Clinical Epidemiology* 62 (2): 138–42. https://doi.org/10.1016/j.jclinepi.2008.08.002.

Munthe-Kaas, H. M., A. Booth, I. Sommer, S. Cooper, R. Garside, K. Hannes, J. Noyes, and The CAMELOT Development Group. 2024. “Developing CAMELOT for assessing methodological limitations of qualitative research for inclusion in qualitative evidence syntheses.” *Cochrane Evidence Synthesis and Methods* 2 (6): e12058. https://doi.org/10.1002/cesm.12058.

Nicaise, I., and T. Bircan. 2024. *Anticipating and Managing Migration Flows towards the EU: HumMingBird Synthesis Report (Deliverable 10.4)*. HIVA – KU Leuven (HumMingBird project, Horizon 2020, No 870661). https://hummingbird-h2020.eu/images/projectoutput/d10-4-synthesis-report-end.pdf.

OECD. n.d. *Migration-related activities in official development assistance (ODA)*. OECD website. https://www.oecd.org/en/topics/sub-issues/oda-eligibility-and-conditions/migration-related-activities-in-oda.html.

Polanin, J. R., and B. Snilstveit. 2016. “Converting between effect sizes.” *Campbell Systematic Reviews* 12 (1): 1–13.

R Core Team. 2022. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org.

Rodríguez-Sánchez, A., and J. Tjaden. 2025. ‘Estimating Irregular Migrant Stocks Using Social Media Data and Machine Learning’. https://doi.org/10.5281/ZENODO.14808984.

Slootjes, J., and R. Sohst. 2024. *Towards the more effective use of irregular migration data in policymaking*. MPI Europe Policy Brief. Migration Policy Institute Europe. https://www.migrationpolicy.org/research/irregular-migration-data-policymaking.

Snilstveit, B., R. Bhatia, K. Rankin, and B. Leach. 2017. *3ie evidence gap maps: A starting point for strategic evidence production and use*. New Delhi: International Initiative for Impact Evaluation (3ie). https://www.3ieimpact.org/sites/default/files/2019-01/3ie_evidence_gap_maps.pdf.

Snilstveit, B., M. Vojtkova, A. Bhavsar, J. Stevenson, and M. Gaarder. 2016. "Evidence & gap maps: A tool for promoting evidence informed policy and strategic research agendas." *Journal of Clinical Epidemiology* 79: 120–29. https://doi.org/10.1016/j.jclinepi.2016.05.015.

Sterne, J. A. C., J. Savović, M. J. Page, R. G. Elbers, N. S. Blencowe, I. Boutron, C. J. Cates, et al. 2019. "RoB 2: A revised tool for assessing risk of bias in randomised trials." *BMJ*, 28 August 2019, l4898. https://doi.org/10.1136/bmj.l4898.

Sterne, J. A. C., M. A. Hernán, B. C. Reeves, J. Savović, N. D. Berkman, M. Viswanathan, D. Henry, et al. 2016. "ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions." *BMJ*, 12 October 2016, i4919. https://doi.org/10.1136/bmj.i4919.

Schewel, K., and A. Debray. 'Global Trends in South–South Migration'. In *The Palgrave Handbook of South–South Migration and Inequality*, edited by H. Crawley and J. Kofi Teye, 153–81. Cham: Springer International Publishing, 2024. https://doi.org/10.1007/978-3-031-39814-8_8.

Tafani, I., and M. Riccaboni. 2025. "The impact of the EU–Turkey agreement on the number of lives lost at sea." *Humanities and Social Sciences Communications* 12 (1): 869. https://doi.org/10.1057/s41599-025-04900-1.

The Government of the Netherlands. 2024. *Voluntary national review GCM – the Netherlands (GCM voortgangsrapportage 2024)*. United Nations Network on Migration. https://migrationnetwork.un.org/system/files/docs/NL%20GCM%20voortgangsrapportage%202024%20-%20final.pdf.

Thomas, J., S. Graziosi, J. Brunton, Z. Ghouze, P. O'Driscoll, M. Bond, and A. Koryakina. 2023. *EPPI-Reviewer: Advanced software for systematic reviews, maps and evidence synthesis*. EPPI Centre, UCL Social Research Institute, University College London. https://eppi.ioe.ac.uk/cms/About/AboutEPPIReviewer/tabid/2967/Default.aspx.

Tjaden, J., and F. A. Dunsch. 2021. "The effect of peer-to-peer risk information on potential migrants: Evidence from a randomized controlled trial in Senegal." *World Development* 145: 105488. https://doi.org/10.1016/j.worlddev.2021.105488.

Trauner, F., J. P. Brekke, I. Adam, O. N. Cham, and K. Thorbjørnsrud. 2024. "'It's not about the information, it's about the situation': Understanding the misalignment between EU deterrence messaging and migrants' narratives." *Journal of Ethnic and Migration Studies* 50 (14): 3379–95. https://doi.org/10.1080/1369183X.2024.2333865.

Triandafyllidou, A., L. Bartolini, and C. F. Guidi. 2019. *Exploring the links between enhancing regular pathways and discouraging irregular migration: A discussion paper to inform future policy deliberations.* Geneva: International Organization for Migration (IOM). https://publications.iom.int/books/exploring-links-between-enhancing-regular-pathways-and-discouraging-irregular-migration.

UN. 'Global Compact on Refugees'. New York: UN, 2018. https://www.unhcr.org/5c658aed4.pdf.

UNHCR. 'Budget and Expenditure'. Planning, Funding and Results, 2025. https://www.unhcr.org/about-unhcr/planning-funding-and-results/budget-and-expenditure?year=2019.

———. 'Budget and Expenditure, Global Focus', 2022. http://reporting.unhcr.org/financial.

———. 'Global Trends 2024'. UNHCR Global Trends Report. UNHCR, 2025. https://www.unhcr.org/global-trends-report-2024.

UNSD. 'Global Indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development'. United Nations, 2022. https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202022%20refinement_Eng.pdf.

Valentine, J. C., A. M. Aloe, and T. 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, W., and M. W.-L. Cheung. 2010. "Outlier and influence diagnostics for meta-analysis." *Research Synthesis Methods* 1 (2): 112–25. https://doi.org/10.1002/jrsm.11.

Vutha, H., L. Pide, and P. Dalis. 2011. "Irregular migration from Cambodia: Characteristics, challenges, and regulatory approach." https://www.econstor.eu/handle/10419/126861.

Waddington, H., B. Snilstveit, J. Hombrados, M. Vojtkova, D. Phillips, P. Davies, and H. White. 2014. "Farmer field schools for improving farming practices and farmer outcomes: A systematic review." *Campbell Systematic Reviews* 10 (1). https://doi.org/10.4073/CSR.2014.6.

Waddington, H., H. White, B. Snilstveit, J. G. Hombrados, M. Vojtkova, P. Davies, A. 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.

Yayboke, E. K., and C. G. Gallego. 2019. *Out of the shadows: Shining a light on irregular migration*. CSIS Project on Prosperity and Development. https://reliefweb.int/report/world/out-shadows-shining-light-irregular-migration.

Xanthopoulou, E. 2024. 'Mapping EU Externalisation Devices through a Critical Eye'. *European Journal of Migration and Law* 26 (1): 108–35. https://doi.org/10.1163/15718166-12340170.