EVIDENCE GAP MAP

CASH TRANSFERS AND CASH PLUS PROGRAMS IN LOW- AND MIDDLE-INCOME COUNTRIES

FOR

DEUTSCHES EVALUIERUNGSINSTITUT DER ENTWICKLUNGSZUSAMMENARBEIT (DEVAL)

CENTER FOR EVALUATION AND DEVELOPMENT (C4ED)

MANNHEIMER ZENTRUM FÜR EVALUATION UND ENTWICKLUNGS-FORSCHUNG (MZEEF)

CENTRE POUR L'ÉVALUATION ET LE DÉVELOPPEMENT

CENTRO PARA LA EVALUACIÓN Y EL DESARROLLO

WWW.C4ED.ORG

DEval GERMAN INSTITUTE FOR DEVELOPMENT EVALUATION

With financial support from the Federal Ministry for Economic Cooperation and Development

Center for Evaluation and Development

21 MARCH 2024

List of authors

Full name	Institution
Mathilda Featherston-Lardeux	Center for Evaluation and Development
Atika Pasha	Center for Evaluation and Development
Christiaan de Swardt	Center for Evaluation and Development
Viviana Urueña	Center for Evaluation and Development
Denise Hörner	German Institute for Development Evaluation (DEval)
Stefanie Knoll	German Institute for Development Evaluation (DEval)
Kathrin Wolf	German Institute for Development Evaluation (DEval)

Please cite this report as follows:

Pasha, A., V. Urueña, S. Knoll, K. Wolf, C. De Swardt, M. Featherston-Lardeux and D. Hörner (2023), Cash Transfers and Cash Plus Programs in Low-and Middle-Income Countries: Evidence Gap Map, Center for Evaluation and Development (C4ED) and German Institute for Development Evaluation (DEval), Bonn.

The interactive EGM tool can be found here: EGM: Cash Transfers and Cash Plus Programs in Low- and Middle-Income Countries | socialprotection.org. Please note that the interactive tool uses unconditional versus conditional cash transfers as a filter option, whereas this report draws on a previous version that differentiated between those types of transfer as part of the main framework.

ACKNOWLEDGEMENTS

The research team acknowledges input from Sarah Young on the search strategy. They also acknowledge Ankhbayar Delgerchuluun, Zahra Kamal, Dania Natalia Ochoa Ruiz, and Hanne Schwabe for their support in the screening and data extraction process and Irene Furlani for structuring the first visualization of the EGM. Finally, we would like to thank Alexander Prior for his overall support on the administrative ordeals related to EPPI Reviewer.

EXECUTIVE SUMMARY

Cash transfers are among the primary tools in social protection aimed at reducing poverty and enhancing people's lives. These transfers have the dual purpose of alleviating poverty among vulnerable populations and strengthening their ability to withstand unexpected hardships and shocks. To enhance the impact of cash transfers for a given transfer amount, they are sometimes linked to recipients' compliance with certain conditions, such as sending their children to school or giving birth in a health clinic. There is also a growing trend of combining transfers with additional interventions or services. Examples of those additional interventions are transferring information (e.g., encouraging mothers to increase their babies' protein intake, parenting skills classes, or information on nutrition), offering psychosocial support, or providing non-cash items in the form of food or other in-kind transfers. When cash transfers are coupled with such complementary support, the interventions are referred to as cash plus programs.

In recent years, the body of evidence on the effects of cash transfers and cash plus interventions has increased significantly, particularly for low- and middle-income countries. The range of studied outcomes spans diverse fields such as health, education, consumption, and women's empowerment. As more studies and information become available, the need for synthesis and evidence mapping becomes ever greater.

This evidence gap map aims to facilitate evidence-informed decision-making by enabling easy access to the existing rigorous evidence on cash transfers and cash plus programs in low- and middle-income countries. It includes a total of 709 impact evaluations and 33 systematic reviews and provides a visual and interactive representation of these studies in a framework of cash transfer types and a broad range of outcomes. Moreover, it pinpoints critical areas with limited or no evidence, highlighting areas in which more impact evaluations or syntheses are needed.

This evidence gap map serves as a comprehensive repository of knowledge, consolidating robust evidence on cash transfers across diverse contexts and intervention characteristics. Through its various filter options, it allows development practitioners to gain insights into specific aspects of cash transfer interventions of interest and to select appropriate studies. The filter options range from the study region to the characteristics of the target population, such as beneficiaries' age and sex, as well as targeting approaches and payment methods.

A considerable body of evidence examines conditional cash transfer interventions, with or without additional components. Conditional cash transfer interventions not only outnumber unconditional cash transfer interventions but also are more likely to incorporate plus components.

In terms of plus components, information, nudges, or behavioral change communication are the most studied plus components in both impact evaluations and systematic reviews. For impact evaluations, the second largest focus of the available evidence is on training, while for systematic reviews, food transfers are as frequently investigated as information, nudges, or behavioral change communication. Psychosocial support is the least commonly studied plus component in impact evaluations and systematic reviews.

Health is the most extensively examined outcome category, followed by living standards and consumption, employment and entrepreneurship, and education. At least some

evidence is available across all intervention-outcome combinations, with a minimum of 17 impact evaluations available for each combination. The two least studied outcomes are related to social cohesion on the one hand and gender equality and empowerment on the other. Here, more impact evaluations are needed. In terms of systematic reviews, health outcomes are also the most studied, followed by education outcomes and outcomes related to employment and entrepreneurship. This does not imply that these areas are saturated in terms of systematic reviews.

There is a great potential for future syntheses for outcomes assessing living standards and consumption, employment and entrepreneurship, and agricultural production. Many impact evaluations measure these outcomes. However, there are very few accompanying systematic reviews synthesizing the results. There are further synthesis gaps for outcomes related to financial inclusion, savings, and insurance, and gender equality and empowerment, as they are examined in only one systematic review each. Furthermore, no systematic review reports on outcomes related to social cohesion due to the low number of impact evaluations. More impact evaluations for these outcomes are needed to allow for evidence syntheses.

Looking at different populations, there exists ample evidence for women and vulnerable age groups. Numerous studies examine the effects of cash transfers on women and girls, children, adolescents, and the elderly. Among these, conditional cash transfers are more frequently studied for women and girls, children, and adolescents, while unconditional cash transfers are more commonly investigated for the elderly. Health outcomes are predominantly studied for women, girls, and the elderly, whereas education outcomes take precedence for children and adolescents. However, there are notable evidence gaps for indigenous people and local communities, persons with disabilities, and there are no impact evaluations studying the effects on members of the LGBTQ+ community.

The majority of evidence assesses interventions implemented entirely or partially by governments. Of the 709 impact evaluation studies, 66% evaluate interventions exclusively implemented by governments, while 20% assess interventions implemented solely by non-government actors. Governmental interventions predominantly consist of conditional cash transfers (58%) and often have more than 10,000 beneficiaries (76%). The distribution of evidence in terms of outcomes for governmental interventions is consistent with the overall patterns. In addition, governmental interventions tend to have a longer duration and are more likely to focus on improving access to basic services and providing healthcare as plus components than non-governmental interventions.

Geographically, the majority of studies are concentrated in Sub-Saharan Africa and Latin America and the Caribbean. In Sub-Saharan Africa, Kenya, Malawi, and South Africa are the most studied countries, while in Latin America and the Caribbean, Mexico, Brazil, and Colombia receive the most attention. Evidence for cash transfers is limited in the Middle East and North Africa and especially in Europe and Central Asia. Evidence is more abundant in contexts that are assessed as less politically fragile ("Warning" contexts) than in contexts that are assessed as more politically fragile ("Alert" contexts). 84% of studies are set in "Warning" contexts, while only 14% are set in "Alert" contexts. Interventions implemented in "Warning" contexts tend to focus on conditional cash transfers, while in "Alert" contexts or humanitarian settings, the evidence is more abundant for unconditional cash transfer interventions. Only 6% of studies investigate the impact of cash transfers in humanitarian settings, with more than half of these studies set in "Alert" contexts.

The available evidence primarily includes evaluations performed shortly after the start of the intervention, meaning three years or sooner. Approximately 65% of studies fall into this category, while in 25% of studies evaluations are performed four to nine years after the start and in 7% ten years or more after the start of the intervention.

Substantial evidence gaps exist concerning the best methods for targeting beneficiaries and the mode of delivery of CTs. Only five impact evaluations (and no systematic review) investigate targeting approaches, comparing different methods of identifying eligible beneficiaries. Some studies compare community-based targeting, where a community jointly decides who should receive the cash transfer, vs. alternative methods that define categories of eligible beneficiaries (e.g. individuals being 55 years or older). Similarly, evidence gaps exist for studies investigating whom to target (e.g., targeting men or women or targeting children or parents), with 15 impact evaluations and one systematic review addressing this issue. These studies predominantly present results for education and health outcomes. In addition, few studies report enough information to assess whether the CT was provided through digital or physical means.

TABLE OF CONTENTS

ABBREVIATIONS

BCC	Behavior Change Communication
C4ED	Center for Evaluation and Development
CBT	Community-based Targeting
CTs	Cash Transfers
CTs+	Cash Transfers Plus
CCTs	Conditional Cash Transfers
CCTs+	Conditional Cash Transfers Plus
DEval	Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit
ECA	Europe and Central Asia
EGM	Evidence Gap Map
GDC	German Development Cooperation
HIC	High-Income Country
IEs	Impact Evaluations
LAC	Latin America and Caribbean
LGBTQ+	Lesbian, Gay, Bisexual, Transgender, Queer/Questioning and others
LMIC	Low- and Middle-Income Country
MENA	Middle East and North Africa
PICOS	Population, Intervention, Comparison, Outcome, Study design
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PMT	Proxy-means test
RQ	Research Question
SRs	Systematic Reviews
SRH	Sexual and reproductive health
SSA	Sub-Saharan Africa
SURE	Supporting the Use of Research Evidence
ToC	Theory of Change
UCTs	Unconditional Cash Transfers
UCTs+	Unconditional Cash Transfers Plus

LIST OF FIGURES

LIST OF TABLES

1. RATIONALE AND CONCEPTUAL FRAMEWORK

The Center for Evaluation and Development (C4ED) was commissioned by the German Institute for Development Evaluation (DEval) to develop an evidence gap map (EGM) on the effects of cash transfers (CTs) and cash transfer plus programs (CTs+) on a broad set of development outcomes. The design and effects of different types of CT interventions in various settings are of particular interest to the social protection sector of the German Development Cooperation (GDC). This report presents the assignment's rationale, objectives, methodology, and results.

1.1. RATIONALE AND OBJECTIVES

CTs are among the primary tools in social protection aimed at reducing poverty and enhancing people's lives. They are intended to expand beneficiaries' opportunities for spending and investment or incentivize behavioral changes expected to improve long-term human capital. They consist of direct and predictable monetary transfers that are either provided unconditionally or subject to conditions, that are geared towards incentivizing behavioral changes (Carter et al., 2019; Little et al., 2021). CT programs with conditions aim to break intergenerational cycles of poverty by linking programs to enduring investments, such as a child's health or education. Research indicates that CTs can contribute to various development outcomes directly (see e.g., Bastagli et al., 2016 for effects of CTs on recipients) or indirectly by affecting the local economy and thus also non-recipients (e.g., Egger et al., 2022).

To enhance the impact of CTs for a given transfer amount, there is a growing trend of combining them with additional interventions or services. Examples of those additional interventions or services are providing information, offering psychosocial support, or providing non-cash items in the form of food or other in-kind transfers. When CTs are coupled with such complementary support, the interventions are referred to as CT+ programs (Roelen et al., 2017).

The multitude of studies makes it increasingly difficult for stakeholders in development research, policy, and practice to maintain an overview of the existing evidence or to use the available evidence to inform policymaking. In recent years, the body of evidence on the effects of CT and CT+ interventions on a broad range of outcomes increased significantly. The range of studied outcomes spans diverse fields such as health, education, consumption, and women's empowerment, with a particular focus on the context of low- and middle-income countries (LMICs). As more studies and information become available, the need for synthesis and evidence mapping becomes ever greater.

This EGM aims to facilitate evidence-informed decision-making in the social protection sector by enabling easy access to the existing rigorous evidence on the effects of CTs and CTs+ regarding a broad range of outcome indicators in LMICs. It provides a visual and interactive representation of completed systematic reviews (SRs), meta-analyses, and impact evaluations (IEs), structured around a framework of interventions and outcomes.[1] Moreover, it

[1] An SR is a rigorous and structured research methodology used to comprehensively and impartially summarize, evaluate, and synthesize existing research (quasi-)experimental studies and evidence on a specific research question or topic. It involves a systematic and well-documented process to locate, assess, and analyze all relevant studies available in the literature. A meta-analysis follows similar steps as the SR, but besides synthesizing the evidence it also provides an overall estimate of the effect or relationship being studied, along with a measure of its precision. Researchers interpret these results and draw conclusions about the strength, direction, and significance of the effect.

pinpoints critical areas with limited or no evidence from IEs and SRs, and suggests potential areas for future research.

The EGM serves two main objectives:

1. Supporting policymakers in **designing evidence-informed strategies, policies, and interventions** by easily **identifying relevant IEs, meta-analyses, and SRs** for areas of policy interest.
2. Supporting researchers and policymakers in **building strategic research agendas by**:
    a. Identifying **absolute gaps** of evidence where few or no primary studies exist. In these areas, conducting or commissioning **new IEs** could be a priority for future research projects.
    b. Identifying **synthesis gaps** of evidence where many IEs, yet no or few SRs exist. In these areas, conducting or commissioning **evidence syntheses** (in the form of SRs or meta-analyses) might be particularly useful.

To optimize the utility of the EGM and complement the work of other researchers who are concurrently developing an EGM focused on social protection, this EGM primarily concentrates on a core category of social protection interventions, namely CTs and CTs+. Simultaneously, a wider range of outcome categories are considered, discussed in more detail in Section 3.1.4.

1.2. CONCEPTUAL FRAMEWORK

The conceptual framework was initially drafted based on a comprehensive literature review and refined through input received during and following an expert workshop. The online expert workshop had the dual purpose of fostering collaboration and knowledge sharing, and enriching the first draft of the conceptual framework with new insights and knowledge gaps. It was attended by stakeholders from various sectors, including GDC, DEval, academia, and the broader international development community.[2] The conceptual framework was shaped as a result of those interactions. It delineates the theoretical underpinnings of the EGM and provides definitions for key concepts.

This EGM focuses exclusively on CT and CT+ interventions targeted to households or individuals in LMIC settings, with the aim of reducing their vulnerabilities. CTs are defined as direct and predictable monetary transfers[3] and may take one of the following forms: i) Unconditional cash transfers (UCTs), ii) conditional cash transfers (CCTs), and iii) CT+ programs, which can be subdivided into unconditional (UCTs+) and conditional (CCTs+). UCTs/UCTs+ do not require that potential recipients meet any pre-defined conditions (apart from belonging to the target population), whereas CCTs/CCTs+ are given with the requirement that the beneficiaries meet certain conditions – often related to human capital development, or health-related behavior. CTs+ combine UCTs or CCTs with one or more types of complementary support (Roelen et al.,

[2] The minutes from this workshop can be found in the supplementary material.
[3] In-kind assistance, e.g., school feeding programs and vouchers are sometimes also categorized under social transfers but do not fall into the definition of CTs used in this report. Similarly, matched savings programs or similar interventions are not included. Lottery incentives (e.g., conditional on being STI negative; see Stoner et al., 2021) are not predictable and, therefore, do not constitute CTs according to the definition used in this report.

2017). The additional components may consist of i) information, nudges, or behavior change communication (BCC) [4], ii) food transfers, iii) other in-kind transfers, iv) health care, v) psychosocial support, vi) training, vii) provision or facilitation of access to services, or, viii) other support. Plus categories are mostly defined following the categorization presented in Little et al. (2021).[5] To illustrate these intervention types, Table 1 provides examples from included studies in this EGM.

Table 1: Examples of interventions

General type of CT intervention	Specific type of intervention	Examples
CCT	CCT without plus component	The Bolsa Familia Program in Brazil (e.g., Olson et al., 2019), provides a CT to poor households with a pregnant woman or children. However, to be eligible for the transfer, children must complete vaccine schedules and/or have at least 85% school attendance, and pregnant women are required to attend prenatal appointments.
	CCT+	The Pantawid Pamilyang Pilipino Program (e.g., Derviservic et al., 2021) is a governmental program in the Philippines that offers CTs to households and organizes family development sessions on topics such as health, nutrition and responsible parenting, the attendance of which is one of the conditionalities for the CT. Households can receive additional amounts if all conditionalities are met. The conditions include health-seeking behavior for children and pregnant women, school attendance and the attendance of the family development sessions by parents or guardians.
UCT	UCT without plus component	South Africa's Child Support Grant (e.g., Coetzee, 2013) provides a monthly CT of 500 South African Rands to the caregivers of children under the age of 18, who belong roughly to the poorest 30% of the population.
	UCT+	Harris-Fry et al., 2018, describe a study in which pregnant women were allocated to different treatment arms, all having monthly "Participatory Learning and Action" women's groups. One of these treatment arms combined this women's group with a monthly CT of 750 Nepalese rupees without needing to meet any conditions.
Combinations of CTs	Either conditional or unconditional CTs and CTs+	Ujana Salama is a program attached to the Productive Social Safety Net in Tanzania. Additional to a fix amount (UCT), beneficiaries had the opportunity to receive a CCT conditional on child health and school enrolment. The intervention targeted adolescents in beneficiary households who, besides the CT, received in-person

[4] BCC is a communication strategy that encourages individuals or communities to change their current behavior to a desired behavior. Examples of BCC include home visits from health workers and community sensitization meetings.
[5] Please note that one or multiple plus categories can be combined with the CT. For example, a CT+ program can provide an asset transfer, and also offer training on productively employing the asset. Such CT+ programs therefore will fall into both the "other in-kind transfers" and "training" categories.

		livelihood and life skill training, mentoring, grants, and health services (Waidler et al., 2022).

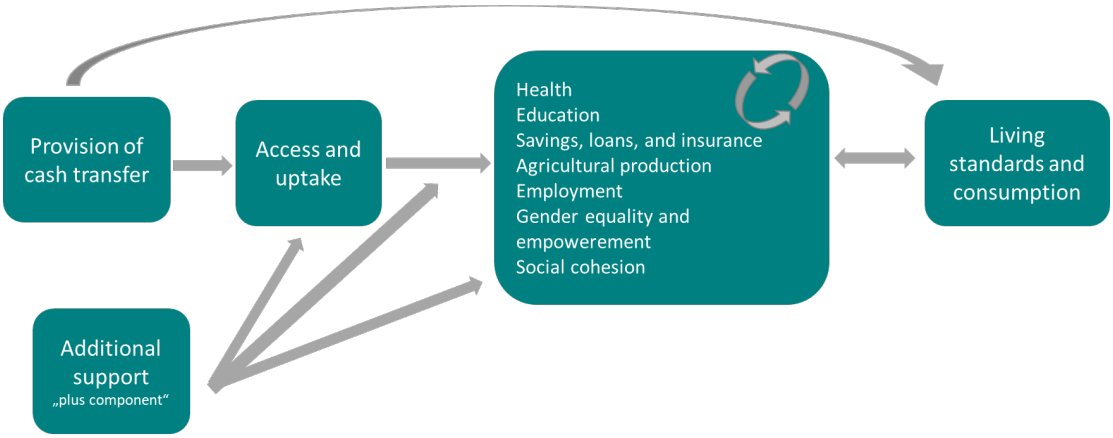
Source: Own represenation.

The EGM considers a broad range of poverty-related outcomes, reflecting the fact that CTs have the potential to affect a variety of different areas. This includes outcome categories directly related to multi-dimensional poverty, such as health, education, and living standards (OPHI & UNDP, 2022), as well as outcomes that can reduce or prevent poverty and improve health and living standards in the longer term. The latter include outcomes like knowledge and skills, ownership of (productive) assets, employment and income, as well as savings and insurance. Since the provision and targeting of resources may also affect household and societal dynamics, outcomes such as gender equality, empowerment, and social cohesion are also included. Furthermore, the EGM includes studies that consider how beneficiaries are targeted and which beneficiaries should be targeted to achieve the desired outcomes most effectively.

Figure 1 depicts a simplified representation of a theory of change (ToC) that connects CT interventions to various outcome areas. Assuming that the appropriate target population is correctly identified, has access to the CT, and chooses to participate in the intervention, CTs can potentially bring about positive effects in different aspects of development. For instance, households may utilize the extra cash to cover school fees, purchase educational materials for the children, or to spend on transport to local clinics or hospitals. Specific conditions, as seen in CCTs or additional components, in the case of CT+ interventions, can strengthen these effects. For example, with CCTs, the payment might be contingent upon households ensuring that the children attend school regularly or that pregnant women give birth in a healthcare facility. For CTs with plus programs, the transfer could be accompanied by educational information campaigns encouraging children to attend school more regularly, or additional health interventions such as vaccinations.

Short-term and long-term impacts in various domains interact and mutually reinforce one another in a positive cycle. This is shown in the two rightmost blocks of the ToC illustrated in Figure 1. While increased school enrolment and regular attendance can bring short-term benefits to individual well-being, allocating financial resources to education represents an investment in building human capital. Consequently, it can yield long-term improvements in a household's standard of living and its capacity for higher future consumption, which extends to subsequent generations.

Figure 1 Theory of change (simplified illustration)



Sources: Own illustration.

2. RESEARCH QUESTIONS

To meet the objectives described in Section 1.1, two overarching research questions (RQs) need to be answered.

1. What is the available rigorous evidence[6] on CT and CT+ interventions in LMICs?
2. What are important rigorous evidence gaps on CT and CT+ interventions in LMICs?

To address these in a comprehensive manner, this report will answer the following four specific RQs:

1. What is the available rigorous evidence and what are the evidence gaps on the effects[7] of CT interventions in LMICs?
2. What is the available rigorous evidence and what are the evidence gaps on the effects of CT interventions on the following population groups in LMICs?[8]
    a. Women and girls
    b. Children
    c. Adolescents
    d. Elderly
    e. Persons with disabilities
    f. Migrants, migrant workers, or refugees
    g. Lesbian, Gay, Bisexual, Transgender, Queer/Questioning and others (LGBTQ+)
    h. Smallholder farmers
    i. Indigenous peoples and local communities
3. What is the available rigorous evidence and what are the evidence gaps on the effects of CT interventions in LMICs with the following characteristics?
    a. Governmental or non-governmental intervention[9]
    b. Humanitarian/emergency response
    c. Fragile status of the context [10]
    d. Rural or urban setting
    e. Number of beneficiaries (<500, 501-1,000, 1,001-5,000, 5,001-10,000, more than 10,000 beneficiaries)
    f. Targeting specific population groups
    g. IT-based administration of CTs
    h. Timing of impact measurement
4. What studies focus on how to identify eligible beneficiaries or test which beneficiaries to target?

[6] "Rigorous evidence" is defined in Section 3.1.5.
[7] Here, "effects" refer to effects on outcomes that fall into the outcome categories discussed in Section 3.1.4.
[8] Note that this is based on reported effect estimates and not necessarily the targeting of the intervention. E.g., a study that does not specifically target women but discusses heterogenous effects based on sex is marked as such. Studies that exclusively target specific subpopulations are also marked accordingly.
[9] Governmental interventions are often part of a broader public social protection system. Non-governmental interventions are usually supplementary interventions that researchers, non-governmental organizations, or other non-governmental parties implement.
[10] In line with BMZ (https://www.bmz.de/de/themen/fragile-staatlichkeit/definition-18924, accessed 7.11.2022), country contexts are categorized according to the Fragile States Index (https://fragilestatesindex.org/). Due to the number of studies and contexts included in the EGM and because no information was extracted on the year in which an intervention started, it was infeasible to assign fragility status per study. Instead, contexts were assigned a status based on their most recent (2023) fragility score (https://fragilestatesindex.org/country-data/). Contexts with index values between 60.1 and 90 are categorized as "Warning," while with 90.1 and above, contexts are categorized as "Alert."

3. METHODOLOGY

Having discussed the conceptual framework and the main RQs, this section explains the technical approach followed by this EGM.

To define the scope of the EGM, the PICOS systematic search model is applied – namely by defining the Population, Interventions, Comparators, Outcomes, and Study designs of interest. The basis of the PICOS is the conceptual framework and the ToC presented in Section 1.2. Details of the criteria for inclusion, as well as the exclusion of studies, are provided in Sections 3.1 and 3.2. Section 3.3 shortly describes the electronic search over multiple databases and websites of key agencies and research institutes, which is then limited to studies published in English from 2005 onwards. Section 3.4 summarizes the approach to data management and Section 3.5 introduces the procedure for judging the quality of included SRs. Section 3.6 briefly explains how results will be visualized in the accompanying EGM.

3.1. CRITERIA FOR INCLUSION OF STUDIES

The PICOS criteria determine the eligibility of evidence included in this EGM. The PICOS model constitutes the basis for the screening protocol and the data extraction tool (provided in the supplementary materials). The PICOS model is closely based on the conceptual framework presented in Section 1.2.

To be included in this EGM, studies must adhere to all of the following criteria:

3.1.1 Population under study

Individuals and households residing in LMICs (as per Work Bank definition of 2022, see the supplementary material), irrespective of age, gender, income levels, and socioeconomic vulnerabilities, fall under the population of interest for this EGM.

3.1.2 Interventions

Studies on two different categories of interventions are included, namely:

a. CTs: Direct, regular, and predictable transfers that increase income and aim to reduce poverty
    (1) UCTs: Beneficiaries do not need to meet certain conditions in order to receive the transfers
    (2) CCTs: Beneficiaries must meet certain conditions in order to receive the transfers
b. CTs+: CTs combined with supplementary support e.g.,
    (1) Information, nudges or BCC
    (2) Food transfers
    (3) Other in-kind transfers
    (4) Health care (including referal to health care providers)
    (5) Psychosocial support (e.g., psychosocial stimulation for children)
    (6) Training
    (7) Provision or facilitation of access to services
    (8) Other support

3.1.3 Comparator

The inclusion of studies is based on the presence of one of the following comparison groups:

a) CTs vs. pure control[11]
b) CTs+ vs. pure control
c) CTs vs. "plus" component only
d) CTs+ vs. "plus" component only
e) CTs vs. CTs+
f) Different modalities or types of CTs or CTs+ to each other (with or without a pure control group). This includes supplementary interventions aimed at improving targeting, uptake, or the receipt of CTs.[12]

3.1.4 Outcomes

Table 2 lists the eight outcome categories and 37 subcategories that are included.

[11] Pure control means that participants do not receive any kind of CTs or CTs+ interventions. Yet, in IEs, information is collected for this group to serve as "counterfactual", i.e., as comparison group.
[12] Note that this categorization excludes different lengths of exposure to the same modality or type of CT. Phase-in designs of the same CT (the same amount, frequency, conditionality, plus component, and method of disbursement) are therefore excluded from this EGM.

Table 2: Outcome categories and subcategories[13]

Health	Education	Living standards and consumption	Financial inclusion, savings, and insurance	Agricultural production	Employment and entrepreneurship (non-agricultural)	Gender equality and empowerment	Social cohesion
Child health	Learning and achievement (test scores, literacy, cognitive development, completion/graduation)	Housing, electricity and water, sanitation, and hygiene (WASH) infrastructure	Financial literacy	Land ownership and/or used	Vocational training and technical skills	Reduction of gender-gaps	Social capital
Sexual, reproductive, and maternal health	Access to education (enrolment, attendance, dropouts, and truancy)	Household income and expenditures	Financial services	Agricultural assets and investments	Entrepreneurial skills	Production decisions	Creation of networks
Mental health and wellbeing	Attitudes and personal development	Food security	Cash savings	Adoption and knowledge of agricultural technologies and practices	Employment or self-employment	Control over household resources and income	Trust in others and institutions
Access and use of health services		Other non-food consumption, consumption in general, non-productive or household assets	Insurance	Agricultural yield and income	Productive (non-agricultural) assets and investments	Leadership	Attitudes towards other groups
Nutrition					Entrepreneurial income	Time-use	Civic engagement
Other health outcomes					Child labor or work		

Source: Own representation.

[13] Examples for some of the outcomes in the subcategories can be found in Appendix C.

3.1.5 Study design

Three different types of rigorous study designs are considered, namely

i) Experimental designs, such as cluster and individual randomized controlled trials
ii) Quasi-experimental designs; this EGM includes difference-in-difference, instrumental variables and regression-discontinuity-designs. It also includes matching methods, such as propensity score matching and synthetic control methods.
iii) Systematic evidence syntheses (SRs and meta-analyses) that only include quantitative studies following one of the two study designs described above.[14]

The minimum sample size for inclusion of a study is 30 per treatment or comparison group.[15]

3.1.6 Setting of the intervention

Only interventions taking place in LMICs, according to the World Bank criteria for 2022, are considered.

3.1.7 Timing of the outcome measurement

Outcomes must not be measured immediately after the intervention (as is usually the case with lab experiments) but with a minimum delay of one month.

3.1.8 Language of publication

Limitation to evidence presented in English.

3.1.9 Publication date and type

Limitation to evidence published from 2005 onwards. Studies of the following type are considered eligible for inclusion:

i) Published peer-reviewed articles
ii) Reports/Grey literature (e.g., discussion papers/working papers)
iii) Theses (Bachelor's, Master's, PhD)

3.2. CRITERIA FOR EXCLUSION OF STUDIES

Studies published before 2005 are excluded, as well as those which are not presented in English. In addition, this EGM excludes evidence with any of the following characteristics:

3.2.1 Population under study

Studies including only high-income countries (HICs), as well as studies including both HICs and LMICs (as defined above), but not presenting disaggregated data on LMICs.

3.2.2 Intervention

Studies of social protection programs that do not include a CT component.

[14] SRs where quantitative evidence can be clearly distinguished from qualitative evidence are also included in this EGM.
[15] A commonly referred general rule for central limit theorems to hold is a sample size of 30 (see e.g., Chang et al., 2006). Further, the sample size required to detect an effect size of one standard deviation is approximately 30.

3.2.3 Comparator

Studies without an explicit comparator (see Section 3.1.3), even if the study design is valid and/or the outcomes and interventions are relevant.

3.2.4 Outcomes

Studies that assess effects on outcome categories other than those stated in Section 3.1.4, specifically outcomes such as crime or political attitudes (as far as these are not related to or discussed in the context of social cohesion).

3.2.5 Study design

Studies with outcomes measured immediately or less than one month after the intervention (e.g., lab experiments). In addition, there are five overarching categories for excluded methodologies:

a) Quantitative designs that do not use the (quasi-)experimental designs as defined in Section 3.1.5, including:
   i) Granger causality
   ii) Correlation analysis
   iii) Cross-sectional studies
   iv) Cohort designs
   v) Random effects
   vi) Input-output models
   vii) General equilibrium models
   viii) Theoretical, modeling, and simulation studies
   ix) Case-control studies, controlled before and after studies
   x) (Interrupted) time series designs
b) Qualitative studies, including but not limited to:
   i) Ethnography
   ii) Grounded theory
   iii) Phenomenology
   iv) Qualitative case studies
c) Traditional narrative reviews
d) Opinion pieces, editorials, perspectives
e) Non-SRs

3.3. SEARCH STRATEGY

The electronic search for relevant evidence was based on the PICOS model described via the inclusion and exclusion criteria in the previous section (Section 3.1). Multiple search terms were combined using Boolean logic: "OR" is used between different terms within the same category, while "AND" combines different categories of search terms to form a single query.[16] Searches were performed over the following databases and agency/research institute websites:

[16] An overview of the search strategy employed can be found in Appendix A, and the exact search strategy for each database can be found in the supplementary materials.

Databases:

- EconLit (via EBSCO)
- Web of Science (Social Sciences Citation Index)
- Scopus
- Cochrane Database of Systematic Reviews
- JSTOR
- Google Scholar
- International Initiative for Impact Evaluation: 3ie Development Evidence Portal

Websites of agencies and research institutes:

- Campbell Collaboration
- National Bureau of Economic Research

These databases and websites differ regarding their requirements for search queries. Given these differences, the search terms were customized for each search engine, aiming to maintain a high degree of comparability between the individual searches. The search terms used for each search engine, as well as their limitations, are described in the supplementary material and Appendix A respectively. The search was limited by language and time period, meaning that only evidence published in English from 2005 onwards was searched for.

The assessment of the relevance of the search strategy regarding the search objective was based on the inclusion rate of benchmark studies in the final search results. The list of benchmark studies was proposed by C4ED (with the support of sector experts) and approved by DEval. Refinement of the search strategy continued until any of the following conditions are satisfied: (i) among a maximum number of 10,000 studies retrieved from all databases, at least 2/3 of benchmark studies were found, or (ii) major refinements in the search have been made and at least ½ of benchmark studies were found within the aforementioned number of studies.

3.4. DATA MANAGEMENT

There are three stages to data management, which are described below. The results of the search strategy described in Section 3.3 were uploaded to the EPPI Reviewer 4 software, which allowed for easy collaboration between reviewers.

In the pilot phase, 131 titles and abstracts were double-screened. Reviewers attained an overall agreement rate of 85% (more details are provided in Appendix B). All disagreements were resolved through discussion among reviewers and third-party arbitration. Given the satisfactory agreement rate, single screening was conducted until saturation, defined as including not more than one paper for every 100 screened papers. The remaining titles and abstracts were discarded, assuming that the EPPI priority screening algorithm moved the least relevant studies to the end of the screening sample.

The full texts of the titles and abstracts included in the previous stage were retrieved and uploaded to EPPI Reviewer 4. At this stage, double screening of full-texts was conducted for 59 papers. The intention of this piloting phase was to ensure a common understanding of the inclusion criteria. As shown in Appendix B, reviewers achieved an agreement rate of 83%. All reviewer differences were resolved through discussion and, if necessary, third-party involvement.

All further papers after this pilot were single-screened by one reviewer per paper. The screening protocols for title and abstract and full-text can be found in the supplementary material.

Following full-text screening, included studies underwent a data extraction phase, whereby the information required to answer all RQs were extracted.[17] In total, 21 IEs were double-extracted to ensure consistency in using the data extraction tool. All disagreements were discussed and resolved among reviewers. Following this piloting phase, all remaining studies were single-extracted. Any uncertainties that reviewers faced after the piloting phase were jointly discussed and resolved by all reviewers. Similarly, 13 SRs were double-extracted to ensure a consistent answering of the quality assessment tool, after which the remaining SRs were single-extracted.

3.5. QUALITY ASSESSMENT

Given the broad scope of this EGM, the quality assessment is only applied to the SRs. The confidence in the quality of each SR was appraised using the 3ie Supporting the Use of Research Evidence (SURE) checklist. The SURE checklist for SRs consists of three sections. Section A assesses the level of confidence in an SR's search strategy (high, medium, or low confidence), whereas Section B assesses the level of confidence in an SR's methods of analysis. Section C enquires after various additional negative and positive factors that could mediate the overall rating of the SR. Sections A and B consists of multiple subsections that pose various screening questions and checklists. The overall assessment of a review's reliability (again, high, medium, or low confidence) was determined based on the confidence levels of Sections A and B, as well as various mitigating factors considered in Section C. The supplementary material presents the full checklist and the method by which overall confidence levels were assigned to SRs.

3.6. VISUALIZATION OF RESULTS

The EGM provides a visual representation of the studies that made it through the full-text screening and data extraction process. This map is two-dimensional, with rows representing different types of interventions and columns representing various outcome subcategories as outlined in the PICOS criteria. Each cell on the map corresponds to a specific combination of intervention and outcome subcategories. In the cell there are up to four bubbles: one for IEs and one each for SRs of low, medium and high quality. The size of the bubble reflects the quantity of available evidence for that combination. Users can apply filters to the map, which means the map will only display studies that meet the chosen criteria, such as particular study characteristics or the reporting of effects for specific subpopulations. Additionally, both intervention and outcome subcategories can be collapsed to only show the aggregation at the higher category level. It is possible to collapse individual categories without collapsing all, and studies that are present in more than one subcategory of a given category will only appear once when collapsed.

[17] The employed data extraction form can be found in the supplementary material.

4. RESULTS

This EGM report, supplementing the EGM itself, presents a brief descriptive analysis addressing the RQs. This section summarizes the characteristics of the body of evidence (Section 4.1) and addresses each RQ in turn (Sections 4.2 to 4.5). The descriptive analysis will pool UCT and UCT+ interventions (referred to as "UCT") and CCT and CCT+ interventions (referred to as "CCT") and focus on the eight outcome categories (cf. Table 2). In addition, this section includes a short descriptive assessment of the quality of the included SRs (Section 4.6).

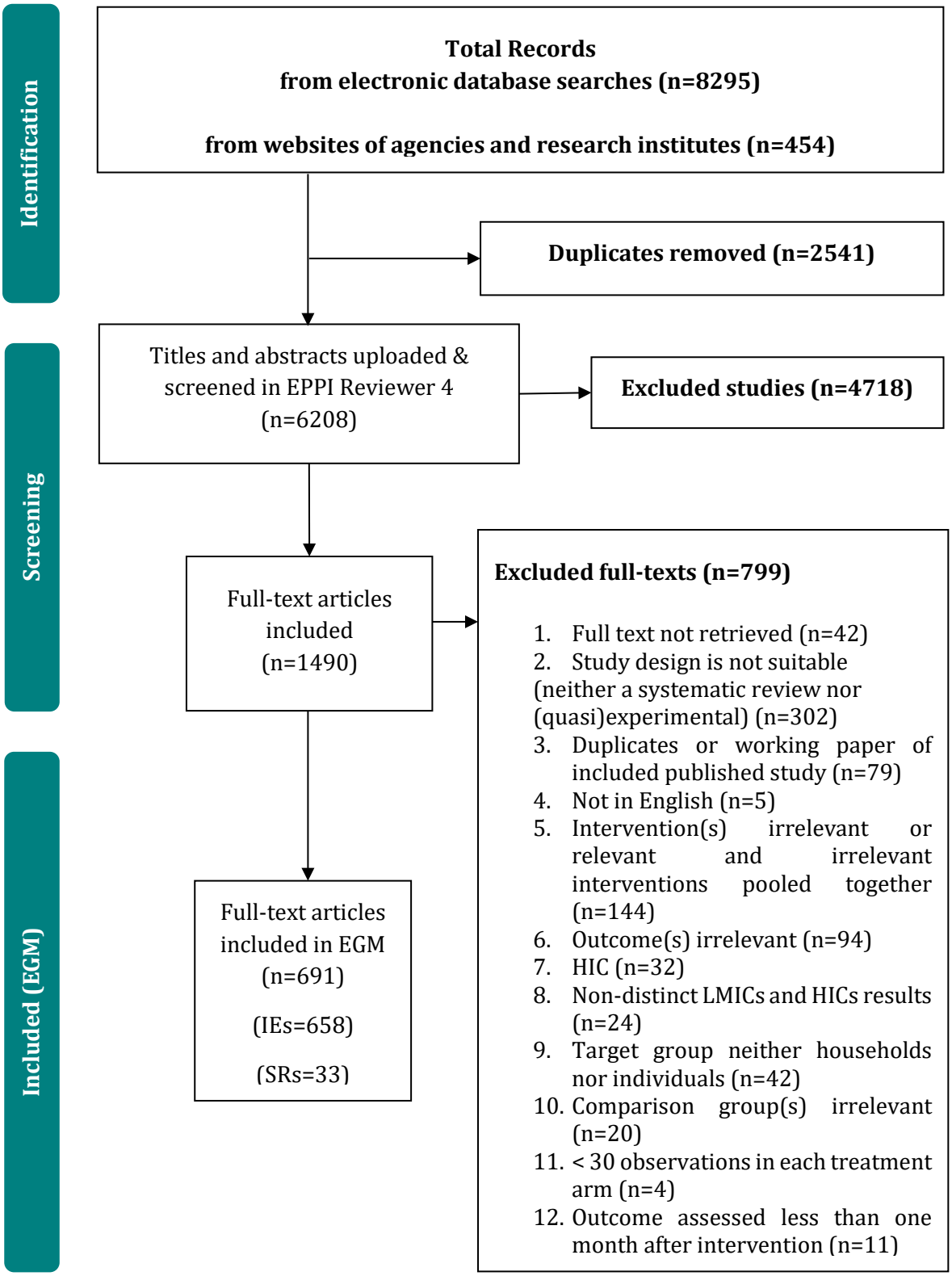
4.1. SEARCH RESULTS

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram, presented in Figure 2, depicts the number of studies at the different screening stages. Across all databases and websites, the search round presented 8,749 records, which was reduced to 6,208 records after deduplication.[18] These records were imported into EPPI Reviewer 4 to facilitate a collaborative screening and data extraction. From those, 4,718 records were excluded during title and abstract screening. Following title and abstract screening, 1,490 records remained for full-text screening. 799 further records were excluded at this screening stage. Following the full-text screening, 691 studies – thereof 658 IEs and 33 SRs – were included in the final set of records. The reasons for exclusion at each stage are presented in Figure 2.

Some IE records contain more than one evaluation. This usually occurs when the article evaluates multiple programs, often each in a different country. Therefore, each of these investigations of a program was considered a separate evaluation. This means that, while 658 IEs are included in this EGM, the evaluations included amount to a larger number with 709 evaluations. Because the EGM showcases these 709 evaluations and 33 SRs, the 709 evaluations will be referred to as "studies" throughout the remainder of the report.

[18] The precise search strategy employed and the number of studies found from running the search strategy over each database and website (listed in Section 3.3) are presented in Appendix A.

Figure 2: PRISMA diagram



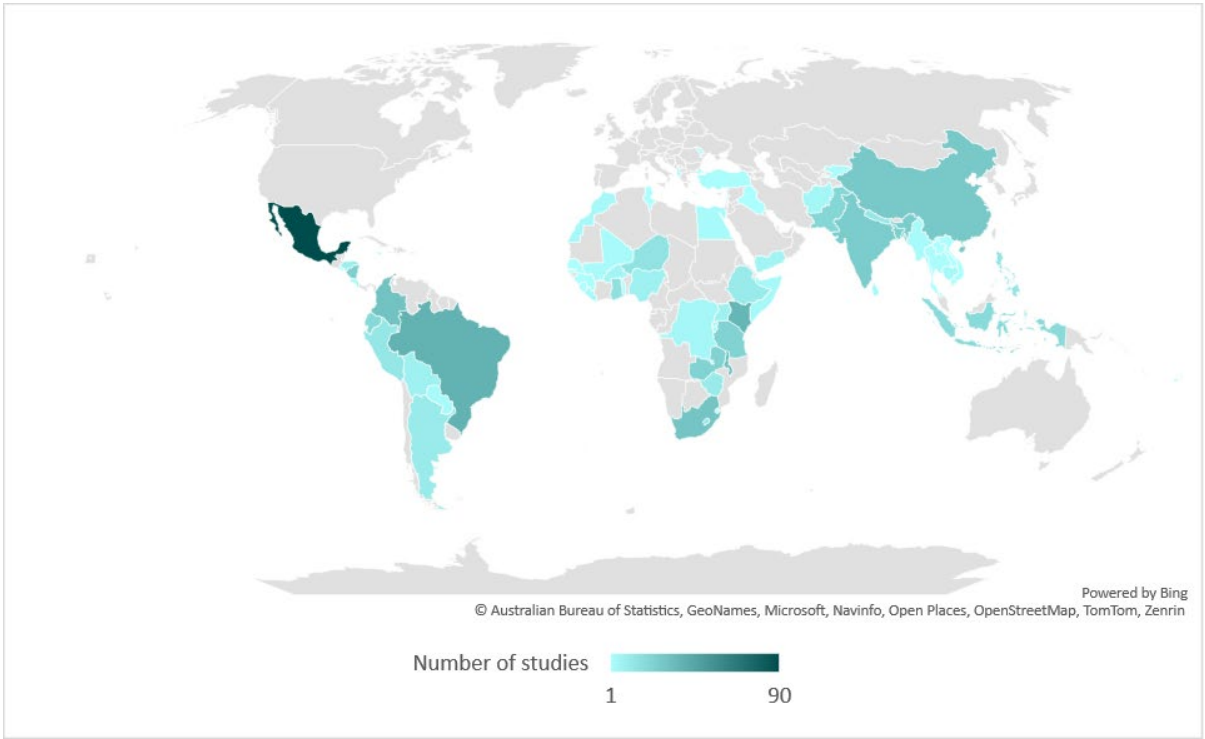
From Page et al., (2021). For more information, visit: http://www.prisma-statement.org/.
Sources: Own review.

4.2. AVAILABLE EVIDENCE ON CT INTERVENTIONS (RQ 1)

4.2.1 Geographic distribution of studies

As shown in Figure 3, most of the studies are conducted in either Sub-Saharan Africa (SSA) (n=263) or Latin America and the Caribbean (LAC) (n=255).[19] Within SSA, Kenya (n=37), Malawi (n=32), and South Africa (n=30) are particularly well-studied. Within the LAC region, Mexico (n=90), Brazil (n=39) and Colombia (n=31) are the most-studied countries. 93 studies were conducted in South Asia, quite evenly distributed between India (n=26), Bangladesh (n=25), and Pakistan (n=21) and 73 studies were done in East Asia and the Pacific (EAP), with 28 studies conducted in China, 19 studies in Indonesia, and 14 studies in the Philippines. The scope of evidence on the impact of CTs is smaller for both the Middle East and North Africa (MENA) (n=20) and especially Europe and Central Asia (ECA) (n=7). Within the MENA region, most studies are conducted in Yemen (n=6), closely followed by Lebanon (n=5). Within the ECA region, two studies each are conducted in Albania, the Kyrgyz Republic, and Turkey, with one study conducted in Moldova. LMICs without any rigorous evidence on CTs include Malaysia in EAP, Georgia in ECA, the Caribbean islands in LAC (aside from Jamaica), Algeria in MENA, Bhutan in South Asia, and Mozambique in SSA.

Figure 3: Geographic distribution of studies



Source: Own illustration. The map does not include the SRs as they include studies from more than one country/region.

The SRs follow a similar regional distribution as the IE studies. 29 SRs include studies from LAC and 28 SRs comprise studies from SSA.[20] South Asia is relatively well-represented as well,

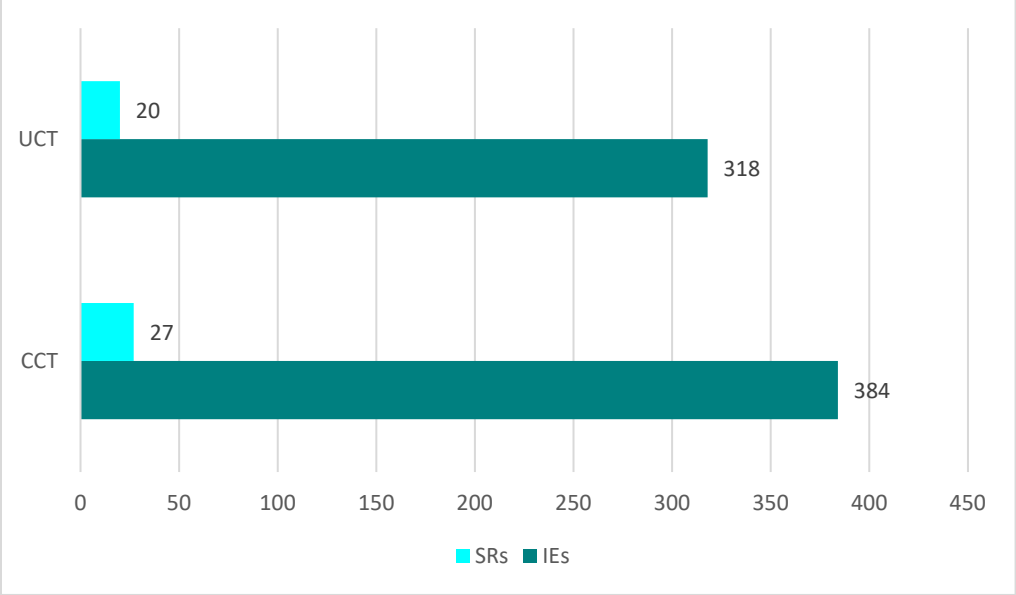
[19] Note that one of the included studies (Dwyer et al., 2022) presents pooled results from three countries and is counted three times. The geographic distribution of studies, therefore, includes 711 entries.
[20] Note that these numbers add up to more than 33 because one SR can include studies from more than one region.

with 23 SRs from this region. Other regions are less represented; 15 SRs contain studies from EAP, eight SRs include studies from MENA, and five SRs include studies from ECA.

4.2.2 Intervention characteristics[21]

Overall, CCT interventions – that is CCTs with or without a plus component – are the most frequently studied intervention type. They also occur more frequently than UCT interventions in the SRs**.** As some studies do not specify the nature of the CT intervention, a separate category is created ("CT without specification"). However, due to the unspecified nature of these interventions, they are not further discussed here.[22] Figure 4 highlights how UCT and CCT interventions are distributed across both IEs and SRs.

Figure 4: Distribution of intervention types



Source: Own illustration. The figure omits CTs without specification, as these CTs are not analyzed further in the text.

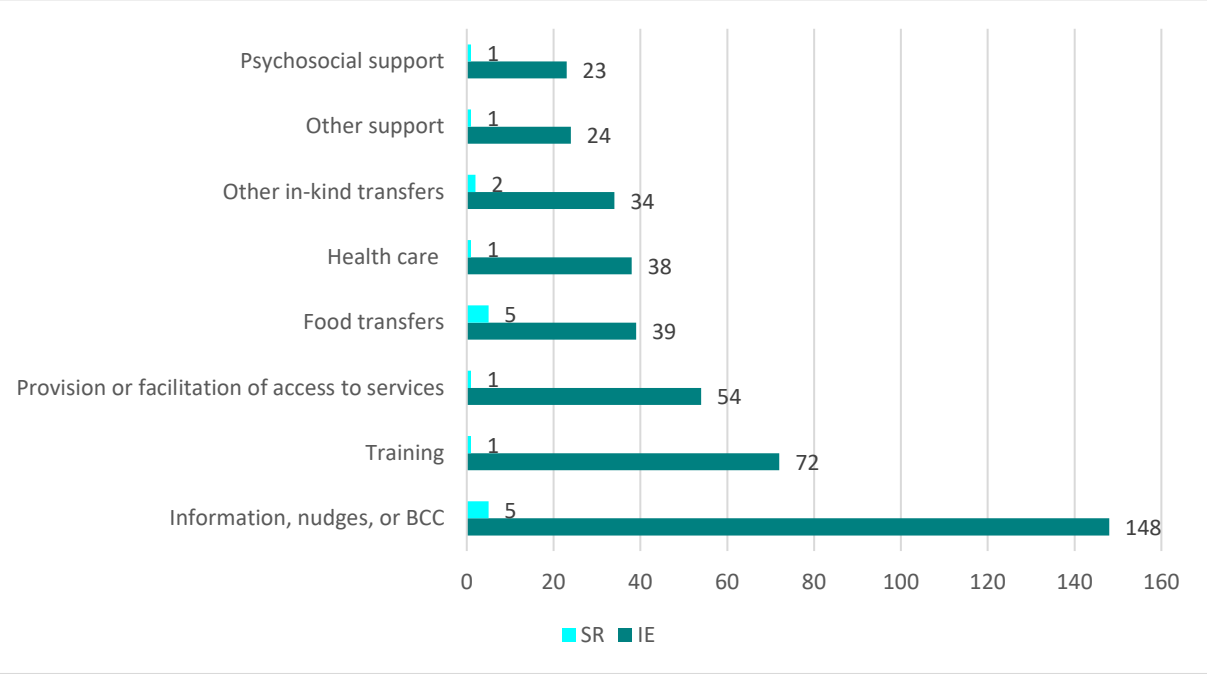
CCTs are more likely to be provided with plus components than UCTs. In total, 276 studies and nine SRs contain interventions with at least one plus component.[23] While 93 studies investigate UCTs with at least one plus component, the number of CCTs with plus components is nearly twice as high at 173 studies. Figure 5 shows the number of studies investigating different types of plus components.[24]

"Information, nudges, or BCC" is the most common plus component provided (n=148) for the IEs, followed by "Training" (n=72). "Psychosocial support" (n=23) is provided the least frequently. Some differences in evidence patterns occur for SRs, as compared to IEs. Of these, "Information, nudges, or BCC" and "Food transfer" are the most common plus components (n=5 each). Therefore, CTs with food transfers seem to be particularly well-studied in SRs compared to their relative frequency in IEs. Furthermore, "Training" and "Provisioning or facilitation of access

[21] As per the inclusion criteria, all IE studies contain a CT intervention, either in the form of a CCT or a UCT (or "unspecified" when this could not be established), with or without a plus component. As each study can report on more than one type of CT, the total count is larger than the number of studies.
[22] Only a small number of studies report on CT without specification interventions (n=29).
[23] One SR can contain more than one plus component, meaning that the total number of plus components do not add up to nine.
[24] As some interventions use several plus components, the total number of plus components in Figure 5 (n=432) exceeds the total number of papers with a plus component (n=276).

to services" components, which are fairly well-studied in the IEs, are only represented in one SR each.

Figure 5: Distribution of plus components



Source: Own Illustration.

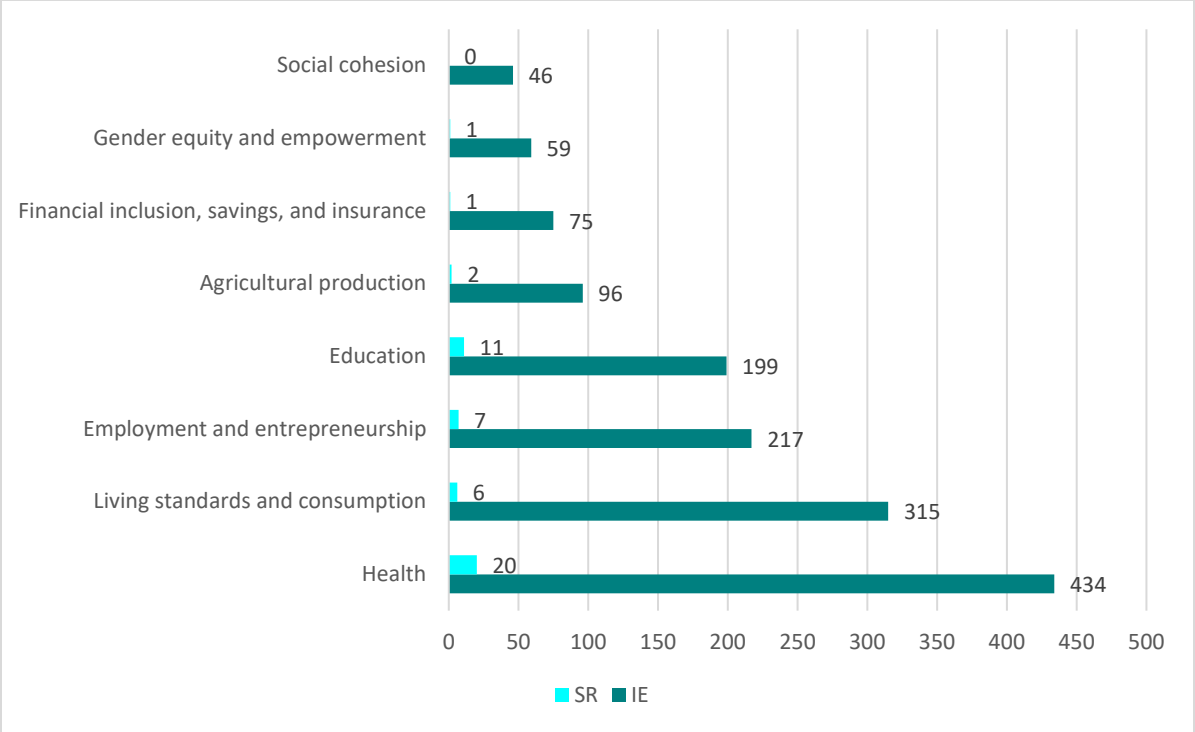
4.2.3 Outcome characteristics

"Health" (n=434) and "Living standards and consumption" (n=315) are the categories most frequently studied in the IEs.[25] These two outcome categories are followed by "Employment and entrepreneurship" (n=217) and "Education" (n=199). Evidence is relatively sparse (compared to other outcome categories) for "Financial inclusion, savings, and insurance" (n=75), "Gender equality and empowerment" (n=59), and "Social cohesion" (n=46) (see Figure 6).

There are only small differences in the evidence pattern between IE studies and SRs. Eleven SRs report "Education" outcomes, which means that this category is relatively well-studied in SRs in comparison to IEs.[26] Reporting on other outcomes follows the same pattern for SRs and IEs. Most of the SRs include effects on "Health" outcomes (n=20), making it a particularly well-studied outcome category. Seven and six SRs report effects on "Employment and entrepreneurship" and "Living standards and consumption", respectively, making these two categories well-studied in both IEs and SRs (although less so than the "Health" category). As with IEs, evidence is very limited for "Financial inclusion, savings and insurance" and "Gender equality and empowerment" with each outcome category only reported in one SR. No SRs report on outcomes in the "Social Cohesion" category.

[25] One study (either IE or SR) can report on more than one outcome category, meaning that the number of outcome categories reported does not sum up to the number of studies.
[26] This difference in SRs' thematic frequencies (compared to the IEs) may also explain the discrepancy in the plus intervention components studied in SRs. It might be that SRs are aligned to interest in particular thematic areas, where only particular types of intervention bundles are found.

Figure 6: Distribution of outcomes



Source: Own Illustration.

4.2.4 Combinations of interventions and outcomes

Within IEs, "Health" is the most studied outcome category across UCTs and CCTs (n=198 and n=234, respectively). "Living standards and consumption" is the second-most studied outcome for UCTs (n=198). For CCTs, "Education" (n=131) is the second most frequently studied outcome category after "Health" (n=234).[27] This order of frequency for CCTs is also reflected in the SRs, which are most frequently found for outcome categories "Health" and "Education", with 20 and eleven SRs, respectively.

At least some evidence exists for all intervention-outcome combinations and most of the CCT interventions are combined with "Health" and "Education" outcomes, meaning that this intervention is paired less often with other outcomes than UCTs. The biggest gaps in evidence concern the effects of UCTs and CCTs on "Social cohesion" (n=17 for CCTs and n=25 for UCTs). These are followed by the respective effects on "Gender equality and empowerment" (n=29 for CCTs and n=30 for UCTs). In a similar range are the effects of CCTs on "Agricultural production" and "Financial inclusion, savings and insurance", which are studied by 30 and 35 IEs, respectively.

[27] UCTs are more about combatting poverty, while CCTs are designed to augment social development. Income, expenditures and food security all fall under "Living standards and consumption", while conditionalities of CTs often surround health-seeking behavior and education for children as these are generally seen as very important levers to break the cycle of intergenerational poverty transmission.

Table 3: Cross-frequencies of impact evaluation studies

Outcome / Intervention	Health	Education	Living standards and consumption	Financial inclusion, savings and insurance	Agricultural production	Employment and entrepreneurship	Gender equality and empowerment	Social cohesion	**Total**
Unconditional cash transfer	198	76	198	44	59	115	30	25	318
Conditional cash transfer	234	131	109	30	35	105	29	17	384
Cash transfer without specifications	16	3	14	2	7	4	2	4	29
Total	434	199	315	75	96	217	59	46	**709**

Table 4: Cross-frequencies of systematic reviews

Outcome / Intervention	Health	Education	Living standards and consumption	Financial inclusion, savings and insurance	Agricultural production	Employment and entrepreneurship	Gender equality and empowerment	Social cohesion	**Total**
Unconditional cash transfer	12	6	4	1	1	5	0	0	20
Conditional cash transfer	16	9	4	0	0	5	1	0	27
Cash transfer without specifications	6	2	1	0	1	2	0	0	9
Total	20	11	6	1	2	7	1	0	**33**

Sources: Own representation.

4.3. AVAILABLE EVIDENCE ON CT INTERVENTIONS FOR VARIOUS POPULATIONS (RQ 2)

The following subsection describes the evidence on CT interventions for different types of vulnerable populations. This evidence is derived from the IEs and was not extracted from the SRs.[28] As the information was not extracted for SRs, only the existing absolute evidence gaps for IEs are discussed but not any synthesis gaps.

There is abundant evidence for women and for vulnerable age groups. Plenty of studies report the effects of CT interventions for women and girls (n=360), children (n=209), adolescents (n=250), and the elderly (n=90). For women and girls, children, and adolescents, more evidence exists on CCTs, while more evidence exists for the effects of UCTs on the elderly. "Health" outcomes are most frequently reported for women and girls and the elderly, whereas "Education" outcomes are most frequently reported for children and adolescents.

There are absolute evidence gaps for indigenous peoples and local communities (n=10), persons with disabilities (n=8), and for the LGBTQ+ community (n=0). Other vulnerable population groups for which there is relatively more evidence include smallholder farmers (n=16) and migrants, migrant workers, and refugees (n=13).

4.3.1 Women and girls

More than half of the included studies report outcomes for women and girls. 360 of the 709 studies report outcomes measured for women and girls exclusively, or present results disaggregated by gender. From those 360 studies, 233 studies report outcomes for girls younger than 18 years old, whereas 159 studies report outcomes for women 18 years or older.[29] 171 studies report results for women in pooled age groups.

Evidence for the impact of CT interventions on women and girls is larger for CCTs than for UCTs and concentrates on "Health" and "Employment and entrepreneurship" outcomes. Within the 360 studies reporting outcomes for women and girls, the most evidence exists for CCTs, present in 64% of studies (n=229), followed by UCTs, which occur in 40% of the studies reporting outcomes for women and girls (n=145).[30] CCTs are therefore more represented for women and girls than in the overall sample of studies, with 54% of studies reporting on CCTs (384 out of 709 studies). UCTs are underreported for women, given that 45% of the overall sample report on this intervention type (318 out of 709 studies). 63% of studies measuring the effects on women and girls inspect "Health" outcomes (n=228). This is mostly driven by reporting on the "Sexual and reproductive health (SRH)" outcome subcategory (n=131). The second most abundant outcome category is "Employment and entrepreneurship", present in 36% of the studies (n=130) that report outcomes for women and girls.

There are also notable evidence gaps for this subpopulation. The least evidence exists on "Financial inclusion, savings and insurance", and "Social cohesion" outcomes. Only 10% and 8% of studies report these outcomes, corresponding to 37 and 30 studies, respectively.

[28] SRs aggregate findings from a wide array of primary studies, who do not always report outcomes for specific population groups. SRs report the pooled results across the wide range of population groups considered in the primary studies. Extracting information for specific population groups for SRs is, therefore, outside the scope of this project.
[29] Note that these two numbers do not add up to 360 as one study could report results for girls that fall into both age categories, i.e., girls above and below the age of 18.
[30] Note that the number of interventions does not add up to 360 because one study could contain both CCT and UCT intervention arms.

4.3.2 Children

There is an abundance of evidence for the impact of CCT interventions on "Education" and "Health" outcomes for children. Outcomes for children (6-13 years old) are measured in 209 studies, the largest proportion of which come from CCTs (64%, n=133). CCTs are therefore over-represented for this group compared to the frequency with which CCTs are studied in the overall sample (54%, or 384 out of 709 studies, as shown in Figure 4). As expected, many CCT programs set conditionalities on school attendance and clinic visits for children, meaning that the impact of the CCTs on adherence to these conditionalities are particularly well-studied areas.[31] A large majority of studies (71%, n=149) report "Education" outcomes, driven by the "Access to education" outcome subcategory (n=136). Additionally, 50% of studies report effects on "Health" outcomes (n=104), particularly "Child health" (n=64) which is not surprising given the population type.

Evidence on UCTs (compared to CCTs) on children is relatively scarce. The same is true for outcomes that are not directly relevant for children. 41% of studies (n=86) report on UCT interventions. In terms of outcomes, it is not surprising that there is a relative lack of evidence for "Financial inclusion, savings, and insurance" (11% of studies), and "Social cohesion" (5% of studies), corresponding to 24 and ten studies, respectively. However, evidence is also lacking for the "Gender equality and empowerment" outcome (9%, n=19), which may be of more relevance to children, and to girls in particular.

4.3.3 Adolescents

Evidence for the adolescent subpopulation follows the same pattern as the evidence for children. 250 studies report outcomes measured for adolescents (14-17 years old). CCTs are well-represented, with 62% of studies reporting results for adolescents using CCTs (n=154). CCTs are therefore over-represented also for this subpopulation compared to the overall sample (occurring in 54%, or 384 of the 709 studies). In terms of outcomes, most studies again describe 'Education' (60%) and "Health" outcomes (56%), corresponding to 151 and 139 studies, respectively. As with children, "Education" outcomes are predominantly made up of the "Access to education" outcome subcategory (n=136). "Health" outcomes are predominantly made up of the "SRH" outcome subcategory (n=75), which is not surprising as SRH studies largely focus on the adolescent population. The "Access to education" outcome subcategory is also well-studied for this subpopulation, whereas "Health" outcomes start shifting to outcome subcategories other than "Child health", notably "SRH".

The evidence gaps for adolescents are similar to those for children. In terms of interventions, UCTs are relatively understudied for this subpopulation, included in 42% (n=106) of the studies reporting results for adolescents. This compares with 45% of studies in the overall sample that includes evidence on UCTs (318 out of 709 studies, see Table 3). Outcome categories "Gender equality and empowerment" (10%) and "Social cohesion" (6%) are also understudied for this subpopulation, corresponding to 26 and 14 studies, respectively.

4.3.4 Elderly

There is a sizeable amount of evidence for interventions and outcomes that are directly relevant for the elderly subpopulation. 90 studies report outcomes measured for the elderly

[31] Examples are the *Familias en Acción* CCT in Colombia, conditioned on regular clinic visits for children up to seven years old (Attanasio et al., 2015). PROGRESA/Oportunidades/Prospera CCT in Mexico is conditioned on regular clinic visits for infants and on children's regular primary and secondary school attendance (Avitabile, 2021).

(65+ years old). In terms of interventions, UCTs are particularly well-studied. 62% of the studies reporting outcomes for this group report the outcomes of UCT interventions (n=56). The high percentage for this subpopulation is related to the receipt of pensions, which tend to be provided unconditionally.[32] Only one of the UCT interventions includes a plus component, indicating an absolute evidence gap in this respect. In terms of outcomes, most studies (61%) examine "Health" outcomes (n=55). Within this outcome category, most studies investigate CTs' impact on the elderly's "Access and use of health services" (n=31). 54% of studies report outcomes in the "Employment and entrepreneurship" category for this subpopulation (n=49). Of these studies, almost all investigate the impact of CTs on "Employment or self-employment" outcome subcategory (n=48).

Evidence gaps occur for this subpopulation in reporting on CCTs and reporting on the same outcome categories as for women and girls, children, and adolescents. CCTs are investigated in 36% of studies reporting outcomes for this subpopulation (n=32). In terms of outcomes, "Social cohesion" is reported in 10% of studies (n=9), while "Gender equality and empowerment" is reported in 8% of studies (n=7).

4.3.5 Persons with disabilities

Few studies report results for persons with disabilities. Overall, only eight studies report outcomes for this particular subpopulation. Studies that do report on this subpopulation mostly investigate UCT interventions (n=7). One other study inspects a CT without specification. Regarding the outcomes being investigated, there are no clear patterns. Half of the included studies present results on "Health", "Education", and "Living standards and consumption" (n=4 each). Three studies report outcomes in "Employment and entrepreneurship" (38%). One study each explore effects on "Agricultural production" and "Social cohesion" (13%) outcomes.

There are absolute evidence gaps for the impact of CCTs and for the effects on a range of different outcomes for persons with disabilities. No studies with outcomes for this subpopulation investigate CCTs, indicating a total lack of evidence for these interventions. In addition, no studies report the effects on "Financial inclusion, savings, and insurance", or "Gender equality and empowerment" for persons with disabilities. This indicates that there are absolute evidence gaps on these outcomes for the persons with disabilities subpopulation.

4.3.6 Migrants, migrant workers, or refugees

Evidence is lacking for the subpopulation of migrants, migrant workers, or refugees (n=13). The little available evidence for this subpopulation centers around UCT interventions (85%, n=11) and "Living standards and consumption" (85%, n=11), and "Health" (69%, n=9). While only consisting of four studies in absolute terms, 31% of these studies report results for "Social cohesion", which is notably higher than the 6% of studies in the overall sample of included studies reporting on social cohesion (46 of 709 studies). Two studies each present outcomes in the "Trust in others and institutions" and "Attitudes towards other groups" subcategories.

There are glaring shortages of evidence in CCTs and in gender and agricultural outcomes for this subpopulation. Only three studies consider CCT interventions (23%). Evidence for this intervention type is therefore lacking for migrants, migrant workers, or refugees, relative to the evidence for CCTs in the overall sample of studies. With regards to outcomes, only two studies

[32] Some (non-exhaustive) examples include the Old Age Pension in South Africa (see e.g., Gelo et al., 2023), the Programa 70 y Más in Mexico (see e,g., Aguila et al., 2020), and the New Rural Old Age Insurance in China (see e,g,, Chen et al., 2020).

explore "Gender equality and empowerment" outcomes (15%), and only one study inspects "Agricultural production" (8%).

4.3.7 LGBTQ+

The EGM includes no study reporting outcomes for people identifying as LGBTQ+, meaning there is an absolute evidence gap for this subpopulation. This could be due to the politically sensitive nature of asking individuals whether they belong to this subpopulation.

4.3.8 Smallholder farmers

Relatively little evidence exists for smallholder farmers, with only 16 studies reporting on this subpopulation. While some of the populations under study may have comprised mostly smallholder farmers in some studies, particularly in rural areas, they were not clearly and explicitly identified as such, which would lead to these studies not being included under this population category.

This subpopulation has a different distribution of evidence than the overall sample, especially related to the outcomes reported. Of the 16 studies reporting results for smallholder farmers, 69% use UCTs (n=11), making this the most studied intervention (whereas CCTs are the most studied in the overall sample). As expected for this subpopulation, "Agricultural production" is the most frequent outcome category, being reported in 81% of studies (n=13), as opposed to the overall sample, in which outcomes for "Agricultural production" are only rarely reported (14%, or 96 out of 709 studies). For smallholder farmers, reporting in this outcome category is mostly driven by the "Agricultural assets and investment" outcome subcategory (n=12). 69% of studies delve into "Living standards and consumption" (n=11) outcomes. Evidence is also highly concentrated in certain regions, with 69% of the studies investigating outcomes for smallholder farmers set in SSA (n=11).

There are some noticeable evidence gaps for this subpopulation, especially for outcomes related to education. The outcome categories "Education" and "Social cohesion" are each represented by two studies (13%). Only one study describes "Gender equality and empowerment" outcomes (6%). The absence of evidence for "Education" is noticeable, given that this outcome is reported in 28% of the studies in the overall sample (199 out of 709 studies. See Table 3). Furthermore, evidence for this subpopulation is lacking outside of the SSA region.

4.3.9 Indigenous peoples and local communities

Overall, the evidence is scarce for indigenous people and local communities (n=10). Contrary to the authors' expectations, CCTs are mostly studied for this subpopulation. Within the ten studies reporting results for this subpopulation, 70% (n=7) make use of CCT interventions. Given the high vulnerability status of this population and that monitoring and verification processes to ensure compliance with the conditionalities can be quite costly for this specific subpopulation, this finding is striking. Similarly, 70% of studies report "Education" outcomes for indigenous peoples and local communities (n=7), indicating that this is a well-researched outcome, relative to the frequency with which "Education" appears in the overall sample, and relative to other outcomes reported for this subpopulation. Half of the studies describe "Health" outcomes (n=5).

There are absolute evidence gaps for this subpopulation in terms of outcome categories, and there is no evidence outside the LAC region. When observing the outcome categories, no study reports outcomes in the "Financial inclusion, savings, and insurance", "Agricultural

production", "Gender equality and empowerment", or "Social cohesion" categories. Additionally, UCT interventions are understudied for this subpopulation, present in 30% of studies (n=3). Notably, all studies reporting outcomes for this subpopulation are set in the LAC region, with no studies from other regions measuring effects specifically for indigenous peoples and local communities.

4.4. AVAILABLE EVIDENCE ON CT INTERVENTIONS WITH DIFFERENT IMPLEMENTATION-RELATED CHARACTERISTICS (RQ 3)

In the following subsection, the evidence on effects of CT interventions is described for different implementation-related characteristics. As with the previous subsection, this information was not extracted from the SRs, so the following discussion focuses exclusively on the included IEs. The discussion will centre around the existing evidence and absolute evidence gaps for IEs, rather than existing synthesis gaps.

4.4.1 Governmental and non-governmental intervention

The majority of evidence presented in the EGM is from interventions either exclusively or partly implemented by governments. 66% of the 709 IEs evaluate interventions implemented exclusively by governments (n=473), while 20% evaluate interventions implemented exclusively by non-government actors (n=142). 12% of studies had both government and non-government agents involved in implementation (n=82).[33]

Governmental interventions more frequently consist of CCTs than of UCTs, (58% and 40%, respectively, with 271 studies on CCTs, and 185 studies on UCTs). At the same time, there are no differences in the pattern of evidence on outcomes. In line with the overall pattern, most studies report "Health" (54%, n=254) and "Living standards and consumption" outcomes (40%, n=187), while there is less evidence for "Gender equality and empowerment" (6%, n=28), and "Social cohesion" (4%, n=20).

For exclusively non-governmental interventions, there is more evidence for UCTs compared to CCTs. More specifically, 74% of studies investigate UCTs (n=105), which is much higher than the percentage with which UCT interventions occur in the overall sample (45%, 318 out of 709 studies). In comparison, CCTs are the least-studied intervention type, with only 47% of studies (n=67). The outcomes studied are heavily concentrated in the "Health" category (80%, n=114). This outcome category is driven by a lot of evidence in the "SRH" (n=46) and "Nutrition" (n=43) outcome subcategories. Less evidence exists for "Agricultural production" (14%, n=20) and "Social cohesion" (12%, n=17).

Evidence is evenly spread between CCTs and UCTs for interventions implemented by both governmental and non-governmental actors, with the usual evidence patterns for outcomes being replicated. Of interventions implemented by both governmental and non-governmental actors, CCT interventions are the most studied (66%, n=54). 49% of studies investigate UCTs (n=40). The most evidence exists for "Health" (63%, n=52), and "Living standards and consumption" (56%, n=46) outcomes. The least evidence exists for "Social cohesion" (11%, n=9), and "Gender equality and empowerment" (9%, n=7).

Interventions implemented exclusively by governments tend to have a longer duration than interventions implemented by non-governmental actors and are marginally more

[33] 18 studies did not provide enough information to assess who the implementing agent is.

focused on increasing access to basic services and providing healthcare. Overall, 61% of IEs report on interventions with a duration of five years or longer (n=433). However, 390 of the 467 interventions implemented exclusively by governments have a duration of five years or longer (84% of interventions implemented by governments). Only four of the 142 interventions implemented by exclusively non-governmental actors have a duration of five years or longer (3% of interventions implemented by non-governmental actors). 22% of governmental interventions with plus components focus on increasing access to basic services (27 out of 121 studies), while only 16% of non-governmental interventions focus on this (14 out of 88 studies). Similarly, 18% of governmental interventions with plus components provide healthcare as well (22 out of 121 studies), compared to only 14% of non-governmental interventions with plus components (twelve out of 88 studies).

4.4.2 Humanitarian and emergency response

Evidence for CTs in humanitarian/emergency situations is concentrated in UCT interventions and in the "Health" and "Living standards and consumption" outcome categories. In the overall sample, only 6% of studies investigate the impact of CTs in humanitarian settings (n=46). Of these studies, most make use of UCTs, making up 76% of studies (n=35). This finding is not surprising as UCTs can be quickly deployed in emergency situations, providing immediate relief to those in need without the need for complex conditionality structures. The most studied outcomes are "Health" (72%) and "Living standards and consumption" (65%). This translates to 33 and 30 of the studies implemented in humanitarian settings, respectively.

Some evidence gaps remain for interventions implemented in humanitarian settings. Evidence is sparse for "Social cohesion" outcomes (13%, n=7), and especially for "Financial inclusion, savings, and insurance", and "Gender equality and empowerment" (11%, n=5 each).

4.4.3 Fragility status of the context

As with humanitarian/emergency responses, evidence for studies set in politically fragile, "Alert" contexts[34] is more lacking than evidence set in politically less-fragile "Warning" contexts.[35] 84% of studies are set in "Warning" contexts (n=594), with 14% of contexts set in "Alert" contexts (n=98). The rest of studies are set in contexts with a "non-fragility" status (n=17). Evidence for "Alert" contexts is heavily concentrated in specific regions, with 76% of the studies set in "Alert" contexts occurring in SSA. Unsurprisingly, most studies investigating interventions in humanitarian/emergency settings occur in "Alert" contexts. More than half of the studies investigating CTs in humanitarian settings are conducted in "Alert" contexts (54%, n=25), with 46% of studies in humanitarian settings conducted in "Warning" contexts (n=21).

In "Warning" contexts, evidence is more abundant for CCTs, whereas UCTs are more frequent in in "Alert" contexts. For "Warning" contexts, CCTs account for 58% of interventions studied (n=345). However, studies in "Alert" contexts commonly use UCTs, with 77% of studies in these contexts making use of this intervention type (n=75). CCTs are the least studied intervention in studies set in "Alert" contexts (34%, n=33). Therefore, there is also relatively little evidence for the effects of UCTs in "Warning" contexts, and for CCTs in "Alert" contexts.

[34] Contexts with a Fragile State Index value of 90.1 or higher in 2023.
[35] Contexts with a Fragile State Index value of between 60.1 and 90 in 2023.

Most studies set in "Warning" and "Alert" contexts look at "Health" outcomes, with 60% and 74% of studies reporting on this outcome, respectively. This translates to 357 studies in "Warning" contexts, and 73 studies in "Alert" contexts. For "Warning" contexts, this is driven by studies reporting on the "Child health" outcome subcategory (n=127), while it is driven by studies reporting "Nutrition" outcome subcategory for "Alert" contexts (n=34). The second most common outcome in both fragility categories is "Living standards and consumption" (41%, n=244 and 66%, n=65, respectively). For "Warning" contexts, most studies examining effects on this outcome consider indicators belonging to the "Income and expenditures" outcome subcategory (n=185). For "Alert" contexts, studies within this outcome category mostly report "Food security" (n=47). For both context categories, evidence is scarce on "Gender equality and empowerment", and "Social cohesion". For "Warning" contexts, only 8% and 6% of studies report these outcome categories, respectively (n=48 and n=36). For "Alert" contexts, the percentages are slightly higher, at 11% and 10%, respectively (n=11 and n=10).

4.4.4 Rural or urban setting

Most studies report results for predominantly rural areas. 52% of studies were conducted in a predominantly rural setting (n=371). 21% take place in a mixed rural and urban setting (n=150), while only 16% of studies report results from a predominantly urban setting (n=113). This finding is not surprising as most of the vulnerable population is usually located in predominantly rural areas and hence more subject to be a target group of social protection programs. The remaining 26% of studies do not provide enough information to accurately assess which setting results are reported from (n=184).[36]

Minor discrepancies from the overall pattern of evidence for outcome categories emerge when considering studies reporting results for predominantly rural areas. In terms of interventions, the most evidence still exists for CCTs, with 57% of studies reporting results for this modality (n=213). UCTs are reported for 49% of studies reporting results in predominantly rural areas (n=180). Again, evidence mostly exists for "Health" (60%, n=222) and "Living standards and consumption" outcomes (46%, n=170). Quite obviously, evidence for the "Agricultural production" outcome is relatively more abundant among studies reporting outcomes for predominantly rural areas (19%, n=71). The least evidence again exists for "Gender equality and empowerment" (11%, n=39), and "Social cohesion" (6%, n=23).

For studies reporting results in predominantly urban areas, most studies focus on CCTs. 70% of studies reporting results in predominantly urban areas use CCTs (n=78). This compares to 32% of studies using UCTs (n=35). Evidence on the impact of UCTs in predominantly urban areas is therefore relatively scarce, when compared with the percentage of UCTs in the overall sample of studies (45%, 318 of 709 studies), while evidence for CCTs is relatively abundant, compared to the overall sample (54%, 384 of 709 studies). For outcomes, the most evidence again exists for "Health" (61%, n=68), and "Living standards and consumption" (37%, n=41). The least

[36] Note that the percentages do not add up to 100% because of the way rural and urban settings were coded. If a study reported results pooled across rural and urban areas, and results for rural areas separately, it would be coded as recording both pooled results and results in predominantly rural areas. If results were reported separately for urban and rural areas (with results never pooled together), it would be coded as presenting both urban and rural results, but not pooled results. One study could therefore have results reported for multiple locations.

evidence is available for "Gender equality and empowerment" (5%, n=6) and, as expected, on "Agricultural production" (3%, n=3).

4.4.5 Number of beneficiaries

Most evidence exists for interventions with more than 10,000 beneficiaries.[37] 57% of the included studies evaluate interventions that reach more than 10,000 beneficiaries (n=407). Furthermore, 97% of these interventions are government programs (either implemented alongside non-governmental actors, or exclusively by government) (n=394). This is not surprising as governments can link social protection programs to national policies and have more resources to scale-up intervention. The second most evidence exists for interventions with between 1,001 and 5,000 beneficiaries (13%, n=90), with more than half of these interventions implemented by exclusively non-governmental actors (n=53). The least evidence exists for interventions with between 5,001 and 10,000 beneficiaries, with only 1% of studies reporting on interventions of this size (n=8). Potentially, this could represent a divide between larger interventions implemented by governmental parties (more than 10,000 beneficiaries) and smaller interventions implemented by non-governmental parties (less than 5,000 beneficiaries). 18% of studies did not provide enough information to assess how many beneficiaries benefitted from the intervention (n=129).

4.4.6 Targeting specific population groups

An abundance of evidence exists for interventions targeting women and girls, reported in 45% of the included studies (n=321). The majority of those interventions correspond to CCT programs (71%, n=227). 37% of the interventions targeting women and girls are UCTs (n=120). The interventions of 50% of included studies did not apply any gender targeting (n=354), whereas only 2% of studies investigate interventions targeting men or boys (n=14). The remaining studies did not provide enough information to accurately assess the sex of the target population group, if any.

Abundant evidence exists for CCTs targeting children and adolescents, and for UCTs targeting the elderly. Of the 289 studies with interventions targeted to children, 80% are CCTs (n=231). Similarly, 79% of the 282 studies investigating interventions targeted to adolescents are CCTs (n=222). Comparatively less studies examine interventions targeted to the elderly (n=93) and, of these, 72% are UCTs (n=67). This also means that less evidence of UCTs targeted towards children or adolescents, and of CCTs targeted towards the elderly exists.

Few studies target persons with disabilities, migrants, migrant workers, or refugees, smallholder farmers, and indigenous peoples and local communities and there is no study targeting the LGBTQ+ community. 44 studies examine interventions targeting persons with disabilities, with 82% of these interventions being UCTs (n=36), and 23% being CCTs (n=10).[38] Only 15 studies describe interventions targeted to migrants, migrant workers, or refugees, of which 73% are UCTs (n=11), and 27% are CCTs (n=4). Large evidence gaps also exist for interventions targeting smallholder farmers (n=9, seven UCTs and six CCTs), and interventions

[37] Note that this is not the same as a sample size. A sample size is the number of individuals or households included in the IE, while the number of beneficiaries is the number of individuals or households that received CTs from the interventions being evaluated.

[38] Note that these percentages do not add up to 100% because one study could contain both CCTs and UCTs targeted at persons with disabilities.

targeting indigenous peoples and local communities (n=3, two UCTs and one CCT). There is no study in the EGM reporting on interventions targeting the LGBTQ+ community.

4.4.7 IT-based administration of the cash transfers

Information on exactly how cash is transferred to beneficiaries is lacking for most studies, although some studies report providing CTs through digital means. From the included studies, 100 provided beneficiaries with cash digitally (through mobile money, ATM cards, etc.) and 113 studies provide beneficiaries with cash through physical means. However, the direct mode of payment was not always reported by the authors, meaning that this information is lacking for the remaining 496 studies.

Within the 100 studies known to deliver payments digitally, more evidence exists for UCTs and the pattern of evidence for outcomes tends to follow the pattern for the overall sample. UCT interventions account for 62% (n=62) of all the interventions used in studies known to provide cash through digital means. This corresponds to a relatively larger amount of evidence than in the overall sample (where UCTs only occur in 45% of studies). The outcomes for which most evidence exists are "Health" (70%) and "Living standards and consumption" (55%), corresponding to 70 and 55 studies, respectively.

4.4.8 Timing of impact measurement

The available evidence is skewed towards evaluations performed shortly after the start of the intervention (three years or less). 65% of studies in the EGM (n=459) report results shortly after the start of the intervention. The corresponding numbers for evaluations performed between four and nine years and ten years or more are 25% of studies (n=175), and 7% of studies (n=47), respectively. The remaining 4% of studies did not provide enough information to assess the timing of the evaluation after the start of the implementation (n=28).

Studies reporting evidence shortly after the intervention implementation tend to replicate the evidence pattern of the overall sample. Studies that report impact measurements shortly after implementation most often contain CCT interventions (56%, n=256), with 51% of studies containing UCTs (n=232). The pattern of outcomes evaluated shortly after implementation reflects the pattern of outcomes included in the EGM. The most evidence exists for "Health" (65%, n=299) and "Living standards and consumption" (49%, n=225). The least evidence exists for "Gender equality and empowerment" (10%, n=44) and "Social cohesion" (8%, n=35).

More evidence exists for CCTs evaluated four to nine years after the start of the intervention than UCTs, while the evidence pattern for outcome categories remains unchanged. Studies that evaluate results four to nine years after the intervention's start predominantly include CCTs (58%, n=101), followed by UCTs (42%, n=73). This distribution compares to the distribution of CCT and UCT interventions in the overall sample (54% and 45%, respectively). Consequently, one can conclude that more evaluations conducted in the four to nine years after implementation exists for CCTs than for UCTs. The evidence pattern for outcomes remains stable. The most evidence is available for "Health" outcomes (54%, n=95), followed by "Living standards and consumption" (37%, n=95). Evidence is lacking for "Gender equality and empowerment" (6%, n=10), and "Social cohesion" in particular (3%, n=5).

For evaluations performed ten years or more after the intervention's start, relatively substantial evidence exists for education outcomes. The interventions contained in these studies are evenly spread between CCTs and UCTs. 49% of studies contain CCT interventions

(n=23), while 47% contain UCT interventions (n=22). "Health" remains the category that results are most often reported for (62%, n=29). However, the "Education" outcome category replaces "Living standards and consumption" to become the second most reported outcome category, with 40% of studies reporting results for this outcome (n=19). This result is evenly driven by studies reporting on "Learning" and "Access" outcome subcategories (n=15 and n=13, respectively).[39] Again, the least evidence exists for "Social cohesion" (6%, n=3), "Financial literacy, savings, and insurance", and "Gender equality and empowerment" (4%, n=2 each).

4.5. EVIDENCE ON HOW TO IDENTIFY ELIGIBLE BENEFICIARIES OR WHOM TO TARGET (RQ 4)

There are absolute evidence gaps on how best to target beneficiaries. Only five of the included IEs, and no SR, investigate how to identify the target population, meaning that they compare different targeting approaches. Of these, two study the targeting of both UCTs and CCTs, and one studies the targeting of a CT without specification. For outcomes, three of the five studies report outcomes in the "Living standards and consumption" category. The categories of "Health", "Employment and entrepreneurship", and "Social cohesion" also each have results presented in one study. This means that no studies provide evidence on the comparisons in targeting methods when the desired outcomes are "Education", "Financial inclusion, savings and investment", "Agricultural production", or "Gender equality and empowerment".

While only few studies compare targeting strategies, it is interesting to note which comparisons are made. Two studies compare community-based targeting (CBT) methods with alternative approaches. With a CBT targeting approach, the community jointly decides who should be eligible for CTs. One of these two studies compares CBT targeting methods with both targeting based on household dependency ratios (the ratio of working-age adults to children and the elderly in a household) and targeting based on old age, assessing the impact on "Living standards and consumption" outcomes. The other study also assesses "Living standards and consumption" outcomes of CBT targeting methods, compared to either a proxy-means test (PMT) targeting method or to an alternative formula for identifying food insecure households. A PMT refers to a targeting approach whereby households or individuals are targeted based on some easily identifiable measures that are correlated with household welfare, such as the materials that their house is made of. Two further studies compare universal targeting (whereby everyone is eligible for the CT) against another form of targeting on "Living standards and consumption" outcomes. For one of these studies, the comparison is made to targeting potential beneficiaries based on whether they had prior experience in non-agricultural activities or owned certain household assets. The other study compares "Health" outcomes from universal targeting to a means testing targeting method, whereby only households or individuals below a certain income level are targeted. One study compares a discretionary targeting method (which refers to parliamentarians deciding who is eligible for the CT) and a PMT targeting method on outcomes of "Social cohesion".

Similar evidence gaps exist for studies investigating whom to target to best achieve the desired results. There are 15 IEs and one SR investigating whom to target with CT interventions to best achieve the desired results (i.e., to have the largest possible desired effect). Of these IEs, 73% present results of an UCT (n=11) and 67% present results of a CCT (n=10). For these studies, the most evidence exists for "Education", with 60% of studies reporting on this outcome (n=9).

[39] Note that these two numbers do not add up to 19 because one education outcome category can contain results for multiple outcome subcategories.

"Health" is reported in 53% of studies (n=8), and both "Living standards and consumption" and "Employment and entrepreneurship" are reported in 47% of studies (n=7). On the flip side, evidence is lacking for "Financial inclusion, savings and insurance", "Agricultural production" and "Gender equality and empowerment" outcomes (13%, n=2 each), and "Social cohesion" (7%, n=1). This finding was not expected as the first two outcome categories are related to the use of resources and therefore, a more substantial body of evidence was anticipated given the crucial role of identifying the most adequate recipients. The SR covers all types of interventions but reports outcomes only in the "Education" outcome category.

The comparisons of whom to target to best achieve the desired outcomes fall into four (sometimes overlapping) groups. The largest group refers to studies that test whether to target men or women. Eight of the 15 IEs that assess whom to target fall within this group. Overlapping with this group is the group of studies testing whom to target within a household. Five studies fall within this group, with four studies assessing whether to target mothers or fathers within a household and one study assessing whether to target schoolgirls or their parents.[40] The third group, containing three studies, assesses whether it is best to target poorest potential beneficiaries, or less poor potential beneficiaries. The remaining group contains studies assessing the targeting of miscellaneous groups. One study assesses whether it is best to target individuals with or without more intervention beneficiaries in their area. Another study assesses whether it is best to target girls who were still in school or girls who had dropped out of school, while a third study assesses whether it is better to target pregnant women directly, or community healthcare workers instead.

4.6. QUALITY OF STUDIES – SURE ASSESSMENT FOR SYSTEMATIC REVIEWS

The overall confidence in the findings of the SRs included in this EGM is rated as low to medium. Of the 33 included SRs (21%) are rated as High Confidence (n=7). 33% of the total included SRs are rated as Medium Confidence (n=11), while the remaining 45% of SRs are rated as Low Confidence (n=15). Confidence in both SRs' search strategies (Section A) and method of analysis (Section B) is low, with both sections containing 24 out of 33 studies (73%) rated as Low Confidence. In Section A, ratings of Low Confidence are mostly awarded because SRs did not report independent screening by at least two reviewers (n=20 out of 33). Authors also often did not report using a criterion to assess the quality or risk of bias of included studies (n=18 out of 33). Ratings of Low Confidence in Section B are mostly applied because authors do not report independent data extraction by at least two reviewers (n=21 out of 33). Regarding mediating factors, most SRs acknowledge their limitations (n=25 out of 33) and refrain from drawing strong policy conclusions (n=29 out of 33). For those studies in where both mediating factors occur, the confidence rating increases by one level (e.g., from Low to Medium).

[40] Note that the studies assessing whether to target mothers or fathers also fall into the group of studies testing whether to target men or women. One study assesses both; whether to target men or women (where men and women are the household head of different households) and whether to target mothers or fathers (where mothers and fathers are in the same household).

5. LIMITATIONS OF REVIEW METHODS

5.1. TIME BIAS

There are a couple of factors concerning time that could introduce bias to the EGM. The first pertains to the restriction of utilizing evidence published from 2005 onwards. Relevant studies published before 2005 are excluded from the EGM, although the restriction implies that interventions initiated before 2005, but evaluated after 2005, are included.

The second aspect is associated with selecting LMICs based on the World Bank's 2022 criteria. This approach establishes a definitive boundary determining which countries are eligible or ineligible for potential intervention implementation. Nonetheless, it fails to account for the possibility that the income status during intervention implementation or publication may have changed. Countries that had initially implemented CTs while still officially categorized as LMIC but subsequently transitioned to HIC in 2022 have been excluded from this analysis. Examples of such countries include Chile, Panama, and Uruguay.

5.2. LANGUAGE BIAS

Since this EGM excludes evidence not presented in English, there is a risk of losing relevant findings, for e.g., Francophone Africa or LAC. However, this limitation is not particularly troubling as according to the results presented, the LAC region is well-represented among the studies included in this EGM. In addition, there is a tendency that English language journals are more likely to publish statistically significant results compared to non-significant results (Higgins et al., 2019), possibly leading to some degree of language bias in this EGM.

5.3. EVIDENCE TYPES

Based on the criteria regarding the study design, the EGM likely excludes a substantial number of studies related to the interventions of interest that do not follow (quasi-)experimental designs. Therefore, reported evidence gaps need to be interpreted as gaps in (quasi-)experimental evidence rather than as gaps in evidence in general. Note that for some outcome types, such as "Gender equality and empowerment" or "Social cohesion" it might be fundamentally more challenging to generate (quasi-) experimental evidence. Furthermore, this EGM is not able to provide information on the differential impacts by time of exposure to CTs (i.e., by comparing impacts of the CTs on individuals that have been exposed for ten years, versus individuals that have been exposed for three or one year). It excludes studies that phase-in CT interventions and which by definition do not have a pure control group. This means that some studies estimating the long-term impacts of CTs, through using phase-in designs, are excluded.

6. CONCLUSION

This report contributes to evidence-informed decision-making by describing the extent of available evidence regarding CT and CT+ programs for a wide variety of outcomes, ranging from health to social cohesion. It accompanies the interactive EGM tool by summarizing existing IEs and SRs with regard to various intervention characteristics (e.g., mode of delivery or targeted subpopulation) and outcomes as well as by highlighting evidence gaps in the literature.

The EGM aims to facilitate evidence use by enabling easy access to the existing rigorous evidence on CT and CT+ programs in LMICs. It includes a total of 709 IEs and 33 SRs and provides a visual and interactive representation of these studies by mapping types of CT programs and a broad range of outcomes on to an intervention-outcome matrix. In doing so, it also pinpoints critical areas with limited or no evidence, and hence highlights in which areas more primary research or syntheses are needed.

In terms of evidence on interventions and outcomes, more evidence exists for CCTs than for UCTs. Most evidence is concentrated in the areas of health and living standards and consumption. The finding that health outcomes are the most represented in the EGM may be related to IEs being very common in the area of health in general (Cameron et al., 2016). The amount of evidence per outcome category follows a similar pattern in the included SRs, except for a notably higher amount of evidence on education outcomes in the SRs

CCTs are more likely than UCTs to be provided with plus components. Within IEs, plus components are most commonly studied in the form of information, nudges, or BCC, followed by training as the second most frequently studied plus component. Psychosocial support is the least commonly studied. The frequency of plus components in SRs follows the same pattern as in IEs, aside from the fact that SRs provide relatively more evidence on food transfers than the IEs.

The greatest potential for future evidence syntheses exists in the area of living standards and consumption, followed by outcomes related to employment and entrepreneurship, and agricultural production. For example, 44% of the overall IEs study outcomes related to living standards and consumption, but only 18% of overall SRs cover this area.

The most substantial evidence gaps, and hence potential for primary research, exist in the areas of social cohesion, gender equality and empowerment, and financial inclusion, savings, and insurance. In comparison to the other outcomes explored in this EGM, there is a relative lack of primary research for these outcomes. While there is a larger number of IEs investigating financial inclusion, savings, and insurance outcomes compared to social cohesion and gender equality and empowerment, the existing evidence is still unlikely to lend itself to a robust synthesis due to variations in intervention designs, which means that too few studies are similar enough for synthesis.

Overall, the amount of evidence on the effects of CTs differs widely across different subpopulations, revealing important areas for future research. Many studies investigate effects for women and girls, with more than half of studies reporting outcomes for this subpopulation. Evidence is also abundant for adolescents and children. However, evidence is especially lacking for smallholder farmers, migrants, migrant workers, or refugees, indigenous peoples and local communities, persons with disabilities, and the LGBTQ+ community, with no studies reporting on the last subpopulation type. Notable evidence gaps for subpopulations exist within intervention categories as well. There is a relative lack of evidence on UCTs for women and girls, children, and indigenous people and local communities. Regarding CCTs, there is little evidence for the elderly and smallholder farmers, and none for persons with disabilities.

Predominantly, studies report results for interventions targeting more than 10,000 beneficiaries, based in rural areas, and implemented by governmental parties. The least evidence is found for interventions implemented exclusively by non-government actors, as opposed to interventions implemented exclusively by government actors or jointly by government and non-government actors.

Characteristics such as fragility status, humanitarian/emergency response, mode of delivery of CTs, and timeframe of the evaluation should be studied further. With respect to the fragility status of the country, most evidence exists for interventions in "Warning" contexts, with less evidence for interventions in "Alert" contexts. Few studies investigate CTs in humanitarian settings, indicating a substantial evidence gap in emergency response settings, which could possibly be due to the ethical considerations and implementation challenges in such cases. Evidence is also lacking on the mode of delivery of CTs. Few studies report enough information to assess whether the CT was provided through digital or physical means. This means that policymakers can only draw on limited evidence regarding the mode of delivery. Finally, most of the impact measurements are performed shortly after interventions have started (three years or less after the start of the intervention).

Lastly, more evidence is required on both how to target potential beneficiaries, and which potential beneficiaries to target to make CT interventions most effective. Only five studies assess the effectiveness of different methods of targeting eligible potential beneficiaries although the way recipients are targeted may directly influence the program's effectiveness. Which potential beneficiaries to target is another understudied area, with only 15 IEs and one SR reporting on this matter. Most of the existing studies compare the difference in effects when targeting men versus women; others look at the effects of targeting poorer and less poor beneficiaries.

BIBLIOGRAPHY

Aguila, E., & Casanova, M. (2020). Short-Term Impact of Income on Cognitive Function: Evidence From a Sample of Mexican Older Adults. *Journal of Aging and Health, 32*(7-8), 591-603. https://doi.org/10.1177/0898264319841155.

Alkire, S., Meinzen-Dick, R., Peterman, A., Quisumbing, A.R., Seymour, G., & Vaz, A. (2013). The Women's Empowerment in Agriculture Index. *World development, 52*, 71-91. https://doi.org/10.1016/j.worlddev.2013.06.007.

Attanasio, O.P., Oppedisano, V., & Vera-Hernandez, M. (2015). Should Cash Transfers Be Conditional? Conditionality, Preventive Care, and Health Outcomes. *American Economic Journal: Applied Economics, 7*(2), 35-52. https://doi.org/10.1257/app.20130126.

Avitabile, C. (2021). Spillovers and Social Interaction Effects in the Demand for Preventive Healthcare: Evidence from the *PROGRESA* program. *Journal of Health Economics, 79*, 102483. https://doi.org/10.1016/j.jhealeco.2021.102483.

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T., & Pellerano, L. (2016). *Cash transfers: what does the evidence say? A rigorous review of impacts and the role of design and implementation features.* ODI Research Reports. https://cdn.odi.org/media/documents/11316.pdf (Accessed on 20.09.2023).

Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) (2017). *Social protection for equitable development.* BMZ Position Paper 09|2017. https://www.bmz.de/resource/blob/23762/c115e74e39231821752d303ae1f246ca/strategiepapier420-social-protection-data.pdf (Accessed on 12.12.2022)

Cameron, D., Mishra, A., & Brown, A. (2016). The growth of impact evaluation for international development: how much have we learned*? Journal of Development Effectiveness, 8*(1), 1-21. https://doi.org/10.1080/19439342.2015.1034156.

Carter B., Roelen, K., Enfield, S., & Avis, W. (2019). *Social Protection Topic Guide*. Knowledge, Evidence and Learning for Development (K4D). https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/14885/Social_Protection_Topic_Guide_online.pdf?sequence=1&isAllowed=y (Accessed on 12.12.2022)

Chang, H., Huang, K., & Wu, C. (2006). Determination of sample size in using central limit theorem for weibull distribution*. International journal of information and management sciences*, *17*(3), 31-46. Available: http://163.13.238.245/IJIMS/files/recruit/569_76fb6a86.pdf.

Chen, S., Chen, X., Law, S., Lucas, H., Tang, S., Long, Q., Xue, L., & Wang, Z. (2020). *Pension and Health Services Utilization: Evidence from Social Pension Expansion in China*. IZA Discussion Paper No. 13876. Available: https://docs.iza.org/dp13876.pdf.

Coetzee, M. (2013). Finding the Benefits: Estimating the Impact of The South African Child Support Grant. *South African Journal of Economics, 81*(3), 427-450. https://doi.org/10.1111/j.1813-6982.2012.01338.x.

Dervisevic, E., Perova, E., & Sahay, A. (2021). *Long-Term Impacts of Short Exposure to Conditional Cash Transfers in Adolescence: Evidence from the Philippines.* World Bank Policy Research Working Paper No. 9617. http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-961.

Dwyer, R.J., & Dunn, E.W. (2022). Wealth redistribution promotes happiness. *Social Sciences, 119*(46), e2211123119. https://doi.org/10.1073/pnas.2211123119.

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., & Walker, M. (2022). General Equilibrium Effects of Cash Transfers: Experimental Evidence From Kenya. *Econometrica, 90*(6), 2603-2643. https://doi.org/10.3982/ECTA17945.

Gelo, D., Kollamparambil, U., & Jeuland, M. (2023). The causal effect of income on household energy transition: Evidence from old age pension eligibility in South Africa. *Energy Economics, 119,* 106538. https://doi.org/10.1016/j.eneco.2023.106538.

Harris-Fry, H.A., Paudel, P., Harrisson, T., Shrestha, N., Jha, S., Beard, B.J., Copas, A., Shrestha, B.P., Manandhar, D.S., De Costello, A.L.M., Borja, M.C., & Saville, N.M. (2018). Participatory women's groups with cash transfers can increase dietary diversity and micronutrient adequacy during pregnancy, whereas women's groups with food transfers can increase equity in intrahousehold energy allocation. *Journal of Nutrition 148*(9), 1472-1483. https://doi.org/10.1093/jn/nxy109.

Higgins, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (eds) (2019). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Available: https://training.cochrane.org/handbook/archive/v6.

Little, M., Roelen, K., Lange, B., Steinert, J., Yakubovich, A., Cluver, L., & Humphreys, D. (2021). Effectiveness of cash-plus programmes on early childhood outcomes compared to cash transfers alone: A systematic review and meta-analysis in low- and middle-income countries*. PLoS Med 18*(9): e1003698. 1-24. https://doi.org/10.1371/journal.pmed.1003698.

Olson, Z., Clark, R.G., & Reynolds, S.A. (2019). Can a conditional cash transfer reduce teen fertility? The case of Brazil's Bolsa Familia. *Journal of Health Economics 63*, 128-144. https://doi.org/10.1016/j.jhealeco.2018.10.006.

Oxford Poverty & Human Development Initiative and United Nations Development Programme (OPHI and UNDP) (2022). *Unpacking deprivation bundles to reduce multidimensional poverty*. https://ophi.org.uk/wp-content/uploads/G-MPI_Report_2022_Unpacking.pdf (Accessed on 12.12.2022)

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372(*n71). https://doi.org/10.1136/bmj.n71

Roelen, K., Devereux, S., Abdulai, A., Martorano, B., Palermo, T., & Ragno, L. (2017). *How to make 'cash plus' work: Linking cash transfers to services and sectors*. INNOCENTI Working Papers: 2017–10. UNICEF Office of Research Florence. https://www.calpnetwork.org/wp-content/uploads/2020/03/how-to-make-cash-plus-work-1.pdf (Accessed on 12.12.2022)

Stoner, M.C.D., Kilburn, K., Godfrey-Faussett, P., Ghys, P., & Pettifor, A.E. (2021). Cash transfers for HIV prevention: A systematic review. *PLoS medicine*, *18*(11), e1003866. https://doi.org/10.1371/journal.pmed.1003866.

Waidler, J., Gilbert, U., Mulokozi, A., & Palermo, T.A. (2022). "Plus" Model for Safe Transitions to Adulthood: Impacts of an Integrated Intervention Layered onto A National Social Protection Program on Sexual Behavior and Health Seeking among Tanzania's Youth. *Stud Fam Plann 53*(2):233-258. https://doi.org/10.1111/sifp.12190.

APPENDIX

A. SEARCH STRATEGY OVERVIEW

Resource Name	Website/ Platform	Date of search	# of results	Limitations and adaptation of search term
Search 1: Databases with full search term				
Social Science Citation Index-Expanded	Web of Science https://www.webofscience.com/basic-search	March 22, 2023	1,501	No limitations
EconLit	EBSCOhost https://web-s-ebscohost-com.ez.sun.ac.za/ehost/search/advanced?vid=0&sid=074bbffd-57a7-4cc8-8c2c-dfbacba2570a%40redis	March 22, 2023	1,139	No limitations
Scopus	Elsevier https://www-scopus-com.ez.sun.ac.za/search/form.uri?display=basic#basic	March 22, 2023	3,362	No limitations
Total # of results from databases			6,002	
Total # of duplicates removed			1,666	
Total # of records from Search 1			**4,336**	

Search 2: Databases and websites with adapted search terms				
Campbell Collaboration Systematic Reviews	https://onlinelibrary.wiley.com/search/advanced	April 6, 2023	75	The Campbell Wiley platform disseminates high-quality systematic reviews. The platform is not sophisticated enough to handle a complex search (with multiple terms and truncation operators for phrases in quotes) and hence, only key terms related to the intervention were included. Three sets of terms (at title and abstract levels and in all fields) were run in order to capture as many papers as possible. Campbell Systematic Reviews are indexed in WoS, so studies should also be captured with the more complex search in WoS.
Cochrane Database of Systematic Reviews	https://www.cochranelibrary.com/advanced-search	April 12, 2023	90	The Cochrane platform is a database for systematic reviews in health care. This platform can handle complex searches, but the search was limited to SRs in order to exclude research protocols.
International Initiative for Impact Evaluation (3ie) Development Evidence Portal	https://developmentevidence.3ieimpact.org/	March 25, 2023	892	The 3ie portal has a focus on (quasi-)experimental IEs and syntheses thereof in developing countries. Therefore, country and methodology filtering terms are not required. Advanced search is supported with a few restrictions. E.g., when using *exact* phrase searching, all plurals and spelling variants must be included individually, e.g., *exact* searches of "childhood vaccination*" will only retrieve results for "childhood vaccination", not "childhood vaccinations". Therefore, *inexact* search terms were used where feasible, to return all results containing the same stem word.
National Bureau of Economic Research	https://www.nber.org/papers and search of NBER Working Papers on ABI/Inform in ProQuest deduplicated	April 6, 2023	379	NBER Working Papers on ABI/Inform in ProQuest is a collection of scholarly content including full-text scholarly journals, dissertations, and collections of working papers. This database has advanced search functionality which allows using the complete search term but may not include very recent research. To ensure that more recent NBER content is captured, the search was extended to the NBER website. However, as more complex searches that include the geographic or study design limiters are not supported on the NBER website, only the most relevant intervention terms ("cash transfer" OR "cash transfers") were used in order to identify results most directly relevant to this EGM.
JSTOR	https://www.jstor.org/action/showAdvancedSearch	March 25, 2023	472	JSTOR is a search engine for scholarly articles but given its 200-character limit (including spaces) it does not allow for a search with multiple search terms. Hence, it was decided to only include the most relevant intervention terms at the title and

				abstract levels ("cash transfer" OR "cash transfers"), in order to identify results most directly relevant to this EGM.
Google Scholar	www.scholar.google.com	April 2023	839	Google Scholar indexes scholarly literature like peer-reviewed journals, academic books, conference papers, and more, however, the search fields of the simple and advanced search interfaces are limited to a length of 256 characters. In addition, Google Scholar has no truncation operators, hence complete words have to be used. This limits the possibility of complex searches. Consequently, keywords that capture interventions, methods, and LMIC were used. Retrieving entries from Google Scholar involves manual selection of individual studies, which is why the retrieval stage was combined with a quick title screening. Irrelevant titles were not exported.
Total # of results			2,747	
Total # of duplicates removed			292	
Total # after deduplication			2,455	
Total # of duplicates against Search 1 removed			583	
Total # of records from Search 2			1,872	
Total # of records from Search 1 and 2			**6.208**	

The full details of the search strategy can be found in the supplementary material.

B. RESULTS OF SCREENING PILOTS AT TITLE AND ABSTRACT AND FULL-TEXT PHASE

Date	Agenda	Final Team decision
12.04	**Pilot 1.1 – Title and Abstract:** Screening on Title and Abstract (Pilot): Resolving disagreements on Title and Abstracts between coding groups based on the existing protocol: • *Group 1: (agreement of 86%)* o 29 papers reviewed, o 4 disagreements found • *Group 2: (agreement of 90%)* o 29 papers reviewed, o 3 disagreements found • *Group 3: (agreement of 83%)* o 29 papers reviewed, o 5 disagreements found Resolutions made at meeting: ▪ Met targets laid out in protocol. ▪ But decided to do second pilot, to help with training, increase comfort with screening.	**Exclude** ->Cheteni et al. (2023); Tette et al. (2016); Zamboni et al. (2018); Pankomera et al. (2019); Silverio-Murrillo (2021); Kolesar et al. (2017); Masset et al. (2018); **Reasons:** o Study design is unsuitable, o Irrelevant intervention, o Not targeted to households or individuals, **Include** ->Steklov et al. (2016); Bernhard et al. (2017); Almas et al. (2015); Cruz et al. (2014); Spantig (2021) **Reasons:** o Relevant interventions (PROGRESA), o Targeted at households and individuals (entrepreneurs), o Is in LMIC (North Macedonia), o Study design is suitable (theoretical paper with possibly relevant empirical analysis)
13.04	**Pilot 1.2 – Title and Abstract:** Screening on Title and Abstract (Pilot): Resolving disagreements on Title and Abstract Screening between groups based on protocol: • *Group 1 (agreement of 87%)* o 15 studies reviewed, o 2 disagreements found, • *Group 2 (agreement of 87%)* o 15 studies reviewed, o 2 disagreements found,	**Exclude ->**Amaro et al. (2016); Balmori (2020); Galarrga et al. (2014); Roelen (2017); Brollo et al. (2012); **Reasons:** o Not targeted to households or individuals, o Study design is unsuitable, **Include** ->Buller et al. (2018); Graham (2020);

	• *Group 3 (agreement of 79%)* o 14 studies reviewed, o 3 disagreements found, <u>Across both pilots</u>: • *Group 1 (agreement of 86%)* o 44 studies reviewed, o 6 disagreements found, • *Group 2 (agreement of 87%)* o 44 studies reviewed, o 5 disagreements found, • *Group 3 (agreement of 81%)* o 43 studies reviewed, o 8 disagreements found, • *Pooled (agreement of 85%)* o 131 studies reviewed, o 19 disagreements found,	
23.05	**Pilot 2.1 – Full-text:** Full-text screening (Pilot): Resolving disagreements on full-text screening between groups based on protocol: • *Group 1 (agreement of 91%)* o 22 studies reviewed, o 2 disagreements found, • *Group 2 (agreement of 77%)* o 22 studies reviewed, o 5 disagreements found,	**Exclude** ->
Belchoir et al. (2022); MacPherson et al. (2021); Porter (2010); Shi (2016); Stecklov et al. (2016); Tohari et al. (2019);
Reasons: o Study design is unsuitable, o Irrelevant intervention, o Irrelevant outcomes, o Outcome assessed less than one month after intervention, **Include** ->
Barrientos et al. (2013) **Reasons:** o Relevant outcome (employment),
26.05	**Pilot 2.2 – Full-text:** Full-text screening (Pilot): Resolving disagreements on full-text screening between groups based on protocol: • *Group 1 (agreement of 75%)*	**Exclude** ->
Carolina et al. (2021); Dang et al. (2019);
Reasons: o Irrelevant outcomes,

	o 8 studies reviewed, o 2 disagreements found, • *Group 2 (agreement of 86%)* o 7 studies reviewed, o 1 disagreement found,	o Study design is unsuitable,
	Across both pilots: • *Group 1 (agreement of 87%)* o 30 studies reviewed, o 4 disagreements found, • *Group 2 (agreement of 79%)* o 29 studies reviewed, o 6 disagreements found, • *Pooled (agreement of 83%)* o 59 studies reviewed, o 10 disagreements found,	

C. LIST OF INCLUDED OUTCOME CATEGORIES AND SUBCATEGORIES WITH EXAMPLES

The following eight outcome categories with 37 different subcategories are included:

a. Health
   (1) Child health (including nutritional status indicators such as stunting)
   (2) Sexual, reproductive and maternal health
   (3) Mental health and well-being
   (4) Access and use of health services
   (5) Nutrition (e.g., food consumption, infant and young child feeding practices, dietary diversity, nutritional biomarkers)
   (6) Other health outcomes

b. Education
   (1) Learning and achievement (test scores, literacy, cognitive development, completion/graduation)
   (2) Access to education (enrolment, attendance, dropouts, and truancy)
   (3) Attitudes and personal development (e.g., aspirations)

c. Living standards and consumption
   (1) Housing, Electricity and water, sanitation, and hygiene (WASH) infrastructure
   (2) Household income
   (3) Food security
   (4) Other non-food consumption, consumption in general, non-productive or household assets

d. Financial inclusion, savings and insurance
   (1) Financial literacy (e.g., take-up of financial training)
   (2) Financial services (e.g., formal savings account, take-up of loans)

(3) Cash savings
(4) Insurance

e. Agricultural production
   (1) Land ownership and/or used
   (2) Agricultural assets and investments
   (3) Adoption and knowledge of agricultural technologies and practices
   (4) Agricultural yield and income

f. Employment and entrepreneurship (non-agricultural)
   (1) Vocational training and technical skills
   (2) Entrepreneurial skills
   (3) Employment or self-employment (including formality)
   (4) Productive (non-agricultural) assets and investments
   (5) Entreprenurial income
   (6) Child labor or work (incl. Household, unpaid, etc.)

g. Gender equality and empowerment[41]
   (1) Reduction of gender-gaps (e.g., in education or labor market outcomes)
   (2) Production decisions (i.e., decision-making regarding productive activities of the household)
   (3) Control over household resources and income (e.g., ownership of/access to and decision-making power regarding assets, credit, and household income and spending)
   (4) Leadership (e.g., access to/membership in community and social groups, opportunities of speaking in public)
   (5) Time-use (e.g., time allocation between chores and leisure, access to childcare)

h. Social cohesion
   (1) Social capital
   (2) Creation of networks
   (3) Trust in others and institutions
   (4) Attitudes toward other groups
   (5) Civic engagement

[41] Outcomes follow the five domains of women's empowerment as defined in the "Women Empowerment in Agriculture Index" developed by Alkire et al., (2013).

Center for Evaluation and Development

C4ED

O7, 3
68161 Mannheim, Germany

Email: info@c4ed.org

www.c4ed.org

With financial support from the

Federal Ministry
for Economic Cooperation
and Development