Improving food security in humanitarian emergencies: An evidence gap map
ABOUT HAEC

The Humanitarian Assistance Evidence Cycle (HAEC) Associate Award works to increase the efficiency and effectiveness of emergency food security activities funded by the USAID Bureau for Humanitarian Assistance (BHA) by increasing the use of cost-effective and timely impact evaluations in humanitarian contexts. Impact evaluations provide robust evidence to inform technical approaches to improve the impact of humanitarian interventions. HAEC is a three-year activity (2021-2024) funded by USAID/BHA and issued through the IDEAL Leader with Associate Awards Activity, implemented by TANGO International, Save the Children, the International Initiative for Impact Evaluation (3ie), and Causal Design.

ABOUT 3IE

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed, equitable, inclusive, and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesize and quality assured evidence of what works, for whom, how, why, and at what cost.

3IE EVIDENCE GAP MAPS

3ie evidence gap maps (EGMs) are thematic collections of information about impact evaluations and systematic reviews that measure the effects of international development policies and programs. The maps provide a visual display of completed and ongoing systematic reviews and impact evaluations in a sector or sub-sector, structured around a framework of interventions and outcomes. The report provides all supporting documentation for the map, including background information for the theme of the map, the methods applied to systematically search and screen the evidence base, as well as the main results.

ABOUT THIS EVIDENCE GAP MAP

This report presents the findings of a systematic search to identify and map the evidence base of impact evaluations and systematic reviews of interventions that aim to improve food security during humanitarian crises. The online EGM is available here. All content in this report is the sole responsibility of the authors and does not represent the opinions of USAID/BHA, TANGO International, Save the Children, Causal Design, and 3ie, including its donors or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. Please direct any comments or queries to the corresponding author, Cem Yavuz, at: cyavuz@3ieimpact.org.

PHOTO CREDITS

Sacha Myers/Save the Children

AUTHORS

Cem Yavuz, 3ie; Paul Fenton Villar, 3ie and University of East Anglia; Miriam Berretta, 3ie; Ashiqun Nabi, 3ie; Chris Cooper, University of Bristol, University College London; Shannon Shisler, 3ie
RECOMMENDED CITATION


CONTACT:

Humanitarian Assistance Evidence Cycle (HAEC)
Save the Children, Department of Hunger and Livelihoods
899 North Capitol St NE, Suite 900
Washington, DC 20002
haec@savechildren.org
www.fsnnetwork.org

DISCLAIMER

This report is made possible by the generous support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of the Humanitarian Assistance Evidence Cycle (HAEC) Associate Award and do not necessarily reflect the views of USAID or the United States Government.
ACKNOWLEDGMENTS

The authors are grateful for the support provided by Lloyd Owen Banwart (TANGO International) and Christy Lazicky (Causal Design) who were instrumental in organizing consultations and feedback to ensure the scope of this EGM and the report itself was practitioner relevant. We thank Adam Trowbridge (USAID, BHA) and Mara Mordini (USAID, BHA) for collating their own, as well as their colleague’s feedback on the final report. We also thank Birte Snilstveit (3ie) for providing guidance across the EGM project.
# TABLE OF CONTENTS

Acknowledgments ........................................................................................................................................... i  
List of Tables .................................................................................................................................................. iii  
List of Figures .................................................................................................................................................. iii  
Acronyms ....................................................................................................................................................... iv  
Executive Summary ...................................................................................................................................... v  
1. Introduction ............................................................................................................................................. 1  
2. Background: Humanitarian emergencies and food security ................................................................. 4  
3. Humanitarian interventions and food security outcomes .......................................................................... 5  
   Interventions in humanitarian settings ......................................................................................................... 5  
   Food security outcomes ................................................................................................................................ 6  
4. Methods ..................................................................................................................................................... 7  
   4.1 Criteria for including and excluding studies in the EGM ...................................................................... 7  
   4.2 Literature search ................................................................................................................................... 8  
      4.2.1 Search Strategy .............................................................................................................................. 8  
      4.2.2 Search Results ............................................................................................................................... 9  
   4.3 Coding of study characteristics ............................................................................................................ 11  
5. Analysis ..................................................................................................................................................... 11  
   5.1 Analysis of evidence on the effects of humanitarian interventions on food security ......................... 12  
      5.1.1 Growth of the evidence base ........................................................................................................ 12  
      5.1.2 Intervention and outcome coverage ............................................................................................ 13  
      5.1.3 Geographic distribution of impact evaluations .......................................................................... 17  
      5.1.4 Distribution of impact evaluations across types of emergency contexts .................................. 20  
      5.1.5 Distribution of impact evaluations according to emergency phase of implementation ........... 21  
      5.1.6 Evidence of interventions’ effects on vulnerable populations .................................................. 21  
   5.2 Synopsis of impact evaluation methods .............................................................................................. 22  
6. Systematic review critical appraisals and summary of findings from high- and medium-confidence reviews ............................................................................................................................... 26  
   6.1 Results of systematic review critical appraisals .................................................................................. 26  
   6.2 Characteristics of high- and medium-confidence systematic reviews .............................................. 27  
   6.3 Main findings of high- and medium-confidence systematic reviews ................................................ 27  
      6.3.1 Cash, food and other in-kind transfers ....................................................................................... 28  
      6.3.2 Water security interventions ....................................................................................................... 29  
      6.3.3 Nutrition interventions .................................................................................................................. 29  
   6.4 Common implications for research ....................................................................................................... 29  
   6.5 Caveats and synthesis gaps .................................................................................................................. 30  
7. Who is funding and implementing impact evaluations and systematic reviews of humanitarian interventions? .................................................................................................................................................. 31  
8. Conclusion .................................................................................................................................................. 32  
   8.1 Implications for policymakers ............................................................................................................... 33  
   8.2 Implications for future research and summary of evidence gaps .................................................... 34  
9. References ................................................................................................................................................. 35
LIST OF TABLES

Table 1. Types of humanitarian interventions included ................................................................. 6
Table 2. Types of food security outcomes included ........................................................................ 7
Table 3. Summary criteria for studies to be included in the EGM .................................................. 8
Table 4. Emergency contexts of the included impact evaluations .................................................. 20
Table 5. Emergency phase of interventions in included impact evaluations .............................. 21
Table 6. Studies targeting vulnerable groups ................................................................................. 22
Table 7. Frequency of included studies by study designs ............................................................... 23
Table 8. Characteristics of high- and medium-confidence systematic reviews .............................. 27
Table 9. Number of studies by program implementing and funding agency types. ................... 31
Table 10. Top 5 most commonly reported implementing agencies and funders ........................ 32

LIST OF FIGURES

Figure 1. PRISMA Diagram of systematic search and screening .................................................. 10
Figure 2. Number of impact evaluations and systematic reviews identified by year .................. 13
Figure 3. Distribution of impact evaluations and systematic reviews by intervention-outcome type
pairing ...................................................................................................................................................... 15
Figure 4. Trends in the cumulative number of impact evaluations by region ................................ 17
Figure 5. Map of impact evaluations by country ............................................................................. 18
Figure 6. Map of countries at risk from humanitarian crises and food insecurity ....................... 19
Figure 7. Confidence ratings for included systematic reviews ...................................................... 26
# ACRONYMS

<table>
<thead>
<tr>
<th>ACRONYM</th>
<th>FULL NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHA</td>
<td>Bureau for Humanitarian Assistance</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Coronavirus disease 2019</td>
</tr>
<tr>
<td>EGM</td>
<td>Evidence Gap Map</td>
</tr>
<tr>
<td>FAO</td>
<td>Food and Agriculture Organization</td>
</tr>
<tr>
<td>FCDO/ DFID</td>
<td>Foreign, Commonwealth &amp; Development Office (formerly DFID)</td>
</tr>
<tr>
<td>HHS</td>
<td>Household Hunger Scale</td>
</tr>
<tr>
<td>IDP</td>
<td>Internally Displaced People</td>
</tr>
<tr>
<td>L&amp;MIC</td>
<td>Low- and Middle-Income Countries</td>
</tr>
<tr>
<td>NFI</td>
<td>Non-food Items</td>
</tr>
<tr>
<td>PICOS</td>
<td>Populations, Interventions, Comparators, Outcomes and Study Designs</td>
</tr>
<tr>
<td>PRISMA</td>
<td>Preferred Reporting Items for Systematic Reviews and Meta-analyses</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized Controlled Trial</td>
</tr>
<tr>
<td>USAID</td>
<td>United States Agency for International Development</td>
</tr>
<tr>
<td>USD</td>
<td>United State Dollar</td>
</tr>
<tr>
<td>WFP</td>
<td>World Food Programme</td>
</tr>
<tr>
<td>WINGS</td>
<td>Women’s Income Generating Support</td>
</tr>
</tbody>
</table>
EXECUTIVE SUMMARY

Humanitarian emergencies continue to affect a large proportion of the global population, threatening people’s health, livelihoods and food security. With limited resources to respond to these emergencies, there is a growing desire for high-quality evidence to understand the effects of emergency activities on vulnerable populations. This study presents an evidence gap map (EGM) of impact evaluations and systematic reviews on the effects of interventions in humanitarian settings on food security outcomes. Having screened 45,683 studies from 14 academic databases and 44 grey literature sites, we identified 146 impact evaluations and 17 systematic reviews mapped across 8 intervention and 12 outcome types. Examining the characteristics of this evidence base, we find this body of literature has tripled in the last five years, indicating that there is an increasing amount of new knowledge and learning to draw on in this area. However, we also find evidence is spread thinly across intervention categories, highlighting that many evidence gaps still exist, with only a larger cluster of impact evaluations around food, cash and other in-kind transfers and no impact evaluations in some areas (such as on early warning system interventions or outcomes related to food trade). Furthermore, many countries highly vulnerable to emergencies and responses to several types of disasters, such as tsunamis and earthquakes, do not feature in this literature. We also find that the majority of identified systematic reviews in this area are of low quality, but high- and medium-quality reviews are available for food, cash and other in-kind transfers, nutrition interventions and water security interventions. Given the recent growth in the evidence base, we find that even some of the most recent systematic reviews do not include the most recent impact evaluations. This highlights the need for updated and living synthesis projects to ensure stakeholders working in this area have available to them the most up-to-date knowledge and information. The findings from this EGM provide a basis for policymakers to consult rigorous evidence in the formulation of programming. Researchers are able to use the findings to understand where gaps in evidence currently exist with the opportunity to ensure future work contributes to filling these gaps and avoids duplication.
1. INTRODUCTION

Today, the world stands on the brink of unprecedented famines, with an estimated 49 million people experiencing alarming levels of hunger and nearly 193 million people living in a situation defined as a food crisis (FSIN, 2022; Humanitarian Coalition, 2022). Catastrophic events, whose frequency is exacerbated by the climate emergency, are key factors threatening people’s access to food and their food security. Severe weather, including the worst drought in 40 years in East Africa, has decimated crops, livestock and water supplies. Economic crises, made worse by the COVID-19 pandemic, are pushing more and more households into poverty, and conflicts in places such as Ukraine and Ethiopia are also affecting agricultural production and international food prices. These events are exacerbating the demand for humanitarian assistance, with international donors now spending more than USD 30 billion per year on humanitarian support and protection (OCHA, 2021). Furthermore, in 2022, it is estimated more than 274 million people will require humanitarian assistance, a significant increase from the 235 million people in 2021 (which was already the highest figure recorded for several decades) (OCHA, 2021).¹

Amidst the food crisis and rising demand for humanitarian assistance, stakeholder reviews indicate interest is increasing in high-quality evidence of the effects of interventions in humanitarian settings with the aim of informing program implementation and decision-making (Bakrania et al., 2021; Carden, Hanley and Paterson, 2021). Limited resources in the humanitarian sector requires the use of high-quality evidence to ensure what resources are available, are being used effectively (Blanchet et al., 2018).

The purpose of this research is to compile an Evidence Gap Map (EGM) identifying ongoing and existing impact evaluations and systematic reviews of the effects of interventions in humanitarian settings on food security outcomes in low- and middle-income countries (L&MICs). EGMs are thematic collections of information about impact evaluations and systematic reviews structured around a framework of interventions and outcomes that help to make research more accessible and easily identifiable for policymakers, practitioners, researchers and funders of research (Snistveit et al., 2017; White et al., 2020).² In addition, it also highlights evidence gaps and can help inform strategic research agendas.

---

¹ It is estimated that the United Nations and partner organizations will require more than USD 40 billion to assist approximately two thirds of people requiring humanitarian assistance in 2022 (OCHA, 2021).

² Impact evaluations refer to studies measuring changes in outcomes that are attributable to a program or intervention (Gertler et al., 2016). A systematic review is a synthesis of studies compiled using an exhaustive literature search and transparent and reproducible methods to minimize the potential biases associated with appraising the design and results of research (Waddington et al., 2012; Higgins et al., 2022).
Our research objectives aim to:

1) Create and present a publicly accessible index of impact evaluations and systematic reviews on the effects of humanitarian interventions on food security outcomes using systematic methods for information retrieval and identification.

2) Categorize the impact evaluations and systematic reviews according to their study characteristics, thereby indicating thematic, methodological and contextual areas where clusters of evidence currently exist.

3) Identify potential primary evidence gaps, defined as areas where there is a lack of impact evaluations.

4) Identify synthesis gaps, defined as areas where there are either no medium- or high-confidence systematic reviews, or where a review is outdated by five years or more, and/or relevant impact evaluations have since been published.

5) Critically appraise the comprehensiveness of the methods applied in systematic reviews and summarize the findings from medium- and high-confidence systematic reviews on the effects of humanitarian interventions on food security outcomes.

To address these objectives, we have conducted a systematic literature search for impact evaluations and systematic reviews on the effects of seven types of humanitarian interventions including: early warning systems; interventions pre-arranging household finance for disasters; food, cash and other in-kind transfers; agriculture and livestock interventions; nutrition interventions; market-based recovery interventions; and water security interventions. We mapped evidence of the effects of these interventions on twelve types of food security outcomes or outcomes closely related to food security, such as health and economic outcomes. In this paper, we present the results of our findings, examining the frequencies of studies’ key characteristics (e.g. the types of emergencies, interventions, outcomes, geographic location of interventions and study design) and the findings of medium- and high-confidence systematic reviews.

Overall, we identify a growing body of evidence, consisting of 146 impact evaluations and 17 systematic reviews on the effects of interventions aiming to improve food security in humanitarian settings. This body of literature has tripled during the last five years, indicating that there is an increasing amount of new knowledge and learning to help inform programmatic planning and policy in this area. Given that contextual and resource constraints have historically been a bottleneck to the implementation of impact evaluations in humanitarian contexts (Puri et al., 2014; 2017), the growing evidence base provides many new illustrative examples of the application of different study designs and the use of innovative sources of data, such as remotely sensed and other forms of secondary data.

However, we also find that the evidence is spread thinly across different interventions and outcome types and across many contexts vulnerable to climatic and humanitarian emergencies. For example, we

---

3 The interactive online version of our EGM is available at: https://developmentevidence.3ieimpact.org/egm/food-security-in-humanitarian-settings-egm
identified no impact evaluations on the effects of early warning systems interventions and none on outcomes related to food trade. Moreover, for several countries vulnerable to emergencies, such as Burkina Faso and Sudan, we did not identify any studies. The evidence is also limited for several types of disasters, such as earthquakes, tsunamis, volcanic eruptions, landslides and extreme heatwaves. These are among the many primary evidence gaps that exist and much more research is needed in this area to inform and support the humanitarian sector.

Of the evidence available, we find noticeable clusters of impact evaluations related to food, cash and in-kind transfers only. Other areas with noticeable but more modest clusters of impact evaluations include multi-component interventions and market-based recovery interventions. We also find seven high- and medium-confidence systematic reviews across food, cash and other in-kind transfers, water security interventions and nutrition interventions. However, given the recent growth in the evidence base, we find even some of the most recent systematic reviews do not include the most recent impact evaluations. Furthermore, only low-confidence systematic reviews are available for some areas, such as the effect of the direct provision of food on food intake outcomes and cash transfers on employment outcomes. This follows a pattern in this area of low-quality synthesis work, with our appraisals highlighting 10 of the 17 systematic reviews are low-confidence reviews with major deviations from good practices. This highlights the need for updated, expanded and living synthesis projects in this area to ensure stakeholders have available to them the most up-to-date knowledge and information on the effects of interventions.

This EGM fills a significant gap in current humanitarian research, given that there is no extant comprehensive mapping of evidence of the effects of humanitarian interventions to improve food security. It also builds on existing evidence maps by Lane et al. (2022) and Hill, Peredo and Tarazona (2021). Lane et al. (2022) provides an EGM including 2,037 impact evaluations and 182 systematic reviews on the effects of interventions on food security and nutritional outcomes. While some relevant literature is captured by Lane et al. (2022), our EGM focuses specifically on emergencies and as such encompasses a broader range of humanitarian interventions and is derived from a more specific literature search tailored to identifying evidence from these contexts. Similarly, Hill, Peredo and Tarazona (2021) provide a map of evidence on the effects of pre-arranged disaster finance interventions with a truncated search. This EGM expands the search for evidence and again encompasses a broader range of intervention categories. The findings from this EGM provide a basis for policymakers to consult rigorous evidence in the formulation of programming. Researchers are also able to use the findings to understand where gaps in evidence currently exist with the opportunity to ensure future work contributes to filling these gaps and avoids duplication.

The next section provides information on the links between humanitarian emergencies and food security. Section 3 describes the scope of the different types of humanitarian interventions and outcomes we have researched. Section 4 outlines details of the criteria used to determine the inclusion of studies in this EGM and our literature search. Section 5 presents our analysis of the characteristics of identified studies. Section 6 summarizes the results of our critical appraisals of included systematic reviews and the main findings from reviews rated as medium- and high-confidence. Section 7 provides a summary of the data we have retrieved about who has funded and implemented impact evaluations and systematic reviews of humanitarian interventions. Finally, Section 8 contains our concluding remarks, highlighting this study’s key implications for policy and future research.
2. BACKGROUND: HUMANITARIAN EMERGENCIES AND FOOD SECURITY

How to best define food security has been a subject of much debate, with discussions exchanged over many decades about its meaning and definition. Today, the most widely cited definition of food security originates from the 1996 World Food Summit Plan of Action:

“Food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life.”

– (FAO, 1996)

Based on this definition, eradicating hunger is recognized as an issue that requires more than simply the provision of food supply. It highlights four interrelated dimensions to achieving food security that international development agencies and organizations (e.g. FAO, WFP, USAID, DFID/FCDO) have also structured their own work on this issue around. The four dimensions include: **food availability** in terms of quantities and production; **food accessibility** regarding the ability of individuals to purchase and access food; **food utilization** which indicates whether an individual follows an adequately nutritious diet and has access to clean water; and **food stability** indicating whether food availability, access and utilization are adequate and sustainable over time (Fahy 2021; FAO 2006; Gibson 2012; USAID 2022).

Although food security issues can arise outside of a humanitarian setting, humanitarian emergencies can perpetuate the challenges or the conditions that cause food insecurity and present major obstacles to food availability. Humanitarian emergencies are defined as serious disruptions to the functioning of a community or society causing human, material, economic and environmental losses that exceed its capacity to cope using its own resources (Humanitarian Coalition, 2022; IFRC, 2022; UNICEF, 2022). Also often referred to as ‘covariate shocks’ (Holzmann and Jørgensen, 2001), these shocks encompass different types of emergencies, including both natural hazards (e.g. earthquakes and drought), and man-made emergencies (e.g. war, environmental degradation and pollution) (Anderson and Gerber, 2018; Bang and Balgah, 2022; Garber et al., 2020; Sivaraman and Varadharajan, 2021).

There are several reasons why households’ food security may decline during humanitarian emergencies. For example, disasters can impact food production when they cause crop and livestock losses and when they prevent farmers from accessing land, receiving agricultural inputs or cause shortages in the supply of agricultural labor (Bora et al., 2010; Pietzsch, Talley, and Navarro-Colorado, 2018; FAO et al., 2021; Stephens et al., 2020). Similarly, when disasters compromise transportation networks this can also interrupt the flow of trade and goods, contributing to food shortages (Pietzsch, Talley, and Navarro-Colorado, 2018; Stephens et al., 2020). Shortages in food supplies can also increase food prices, making staple foods unaffordable (Aday and Aday, 2020; Davis, Downs, and Gephart, 2021; Weil and Zachmann, 2022), and the negative economic impacts many humanitarian emergencies have on people’s livelihoods can affect households' ability to purchase the nutritious foods they require (Maxwell et al., 2008; European Commission, 2013; Pietzsch, Talley, and Navarro-Colorado, 2018).
Beyond the immediate effects of a humanitarian crisis, research on ‘poverty traps’ also emphasizes that negative shocks can persist and potentially push households into a downward spiral of destitution (a low-level equilibrium) from which they do not fully recover (Carter and Barrett, 2006; Hallegatte et al., 2017; Kovacevic and Pflug, 2011). In this view, the inability of a community to respond to emerging and future humanitarian shocks can increase households’ risk of long-term food insecurity (Barrett, 2010).

In light of the continued food insecurity issues caused by emergencies, international actors have reiterated their focus on ensuring all populations remain food secure. The most recent iteration of this doctrine comes from the 2015 Sustainable Development Goals (SDGs). Specifically, SDG 2 which aims to: “end hunger, achieve food security and improve nutrition and promote sustainable agriculture” (United Nations, 2022). Food insecurity in emergency settings has also been brought to the forefront of the development agenda by organizations bringing a specific focus to these contexts. For example, the World Food Program (WFP) takes a leadership role in the delivery of food assistance in emergencies and working with communities to improve nutrition and build resilience. Its specific focus on emergency assistance illustrates the recognition of the specific needs and challenges encountered by populations in humanitarian settings (WFP, 2022). However, the increasing number of people requiring humanitarian assistance each year and the limited available resources to address these issues highlights the need to ensure the efficient use of available resources.

3. HUMANITARIAN INTERVENTIONS AND FOOD SECURITY OUTCOMES

An EGM is a thematic collection of evidence on the effects of development policies and programs in a particular sector or thematic area (Snistveit et al., 2017). EGMs are designed around a framework of interventions and associated outcomes. This section describes the scope of the different types of humanitarian interventions and the outcomes we focus on in this research.

Interventions in humanitarian settings

Given that there are different causes of food insecurity in humanitarian settings, there are also many established ways food security may be addressed in these contexts. The scope of the interventions selected for this EGM has been developed in consultation with USAID’s Bureau for Humanitarian Assistance (BHA), as well as other relevant practitioners and experts on humanitarian emergencies and food security. We identified interventions through a process where we invited sectoral experts to make suggestions and comment on the priority interventions that they would like included in this evidence-mapping exercise. We then grouped the prioritized suggestions into seven types of interventions (see Table 1). Appendix 1 presents a detailed account of how the EGM’s scope was developed, as well as the sub-categories of interventions for each intervention type. The sub-categories reflect the different ways these types of interventions may appear in practice.
Table 1. Types of humanitarian interventions included

<table>
<thead>
<tr>
<th>Intervention type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early warning systems</td>
<td>Early warning systems are integrated systems for monitoring, collecting data, analyzing, interpreting and communicating information about the likelihood and risk of meteorological and other events.</td>
</tr>
<tr>
<td>Interventions pre-arranging household finance for disasters</td>
<td>Interventions that create financing arrangements that trigger or can be used during a humanitarian emergency or disaster.</td>
</tr>
<tr>
<td>Food, cash and other in-kind transfers</td>
<td>Interventions directly providing households with food, cash or other in-kind, non-food transfers. This includes the provision of vouchers and non-food items (NFIs).</td>
</tr>
<tr>
<td>Agriculture and livestock interventions</td>
<td>Interventions providing inputs and support to the agricultural and livestock sectors. These interventions assist the cultivation and production of food during an emergency, as well as the management of livestock (e.g. through commercial and slaughter destocking).</td>
</tr>
<tr>
<td>Nutrition interventions</td>
<td>The direct provision of nutritious foods, beyond the normal ration of home diets, and supplements or interventions disseminating information or educating people about nutrition, as well as the provision of treatment for malnutrition.</td>
</tr>
<tr>
<td>Market-based recovery interventions</td>
<td>The provision of cash, training and inputs to traders, and suppliers working outside of the agriculture and livestock sectors.</td>
</tr>
<tr>
<td>Water security interventions</td>
<td>Interventions which directly provide water for human consumption as well as water management interventions.</td>
</tr>
</tbody>
</table>

Food security outcomes

Our consultation process also captured outcomes that experts and practitioners prioritized and suggested for this EGM. In Table 2, we have grouped these outcomes according to the four dimensions of food security. Each group contains direct measures of food security and its different dimensions, as well as factors related to these issues such as economic or agricultural production outcomes. These are approximate groupings, reflecting that the different dimensions and factors determining food security are interrelated. For example, economic outcomes will primarily affect food access but also food utilization and stability. We also include additional outcomes related to health, given the intrinsic links between the two issues (as highlighted during our consultation process), as well as composite measures of food security.
Table 2. Types of food security outcomes included

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Outcome type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food availability</td>
<td>Agricultural production</td>
<td>Measures of agriculture and livestock production and the use of improved technologies and skills.</td>
</tr>
<tr>
<td>Food availability</td>
<td>Food trade</td>
<td>Measures of whether food is being stocked by traders or whether it is being imported into the community.</td>
</tr>
<tr>
<td>Food availability</td>
<td>Markets</td>
<td>Outcomes that measure the ability of markets to supply consumers.</td>
</tr>
<tr>
<td>Food access</td>
<td>Food decisions</td>
<td>Measures coping strategies and decision-making concerning food consumption.</td>
</tr>
<tr>
<td>Food access</td>
<td>Hunger</td>
<td>Measures of household hunger including the Household Hunger Scale (HHS).</td>
</tr>
<tr>
<td>Food access</td>
<td>Economic</td>
<td>Economic indicators related to household purchasing power, including wealth, income and employment.</td>
</tr>
<tr>
<td>Food utilization</td>
<td>Food safety</td>
<td>Measures of whether food is being prepared safely.</td>
</tr>
<tr>
<td>Food utilization</td>
<td>Nutrition</td>
<td>Prevalence and incidence of malnutrition, micronutrient status, anthropometric measures and nutritional knowledge and behaviors.</td>
</tr>
<tr>
<td>Food utilization</td>
<td>Food intake</td>
<td>Measures of food consumption, dietary diversity and food expenditure.</td>
</tr>
<tr>
<td>Food stability</td>
<td>Sustainability of food security</td>
<td>Poverty and inequality and whether food production practices follow best practices for land use sustainability.</td>
</tr>
<tr>
<td>Health</td>
<td>Health</td>
<td>Morbidity, mortality and health knowledge and behavior.</td>
</tr>
<tr>
<td>Composite measures of food security</td>
<td>Composite</td>
<td>Food security measures that combine different aspects of food security into a single indicator.</td>
</tr>
</tbody>
</table>

4. METHODS

We adopted common standards and methods for compiling an EGM (Snistveit et al., 2016; 2017; White et al., 2020) and have outlined in a detailed protocol document and its technical appendix our methods for including, searching, screening and coding literature (see Yavuz et al., 2022). Below we summarize the key aspects of our approach for identifying relevant literature and details of the results of our search.

4.1 Criteria for including and excluding studies in the EGM

Table 3, we summarize key criteria determining whether we included a study in this EGM. This describes the populations, interventions, comparators, outcomes and study designs (PICOS) of studies included in the map, as well as other relevant characteristics.
Table 3. Summary criteria for studies to be included in the EGM

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Studies implemented in a humanitarian emergency consisting of participants residing in low- and middle-income countries. We also included studies consisting of participants residing in a selection of high-income countries, so long as the intervention was implemented in a refugee camp.*</td>
</tr>
<tr>
<td>Interventions</td>
<td>We included the following seven intervention categories: early warning systems; interventions pre-arranging household finance for disasters; food, cash and other in-kind transfers; agriculture and livestock interventions; nutrition interventions; market-based recovery interventions; water security interventions.</td>
</tr>
<tr>
<td>Comparison</td>
<td>A study must have included a comparison group, though there were no exclusion criteria based on the comparison condition of a control group.</td>
</tr>
<tr>
<td>Outcomes</td>
<td>We included the following types of outcomes: agricultural production outcomes; food trade; markets; food decisions; hunger; economic; food safety; nutrition; food intake; sustainability of food security; health; composite measures of food security.</td>
</tr>
<tr>
<td>Study designs</td>
<td>We included both impact evaluations and systematic reviews. For impact evaluations, we included studies using an experimental or quasi-experimental design. Systematic reviews which synthesized the effects of an intervention on outcomes were included (detailed descriptions of included study designs are available from the technical appendix of this study’s protocol).</td>
</tr>
<tr>
<td>Language</td>
<td>Studies in any language were eligible, although search terms used were in English.</td>
</tr>
<tr>
<td>Publication date</td>
<td>All studies published from 2000 onwards.</td>
</tr>
<tr>
<td>Status of studies</td>
<td>We included ongoing⁴ and completed impact evaluations and systematic reviews. This includes prospective study records, protocols and trial registries.</td>
</tr>
</tbody>
</table>

Notes: * Studies of interventions in refugee camps in the following high-income countries are included: Austria, Belgium, Cyprus, France, Germany, Greece, Lithuania, Italy, Slovak Republic and Spain.

⁴ Ongoing studies, such as protocols, were included when they provided sufficient information to meet all criteria. This includes an explanation of primary and secondary outcomes, as well as the intervention to be evaluated.
Improving food security in humanitarian emergencies: An evidence gap map

databases, repositories and institutional websites with sufficient search functionality. We also developed a list of databases, repositories and websites to search. This list was informed by suggestions from sector experts and our information specialist, as well as other systematic reviews and evidence gap maps. To reduce the risk of publication bias, these sources have been selected to cover a range of publication types, including academic journal articles, working and discussion papers, conference proceedings, theses and dissertations, and institutional reports. A complete list of evidence sources and search terms is available in Appendix 2.

We complemented these electronic database searches using forwards and backwards citation tracking techniques, while also publishing an online blog calling for suggestions for includable studies from academic and other evidence communities. Backward citation tracking involves manually screening the reference lists or bibliography of included studies for other eligible studies cited in its text. Forward citation tracking consists of searching for eligible studies that have cited the original included study we identified. For forward citation tracking we utilized Google Scholar Citations.

4.2.2 Search Results

Overall, we identified 70,106 records after searching 14 academic databases and 44 grey literature sources in July 2022 and conducting citation tracking in September 2022. After removing duplicate records, we identified 45,683 records using our search strategy (see the PRISMA diagram reported in Figure 1. We used a combination of techniques to identify includable studies from this large sample of records. This included using manual screening, machine learning models and 3ie’s repository of previously screened literature to prioritize and remove irrelevant records (see Appendix 3 for more details). In total, machine learning technology was used to exclude 26,800 records and independent reviewers single-screened the titles and abstracts of 16,580 further records against the inclusion criteria. 3,366 records were included for screening at full text. The full text of 55 records could not be identified, but the remaining 3,311 records were screened at full text by two independent reviewers. Any disagreements between reviewers were reconciled with the input of a third reviewer when necessary. Appendix 4 provides examples of records excluded on different criteria.

5 The online blog is available at: https://www.3ieimpact.org/blogs/mapping-evidence-gaps-interventions-improving-food-security-humanitarian-settings
Figure 1. PRISMA Diagram of systematic search and screening
Ultimately, 227 records satisfied our inclusion criteria. This includes 163 unique studies (146 impact evaluations and 17 systematic reviews) and 64 linked papers. Over half of the included studies were journal articles (n = 88), with the rest being published reports (n = 39), working papers (n = 25) and theses (n = 6). The remaining 5 studies are protocols or ongoing impact evaluations. A full list of included studies is available in Appendix 5.

### 4.3 Coding of study characteristics

Having identified impact evaluations and systematic reviews on the effects of humanitarian interventions on food security outcomes, we extracted data on study characteristics. This includes factors such as the year of publication, country of the intervention, emergency type, study design and so on. Data extraction was conducted independently by trained reviewers. Reviewers requested a second opinion when questions or uncertainties arose. When extracting data, we created a single entry combining the information of a main study and its linked publications. Drawing on all linked publications ensured that the information available was as comprehensive as possible for each study reported. Appendix 6 presents the data extraction templates and systematic review critical appraisal criteria.

We also conducted critical appraisals of all 17 systematic reviews using an adapted version of the SURE checklist, which is used for making objective judgements about how much confidence can be placed in a systematic review. Appraisals for reviews that already existed within 3ie’s Development Evidence Portal were double-checked by a core team member, and appraisals for newly identified reviews were completed by a senior staff member with expertise in synthesis methodologies. Through this appraisal, we assessed the rigor of the methods applied in systematic reviews against gold standard practices relating to the search, screening, data extraction and synthesis methodologies each review employed. We rated each systematic review as low, medium or high confidence drawing on guidance provided by Snilstveit et al. (2017). We do not critically appraise impact evaluations. Such assessments are time and resource intensive, and this activity is typically beyond the scope of EGMs.

### 5. ANALYSIS

In this section, we examine the characteristics of the 146 impact evaluations and 17 systematic reviews we have identified on the effects of humanitarian interventions on food security outcomes. We

---

6 We considered two or more papers linked when they were published by the same authors studying the same intervention and research question. In many cases, the linked papers would have the same sample and analysis but a different publication status (e.g. working paper vis-à-vis journal article). In other cases, papers used a subset of the dataset shared by the main paper(s), and/or included additional analysis and/or focused on additional outcome measures. See Appendix 3 for further information on how we identified the main and linked papers.

7 Within our findings section, ongoing studies are classified and grouped as impact evaluations. The online EGM differentiates these and allows users to view where upcoming research is focusing.


9 3ie’s Development Evidence Portal (DEP): https://developmentevidence.3ieimpact.org/
summarize our main findings and then describe some of our observations about the methods featured in the impact evaluations we identified.

5.1 Analysis of evidence on the effects of humanitarian interventions on food security

5.1.1 Growth of the evidence base

Humanitarian settings are challenging contexts in which to conduct evaluations of interventions. They are often complex settings with a multiplicity of actors and practical and operational challenges. For example, the need for rapid action and the low predictability of disasters often means that very little preparation for an evaluation is possible in advance of an emergency occurring and logistical issues regularly stymie attempts to locate participants when areas become unstable and cause migration or displacement. Such operational problems have historically limited the implementation of impact evaluations of humanitarian interventions (Puri et al., 2014; 2017).

However, we find that progress is being made to increase the evidence available on the effects of humanitarian interventions and that there has been a significant increase in the number of studies during the last five years. Three-quarters of impact evaluations (n = 110) have been published since the beginning of 2017. This accounts for a three-fold increase in the total number of impact evaluations on this topic during the past five years (increasing from 36 studies in 2016 to 146 in 2022).

Figure 2 further presents the distribution of studies identified by year. It displays the upward trend in the number of studies published each year since 2017, with a peak of 24 impact evaluations identified from 2020 and 2021. The figure also shows a potential decrease in the number of impact evaluations identified from 2022 (n = 17), though it is still too early to conclude that this represents a true trend in the data since this is a mid-year figure. Compared to the historical volume of impact evaluations identified in previous years, this is also still a relatively large number of studies.

We also find that there has been a corresponding increase in the number of systematic reviews published during the last five years. We identify four systematic reviews published before 2017, including those by Hall, Blankson and Shoham (2011) on nutrition interventions, Pega et al. (2015) on cash transfers, Ramesh et al. (2015) on water security and Doocy and Tappis (2016) on cash-based approaches. A further 13 systematic reviews have been published since 2017. For example, Blanchet et al. (2017) on emergency public health interventions and Carroll et al. (2017) and Balhara et al. (2017) on nutrition interventions (see Appendix 5 for the complete list of systematic reviews identified).
5.1.2 Intervention and outcome coverage

Despite the recent growth of the evidence base, we find that evidence is still spread thinly across different interventions and outcome types. Figure 3 presents a bubble graph showing the distribution of impact evaluations and systematic reviews by intervention-outcome (I-O) type pairing. The size of each bubble reflects the number of studies in each I-O pairing and the values next to each bubble report the number of impact evaluations (left) and systematic reviews (right) in each category. This figure presents our EGM at a higher level of aggregation than the online map for ease of presentation. Readers are encouraged to view the online map to dive deeper into the intervention and outcome sub-categories and peruse the finer details of this evidence base according to their own specific thematic interests.

However, evidence is still spread thinly across interventions and outcomes, with only a more noticeable cluster of studies related to food, cash and in-kind transfers.

---

Notes: The values for 2022 represent a mid-year figure reflecting that the literature search was completed in July 2022.

10 https://developmentevidence.3ieimpact.org/egm/food-security-in-humanitarian-settings-egm
We see that large areas of the graph are characterized by low numbers of studies, as well as absolute evidence gaps (i.e. areas where we did not identify impact evaluations or systematic reviews on these topics). Given the broad nature of this map, it was not anticipated that all interventions would include evidence on all outcomes. For instance, agriculture production outcomes are unlikely to be evaluated in evaluations of nutrition programs. However, it is striking that areas such as early warning systems include no studies across any of the included outcome types. It is also noticeable that there are some outcome types with very few or no impact evaluations. For example, none of the studies we identified examined the effects of interventions on outcomes related to food trade and very few on market outcomes (n = 5) or the sustainability of food security (n = 7).

This point about the scarcity of evidence across different possible interventions and outcome types becomes even more apparent if we are to look at the evidence on the effects of different variations of implemented interventions. For example, no impact evaluations were identified for the following sub-categories of interventions: savings interventions, contingent credit, veterinary support, provision and access to emergency livestock feed, capacity building for animal husbandry, pest and disease control, treatment of malnutrition and monitoring of water availability, demand and quality. Furthermore, even in areas where there are clusters of evidence, many sub-categories of interventions consist of a limited number or no studies. For example, for market-based interventions (n = 17), the intervention category with the third largest number of impact evaluations, only three of the six interventions which make up this category have more than one identified study.

Looking at the areas of Figure 3 that do appear more populated with studies, we find three intervention types with larger clusters of impact evaluations. This includes food, cash and in-kind transfers and to a lesser degree multi-component interventions and market-based recovery interventions. The area concerning food, cash and in-kind transfers particularly stands out among this group. For example, 8 of the 96 I-O pairings in the chart include 10 or more impact evaluations. Three-quarters of these clusters, 6 of the 8 I-O pairings, are related to the effects of food, cash and in-kind transfer interventions.

Examining these larger clusters of evidence further, we find food, cash and in-kind transfers consist of the largest number of impact evaluations overall (n = 79 impact evaluations). The largest concentration of impact evaluations evaluates their effects on outcomes related to food intake (n = 47), economics (n = 30), nutrition (n = 22) and food decision making (n = 15). There are also smaller clusters of impact evaluations on their effects on composite indices of food security (n = 13), health (n = 7), food safety (n = 6) and hunger (n = 5) (see Figure 3). Within this intervention category, the largest number of impact evaluations concerns cash transfers (n = 57) and contains a mix of unconditional transfers (e.g. Premand and Stoeffler, 2020), as well conditional and restricted transfers. Examples of conditions required for receipt of cash include attending nutrition education (e.g. Bliss et al., 2018) and child school attendance (e.g. Clavijo, 2020). Restricted transfers allowed recipients to spend cash on a select number of goods such as food (e.g. Bedoya et al., 2019) and heating equipment (e.g. Lehmann and Masterson, 2014).
Figure 3. Distribution of impact evaluations and systematic reviews by intervention-outcome type pairing

Notes: The size of each bubble reflects the number of studies in each Intervention-Outcome category and the values next to each bubble report the number of impact evaluations (left) and systematic reviews (right) in each category. The total of the values displayed will be larger than the number of studies identified, reflecting a study may contain multiple interventions and outcomes.
The *multi-component intervention* category includes the second-largest number of impact evaluations overall (n = 38 impact evaluations). This category contains interventions implemented using a combination of intervention or sub-intervention types, such as cash transfers and nutrition education as well as the combination of both food and cash. Studies of multi-component interventions evaluate a range of different outcomes, but the most common are *economic* (n = 18), *food intake* (n = 17) and *nutrition* (n = 11). Based on the studies we identified, we have further grouped the combinations of interventions into four multi-component sub-categories. This includes (i) *multi-component cash, food or other in-kind transfers* (n = 8), which includes interventions providing a combination of cash, food or other in-kind transfers (e.g. Pangaribowo, 2012); (ii) *cash, food or other in-kind transfers + market recovery interventions* (n = 7) (e.g. Mercy Corps and Causal Design, 2015) and (iii) *cash, food or other in-kind transfers + nutrition interventions* (n = 5) (e.g. Tanner, Hayashi and Li, 2015). The fourth group is the *multi-component mix* (n = 18) sub-category, which contains all other combinations with a very low frequency of studies.

*Market-based interventions* include the third largest number of impact evaluations overall (n = 17 impact evaluations), it is also the final intervention category to contain an I-O combination of greater than 10 studies. The largest concentration of impact evaluations evaluate *economic* outcomes (n = 13) followed by *food intake* (n = 7) and *agricultural production* (n = 4). Within *market-based interventions* the largest sub-category is *microfinance, community-based organizations and collectives* (n = 10) (e.g. Karlan et al., 2017; Deininger and Liu, 2013).

Finally, examining the clusters of systematic reviews that exist, we find *food, cash & in-kind transfers* as the intervention category with the largest volume of systematic review evidence (n = 12), with the most common outcomes covered by these reviews being *nutrition* (n = 11), *health* (n = 8) and *food intake* (n = 6). Smaller clusters of systematic reviews also exist for *nutrition interventions* (n = 8) and *water security* (n = 5), and we identified one systematic review within the *multi-component interventions* category covering evidence on multiple outcome types.

Figure 3 also allows for the identification of where primary evidence exists, but no systematic reviews are yet available. This is also known as an absolute synthesis gap. Combinations of intervention and outcomes with greater than 10 impact evaluations but no systematic reviews include: *food, cash and in-kind transfers – composite indices* (n = 13), *market-based interventions – economic* outcomes (n = 13) and *multi-component interventions – agricultural production* outcomes (n = 13). This highlights three potential areas suitable for future synthesis.

However, an important caveat of the visual analysis of synthesis gaps in Figure 3 is that areas with clusters of systematic reviews do not preclude opportunities for further synthesis. For instance, there are sub-categories of interventions not covered by existing reviews. For example, *microfinance, community-based organizations and collectives* (n = 10) in the *market-based interventions* category. Again, interested readers can view the online interactive chart for details of the distribution of studies according to sub-categories of intervention. There may also be reasons to update existing reviews, either to improve quality or include new studies We provide an in-depth analysis of the systematic review evidence and consider these issues further in Section 6 below.
5.1.3 Geographic distribution of impact evaluations

Of the 146 impact evaluations we identified, we find evaluations of interventions implemented in 45 countries across the world. More than half of the interventions have been implemented in Sub-Saharan Africa (n = 79), but smaller clusters of impact evaluations are also available on interventions in the Middle East and North Africa (n = 26), South Asia (n = 21), East Asia and the Pacific (n = 11), Latin America and the Caribbean (n = 10) and Europe and Central Asia (n = 4). As four studies (Mekonnen et al., 2022; Roxin et al., 2021; Karlan et al., 2017; Burnham et al., 2005) included evaluations in multiple countries, the total number of countries for which evaluations are identified is greater than the total number of impact evaluations. A full breakdown of impact evaluations identified for each region and country is available in Appendix 7.

There is limited evidence on the effects of humanitarian interventions in many countries vulnerable to climatic and humanitarian emergencies.

Figure 4 also presents the trend in the cumulative number of impact evaluations by region. This highlights that the recent growth in the number of impact evaluations on humanitarian interventions is also largely explained by an increase in the number of studies on interventions in Sub-Saharan Africa. However, the availability of evaluations from other regions is also increasing but at a more moderate rate.

Figure 4. Trends in the cumulative number of impact evaluations by region

Figure 5 illustrates the distribution of the location of interventions in impact evaluations by country. From this map, it is evident that the largest concentration of studies on interventions in Sub-Saharan Africa is from eastern Africa. Countries with the largest number of impact evaluations in eastern Africa include Kenya (n = 15), Ethiopia (n = 11) and Uganda (n = 10). We also identified clusters of impact evaluations of interventions implemented in Lebanon (n = 12), Congo, Dem. Rep. (n = 7), Yemen (n = 6), Bangladesh (n = 6), India (n = 6), Malawi (n = 5), Nigeria (n = 5) and Somalia (n = 5). However, these are not particularly
large clusters of impact evaluations when considering the substantive scope of this EGM. We have identified relatively few or no impact evaluations that meet our inclusion criteria on interventions in other countries.

*Figure 5. Map of impact evaluations by country*

Examining the geographic distribution of evidence further, we find that very few or no impact evaluations are available for many countries vulnerable to climatic and humanitarian emergencies. Figure 6 presents data on countries at risk of climatic and humanitarian emergencies based on the INFORM Risk index (INFORM, 2022) and on countries' levels of food insecurity using the Food Security Index from the Global Food Security Q2 2022 report (Deep Knowledge Analytics, 2022). For illustrative purposes, we also created a composite indicator combining this data and categorizing countries according to whether they were high or very highly vulnerable to humanitarian emergencies and food insecure, vulnerable and food secure, less vulnerable and food insecure, and less vulnerable and food secure. Appendix 7 provides further details of risk and food insecurity scores for each country, as well as further information about how countries were categorized.

Comparing data from the graphs in Figure 5 and 6 we identify 13 countries that are both highly vulnerable to emergencies and food insecure, but with no impact evaluations identified; these are Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Congo, Guatemala, Honduras, Libya, Madagascar, Papua New Guinea, Sudan and Venezuela. For a further eight food-insecure and vulnerable countries, Haiti, Iraq, Mali, Mozambique, Myanmar, Philippines, Syria and Tanzania, we only identified one or two impact evaluations. There is also a broader group of countries with relatively few impact evaluations but, while data shows they are relatively food secure, they also face a high or very high risk of humanitarian crises. For example, this includes countries such as Colombia (n = 3), which faces a high vulnerability to conflict and influx of a refugee population, Ukraine (n = 0) which is currently in a state of conflict and Azerbaijan (n = 0) which is highly vulnerable to both earthquakes and high-intensity conflict. Such crises threaten the future overall food security status of these countries.
Figure 6. Map of countries at risk from humanitarian crises and food insecurity

Index for risk from humanitarian crises and disasters

Food Insecurity Index

Composite indicator for risk from humanitarian crises and disasters and food insecurity

Notes: The vulnerability score is derived from the INFORM Risk Index 2023, which measures countries' risk from humanitarian crises and disasters (INFORM, 2022). Countries were categorized as vulnerable if they scored as either high or very high risk (above a score of 5 on the index). The food security score for each country is derived from the Access to Food Index reported in the Global Food Security Q2 2022 report (Deep Knowledge Analytics, 2022). Countries were classified as food secure when they score above the median score (6.26) and those which scored 6.26 or below were categorized as food insecure.
5.1.4 Distribution of impact evaluations across types of emergency contexts

We find impact evaluations of interventions implemented during the aftermath of a range of emergency types (Table 4). This includes rapid/ sudden onset emergencies (n = 15), slow onset emergencies (n = 54) and emergencies with variable onset periods (n = 71). Other emergency contexts (n = 6) include evaluations conducted following unspecified weather shocks (n = 4) and extreme winters (n = 2).

Table 4 shows impact evaluations are most commonly performed for interventions implemented in conflict-afflicted settings. Approximately 40 percent (n = 63) of interventions were implemented in conflict-affected settings, such as Yemen (e.g. Ecker and Maystadt, 2021). This category includes refugee settings with displacement as a result of conflict, for example Turkey (e.g. Özler et al., 2021) and Lebanon (e.g. de Hoop et al., 2018). Other settings impact evaluations have been conducted include during droughts (n = 26) and epidemics (n = 19) such as COVID-19 (e.g. Brooks et al., 2022) and Ebola (e.g. Rosas, Acevedo and Zaldivar, 2017). A more limited number of impact evaluations also examine the effects of interventions in the aftermath of floods (n = 8), complex crises (n = 8), such as the occurrence of both conflict and droughts concurrently in Somalia and Mali and tropical cyclones (n = 7).

Table 4. Emergency contexts of the included impact evaluations

<table>
<thead>
<tr>
<th>Emergency</th>
<th>No. of studies</th>
<th>Emergency</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid/ sudden onset emergencies</td>
<td>15</td>
<td>Slow onset emergencies</td>
<td>54</td>
</tr>
<tr>
<td>Earthquake</td>
<td>0</td>
<td>Cold wave</td>
<td>0</td>
</tr>
<tr>
<td>Flood</td>
<td>8</td>
<td>Drought</td>
<td>28</td>
</tr>
<tr>
<td>Landslide</td>
<td>0</td>
<td>Epidemic</td>
<td>19</td>
</tr>
<tr>
<td>Tropical cyclone</td>
<td>7</td>
<td>Heatwave</td>
<td>0</td>
</tr>
<tr>
<td>Tsunami</td>
<td>0</td>
<td>International displacement</td>
<td>0</td>
</tr>
<tr>
<td>Volcano</td>
<td>0</td>
<td>Political and economic</td>
<td>7</td>
</tr>
<tr>
<td>Emergencies with variable onset periods</td>
<td>71</td>
<td>Other</td>
<td>6</td>
</tr>
<tr>
<td>Complex crisis</td>
<td>8</td>
<td>Unspecified weather shock</td>
<td>4</td>
</tr>
<tr>
<td>Conflict</td>
<td>63</td>
<td>Extreme winter</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Technological disaster</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wildfire</td>
<td>0</td>
</tr>
</tbody>
</table>

Notes: Studies may have included multiple interventions implemented in different contexts or interventions may have been responding to multiple, distinct, emergencies, hence the total of values is larger than the number of included studies.

However, we also find an absence of evidence related to several types of emergencies. For instance, we find no impact evaluations of the effects of interventions in the aftermath of disasters, such as earthquakes, tsunami’s, volcanic eruptions, landslides and extreme heatwaves. Some of these emergency
settings can represent more extreme disasters and can create uniquely challenging operational circumstances and societal problems. For example, tsunamis, such as the one in the Indian Ocean in 2004, can create extreme levels of destruction and displacement. A priority for future research should be to ensure that our understanding grows about interventions implemented in different contexts, under different levels of severity of disasters and that it is focused on different types of challenges caused by different types of disasters.

5.1.5 Distribution of impact evaluations according to emergency phase of implementation

We find that the vast majority of impact evaluations we identified (n = 115, 79%) have evaluated interventions purposefully designed to mitigate an emergency (i.e. during the emergency response or early recovery phase of a crisis). The other 31 studies exploited the occurrence of an emergency during a pre-existing program to analyze the interventions mitigating effect. For example, Gunnsteinsson et al. (2022) analyzed the mitigating effect of a supplementation program after a tornado occurred in their study area.

The body of evidence related to the effects of anticipatory action is growing but overall it still relatively small.

We also identified a small and growing body of evidence for interventions taking place as anticipatory action (n = 23) (Table 5). Of the 23 studies, 18 have been published since 2015. This aligns with the findings of recent reviews on pre-arranging household finance for disasters, which also highlight a growing interest and number of studies in this area (Weingärtner, Pforr and Wilkinson, 2020; Hill, Peredo and Tarazona, 2021).

**Table 5. Emergency phase of interventions in included impact evaluations**

<table>
<thead>
<tr>
<th>Emergency phase</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-emergency</td>
<td></td>
</tr>
<tr>
<td>Anticipatory action</td>
<td>23</td>
</tr>
<tr>
<td>Post-Emergency</td>
<td></td>
</tr>
<tr>
<td>Emergency response</td>
<td>109</td>
</tr>
<tr>
<td>Early recovery</td>
<td>15</td>
</tr>
</tbody>
</table>

Note: One study (Dietrich and Schmerzeck, 2019) evaluated the effects of a cash transfer during both the emergency response and early recovery phase, hence the total number of studies here is greater than number of included impact evaluations.

5.1.6 Evidence of interventions’ effects on vulnerable populations

Given the thematic focus of our EGM, all populations targeted by the emergency interventions included can be classed as vulnerable. Yet certain populations, such as women (Fuhrman et al., 2020; Davies and Bennett, 2016; Aoláin, 2011), children (Masten and Motti-Stefanidi, 2020; Kousky, 2016), older people (Kaga and Nakache, 2019; Duault, Brown and Fried, 2018; Karunakara and Stevenson 2012), people with disabilities (Kent and Ellis, 2015; Rohwerder, 2013; Hemingway and Priestly, 2006) and refugees and internally displaced people (IDPs) (Lejano, Rahman and Kabir, 2020; Garry and Checci, 2020; Pauvert,
Twigg and Sagramola, 2017), have been historically neglected in the delivery of humanitarian aid, with their specific needs often being overlooked.

There is a growing evidence base on the effects of interventions targeted towards vulnerable groups during emergencies, such as children, refugees, women and IDPs.

We identified a growing evidence base of interventions that specifically targeted each of these vulnerable groups (Table 6). Refugees were the most commonly targeted vulnerable population (n = 23), followed by women (n = 23), IDPs (n = 18), people with disabilities (n = 9) and older people (n = 8).

<table>
<thead>
<tr>
<th>Population targeted by intervention</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>26</td>
</tr>
<tr>
<td>Refugees</td>
<td>23</td>
</tr>
<tr>
<td>Women</td>
<td>23</td>
</tr>
<tr>
<td>Internally displace people (IDPs)</td>
<td>18</td>
</tr>
<tr>
<td>People with disabilities</td>
<td>9</td>
</tr>
<tr>
<td>Older people</td>
<td>8</td>
</tr>
</tbody>
</table>

5.2 Synopsis of impact evaluation methods

In this section we review and summarize our observations about the methods featured in the impact evaluations we identified:

- The growing evidence base provides new illustrative examples of the application of different study designs.

Contextual and resource constraints have historically been a bottleneck to the implementation of rigorous evaluations in humanitarian contexts. Puri et al. (2014; 2017) and Leresche et al. (2020) suggest this partly reflects that evaluations of interventions in humanitarian settings, especially in the immediate aftermath of a disaster, pose unique and important methodological challenges. Furthermore, the lack of impact evaluation experts in the humanitarian sector and the shortage of humanitarian expertise in the impact evaluation community has accentuated these challenges.

The growing literature base now provides many illustrative examples of the application of a range of study designs in humanitarian contexts. This includes randomized control trials (e.g. Quattrochi et al., 2022) and natural experiments (e.g. Siu, Sterck and Rodgers, 2021) and studies using non-experimental designs such as statistical matching (e.g. Jamaluddine et al., 2022) and difference-in-difference and fixed effects regressions (e.g. Kalid et al., 2022) (among others). Table 7 presents the frequency of included study designs. Of the 146 included impact evaluations, 42 percent of studies used an experimental study design (n = 62), and 58 percent used a quasi-experimental design (n = 84). Across different designs there are seldom trends in the methods used to evaluate particular interventions. The exception to this is regression discontinuity design where 15 of the 18 studies were within the cash transfers sub-category.
Table 7. Frequency of included studies by study designs

<table>
<thead>
<tr>
<th>Evaluation design</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental impact evaluations</td>
<td>62</td>
</tr>
<tr>
<td>Randomized controlled trials (RCTs)</td>
<td>61</td>
</tr>
<tr>
<td>Natural experiment</td>
<td>1</td>
</tr>
<tr>
<td>Non-experimental impact evaluations</td>
<td>84</td>
</tr>
<tr>
<td>Difference-in-differences &amp; Fixed effects</td>
<td>31</td>
</tr>
<tr>
<td>Statistical matching</td>
<td>25</td>
</tr>
<tr>
<td>Regression discontinuity design</td>
<td>18</td>
</tr>
<tr>
<td>Instrumental variable estimation</td>
<td>9</td>
</tr>
<tr>
<td>Interrupted time series analysis</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: These figures represent a studies dominant estimator. E.g. If statistical matching was used to select a sample and then authors used a difference-in-differences or fixed effects regression, the difference-in-differences or fixed effects has been coded as the dominant estimator. Hence, statistical matching refers to cross-sectional studies only and are characterized by the strong ignorability assumption. The group of difference-in-diffrent and fixed-effect studies are primarily characterized by the parallel slopes assumptions.

- Examples of rarer non-experimental impact evaluation designs, such as instrumental variables and regression discontinuity designs, are starting to appear in this literature.

It is not always possible or feasible to randomize the provision of interventions. Non-experimental designs provide alternative approaches that are capable of establishing causal relationships between interventions and outcomes when carefully executed (Hansen, Klejnstrup and Andersen, 2013; Chaplin et al., 2018; Waddington, Fenton Villar and Valentine, 2022).

However, some designs, such as instrumental variable and regression discontinuity designs, are generally less common in the broader evaluation literature due to the data requirements often necessary to create a convincing application of the approach. For example, identifying interventions with discontinuities or thresholds determining policy assignment can be challenging without having in-depth knowledge of the implementation of policies. This is because not all policies interventions assignments are characterized by such features. Instrumental variables similarly require in-depth knowledge of policies to ascertain whether exogenous factors determine the assignment or uptake of the policy and, even with this knowledge, it may not be possible to identify suitable instrumental variables that explain some degree of variation in a policy variable but has no independent effect on the outcome of interest (Glewwe and Todd, 2022).

Yet novel examples of the application of the regression discontinuity design and instrumental variable approach are starting to emerge in this literature. Examples include using poverty scores (e.g. Ivaschenko et al., 2020), age (Bottan, Hoffmann and Vera-Cossio, 2021) and arrival dates within a refugee camp (MacPherson and Sterk, 2021) as suitable discontinuities. Examples of instrumental variables include conflict intensity along roads pivotal for intervention implementation (Ecker and Maystadt, 2021) and the provision of insurance uptake encouragement mechanisms such as vouchers (Janzen and Carter, 2018; Matsuda, Takahashi and Ikekami, 2019).
• Several illustrative examples of studies using innovative sources of data to inform impact evaluations of humanitarian interventions are also available.

Primary data collection is a significant challenge in humanitarian contexts and a costly aspect of research (Nair et al., 2022). We identify several examples of impact evaluations using alternative sources of data to measure the effects of humanitarian interventions. For example, 11 studies rely solely on secondary data. One source of secondary data includes general public surveys, such as the Indonesian Family Life Survey (Satriawan, 2006), the Yemen National Social Protection Monitoring Survey (Ecker and Maystadt, 2021) and the Ethiopia Rural Household Survey (Gilligan and Hoddinott, 2007). In other examples, authors use monitoring data and “re-cycled” survey data previously collected for an evaluation of a program to answer new research questions (e.g. Lascano Galarza, 2021; Malaeb and Uzor, 2017).

In recent years, there has also been a significant degree of innovation and a trend toward the use of remotely sensed, geospatial data and big data in impact evaluations and econometric research (Rathinam et al., 2020). From our search and included studies, we’re able to identify two novel uses of satellite data. Four included studies used satellite data to identify weather shocks (Premand and Stoeffler, 2020; Dietrich and Schmerzeck, 2019; Knippenberg and Hoddinott, 2017; Macours, Premand and Vakis, 2012). Another study, while excluded from our EGM for not evaluating an outcome of relevance, used nightlight data from satellite imagery to understand economic recovery post-disaster in Mexican municipalities (del Valle, de Janvry and Sadoulet, 2020).

• Impact evaluations that compare multiple prospective interventions can help to account for ethical concerns, inform investment decisions and reduce the confounding that may bias comparisons of the effects of interventions across studies.

The implementation of randomized control trials and other types of impact evaluations in humanitarian contexts have been viewed with some controversy, with issues mainly anchored on ethical concerns. For example, questions have been raised about the appropriateness of having a pure control group receiving no assistance or whether the most vulnerable should be targeted first rather than assistance being randomized among a population (Quattrochi et al., 2020a). However, arguments also support that the designs of impact evaluations can account for ethical concerns. For instance, Puri et al. (2014; 2017) and Quattrochi et al. (2020a) point to designs that compare different prospective types of interventions (i.e. via multiple treatment arms or a straight comparison of two intervention groups).

Along these lines, we identify 28 impact evaluations using multiple treatment arms to compare the effectiveness of intervention types. This includes comparisons of the provision of cash and food (e.g. Babatunde and Olagunju, 2020; Schwab, 2019; MacPherson and Sterck, 2021), cash and vouchers (e.g. Aker, 2017) and the provision of food and a food-for-work program (Gilligan and Hoddinott, 2007). In addition to addressing ethical concerns, these types of within-study comparisons can help to reduce the confounding factors that may bias comparisons of interventions across studies (Fenton Villar and Waddington, 2019; Waddington, Fenton Villar and Valentine, 2022).

• Mixed methods impact evaluations on humanitarian interventions exist, but they represent a minority of studies.

The idea of complementing rigorous quantitative impact evaluations with qualitative data is gaining currency in the development sector. Qualitative information can inform evaluation designs and provide
valuable insights to better understand and explain findings from quantitative analysis. They may also help us to validate the main findings (White, 2008). We find that approximately a quarter of the included impact evaluations (n = 36) collected and analyzed both qualitative and quantitative data. For example, Avdeenko and Frölich (2019) use qualitative interviews to highlight the mechanisms by which the intervention worked. This again is promising and highlights that examples of the application of impact evaluation methods in the humanitarian context are becoming more available. However, mixed methods studies remain a minority of all impact evaluations and future research would benefit from the richness of information gained from triangulating between quantitative and qualitative data.

- **Very few studies adopt equity-sensitive research and analytical methods.**

Equity sensitive approaches are vital for ensuring vulnerable populations are not harmed during all phases of programming and for ensuring the full effects of a program are discovered. Despite all studies targeting vulnerable populations, we were only able to identify one study which utilized an equity sensitive methodology, and one which utilized an equity sensitive analytical framework.

Lombardini and Mager (2019) evaluated data collected using the SenseMaker approach. This equity sensitive methodology was used to collect narrative information from participants on their experiences and expectations of the project. This approach gives a voice to those who are often unheard and allows for data to be collected in a manner which is more meaningful to participants than a purely quantitative survey. Malaeb and Uzor (2017) conceptualize poverty through an equity sensitive analytical framework. Using the Alkire-Foster poverty measurement approach, authors conceptualized poverty based on societal conditions and against a dual cut-off criterion. Despite finding little evidence on equity considerations, this is not to say that studies did not ensure the safety of their participants during implementation and data collection. Instead, it shows that equity is often not considered in the design of programs and their methodologies.

- **Impact evaluations rarely present cost-analysis.**

Integrating cost information into impact evaluations is key to understanding not only what works, but the cost-effectiveness of interventions. This is particularly important for areas such as the humanitarian sector that are facing considerable resource constraints. We find about 55 percent of impact evaluations reported some form of cost data (n = 86). This mainly consisted of reporting information on the amounts of transfer payments to program participants (n = 69). However, this level of reporting is rarely useful for determining the cost-effectiveness of interventions (see Brown and Tanner, 2019 for an elaborate discussion of these issues). Very few included impact evaluations conducted any form of formal cost-analysis (8%, n = 12). This is also below standards in other development sectors where approximately one in five impact evaluations are integrating cost-analysis into their evaluations (Brown and Tanner, 2019). This issue represents an important area that could be improved on in future research.

- **Improvements in reporting of ethical approval and practices are required.**

Ethics approval and reporting is a key step to ensuring that study participants are not harmed during the implementation and evaluation of a program (Evans, 2021). This is particularly important in humanitarian contexts where people’s welfare is or can become increasingly more vulnerable. Despite this, a relatively small share of impact evaluations report having obtained ethical clearance (n = 42). The remaining 104 studies (71%) did not report obtaining ethical clearance. This does not necessarily indicate that it was not
obtained but improvements in transparency and the standard of ethical reporting in research are required of future work.

6. SYSTEMATIC REVIEW CRITICAL APPRAISALS AND SUMMARY OF FINDINGS FROM HIGH- AND MEDIUM-CONFIDENCE REVIEWS

In this section, we examine the results of our appraisals of the systematic reviews we identified and review the key findings and policy implications from the seven medium- and high-confidence reviews included in this EGM. The first section (6.1) presents the results of the quality appraisals of included reviews. Section 6.2 reports the characteristics of high- and medium-quality systematic reviews. Section 6.3 presents specific key findings by intervention type, while Section 6.4 describes cross-cutting implications for research and practice. Section 6.5 provides information about the caveats related to these findings, as well as the synthesis gaps this further analysis of the systematic review evidence has highlighted.

6.1 Results of systematic review critical appraisals

We identified 17 systematic reviews in total. Of these, we appraised three as high confidence, four as medium confidence and 10 as low confidence (Figure 7). Reviews were most typically assessed as low confidence because they did not conduct independent data extraction (n = 4) and/or independent screening of studies at full text (n = 4). Many reviews also failed to conduct a critical appraisal of included studies to identify potential risks of bias (n = 5), meaning that it is unclear which of their conclusions may be based on low-quality evidence.

*Figure 7. Confidence ratings for included systematic reviews*
6.2 Characteristics of high- and medium-confidence systematic reviews

All seven high- and medium-confidence systematic reviews were published between 2015 and 2022. Of these seven systematic reviews (Table 8), four covered food, cash and other in-kind transfer interventions, two covered water security interventions and one covered nutrition interventions. Outcomes reported across these reviews include health (n = 4) and nutrition (n = 4), food intake (n = 3), agricultural production (n = 2), food safety (n = 1), economic (n = 1) and food decisions (n = 1).

Table 8. Characteristics of high- and medium-confidence systematic reviews

<table>
<thead>
<tr>
<th>Title</th>
<th>Intervention Category</th>
<th>Interventions</th>
<th>Outcomes</th>
<th>Confidence level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akparibo et al. (2017)</td>
<td>Food, cash &amp; in-kind transfers</td>
<td>Direct provision of food</td>
<td>Health; Nutrition</td>
<td>Medium</td>
</tr>
<tr>
<td>Doocy and Tappis (2016)</td>
<td>Food, cash &amp; in-kind transfers</td>
<td>Cash transfers; For-work interventions; Food vouchers</td>
<td>Food intake; Food decisions; Economic; Agricultural production</td>
<td>High</td>
</tr>
<tr>
<td>Horst (2019)</td>
<td>Nutrition interventions</td>
<td>Protection &amp; promotion of nutritional status</td>
<td>Nutrition</td>
<td>Medium</td>
</tr>
<tr>
<td>Pega et al. (2015)</td>
<td>Food, cash &amp; in-kind transfers</td>
<td>Cash transfers</td>
<td>Health; Food intake; Nutrition; Agricultural production</td>
<td>High</td>
</tr>
<tr>
<td>Ramesh et al. (2015)</td>
<td>Water security interventions</td>
<td>Extension &amp; maintenance of water storage capacity; Direct provision of safe water &amp; purification systems</td>
<td>Health</td>
<td>Medium</td>
</tr>
<tr>
<td>van Daalen et al. (2022)</td>
<td>Food, cash &amp; in-kind transfers</td>
<td>Cash transfers</td>
<td>Health; Nutrition; Food intake</td>
<td>Medium</td>
</tr>
<tr>
<td>Yates et al. (2017)</td>
<td>Water security interventions</td>
<td>Direct provision of safe water &amp; purification systems</td>
<td>Food safety</td>
<td>High</td>
</tr>
</tbody>
</table>

Notes: See Appendix 8 for an extended version of this table which also includes low confidence systematic reviews.

6.3 Main findings of high- and medium-confidence systematic reviews

Next, we present the main findings of the seven high- and medium-confidence systematic reviews, arranged by intervention type. Due to the small samples of included studies, high levels of heterogeneity between these studies, and a dearth of quantitative evidence that would allow for causal attributions, no
reviews employed quantitative synthesis techniques (e.g. meta-analysis). Instead, all included reviews utilized narrative synthesis as the analysis method. All references to the bias of a systematic review’s primary evidence base, come from each systematic review author’s own risk of bias assessments, as described in their reports.

6.3.1 Cash, food and other in-kind transfers

Two high-confidence reviews (Doocy and Tappis, 2016; Pega et al., 2015) and two medium-confidence reviews (van Daalen et al., 2022; Akparibo et al., 2017) analyzed cash transfer interventions. The evidence base for the high-confidence reviews was very small (Doocy and Tappis (2016) included five studies while Pega et al. (2015) included three studies) and was prone to high risk of bias. The evidence base for the medium confidence reviews was slightly larger, with 23 and 24 included studies for van Daalen et al. (2022) and Akparibo et al. (2017) respectively.

Cash, food and other in-kind transfers may be more effective at increasing food security in emergency contexts, while direct food transfers may be more effective for increasing caloric intake. While personal security concerns are cited as a potential barrier for cash transfers, they are also more cost-effective and efficient than direct provisions of food, which may be particularly salient when resources are scarce.

The main findings of included reviews were mixed. Doocy and Tappis (2016) found that cash transfer, voucher and in-kind food assistance interventions increased household food security within conflict-affected populations, maintained food security in situations of food insecurity and drought, and led to improvements in dietary quality and diversity. Conversely, direct food transfers were more effective at increasing per capita caloric intake. Similarly, van Daalen et al. (2022) concluded that while cash transfers were related to decreased acute malnutrition compared to in-kind food or vouchers, there was also some evidence that in-kind food was more effective for increased caloric intake than cash transfers. Akparibo et al. (2017) found evidence supporting the conclusion that ready-to-use therapeutic food is effective in fostering recovery from moderate or severe acute malnutrition and reducing mortality in humanitarian emergencies, though the overall weight gain was small. The low quality of studies and lack of consistently reported outcomes across studies generally limited the authors conclusions. Finally, while Pega et al. (2015) found that, in general, unconditional cash transfers resulted in reductions of children’s risk of acute malnutrition, their review is based on three studies with high risk of bias and thus their conclusions should be interpreted with caution.

The included reviews also suggested that there may be some additional factors to consider specific to designing cash or in-kind programs. First, cash and voucher programs may have barriers related to personal security which may deter potential participants from engaging in interventions (Doocy and Tappis, 2016). Suggestions include considering alternatives to cash-in-hand (e.g. electronic transfers or smart cards) or putting measures in place to ensure security. In addition, unintended consequences including community tension, social exclusion or verbal abuse from non-participants has been reported (van Daalen et al., 2022). This may be avoided by sampling geographically separated clusters. Finally, Doocy and Tappis (2016) concluded that cash may be a more cost-efficient and cost-effective approach than in-kind food distribution, a consideration that may be useful when resources are limited.
6.3.2 Water security interventions

We identified one high-confidence (Yates et al., 2017) and one medium-confidence (Ramesh et al., 2015) systematic review analyzing water security interventions. Yates et al. (2017) included a large evidence base (106 studies), largely due to their inclusion of many study designs. However, the inclusion of various non-rigorous study designs led to 77 percent of their evidence being appraised as having a high risk of bias. Ramesh et al. (2015) identified a much smaller evidence base consisting of two high, three medium and one low quality studies.

Yates et al. (2017) found limited evidence that WASH interventions reduced both the risk and transmission of disease in emergency contexts. More specifically, they found that the effectiveness of chlorine tablets is mixed, and that when distributed as part of hygiene kits, use of tablets is low unless paired with an educational component. Ramesh et al. (2015) found that high-quality studies indicate point of use water treatment interventions are effective in controlling both the prevalence and incidence of diarrhoea, but that self-reported diarrhoea, an outcome present in each included study, suffers from potential reporting bias.

6.3.3 Nutrition interventions

We identified one medium-confidence review analyzing interventions conducted by civil society organizations (Horst, 2019). The author identified a mixed-quality evidence base of only five studies. The RCT and cluster RCT received quality ratings of very good, the non-randomized controlled study and the prospective cohort study received a rating of good, and the fifth study was rated as poor quality due to self-assignment to treatment, large baseline differences and potential spillovers.

Only one of the included studies examined a nutrition intervention, thus the author was unable to undertake a synthesis. In the absence of synthesis, the author suggests that impact evaluations should aim to gather evidence that can be integrated into feedback loops that would help civil society organizations improve the effectiveness of their programs and concludes that further rigorous evidence is required in order to make any claims about effectiveness.

6.4 Common implications for research

Though additional findings are presented by intervention type above, high- and medium-confidence systematic reviews had several common recommendations for research and practice that were applicable across intervention types:
• There is a need for higher quality rigorous evaluations in humanitarian emergencies in order to clearly answer questions related to intervention effectiveness in these settings. No included reviews were able to conduct a quantitative synthesis due to a dearth of causal evidence.

• There is a need for better reporting practices for studies implemented in humanitarian emergencies. This includes the prepublication of study protocols and rigorous reporting of allocation methods, baseline differences between groups, complete quantitative data, attrition, missing data and cost considerations.

• It would be useful for studies to report common outcome indicators. This would eliminate some of the heterogeneity among studies and allow for clearer comparisons and synthesis.

• Successful implementation of interventions in humanitarian emergencies can be influenced by context-specific security issues. Implementers should take measures to ensure the security of both the participants and the staff of the implementing agency. Safety concerns may hinder willingness to participate in interventions.

• The ethical challenges of using pure control groups in humanitarian emergencies can be overcome through alternative methods such as stepped wedge designs, retrospective control groups and natural experiments.

• Time, resources and logistics are common barriers. Yet the quality of the intervention implementation and design is more critical to intervention effectiveness than the emergency context. Consider building local capacity that can be leveraged when emergency situations arise, ensure technical capacity of implementers, and utilize existing pipelines when available.

### 6.5 Caveats and synthesis gaps

The analysis of the systematic review evidence above highlights some important findings for policy and future research. However, it is important to reiterate the caveat that the number of studies each systematic review draws on is relatively small and the heterogeneity between these studies is large. Furthermore, even in areas where there are medium- or high-confidence systematic reviews, there is reason to believe that for some topics recently published reviews would not necessarily include the latest available impact evaluations. For example, despite having the highest number of extant reviews, cash transfers is also the intervention category with the most prevalent synthesis gap.

To explain this point further, since 2018, 44 studies have been published on cash transfers in humanitarian settings, with 10 studies being published in the first half of 2022 alone. Thus, the review by Akparibo et al. (2017) would not include all of the latest evidence available on this topic. Similarly, the literature search from van Daalen et al. (2022) included studies up to December 2021. Since then, eight of the 10 studies published on cash transfers have evaluated health and nutrition outcomes pertinent to their review. The number of potentially relevant studies would also be larger if we were to include multi-component interventions including cash transfers. Only one low-confidence systematic review on cash transfers completed its search for literature during 2022 but our appraisals highlighted issues concerning the comprehensiveness of its search (e.g., it did not search for grey literature).

The **direct provision of food** is another area with a potential synthesis gap. Only one medium-confidence review evaluated the provision of food (Akparibo et al., 2017). Given this review was published five years ago, there is the potential that new research could be incorporated to ensure the findings remain relevant. Indeed, since 2017, nine impact evaluations have been published evaluating the provision of food.
Further to these limitations of the systematic review evidence, while we previously highlighted in Section 5 some absolute synthesis gaps exist in areas where clusters of impact evaluations are available (such as market-based interventions), our appraisals of existing systematic reviews also show there are several areas where only low-confidence systematic reviews exist. This includes the direct provision of food on food intake outcomes (n = 14) and cash transfers on employment outcomes (n = 11) (see appendix 8). This indicates there are areas that warrant both updating and expanded synthesis efforts to increase our confidence in the findings of reviews.

7. WHO IS FUNDING AND IMPLEMENTING IMPACT EVALUATIONS AND SYSTEMATIC REVIEWS OF HUMANITARIAN INTERVENTIONS?

We also extracted information on the implementers, program funders and research funders of included studies when it was available. The most common program implementers were government agencies (n = 40), made up of governments and government departments within L&MICs, implementing development programs within their own countries. International aid agencies were the second most common implementer (n = 35), made up of bilateral and multilateral agencies such as the World Food Programme (WFP). Non-profit organizations are the final group of implementers identified in more than 10 studies (n = 34). This includes groups such as the International Rescue Committee (IRC) and the Bangladesh Rural Advancement Committee (BRAC). Overall, the most commonly reported implementing organizations of evaluated programs are the World Food Programme (n = 17), and to a lesser extent the United Nation’s Children Fund (UNICEF) (n = 5), Oxfam (n = 4), Mercy Corps (n = 4) and Concern Worldwide (n = 4).

<table>
<thead>
<tr>
<th>Program Implementation Implementing Agency</th>
<th>Research Funding Impact Evaluation</th>
<th>Research Funding Systematic Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic institution</td>
<td>5 (3%)</td>
<td>19 (13%)</td>
</tr>
<tr>
<td>Charitable or private foundation</td>
<td>3 (2%)</td>
<td>14 (10%)</td>
</tr>
<tr>
<td>For-profit firm</td>
<td>8 (5%)</td>
<td>3 (2%)</td>
</tr>
<tr>
<td>Government agency</td>
<td>40 (27%)</td>
<td>30 (21%)</td>
</tr>
<tr>
<td>International aid agency</td>
<td>35 (24%)</td>
<td>44 (30%)</td>
</tr>
<tr>
<td>International financial institution</td>
<td>3 (2%)</td>
<td>18 (12%)</td>
</tr>
<tr>
<td>Non-profit organization</td>
<td>34 (23%)</td>
<td>21 (14%)</td>
</tr>
<tr>
<td>Not specified</td>
<td>45 (31%)</td>
<td>56 (38%)</td>
</tr>
</tbody>
</table>

Notes: Where more than one agency is reported by studies, multi-coding is permitted (i.e. a study may include more than one listed funder).
The most common program funders were international aid agencies (n = 29) and government agencies (n = 22) (Table 9). However, a large proportion of studies also do not report the program funding agency (53%, n = 77). The most common reported individual funders of evaluated programs are the World Food Programme (n = 14), the World Bank (n = 10) and Foreign, Commonwealth & Development Office (FCDO) (n = 9). The United States Agency for International Development (USAID) (n = 5) and European Union (n = 3) are also among the most common funders of evaluated programs in this area but, as the values show, are less frequently reported (Table 10).

Table 10. Top 5 most commonly reported implementing agencies and funders

<table>
<thead>
<tr>
<th>Implementing Agency</th>
<th>Program Implementation Funding Agency</th>
<th>Research Funding Impact Evaluation</th>
<th>Systematic Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. WFP (17)</td>
<td>WFP (14)</td>
<td>FCDO (19)</td>
<td>FCDO (4)</td>
</tr>
<tr>
<td>2. UNICEF (5)</td>
<td>World Bank (10)</td>
<td>USAID (15)</td>
<td>Bill &amp; Melinda Gates Foundation (2)</td>
</tr>
<tr>
<td>3. Oxfam (4)</td>
<td>FCDO (9)</td>
<td>World Bank (13)</td>
<td>IDRC (2)</td>
</tr>
<tr>
<td>4. Mercy Corps (4)</td>
<td>USAID (5)</td>
<td>WFP (9)</td>
<td>NORAD (2)</td>
</tr>
<tr>
<td>5. Concern Worldwide (4)</td>
<td>European Union (3)</td>
<td>3ie (6)</td>
<td>UNICEF (2)</td>
</tr>
</tbody>
</table>

Notes: Where more than one agency is reported by studies, multi-coding is permitted (i.e. a study may include more than one listed funder). Acronyms: Foreign, Commonwealth & Development Office (FCDO, formerly DFID), International Development Research Centre (IDRC), International Initiative for Impact Evaluation (3ie), Norwegian Agency for Development Cooperation (NORAD), United Nation’s Children Fund (UNICEF), United States Agency for International Development (USAID), World Food Programme (WFP).

In terms of who funds research, international aid agencies (n = 44) and government agencies (n = 30) were the most common reported funders of impact evaluations. Meanwhile, international aid agencies (n = 7), government agencies (n = 3) and charitable or private foundations (n = 3) were the most common funders of systematic reviews. Again, however, a relatively large proportion of studies did not specify the research funding agency. For impact evaluations, 38 percent (n = 56) did not specify the evaluations funder and 47 percent of systematic reviews (n = 8) did not report this information either (Table 9). The most common reported individual funder of impact evaluations included FCDO (n = 19), USAID (n = 15) and the World Bank (n = 13). This is followed by the World Food Programme (n = 9) and 3ie (6). Similarly, the most common reported funders of systematic reviews are FCDO (n = 4), followed by the Bill & Melinda Gates Foundation (n = 2), International Development Research Centre (IDRC) (n = 2), Norwegian Agency for Development Cooperation (NORAD) (n = 2) and UNICEF (n = 2) (Table 10).

8. CONCLUSION

In this evidence gap map, we identified a growing literature on the effects of humanitarian interventions on food security outcomes, consisting of 146 impact evaluations and 17 systematic reviews, with a significant share of the studies published in the past five years. Despite this encouraging trend, the broad substantive and geographical scope of our study means that the studies are spread thinly across the 8
intervention (40 sub-categories) and 12 outcome (35 sub-categories) types. An exception to this is for the food, cash, and other in-kind transfers category where large clusters of studies are concentrated.

Although the available evidence is limited in its ability to identify clear implications for policy and practice in some areas, we encourage decisionmakers to consider the findings from the studies that are available when designing future policies and interventions. Furthermore, the gaps we have identified also point to an urgent need for investments in additional research. While research in these contexts is challenging for a variety of reasons, results from this EGM highlight that conducting rigorous evaluations in humanitarian emergencies is indeed possible. We would encourage funders, researchers and other stakeholders to begin addressing the most critical evidence gaps in a strategic and coordinated manner.

8.1 Implications for policymakers

Implications from the included medium- and high-confidence reviews are summarized below and in more detail in Section 6, and we suggest readers consult the full reviews for more details. The utility of the findings from these reviews is restricted in many cases due to the small number of included studies, the high risk of bias identified in the primary studies, the heterogeneity between these studies and consequently, the inability to explore effectiveness questions through quantitative synthesis. Implications from existing syntheses are as follows:

- Cash may be a more cost-efficient transfer than food, although the evidence is limited. However, careful thought should be given to security concerns when planning for training and delivery mechanisms for disbursements.
- Intervention choices should be considered with respect to the desired outcome. For example, WASH interventions are promising for reducing the risk and transmission of diseases in emergency contexts, while water treatment interventions are promising for reducing the prevalence and incidence of diarrhea. Cash transfers may be better for improving malnutrition while food or in-kind transfers may be more beneficial for increasing caloric intake.
- The quality of the intervention implementation and design is more critical to intervention effectiveness than the emergency context, so building local capacity that can be leveraged when emergency situations arise may increase the chances of success.

Policymakers can also consult the EGM to determine whether rigorous evidence is available for any programs being considered, keeping in mind the following:

- Where there are absolute evidence gaps for specific interventions or geographies of interest, consider including an impact evaluation within future programming.
- Consult systematic reviews to inform program design, being mindful when considering low-quality reviews that their conclusions should be interpreted with caution.
- For interventions with existing impact evaluations but no recent or high-confidence systematic reviews, consider findings from primary studies in the design of new programs, keeping in mind that conclusions about intervention effectiveness cannot be drawn from single studies, nor by taking a count of how many interventions reported statistically significant positive results. Additionally, single studies may have limited external validity and thus their results may not be applicable across contexts.
- Consider commissioning a systematic review if the cluster of evidence is sufficient for synthesis.
8.2 Implications for future research and summary of evidence gaps

**Primary research**

Conducting rigorous research in humanitarian emergencies is challenging for a host of reasons, including time and resource constraints as well as ethical and security issues. This EGM highlights that there is a growing body of evidence that can be used when considering not only what works to improve food security in humanitarian contexts, but also how to design, implement and analyze interventions so that future studies are able to contribute to this growing evidence base. Yet, the EGM also identifies clear gaps that can be filled through future evaluations. Based on our findings from this work we suggest that when researchers and funders are commissioning and designing new evaluation studies, they consider:

- Prioritizing primary research gaps in intervention categories and sub-categories where no impact evaluations currently exist (such as early warning systems, as well as savings interventions, contingent credit, veterinary support, provision and access to emergency livestock feed, capacity building for animal husbandry, pest and disease control, treatment of malnutrition and monitoring of water availability, demand, and quality). Alternatively, as the evidence base continues to grow, another option could be to consider prioritizing areas where early evidence is promising or interventions which address the largest issues.

- Prioritizing geographical gaps, in particular those countries highly vulnerable to climatic and humanitarian emergencies, with food insecure populations and where no impact evaluations currently exist (e.g. Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Congo, Guatemala, Honduras, Libya, Madagascar, Papua New Guinea, Sudan and Venezuela).

- Prioritize interventions that anticipate or mitigate emergencies for which no impact evaluations were identified (e.g. earthquakes and tsunamis) and interventions implemented as anticipatory action. The two impact evaluations identified on forecast-based finance (Gros et al. 2020; Gros et al. 2019) may be used as illustrative examples of how anticipatory action can be evaluated in the face of natural hazards.

- Incorporating cost-effectiveness or cost-benefit analyses, which may be particularly useful in resource-constrained settings.

- Utilizing the most rigorous methods feasible within the context. This includes adopting alternatives to pure control groups when they are deemed unethical, such as stepped wedge designs, retrospective control groups, natural experiments or ‘small n’ studies if standard econometric approaches are not suitable (see White & Philip 2012).

- Mixed methods which incorporate the collection of qualitative data may also be useful in order to fully understand the mechanisms through which program effects are realized, or to explain limited or no effects.

- Ensuring studies are sensitive to the needs of vulnerable groups by (a) compliance with the universal ethical standards of respect for persons, beneficence, and justice, including receiving and reporting of ethical approvals by relevant review boards and (b) adopting gender and equity approaches in the design of the intervention, collection of data and research process to account for social norms which might reinforce inequities, and make sure data are collected in ways that people, especially women, feel comfortable speaking.

**Synthesis research**

To date, synthesis work in this area has been hindered by the challenges that exist in the underlying evidence base, including a lack of quantitative data on which to base causal attributions and heterogeneity
among primary studies (e.g. a lack of common outcome indicators). This EGM is a positive step towards promoting the use and uptake of rigorous evidence and highlights a number of synthesis gaps. We suggest that when researchers and funders are commissioning and designing new synthesis projects, they consider:

- Prioritizing synthesis gaps identified where no or only low-confidence systematic reviews (such as on the effects of market-based interventions or the direct provision of food on food intake outcomes and cash transfers on employment outcomes).
- Updating existing medium- and high-confidence reviews in areas where new evidence is available (e.g. cash and in-kind transfers on nutrition and health outcomes and the provision of food).
- Commissioning living synthesis projects so that reviews and this map are updated with new evidence, ensuring that decision-makers have access to the most up-to-date evidence.

9. REFERENCES


School Feeding Programme in Lebanon among Lebanese and Syrian Refugee Children', *Public Health Nutrition*, 25(6), pp. 1678-1690. doi: 10.1017/S1368980022000362


