Interventions to reduce anaemia in low- and middle-income countries: An evidence gap map protocol

Charlotte Lane
International Initiative for Impact Evaluation (3ie)

Meital Kupfer
3ie

Ingunn Storhaug
3ie

Diana Cordova-Arauz
3ie

Miriam Berretta
3ie

Daniel López de Romaña
Nutrition International

Mandana Arabi
Nutrition International

Alison Mildon
Nutrition International

Ashiqun Nabi
3ie

Evidence Gap Map Protocol

July 2023

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 and synthesize high-quality 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 EGM protocol provides all the supporting documentation for the map, including the background information for the theme of the map, and details the methods that will be applied to systematically search and screen the evidence base, extract, and analyse data, and develop the EGM report.

About this evidence gap map protocol

This report presents the protocol for a systematic search to identify and map the evidence base of impact evaluations and systematic reviews of interventions that aim to improve anaemia in low- and middle-income countries. The EGM was developed by 3ie with generous support from Nutrition International. The content of this report is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors, or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. Please direct any comments or queries to the corresponding author, Ashiqun Nabi at anabi@3ieimpact.org

Suggested citation: Lane, C, Kupfer, M, Storhaug, I, Cordova-Arauz, D, Berretta, M, Lopez de Romaña, D, Arabi, M, Mildon, A and Nabi, A, 2023. Interventions to reduce anaemia in low- and middle-income countries: An evidence gap map protocol. London: International Initiative for Impact Evaluation (3ie).

Contents

List of tables

List of figures

Acronyms

3ie	International Initiative for Impact Evaluation
AREA	Accelerated Reduction Effort on Anaemia
EGM	Evidence Gap Map
HEME	HEmoglobin MEasurement
HICs	High-Income Countries
IYCF	Infant and young child feeding
L&MICs	Low- and Middle-Income Countries
MAPS	Multi-Sector Anaemia Platform Strengthening
NI	Nutrition International
SDGs	Sustainable Development Goals
SPRING	Strengthening Partnerships, Results, and Innovations in Nutrition Globally
SR	Systematic Review
UN	United Nations
UNICEF	United Nations Children's Fund
UNSCN	United Nations System Standing Committee on Nutrition
USAID	United States Agency for International Development
WASH	Water, Sanitation, and Hygiene
WHO	World Health Organization
WRA	Women of Reproductive Age

1. Background

1.1 Development problem being addressed

Anaemia is a widespread public health problem characterized by low red blood cell volume and haemoglobin concentration. While traditionally conceptualized as relating to iron, anaemia can also be caused by other micronutrient deficiencies, blood loss, and disease, including genetic disorders. Recent World Health Organization (WHO) estimates of the global prevalence of anaemia indicate that 40% (269 million) of children 6-59 months of age, 37% (32 million) of pregnant women, and 30% (571 million) of women of reproductive age are affected by anaemia (WHO, 2021). The WHO regions of Africa and South-East Asia are most affected (WHO, 2021). Trends for the prevalence of anaemia for the years 2000-2019 indicate a slower decline in anaemia prevalence since 2010, as compared to the period 2000-2010 (Stevens et al., 2022). Possible explanations for this decline in progress include a reduction in focus and the complex aetiologies of the remaining cases.

The main determinants of anaemia include nutritional deficiencies, inflammation, genetic haemoglobin disorders, infections (e.g. malaria, schistosomiasis, and hookworm), and chronic diseases (e.g., gynaecological conditions in women, gastrointestinal disease, chronic kidney disease) that lead to blood loss or the destruction of red blood cells (WHO, 2017). These determinants can vary by age and reproductive status. For example, during pregnancy, maternal iron stores are transferred to the fetus. However, breastmilk does not contain iron, so exclusively breastfed infants are at increased risk of iron deficiency as they have almost zero intake. Iron deficiency is the most common nutritional deficiency leading to anaemia (Chaparro & Suchdev, 2019). However, the proportion of anaemic individuals who are iron deficient can range between 30-71 per cent depending on the burden of infection present in the population (Engle-Stone et al., 2017; With et al., 2017).

The immediate effects of anaemia include delayed cognitive and physical development, fatigue, and reduced productivity. Consequently, anaemia is associated with economic losses and preventable negative health outcomes. Women of reproductive age and children are particularly at risk. Anaemia, especially iron-deficiency anaemia, during pregnancy can lead to adverse birth outcomes, such as low birth weight, increased maternal and neonatal morality, and poorer

cognitive development in infancy and early childhood (Rahman et al. 2016). Anaemia is estimated to cause 115,000 maternal and 591,000 perinatal deaths per year (WHO, 2021).

Reduction of anaemia is one of the World Health Assembly Global Nutrition Targets for 2025 (WHO, 2014) and of the Sustainable Development Goals (SDSN, 2012). There has been limited success in reducing anaemia in recent years. The ongoing pandemic and associated global recession have likely further contributed to this stagnation (Osendarp et al., 2021). As of 2022, COVID-19 related disruptions are expected to have resulted in an additional 2.1 million maternal anaemia cases (Osendarp et al., 2021). In 2020, only one country, Guatemala, was on target to reach the World Health Assembly target of reducing anaemia by 50 percent by 2025 (Global Nutrition Report, 2021).

1.2 Policy responses

The World Health Organization has been at the forefront in providing evidence-informed recommended actions to reduce anaemia for decades. It leads surveillance activities and has developed a series of widely adopted tools and resources to support anaemia reduction (WHO, 2023.). Currently, the WHO recommends contextualized, multi-sectoral strategies, with actions implemented synchronously to address the multiple determinants of anaemia (WHO, 2020). It prioritizes interventions to address nutritional and non-nutritional causes of anaemia. The WHO also highlights programs which function within environmental and socioeconomic domains to reduce anaemia. It advocates for broad stakeholder support and a strong policy environment to facilitate the effective prevention and management of anaemia.

Given that nutrition is an essential component of healthcare, with respect to anaemia, agencies often combine interventions with similar goals when devising policies and approaches. As such, a variety of country-level initiatives also combine government, civil society and non-profit actors to further anaemia reduction efforts (Kinkoy et al., 2021; WHO, 2017; WHO, 2020). For example, the Multi-Sector Anaemia Platform Strengthening (MAPS) program is a country-driven effort to develop multi-sectoral anaemia platforms, establish guidelines, and build capacity. However, no nationally scaled, coordinated, and synchronous programs of contextually relevant anaemia reduction activities across the life course have been implemented and evaluated (Moorthy et al., 2020).

Keats and colleagues (2021) offer a four-way classification for maternal and child nutrition interventions and policies, considering direct and indirect interventions which function within and

outside of the health system. Many of the interventions reflected in their framework involve the collaboration between nutrition and health services, such as leveraging antenatal care visits to distribute iron and folic acid supplements, prophylactic deworming, and insecticide treated bed nets. These activities are thought to reflect promising multi-sectoral approaches to anaemia reduction. Keats and colleagues (2021) also highlight family planning and disease prevention and management broadly as approaches to improving nutrition and reducing anaemia. They consider anaemia treatment to be a direct, health-care nutrition intervention. Other programs, such as school meals and cash transfers are also mentioned in their framework, but not directly linked to anaemia. Many anaemia interventions still focus on nutritional iron deficiency, resulting in programs that primarily focus on the delivery of iron receiving higher priority than other approaches.

Various research working groups collaborate to advance science and generate evidence about anaemia, including the HEmoglobin MEasurement (HEME) Working Group and the Biomarkers Reflecting Nutrition Determinants of Anaemia (BRINDA) 2 Project (SPRING, 2023). The United States Agency for International Development (USAID) funded Advancing Nutrition Anaemia Task Force adopts an ecological approach grounded in an appreciation of systems biology (USAID, 2022). The goal of the taskforce is to provide a comprehensive overview of anaemia, understand the biological mechanisms causing anaemia, and translate this knowledge into the development of effective assessment methods and interventions to improve clinical and public health outcomes. An earlier USAID program, the Strengthening Partnerships, Results, and Innovations in Nutrition Globally (SPRING) project, worked to reduce anaemia at local, national and global levels (WHO, 2020). SPRING, WHO, and the United Nations System Standing Committee on Nutrition (UNSCN) started the Accelerated Reduction Effort on Anaemia (AREA) Community of Practice in 2015 to share and collaborate on ways to improve and scale-up strategies to reduce anaemia, including bringing together 700 stakeholders from over 65 countries (SPRING, 2018). Finally, in 2022 WHO and United Nations Children's Fund (UNICEF) convened the Alliance for Anaemia Action, borne out of the Global Food Systems Summit, that aims to provide a platform for countries to learn and leverage their experience on initiatives relating to anaemia reduction. Specifically, the Alliance hopes to ensure all women, adolescent girls, and children are enabled, empowered, and have access to appropriate and timely actions for the prevention and management of anaemia.

1.3 Why is it important to do this evidence gap map?

In 2016, the UN General Assembly declared 2016 – 2025 the “Decade of Nutrition” and set global targets for nutrition improvements. As a result, there is an increasing body of work synthesising the effects of interventions addressing the direct causes of anaemia (Bhutta, 2016, Moorthy et al., 2020, da Silva Lopes et al., 2021, Mithra et al., 2021, Panchal et al., 2022). In addition to traditional fortification and supplementation programs, systematic reviews in the last decade have considered the effects of deworming; antimalarial interventions; *H. pylori* treatment; and water, sanitation, and hygiene initiatives (Mirtha et al., 2021). Evidence on adverse effects, particularly gastrointestinal issues, is becoming increasingly available (da Silva Lopes, 2021; Mirtha, 2021).

However, recent systematic reviews have also highlighted key research gaps around integrating nutrition-specific and nutrition-sensitive programs (Moorthy 2020; da Silva Lopes, 2021) and using multi-sectoral approaches to improve nutrition (Heidkamp et al., 2021). Evaluations of interventions that address fundamental drivers (e.g., politics, ecology, and inequity), underlying risk factors (e.g., low educational attainment, poverty, and cultural norms), and intermediate risk factors (e.g., food insecurity, health and nutrition knowledge, maternal and child care) of anaemia are thought to be relatively uncommon. More evidence on the effects of interventions for specific populations is also needed, including effects among adolescents, non-pregnant women, older women, men, and people with genetic blood disorders (Moorthy, 2020; da Silva Lopes, 2021). In addition, there is a gap in research on the postnatal period around anaemia screening and optimal interventions for treatment (Parker et al., 2012).

Given that there are, on the one hand, key evidence gaps and, on the other, known clusters of evidence, the aggregation and mapping of impact evaluations that address anaemia will help in the prioritization of research and the adoption of evidence-informed policy. Focusing on the effects of multi-sectoral interventions on anaemia may support innovation, cross-sectoral integration and coordination, and the targeting of new research in the field. By highlighting the evidence on the effects of the direct and indirect healthcare sector and other sectoral interventions on anaemia, we aim to align with and reinforce other global initiatives to bring new actors into the conversation about anaemia reduction, potentially highlighting the opportunity for collaboration with the education; agriculture; social protection; and water, sanitation, and hygiene (WASH) sectors in addition to the nutrition and health-care sectors.

Our mapping exercise is an opportunity to move beyond the common conceptualization of anaemia as an issue related primarily to iron deficiency. This conceptualisation ignores the broad spectrum of anaemia related illnesses and morbidities that afflict all genders and ages. It overlooks that anaemia is the result of multiple causes and excludes many potentially valuable partners from the conversation. In addition, it perpetuates gender-based inequalities by approaching women as reproductive bodies rather than independent actors who are themselves significantly affected by anaemia and related issues (Sedlander et al., 2021).

Although 3ie has its Living Food Systems and Nutrition Evidence Gap Map, which maps food systems interventions measuring food security and nutrition outcomes, our previous work does not address the topic of anaemia (Moore et al., 2021). The map excludes studies focused on specific diseases and clinical outcomes, including anaemia. It also does not include interventions which take place outside the food system, such as primary school education or interventions within the health system. So, a mapping exercise is needed to consider interventions addressing the core drivers and determinants of anaemia.

2. Study objectives and questions

This project aims to improve access to evidence on interventions attempting to reduce anaemia in low- and middle-income countries (L&MICs). It will do this by identifying existing and ongoing impact evaluations and systematic reviews of the effects of interventions and presenting the evidence in a clearly structured evidence gap map. The project will facilitate the use of evidence to inform policy decisions, evidence production, and future interventions in L&MICs.

Evidence gap maps (EGMs) are thematically organized collections of rigorous evidence that help policymakers, practitioners and researchers working in a specific thematic area make evidence-informed decisions. EGMs make existing evidence more accessible and ease the prioritization of future research by mapping existing studies in a field on a framework of interventions and outcomes. Figures are developed to describe the volume of evidence for combinations of interventions and outcomes, the type of evidence (completed or ongoing, impact evaluations or systematic reviews), an indication of research gaps, and a confidence rating reflecting the quality of systematic reviews.

The results will be displayed on 3ie's platform, which provides a graphical and interactive display of the evidence in a matrix form. The interactive map will also provide options to filter the evidence

in the EGM by regions, population characteristics such as age groups and place of residence, disease of interest, gender and reproductive status and evaluation methods, among others.

The specific objectives of this EGM are threefold:

1. Identify and describe the characteristics of impact evaluations and systematic reviews considering the effects of interventions to reduce anaemia directly or that address the determinants of anaemia (direct, intermediate, and underlying risk factors) in L&MICs.
2. Summarize the findings from the included high- and medium-confidence systematic reviews on the effects of interventions to reduce anaemia directly or address the intermediate and underlying risk factors for anaemia in L&MICs.
3. Identify potential primary evidence and synthesis gaps.

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

Table 1. Research questions

Research Questions		***Type***
RQ1	What are the extent and characteristics of empirical evidence on the effects of interventions to reduce anaemia directly or address the intermediate, underlying risk factors and direct causes of anaemia in L&MICs.?	Coverage
RQ2	What are the major primary and synthesis evidence gaps in the literature?	Gaps
RQ3	What intervention/outcome areas could be prioritized for primary research and/or evidence synthesis?	Research

3. Conceptual framework and scope

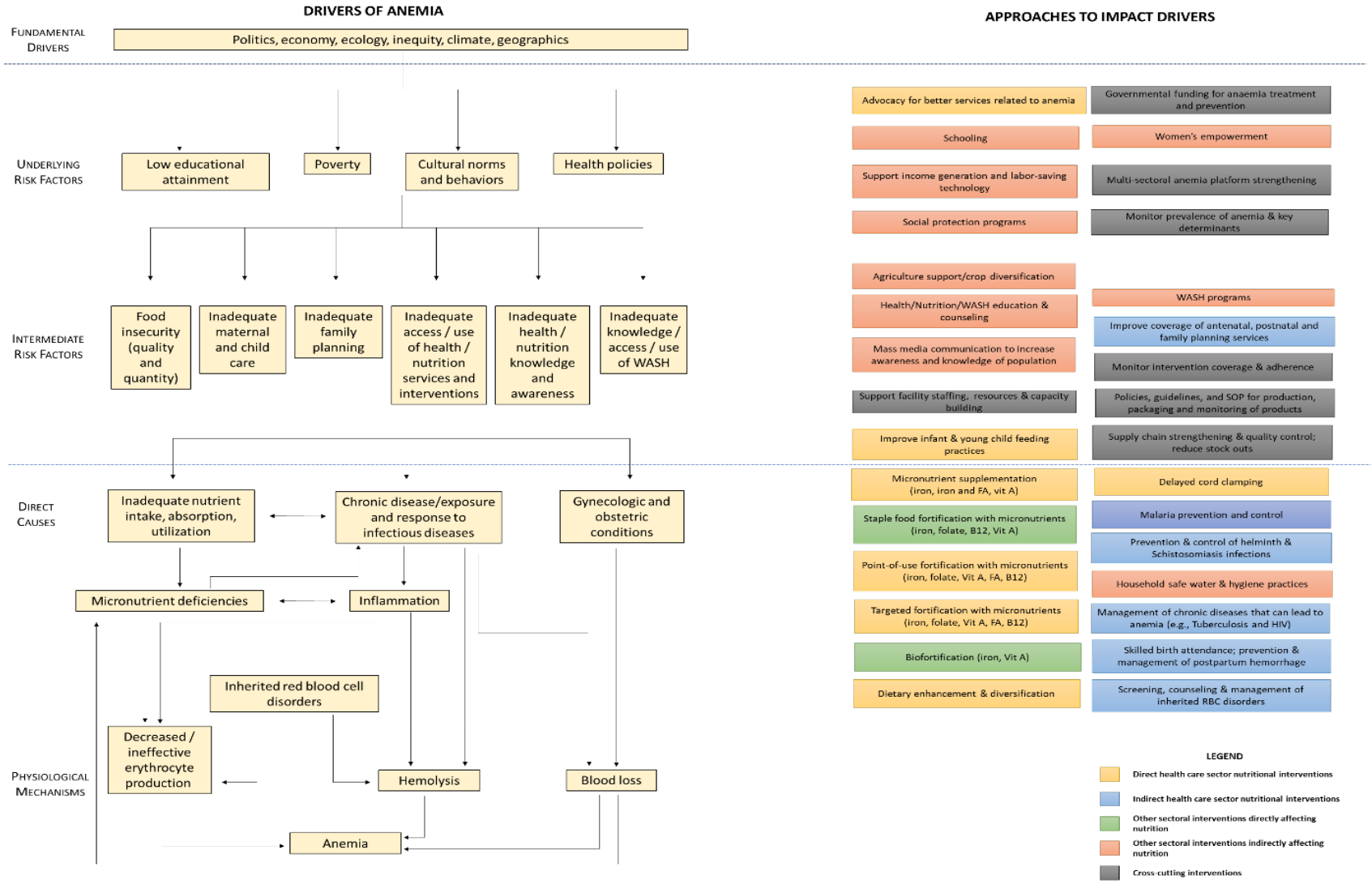
In alignment with growing trends in the field, we approach anaemia as an interdisciplinary issue ultimately driven by politics, the economy, ecology, inequity, climate, and geography (Figure 1, Chaparro and Suchdev, 2019; Hess et al., 2023). These fundamental drivers of anaemia affect the underlying risk factors for anaemia: low educational attainment, poverty, cultural norms and behaviours, and health policies. Underlying risk factors affect intermediate risk factors related to

food insecurity; maternal and child care; family planning; health, nutrition, and WASH knowledge; and access and use of health, nutrition, and WASH services.

Together, these fundamental drivers, underlying risk factors, and intermediate risk factors can aggregate to affect the direct causes of anaemia: inadequate nutrient intake, absorption, and utilization; chronic and infectious disease; and gynaecological and obstetric conditions. When combined with inherited blood disorders, the direct causes can decrease erythrocyte production, induce haemolysis, and lead to blood loss, all of which can cause anaemia.

Our evidence gap map will focus on the three intermediate levels reflected within this conceptual framework: underlying risk factors, intermediate risk factors, and direct causes. Our list of eligible interventions (see 4.3.2 for the full intervention list) is built around these concepts and sets the scope of the EGM to include interventions that have been designed and implemented to address any of these risk factors and direct causes. We will map the evidence in a matrix format with interventions along the vertical axis and outcomes along the horizontal axis. In the map, we will group the interventions under the three abovementioned levels of the conceptual framework. This will help the readers to visualize the clusters and gaps in evidence both at broader levels (i.e. underlying risk factors, intermediate risk factors, and direct causes) as well as for particular intervention outcome pairings.. This decision was made to balance proximity to our outcome of interest with practicality of intervention through development assistance. Intervening at the fundamental driver level may be too distal from the outcome of interest to warrant significant evaluation. Intervening at the physiological mechanisms level would likely require medical intervention, rather than general development assistance.

Figure 1. Conceptual framework



Note: Based on Chaparro and Suchdev, 2019; Hess et al., 2023; Keats et al., 2021. Approaches to impact drivers are meant to be illustrative and not exhaustive.

4. Methods

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

The map will be populated by systematically searching and screening all relevant, completed, and ongoing, impact evaluations and systematic reviews. The included studies will be mapped onto the framework of interventions and outcomes and presented on an interactive, online platform which provides a graphical display of the evidence in a grid-like framework. This visually represents the volume of evidence for intervention-outcome combinations, the type of evidence available (completed or ongoing, impact evaluations or systematic reviews), and a confidence rating for systematic reviews. Additional filters will allow users to further explore the available evidence, for example by global regions, population, or disease of interest.

The interactive map will be accompanied by a report addressing the key research questions, including an analysis of the characteristics of the available evidence and key trends. This will include figures reflecting the number of impact evaluations published over the time, geographical coverage, and interventions and outcomes evaluated.

4.1 Conceptual framework development

The conceptual framework proposed for this work was developed after reviewing a series of published frameworks (Apte et al., 2021; Charparro and Suchdev, 2019; Exemplars in Public Health, 2023; Hess et al., 2023; Namaste et al., 2017; Ngure et al., 2014; Nguyen et al., 2018; Nutrition International, n.d.; Owais et al., 2021; Scott et al., 2022; Sedlander et al., 2021; Shet et al., 2015; SPRING, 2013; WHO, 2020; WHO, 2021; Williams et al., 2020; Yilma et al., 2020) and discussing these with our advisory group to identify those most relevant to this work. The theory of change reflected in our framework is based on the conceptual model of anaemia etiology proposed by Hess and colleagues (2023) who adopted the model from Chaparro and Suchdeve

(2019).[1] The original model was developed based on the determinants identified by Namaste and colleagues (2017), Pasricha and colleagues (2013) and Balarajan and colleagues (2011). These conceptual models group the drivers of anaemia as underlying risk factors (e.g., politics, economy, inequity and geo-climatic factors), intermediate risk factors (e.g., factors associated with physiological vulnerability) and direct causes (e.g., inadequate nutrition, chronic/infectious diseases, gynecologic/obstetric conditions and genetic blood disorders). Our conceptual framework follows the same classification for the drivers of anaemia. The interventions considered for this work were based on this framework and represent the multidisciplinary approach needed to tackle anaemia.

4.2 Search strategy

We will adopt a systematic search strategy following guidelines for systematic literature searching (Kugley et al., 2017). The strategy will address potential publication bias by systematically searching both academic bibliographic databases and grey literature sources such as specialist organisational websites, websites of bilateral and multilateral agencies, and repositories of research in international development. The search strategy will consider sector specific databases, including those related to nutrition, health, education, agriculture, and social protection. We will complement the academic and grey literature search with backward and forward citation tracking. For all included systematic reviews, we will search their list of included studies. To identify additional eligible studies, we will also publish a blog on the 3ie website with an open invitation to share any relevant studies and will contact researchers and experts recommended by the project's external advisory group (members are listed in Appendix A).

We will conduct electronic searches of the following databases of published sources:

- MEDLINE
- EMBASE
- Cochrane Controlled Trials Register (CENTRAL)
- CINAHL

[1] We made minor changes to Hess and colleague's (2023) model to adopt it for our framework. We removed "physiological vulnerability of women, children, and adolescents" from underlying risk factors since this is not something that can be acted or intervened upon. We also modified the intermediate risk factor "limited access to health/nutrition services and interventions" to "Inadequate access to health/nutrition services and interventions " to allow for parallel structure with the other factors in the category.

- CAB Global Health
- CAB Abstracts
- Agricola
- PsychINFO
- Africa-Wide Information
- Academic Search Complete
- Scopus
- Campbell Library

To identify relevant grey literature, we will search the following databases (some of which contain a mixture of published and grey literature):

- Google Scholar
- EconLit
- ENN-Network
- IDEAS/RePEc
- IMMANA grantee database
- WHO Global Index Medicus
- Grey Literature Report
- Social Science Research Network (SSRN)
- Eldis
- Epistemonikos
- 3ie Development Evidence Portal
- Registry of International Development Impact Evaluations (RIDIE)
- Oxfam Policy & Practice

We will also search the following organisational websites:

- AgEcon Search (University of Minnesota
- Innovations for Poverty Action (IPA)
- Abdul Latif Jameel Poverty Action Lab (J-PAL)
- Global Development Network
- World Bank Development Impact Evaluation (DIME) and Impact Evaluation Policy Papers
- Inter-American Development Bank

- Center for Global Development
- Center for Effective Global Action (CEGA)
- DFID Research for Development (R4D)
- USAID
- International Food Policy Research Institute (IFPRI)
- CGIAR
- Food and Agriculture Organization of the United Nations (FAO)
- High Level Panel of Experts on Food Security and Nutrition (HLPE)
- World Food Programme (WFP)
- Action Against Hunger
- UNICEF
- United Nations Evaluation Group
- Asian Development Bank
- World Agroforestry Centre (ICRAF)
- International Livestock Research Institute (ILRI)
- Nutrition International

4.3. Criteria for including or excluding studies

4.3.1 Population

We will include studies of interventions that service any participant type providing they resided in a L&MIC. Income status will be defined based on the World Bank income status classification in the first year of the intervention (Appendix B). Multi-country studies will be included if results are provided for low- or middle-income and high-income countries separately.

Interventions that seek to enhance the performance of a specific niche population, such as athletes, the military, astronauts, or actors/models, will be excluded. However, interventions targeting specific vulnerable populations (e.g., women, persons with disabilities, etc.) will be included. We will also exclude studies of migrants from L&MICs in high-income countries (HICs) or vice versa.

4.3.2 Interventions

Due to practical limitations in scope, we will not be able to fully address the fundamental drivers or physiological mechanisms (Figure 1). Eligible interventions are limited to those addressing

underlying risk factors, intermediate risk factors, and direct causes of anaemia. The specific drivers and corresponding interventions selected for inclusion were identified as priority or emerging approaches in the field, based on recent comprehensive analyses (Hess et al., 2023; Moorthy et al., 2020; Lopez de Romana et al., unpublished; Mildon et al., unpublished). Because these are likely to be of the most interest to our audience, they are the focus of this map. Some of the interventions such as schooling and anti-poverty programs can be very generic in definition and may cover many other social aspects. To ensure that we only include studies that fit the goal of the EGM, which is to map evidence on interventions for reducing anaemia, all studies must measure at least one of the outcomes listed in section 4.3.3.

In some cases, interventions may be considered to act on multiple risk factors and function at multiple levels. For example, interventions related to antenatal, postnatal, and preventative care could be considered as addressing *inadequate access / use of health / nutrition services and interventions* and *inadequate maternal and child care*. The decision of where to classify these interventions is somewhat arbitrary but will not affect which studies are ultimately included in the map. The challenge of grouping specific interventions was most severe with regards to *food security* and *dietary enhancement and diversification.* The majority of interventions listed could be conceptualized as relating to these topics. We severely restrict these definitions with the understanding that there are other interventions which could be included but are reflected elsewhere in the map (Table 2).

Table 2. Interventions and definitions

Intervention group	Sub-groups	Intervention	Definition
Underlying risk factors	Low educational attainment	Schooling	General education or schooling. *Excludes sensitizations and topic-specific education*
	Health policies	Multi-sectoral anaemia platform strengthening	Country-driven efforts to sustain multi-sectoral anaemia platforms, develop anaemia related guidelines, and build capacity
		Advocacy for better services related to anaemia	Lobbying government and other stakeholders for improving services and policies relating to anaemia treatment and prevention
		Product registration and standardization	Registration and standardization of anaemia-products,[1] including policies regarding registration and standardization of anaemia-products and interventions to support the implementation of these policies.
		Governmental funding for anaemia treatment and prevention	Budget allocations to provide anaemia treatment and prevention services. This includes any type of public funding but excludes provision of materials such as supplements.
	Cultural norms and behaviours	Women's empowerment	Initiatives to empower women for decision-making (from Exemplars: "Education and empowerment programs for girls and women can help increase health

Intervention group	Sub-groups	Intervention	Definition
			and nutrition literacy; improve gender equality in marriages; encourage equitable access to food and health services; and shift cultural norms away from early pregnancies and short birth spacing")
		Social and behaviour change communication on gender norms	SBCC focusing on gender norms seeks to address the power dynamics and structures that serve to reinforce gendered inequalities through 1) mass media engagement, 2) advocacy towards legislation and policy 3) working with men and boys 4) community sensitization. *Excludes SBCC related to women's empowerment.*
	Poverty[2]	Social assistance	Non-contributory transfers
		Social insurance	"Contributory schemes providing compensatory support in the event of contingencies such as illness, injury, disability, death of a spouse or parent, maternity/paternity, unemployment, old age, and shocks affecting livestock/crops" (GSDRC, 2019).
		Social care services	"Targeted services provided by the state or non-state actors 'for those facing social risks such as violence, abuse, exploitation, discrimination and social exclusion" (ibid.)
		Labour market programs	Active labour market policies and interventions aim to help the unemployed and the most vulnerable find jobs.
Intermediate risk factors	Food insecurity[3]	Nutrition sensitive agriculture	Homestead food production systems, livestock transfer programs, value chains for nutritious foods, and irrigation programs
	Access / use of health / nutrition services and interventions	Antenatal and postnatal visits	The provision or encouragement of antenatal and postnatal visits at health facilities, such as primary care facilities.
		Resources and staff for health facilities	Provision of monetary and human resources for health facilities to deliver adequate care. This can include training, supervision, and other aspects of resourcing a competent workforce. *Excludes training explicitly related to postpartum haemorrhage.*
		Treatment of moderate or severe acute malnutrition.	Health facility based (in-patient or out-patient) treatment of moderate or severe acute malnutrition. *Excludes regular monitoring activities not linked to treatment programs.*
		Provision of other preventative care	Preventative care for well-being such as annual check-ups, adult immunization, other interventions that do not fall into the below categories. *Excludes preventative treatment of diseases such as malaria and helminths, antenatal care, and routine nutrition monitoring*
		Education and direct support for supply chain management	Provision of training and resources to support the management of stock related to anaemia-prevention products. Generally, this will target the public sector, but supporting private sector supply chains would be included.
			Provision, production, or transport of high-quality anaemia-prevention products to distribution centres (including health clinics). Includes product inspection. Generally, this will target the public sector, but supporting private sector supply chains would be included.

Intervention group	Sub-groups	Intervention	Definition
		Mobilization of public and private sector actors to support anaemia-product value chains	Working with public and private sector actors (government, trade bodies) to facilitate the sale, transport, and production of anaemia-products.[1]
	Inadequate family planning	Family planning and birth spacing counselling	Provision of information on contraceptives and referral to family planning services for unplanned pregnancies, birth spacing and allow individuals to decide if and when to have children.
	Inadequate health / nutrition knowledge and awareness	Breastfeeding and IYCF education and support	Sensitizations, counselling and outreach efforts regarding breastfeeding and complementary feeding practices
		Anaemia education and habit support	Sensitizations and outreach efforts regarding anaemia specifically, including identification of symptoms of anaemia and the use of specific products to reduce anaemia. Includes efforts to increase demand for anaemia and related interventions. Education is likely to be a dominant approach for demand generation.
		Other nutrition education and counselling	Nutrition education and counselling related to consumption of a balanced diet. Only studies which do not fall under dietary enhancement and diversification would be coded here.
		Counselling and management of anaemia due to genetic blood disorders	Targeted provision of care for those with genetic blood disorders known to affect anaemia risk. Genetic blood disorders include sickle cell, haemophilia, G6PD and blood cancers (ex. leukaemia)
	Inadequate access / use of WASH	Water access, sanitation, and hygiene resources	The provision of WASH resources, such as improved water sources, handwashing facilities, and latrines. Other resources would also be included.
		Education on hygiene	Education on hygiene, including food hygiene and handwashing. *Excludes activities related to the prevention of specific communicable diseases below.*
Direct causes	Chronic disease / exposure and response to infectious diseases[5]	Anti-malaria programs	Limited to population level treatment and prevention programs. *Excludes targeted treatment of people with confirmed malaria*
		Routine immunization	Any effort to provide and support the adoption of routine immunization
		Deworming and helminth programs	Limited to population level treatment and prevention programs. *Excludes targeted treatment of people with confirmed worms or helminth*
		Other anti-parasite programs	Limited to population level treatment and prevention programs. *Excludes targeted treatment of people with confirmed parasites*
		HIV programs	Limited to population level treatment and prevention programs. *Excludes targeted treatment of people with confirmed HIV*
		Tuberculosis programs	Limited to population level treatment and prevention programs. *Excludes targeted treatment of people with confirmed tuburculosis*

Intervention group	Sub-groups	Intervention	Definition
	Gynaecological and obstetric conditions	Delayed cord clamping	Training on or the implementation of delayed cord clamping, the practice of waiting at least one minute, and up to three minutes, after delivery to clamp and cut the umbilical cord.
		Management of menses	The use of contraception specific medical procedures to manage menorrhagia (heavy menses). Includes training for these procedures and the provision of resources needed to implement them.
		Management of postpartum haemorrhage	The use of specific medical procedures to manage postpartum haemorrhage. Includes training for these procedures and the provision of resources needed to implement them.
	Inadequate nutrient intake, absorption, and utilization	Supplementation	Provision or use of single and multiple micronutrient supplements. This can be preventative or therapeutic. To distinguish from fortification or direct provision of foods, supplements are defined as calorie-free products which are consumed independently from calorie containing goods (generally in pill form). Multiple micronutrient powders which are mixed into food before being consumed are considered a form of fortification.
		Mass fortification[4]	Broad provision or use of products fortified before they reach the end consumer targeted at the entire population. This would include national fortification programs or other large-scale programs that do not target specific populations.
		Point-of-use fortification[4]	Point-of-use fortification of foods with multiple micronutrient powders, a mixture of vitamins and minerals in powder form that are supplied as small, single-serving packets, the contents of which can be mixed into semi-solid food before consumption. These are often targeted to specific populations but are different from targeted fortification interventions in that whole foods are not provided.
		Targeted fortification[4]	Targeted provision or use of fortified food for specific populations, often infants, young children, and women of reproductive age. Includes the provision of ready to use therapeutic food (RUTF), lipid based nutrient supplement (LNS) and small quantity LNS (SQ-LNS,)
		Biofortification[4]	Education and direct support for process of increasing the density of vitamins and minerals in a crop, through plant breeding or agronomic practices
		Dietary enhancement and diversification	Support for the adoption of food choices, preparation, and processing strategies to enhance the content and / or bioavailability of micronutrients in a daily diet, provision of nutrient rich foods that are not included elsewhere. Includes germination, processing, food-to-food fortification

Notes:

1. We use the term “anaemia-products” broadly to refer to any product aimed to reduce anaemia or iron deficiency. Largely, these products are expected to be fortified foods and supplements. However, iron cooking implements could also be included. We do not impose a restriction on other types of products that would be considered. Mosquito nets could be conceptualized as both a “anaemia-product” and a component of an “anti-malaria program.” In order to keep intervention categories exclusive, we consider them as part of anti-malaria programs.

2. Poverty alleviation interventions considered are limited to social protection programs as defined by Carter and colleagues (2019).
3. The potential list of interventions to address food insecurity is quite long and includes many of those reflected elsewhere in this framework as well as those that were excluded to limit scope. We are limiting intervention to address food insecurity to nutrition sensitive agriculture approaches (Ruel et al., 2018). However, we classify biofortification as an approach to address inadequate nutrient intake, absorption, and utilization. The biological mechanisms linking general agricultural interventions to anaemia and those linking biofortification to anaemia are conceptually distinct. The first assumes a more general wellbeing and health mechansims while the second assumes a more targeted theory of change.
4. Fortification is the practice of deliberately increasing the content of an essential micronutrient, i.e., vitamins and minerals (including trace elements) in a food, to improve the nutritional quality of the food supply and provide a public health benefit with minimal risk to health. We will include fortification approaches which increase total caloric content, such as food-to-food fortification and the provision of LNS here.

We will exclude studies that measure the clinical efficacy of disease treatment approaches on populations with that disease (e.g. comparisons of the effectiveness of malaria treatments among people with malaria). However, we include prophylactic treatments, intermittent preventative treatments, and similar population-based approaches to disease management. In addition, we include the treatment of a diseased population for another disease (e.g. Providing people with HIV routine anti-parasite treatments). These are considered population-based approaches even though the population of interest is restricted.

For studies in which multiple interventions are bundled but not all of them are listed in our intervention framework, to be included, the study must report separate effects for at least one of the interventions that are relevant for this EGM.

Multicomponent interventions

For our purposes, multicomponent interventions have several activities (components) which would fall under different intervention groups within our EGM but the effects of those componets were estimated in a combined way as opposed to seperately. For example, an intervention could distribute supplements while providing antenatal care visits. If the evaluation only quantifies the joint effect of providing the intervention components in combination, we consider the package of intervention components as a multicomponent intervention. If the effects of the two or more components are calculated separately, we consider those to be separate interventions and the study will be coded for each of those interventions.

Multicomponent studies will be categorized based on the intervention components that have been bundled together. Bundled interventions that are evaluated five or more times will be added to the map as new intervention categories. We may also need to create a “mixed” multicomponent bucket for all other combinations where there is no obvious pattern of specific components. This

process will help us ensure that the map avoids artificially inflating the number of included studies and prevents double counting of the same study.

This coding adheres to common principles across all 3ie EGMs:

1. All coding involves categorising studies into ideal types, so some simplification is necessary when describing studies in an EGM.
2. Coding of interventions for the purposes of displaying studies in a typical EGM matrix should aim to describe the evaluative evidence (what the study is testing), rather than intervention components.
3. EGMs may have a secondary objective of describing program components based on the interventions included in EGMs, but the analysis should be clearly labelled as such.

Some of the common cobinations of interventions that we expect to find in multicomponent bundles may include:

- Antenatal care and supplements, deworming, or malaria prevention.
- Schooling and interventions to address inadequate WASH.
- Labour market programs and women's empowerment.
- Nutrition counselling and supplements.
- Breastfeeding and IYCF support and supplementation or fortification.

4.3.3 Outcomes

Outcomes of interest will focus on anaemia as well as other indicators measuring changes to the direct causes of anaemia: inadequate nutrition absorption and utilization; chronic disease / exposure and response to infectious disease; and gynaecological and obstetric conditions (Table 3). Our primary outcome of interest will be anaemia as measured by the WHO cut-points and continuous measures of haemoglobin or non-traditional cut points. Intermediate outcome measures (i.e. nutrient absorption and utilization) will be limited to studies that measure the concentration of key nutrients within the body.Measures of intake will not be included since intake is often disconnected from biologically available nutrients and biologic availability is the key driver of anaemia. Chronic and infectious disease related outcomes will focus on frequency or severity of key diseases which are linked to anaemia. We focus on diseases that lead to blood loss. Other diseases, especially those that can induce anaemia through inflammation, are represented in the filters.

Table 3. Outcomes and definitions

Outcome group	Indicator	Inclusions	Exclusions
Primary outcomes	Anaemia	Categorical measures of haemoglobin status (Hb) that have had traditional WHO cut-off points applied to them to define subjects as healthy, moderately, and / or severely anaemic	Studies that apply non-traditional cut-points.
	Haemoglobin	Continuous measures of haemoglobin levels or those with non-traditional cut-points applied (ex. using quantiles of the observed population)	None
Inadequate nutrient absorption and utilization	Iron	Biomarkers of iron that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include serum ferritin (adjusted and unadjusted for inflammation), total body iron (TBI), and transferrin receptors.	Measures that consider only the consumption of iron without reflecting a biological measure.
	Folate	Biomarkers of folate status that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include serum folate and red blood cell (RBC) folate.	Measures that consider only the consumption of folate without reflecting a biological measure.
	Vitamin B12	Biomarkers of vitamin B12 status that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include serum/plasma vitamin B12 concentration, serum holo transcobalamin concentration, serum methylmalonic concentration, and plasma total homocysteine.	Measures that consider only the consumption of B12 without reflecting a biological measure.
	Vitamin A	Biomarkers of vitamin A that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include serum retinol, serum retinol binding protein, relative dose response, modified relative dose response, retinol isotope dilution, dark adaptation.	Measures that consider only the consumption of vitamin A without reflecting a biological measure.
	Vitamin B6	Biomarkers of vitamin B6 that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include plasma pyridoxal 5'-posphate (PLP).	Measures that consider only the consumption of B6 without reflecting a biological measure.
	Vitamin C	Biomarkers of vitamin C that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indictors include plasma ascorbic acid.	Measures that consider only the consumption of vitamin C without reflecting a biological measure.
	Vitamin D	Biomarkers of vitamin D that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include 25-hydroxyvitamin D.	Measures that consider only the consumption of vitamin D without reflecting a biological measure.
	Vitamin E	Biomarkers of vitamin E that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include plasma alpha-tocopherol.	Measures that consider only the consumption of vitamin E without reflecting a biological measure.
	Copper	Biomarkers of copper that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include serum copper,	Measures that consider only the consumption of copper without reflecting a biological measure.

Outcome group	Indicator	Inclusions	Exclusions
		erythrocyte superoxide dismutase, and/or ceruloplasmin.	
	Zinc	Biomarkers of zinc that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include serum/plasma zinc.	Measures that consider only the consumption of zinc without reflecting a biological measure.
	Selenium	Biomarkers of selenium that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include plasma or serum selenium, glutathione Peroxidase 3 activity and/or seleno-protein P.	Measures that consider only the consumption of selenium without reflecting a biological measure.
	Riboflavin	Biomarkers of riboflavin that can be used to assess status measured as a continuous variable or deficiency measured as a categorical variable. Potential indicators include erythrocyte glutathione reductase activity and urinary riboflavin excretion.	Measures that consider only the consumption of riboflavin without reflecting a biological measure.
Chronic disease / exposure and response to infectious disease	Soil-transmitted helminths	Measures of the frequency or severity of infection with soil-transmitted helminths. This can be referred to as round worm (ascaris lumbricoides), whipworm (trichuris trichiura), and hookworm (necator americanus and ancylostoma duodenale).	Measures of treatment or prevention activities will not be included.
	Schistosomiasis	Measures of the frequency or severity of schistosomiasis. This is the disease caused by infection with schistosoma mansoni, S. haematobium, or S. japonicum. Infection with these parasites will also be included.	Measures of treatment or prevention activities will not be included.
	Malaria	Measures of the frequency or severity of malaria infection. This is the disease caused by infection with plasmodium falciparum or plasmodium vivax.	Measures of treatment or prevention activities will not be included.
	Gastro-intestinal disease	Measures of the frequency or severity of chronic gastrointestinal disease other than those outlined above. This could include H. pylori infection, environmental enteric enteropathy, and ulcers.	Measures of treatment or prevention activities will not be included. Acute gastrointestinal disease will not be included.
	Kidney disease	Measures of the frequency or severity of kidney disease. This could include measures of serum creatine or blood urea nitrogen. However, any study which states that it measures kidney disease will be included.	Measures of treatment or prevention activities will not be included.
Gynaecologic and obstetric conditions	Postpartum haemorrhage	Measures of postpartum blood loss, including mortality due to postpartum haemorrhage	None
	Delayed cord clamping	Measures of the frequency of delayed cord clamping or the time period between birth and cord clamping. While this could be conceptualized as an intervention, the behaviour of delayed cord clamping can be conceptualized as an outcome which is theoretically affected by interventions related to access to health services and interventions.	None

4.3.4 Filters

In its default view, the online map will display all of the included studies. However, the map will have options to filter studies based on certain criteria, allowing users to view a sub-set of the evidence-base. Table 4 shows the filters proposed for this map.

Table 4. Filters to allow for the selection of a sub-set of studies from the overall body of evidence of the map.

Filter	Options for dropdown	Explanation
Region	○ African region ○ Region of the Americas ○ South-east Asia region ○ European region ○ Eastern Mediterranean region ○ Western Pacific region	Menu will allow for the identification of studies which occurred in specific regions based on the WHO classification
Age	○ Infant (0-6 mo) ○ Older infant (7-23) ○ Young child (24-59 mo) ○ Child (60 mo-12 years) ○ Adolescent (13 -19 y) ○ Adult (>19 y) ○ All/not specified	Menu will allow for the identification of studies targeting specific age groups. Age groups are defined based on their nutritional needs. Studies that consider multiple or overlapping age groups will be double-tagged (ex. 12 months to 10 years would be "young child" and "child")
Place of residence	○ Urban ○ Rural ○ Peri-urban ○ All/not specified	Menu will allow for the identification of studies which occur in these settings.
Disease targeted[1]	○ Malaria ○ TB ○ HIV ○ Helminth / worms ○ CRP / AGP ○ Obesity (includes weight) ○ Other parasites ○ Does not target a disease	Menu will allow for the identification of studies which consider variation in impacts on other outcomes based on these moderators. Studies which consider these impacts on disease prevalence will be reflected in the corresponding secondary outcomes columns.

Filter	Options for dropdown	Explanation
Women's empowerment	○ Includes WE aspect ○ Does not include WE aspect	Menu will allow for the identification of studies that explicitly mention a women's empowerment component. Studies which could be conceptualized to relate to women's empowerment without making the explicit connection will be considered not to include a women's empowerment aspect.
Gender and reproductive status	○ Entire population or not specified ○ Men ○ Women (general) ○ Pregnant women ○ Lactating women ○ Non-pregnant/ lactating women ○ Women of reproductive age ○ Adolescent girls	Menu will allow for the identification of studies that target specific populations based on gender and reproductive status. Studies which use gender / reproductive status as an inclusion criteria, moderator, or subgroup for analysis will be reflected here.
Genetic blood disorders	○ β-thalassemia ○ α-thalassemia ○ Sickle cell disorders ○ Glucose-6 phosphate dehydrogenase (G6PD) deficiency	Menu will allow for the identification of studies that target populations with genetic blood disorders or consider genetic blood disorders as moderators of impact. Note that we do not expect any study to affect the prevalence of these disorders, so they will not be considered as outcomes.
Evaluation method	○ Randomized controlled trial ○ Natural experiment ○ Fixed effects (including difference-in-difference) ○ Instumental variable estimations ○ Interrupted time series analysis ○ Regression discontinuity design ○ Statistical matching ○ Narrative/thematic synthesis	Menu will allow for the identification of studies using these evaluation methods
Directly measures haemoglobin	○ Yes ○ No	Menu will allow for the identification of studies that directly measure haemoglobin. While all studies under the anaemia and haemoglobin columns will measure these outcomes, many studies which fall under other outcomes will likely also measure haemoglobin. This would allow for the identification of studies which measure haemoglobin AND the outcome under the

Filter	Options for dropdown	Explanation
		corresponding column to support an understanding of the causal chain.
Cost evidence	○ Yes ○ No	Menu will allow for the identification of studies presenting cost evidence.

Note:

1. Disease will be considered both an outcome and a filter. The outcome coding will be used when the effect of the intervention on disease frequency or severity (ex. The effect of a point-of-use fortification intervention on malaria prevalence) is measured. The filter will be used when variation in the impact of an intervention on another outcome is evaluated (ex. The variation in effect of a fortification intervention on anaemia under different levels of malaria prevalence within a population).

4.3.5 Study designs

In this EGM, we will include both impact evaluations and systematic reviews of the effects of the included interventions. To be included in the map, it is not required for systematic revivews to conduct meta-analysis as such analysis is often not appropriate with highly heterogeneous interventions. We will only include studies if they evaluate the effects of an intervention (Appendix C). We define the requirements for the study design criteria, drawing on commonly accepted standards for impact evaluations (Gertler et al., 2016) and systematic reviews (Waddington et al., 2012).

Impact evaluations

We will include studies using experimental and quasi-experimental study designs to measure a change in eligible outcomes that is attributable to an intervention. A wide range of potential study designs is eligible, including randomized controlled trials, regression discontinuity designs, instrumental variables, fixed-effect regressions, interrupted time series models, matching methods, and the synthetic control method (Appendix C).

We will not exclude studies based on the comparison condition of a control group. A study's control group may consist of participants subject to no intervention, on a wait-list, or in an alternative intervention. However, we will exclude studies that only use simulation or forecast models, ex-ante impact assessments or scenario analyses. Observational studies, evaluations and case studies that do not satisfy the methodological conditions described above, such as before-after studies without a comparison group or cross-sectional studies using designs that do not appropriately address the issues of selection bias or confounding. Finally, we will also exclude studies that only employ qualitative designs, feasibility studies, acceptability studies, and studies

that examine willingness-to-pay for goods, services, and process and business models. We acknowledge that the study types excluded from this map may contain valuable information; however, the focus of this EGM is to map existing rigorous evidence of intervention effectiveness.

Systematic reviews

A systematic review is a synthesis of research evidence on a particular topic obtained through an exhaustive literature search for all relevant studies using widely accepted scientific strategies to minimize errors associated with appraising the design and results of studies. We will include systematic reviews of the effects of interventions if they describe the search, data collection and synthesis methods according to the 3ie database of systematic reviews protocols (Snilstveit et al., 2016). Evidence reviews that do not adopt these methods will be excluded.

Where reviews include a mixture of evidence from both HICs and L&MICs, we will include them if they present disaggregated evidence for L&MICs or if more than 50 percent of the evidence of non-disaggregated results is from L&MICs. When there are no disaggregated results for L&MICs and more than 50 percent of the evidence in the systematic review comes from HICs or it is impossible to ascertain the composition of evidence by income level, studies will be excluded. Reviews that allow for study designs or methods that are not included in this map will be included if at least one included study design is eligible for inclusion in this map.

4.3.6 Other inclusion and exclusion criteria

Language

Studies published in any language will be included, although the search terms used will be in English only.

Publication date

Studies will be included if they were published in 2000 or after. From our experience from developing other EGMs, a very small proportion (less than one percent) of impact evaluations and systematic reviews on interventions implemented L&MICs predate the year 2000. In view of this, considering the small likelihood of missing eligible studies, we will search for and screen studies published after 2000. This will limit the overall breadth of the evidence mapping project and ensures the exercise remains manageable and within our available resource constraints.

Status of studies

We will include ongoing and completed impact evaluations and systematic reviews, both peer-reviewed studies and 'grey' literature. For ongoing studies, we will include prospective study records, protocols, and trial registrations. Providing an indication of the prevalence and characteristics of ongoing evaluations will enrich the analysis of current evidence gaps and support decision-making regarding evidence generation.

4.4 Screening approach

We will document each step in the screening process in detail and graphically present the process in a flow chart in the final report to facilitate replication of the findings. We will manage the selection of studies for data extraction as part of the map using EPPI-Reviewer 4 software (EPPI; Thomas et al., 2010) by implementing the following steps:

4.4.1 Import study records and remove duplicates

We will import all output files (e.g. RIS or .txt files) of the search strategy into EPPI. An automated process within EPPI will be used to remove duplicate files.

4.4.2 Training of screeners

The core project team will train a team of consultants on the protocol. Training will focus on understanding the subject matter and the screening process. Initially, all screeners will screen the same set of studies and be evaluated for consistency with core team decisions. Consultants will achieve an eighty-five percent level of consistency before proceeding to title and abstract screening.

4.4.3 Title and abstract screening

One screener will screen the title and abstract of all imported and de-duplicated studies and will give a judgment of include, exclude, or unsure. Items marked unsure will be screened by a second screener (an approach that has been demonstrated to produce comparable results to double screening at significantly lower cost; Shemilt et al. 2016). Several exclude codes will be available to provide more information on the reasons for exclusion in each case. We will apply screening codes in a hierarchical order so that consistent comparisons can be made about why studies were excluded and at what stage in the screening process. The core team will hold periodic meetings to address studies flagged for a second opinion and make any refinements to the screening

approach. The output of this process will be a set of screened studies that have been put forward for full text screening.

We will use the machine-learning features of EPPI to accelerate the title and abstract screening process. We will begin by screening 100 random abstracts, which will serve as a training set for the construction of a classifier that assigns all remaining abstracts a probability of inclusion based on the training data. We will screen all abstracts with a probability score of 0.3 or greater. We will then screen a random sample of abstracts with lower probability scores to determine if any should be included for full-text screening. If more than 1 percent of this sample is found to be includable in the EGM, we will proceed to screen additional abstracts until this threshold is met.

4.4.5 Full-text screening

We will attempt to retrieve the full text for each study that meets al ofl the title and abstract inclusion criteria. Two reviewers will independently examine each full text in detail against the protocol. Again, we will apply a code to each study that reflects either that the study is included, or why the study is excluded. The output of this stage will be a set of studies deemed suitable to include in the EGM.

4.4.6 Checks for linked publications

The project team will attempt to group publications that focus on the same study (i.e., publications that report on the same intervention and the same study population). This typically occurs in cases where an author group publishes more than one paper in relation to one particular study on a specific population. For each group we will identify one main paper. Descriptive information will only be extracted from the main paper; nevertheless, all linked papers will be checked and any additional information, especially additional outcome measures, will be added to the dataset. This ensures that the extraction is as comprehensive as possible as well as prevents the evidence-base from being artificially inflated. The identification of the main paper – the study that will appear in the map – will be consistent with the approach used by 3ie's Development Evidence Portal (DEP) team. Priority will be given to papers which already exist (and have their data extracted) in the DEP central database. If a potential main paper does not exist on the DEP, priority will be given to the most recent paper.

4.5 Data extraction and critical appraisal

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

> **Basic study and publication information:** This coding will focus on capturing the general characteristics of the study including authors, publication date and status, study location, intervention type, outcomes reported, definition of outcome measures, population of interest, and study and program funders. Effect sizes for evidence synthesis will not be extracted.
>
> **Critical appraisal:** We will critically appraise all included systematic reviews following the practices adopted by 3ie's systematic review database protocol, which draws on work by Lewin and colleagues (2009). This appraisal assesses systematic reviews according to criteria relating to the search, screening, data extraction, and synthesis activities conducted, and covers all the most common areas where biases are introduced. Each systematic review will be rated as low, medium, or high confidence drawing on guidance provided in Snilstveit and colleagues (2017). The tool used for this process is presented in Appendix E. We will not critically appraise impact evaluations, as this is typically beyond the scope of EGMs.

The following processes will be implemented to collect this information:

4.5.1 Develop and refine data extraction tools and codebooks

The draft tools developed for this project will be reviewed and potentially refined in light of any feedback received by the EGM advisory group and insights from project implementation.

4.5.2 Data extraction training and pilot

Coders assigned to each data extraction task will undergo theory- and practice- based training in using the tools provided. Each consultant will code a 'training set' of studies and assessments of inter-rater reliability will be calculated. Coders will start with main-stage extraction once they achieve 85 percent reliability score; otherwise additional training will be provided. .

4.5.3 Main-stage extraction

Two independent coders will extract basic study and publication information. Critical appraisal assessments of systematic reviews will first be single coded and then reviewed by a systematic

review methods expert. Meetings will be held periodically with coders on the project to provide support and resolve queries.

4.5.4 Quality checks

Throughout the data extraction process, the project team will check the extracted data. A core team member will check the consistency of data extracted in duplicate by consultants. We will calculate measures of consistency and use them to inform the checking process. If additional review is warranted, targeted reviews will be conducted. This quality check process is put in place to ensure that the extracted data are accurate and does not assess the quality of the study itself or the evidence presented in the study. Please see section 4.5 for further details on assessment of evidence quality.

4.6 Analysis and reporting

We will conduct a range of descriptive analyses to provide an overview of included studies across the following dimensions:

- Publication year
- Geography
- Interventions
- Outcomes
- Targeted populations
- Methods
- Results of the systematic review critical appraisal

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

Interactive EGM: An interactive evidence gap map that visually presents the current evidence base that is categorized by coverage with respect to the pre-determined intervention-outcome framework, quality and completeness. Filters may be incorporated into the map to enable a more targeted use – for example, by restricting the studies to a specific population. The map will be stored on the 3ie website and shared as a public good.

EGM technical report: The EGM technical report will include a detailed overview of the methods, theory of change and the key results of the EGM. It will provide a high level of analytical detail

and will be supported by technical annexes. This report will conclude by directly addressing the key research questions stated in Section 2 and provide a set of research and policy implications. The technical report will be published by 3ie and shared as a public good.

EGM executive summary: This report will provide a high-level summary of the results and will primarily focus on answering the research questions specified in Section 2 using non-technical language.

5. Engagement and communication plan

Evidence gap maps highlight both absolute gaps, which should be filled with new primary studies, and synthesis gaps, which are ready for new systematic reviews and meta-analyses. They are envisioned as a global public good, and this allows them to be used as a tool which facilitates access to high-quality research. Sharing broadly with external stakeholders can help to validate conclusions and make the findings more likely to be adopted by the development sector broadly.

Therefore, the project team will engage actively with an external advisory group to ensure the utility and dissemination of this work. The project team, in collaboration with Nutrition International, will engage with key stakeholders by seeking academic and practitioner expertise in the field. 3ie will set up an advisory group which will have the aim of providing pro-bono support to the project at several key stages of the project. These stages include developing the project protocol, reviewing the search results produced, reviewing and interpreting emerging findings, and developing and optimizing the analytical outputs produced to aid evidence uptake and use.

Sources of support

This EGM was commissioned with the generous support of Nutrition International.

References

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

Apte, A., Lubree, H., Kapoor, M., Juvekar, S., Banerjee, R., & Bavdekar, A. (2021). Development and implementation of liposomal encapsulated micronutrient fortified body oil intervention for infant massage: an innovative concept to prevent micronutrient deficiencies in children. *Frontiers in public health*, *8*, 567689.

Balarajan, Y., Ramakrishnan, U., Özaltin, E., Shankar, A. H., & Subramanian, S. (2011). Anaemia in low-income and middle-income countries. The Lancet, 378(9809), 2123–2135. doi:10.1016/s0140-6736(10)62304-5

Bhutta, Z. A. (2016). How will the next 'Decade of Nutrition' be different from the past one? *Nature reviews: Gastroenterology & hepatology*, *13*(8), 441-442.

Carter, B., Roelen, K., Enfield, S., & Avis, W. (2019). Social protection topic guide

Chaparro, C. M., & Suchdev, P. S. (2019). Anaemia epidemiology, pathophysiology, and aetiology in low-and middle-income countries. Annals of the New York Academy of Sciences, 1450(1), 15-31.

Coile, A., Wun, J., Kothari, M. T., Hemminger, C., Fracassi, P., & Di Dio, D. (2021). Scaling up nutrition through multisectoral planning: An exploratory review of 26 national nutrition plans. *Maternal & Child Nutrition*, *17*(4), e13225.

da Silva Lopes, K., Yamaji, N., Rahman, M. O., Suto, M., Takemoto, Y., Garcia-Casal, M. N., & Ota, E. (2021). Nutrition-specific interventions for preventing and controlling anaemia throughout the life cycle: an overview of systematic reviews. *Cochrane Database of Systematic Reviews*, (9).

Engle-Stone, R., Aaron, G. J., Huang, J., Wirth, J. P., Namaste, S. M., Williams, A. M., ... & Suchdev, P. S. (2017). Predictors of anemia in preschool children: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *The American journal of clinical nutrition*, *106*(suppl_1), 402S-415S.

Exemplars in Global Health (2023). *What is Anemia?.* Accessed 20 January 2023. Available at: https://www.exemplars.health/topics/anemia-among-wra/what-is-anemia

Fanzo, J., Hawkes, C., Udomkesmalee, E., Afshin, A., Allemandi, L., Assery, O., ... & Schofield, D. (2019). 2018 Global Nutrition Report.

Franklin, N. (2021). *Opinion: Revealing the costs of 'hidden hunger' and combating anaemia*. Devex. Accessed 7 June 2022. Available at:https://www.devex.com/news/opinion-revealing-the-costs-of-hidden-hunger-and-combating-anemia-100049.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2016). Impact evaluation in practice. World Bank Publications.

Headey, D., Heidkamp, R., Osendarp, S., Ruel, M., Scott, N., Black, R., ... & Walker, N. (2020). Impacts of COVID-19 on childhood malnutrition and nutrition-related mortality. *The Lancet*, *396*(10250), 519-521

Heidkamp, R. A., Piwoz, E., Gillespie, S., Keats, E. C., D'Alimonte, M. R., Menon, P., ... & Bhutta, Z. A. (2021). Mobilising evidence, data, and resources to achieve global maternal and child undernutrition targets and the Sustainable Development Goals: an agenda for action. *The Lancet*, *397*(10282), 1400-1418

Hess, S., Tecklenburg, L., & Eichler, K. (2016). Micronutrient fortified condiments and noodles to reduce anemia in children and adults—A literature review and meta-analysis. *Nutrients*, *8*(2), 88.

Hess, S. Y., Owais, A., Jefferds, M. E. D., Young, M. F., Cahill, A., & Rogers, L. M. (2023). Accelerating action to reduce anemia: Review of causes and risk factors and related data needs. *Ann NY Acad Sci., 1523*, 11– 23. https://doi.org/10.1111/nyas.14985

Institute for Health Metrics and Evaluation (IHME) (2019) Global Burden of Disease (GBD). Institute for Health Metrics and Evaluation. http://www.healthdata.org/gbd/2019

International Food Policy Research Institute. (2016). Global nutrition report 2016: from promise to impact: ending malnutrition by 2030.

Kassebaum, N. J., Jasrasaria, R., Naghavi, M., Wulf, S. K., Johns, N., Lozano, R., ... & Murray, C. J. (2014). A systematic analysis of global anemia burden from 1990 to 2010. *Blood, The Journal of the American Society of Hematology*, *123*(5), 615-624.

Keats, E. C., Das, J. K., Salam, R. A., Lassi, Z. S., Imdad, A., Black, R. E., & Bhutta, Z. A. (2021). Effective interventions to address maternal and child malnutrition: an update of the evidence. *The Lancet Child & Adolescent Health*, *5*(5), 367-384.

Keats, E. C., Neufeld, L. M., Garrett, G. S., Mbuya, M. N., & Bhutta, Z. A. (2019). Improved micronutrient status and health outcomes in low-and middle-income countries following large-scale fortification: evidence from a systematic review and meta-analysis. *The American journal of clinical nutrition*, *109*(6), 1696-1708.

Kinyoki, D., Osgood-Zimmerman, A. E., Bhattacharjee, N. V., Kassebaum, N. J., & Hay, S. I. (2021). Anemia prevalence in women of reproductive age in low-and middle-income countries between 2000 and 2018. *Nature medicine*, *27*(10), 1761-1782.

Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A. K., Hammerstrøm, K., & Sathe, N. (2017). Searching for studies: A guide to information retrieval for Campbell systematic reviews. Campbell Systematic Reviews, 13(1), 1–73. https://doi.org/10.4073/cmg.2016.1

Laborde, D., Smaller, C., & Porciello, J. (2020). *What would it cost to avert the COVID-19 hunger crisis?*.

Lakner, C., Yonzan, N., Gerzson Mahler, D., Castaneda Aguilar, R., & Wu, H. (2022). *Updated estimates of the impact of COVID-19 on global poverty: Looking back at 2020 and the outlook for 2021*. World Bank. Accessed 10 June 2022. Available at: https://blogs.worldbank.org/opendata/updated-estimates-impact-covid-19-global-poverty-looking-back-2020-and-outlook-2021.

Launching a data revolution for the Sustainable Development Goals. Available at: https://indicators.report/targets/2-2/

Lewin, S., Oxman, A. D., Lavis, J. N., & Fretheim, A. (2009). SUPPORT Tools for evidence-informed health Policymaking (STP) 8: Deciding how much confidence to place in a systematic review. Health Research Policy and Systems, 7(S1), S8. https://doi.org/10.1186/1478-4505-7-S1-S8

Loped de Romana, D., Golan, J., Mildon, Al, Jefferds, MEL., Rogers, LM., Arabi, M. Comprehensive framework for integrated action on the prevention and management of anaemia: review of preventive and therapeutic interventions to treat and prevent anemia. Not Published.

Mithra, P., Khatib, M. N., Sinha, A. P., Kumar, N., Holla, R., Unnikrishnan, B., ... & Quazi Zahiruddin, S. (2021). Interventions for addressing anemia among children and adolescents: an overview of systematic reviews. *Frontiers in pediatrics*, *8*, 549549.agendas. Journal of Clinical Epidemiology, 79, 120–129. https://doi.org/10.1016/j.jclinepi.2016.05.015

Mildon, A., Lopez de Romana, D., Jefferds, MED., Rogers, LM., Arabi, M. Integrating and coordinating programmes for the management of anaemia across the life course. Not Published.

Moore, N., Lane, C., Storhaug, I., Franich, A., Rolker, H., Furgeson, J., ... & Snilstveit, B. (2021). The effects of food systems interventions on food security and nutrition outcomes in low- and middle-income countries, International Initiative for Impact Evaluation.

Moorthy, D., Merrill, R., Namaste, S., & Iannotti, L. (2020). The Impact of Nutrition-Specific and Nutrition-Sensitive Interventions on Hemoglobin Concentrations and Anemia: A Meta-review of Systematic Reviews. *Advances in Nutrition*, *11*(6), 1631-1645.

Namaste, S. M., Aaron, G. J., Varadhan, R., Peerson, J. M., & Suchdev, P. S. (2017). Methodologic approach for the biomarkers reflecting inflammation and nutritional

determinants of anaemia (BRINDA) project. *The American journal of clinical nutrition*, *106*(suppl_1), 333S-347S.

Ngure, F. M., Reid, B. M., Humphrey, J. H., Mbuya, M. N., Pelto, G., & Stoltzfus, R. J. (2014). Water, sanitation, and hygiene (WASH), environmental enteropathy, nutrition, and early child development: making the links. Annals of the New York Academy of Sciences, 1308(1), 118-128.

Nguyen, P. H., Scott, S., Avula, R., Tran, L. M., & Menon, P. (2018). Trends and drivers of change in the prevalence of anaemia among 1 million women and children in India, 2006 to 2016. *BMJ global health*, *3*(5), e001010

Nutrition International (n.d.). Iron-folic Acid Supplementation in Pregnancy. Program Logic Model. Available at: https://www.nutritionintl.org/learning-resource/iron-folic-acid-supplementation-in-pregnancy-program-logic-model-and-key-considerations-for-program-success/

Osendarp, S., Akuoku, J. K., Black, R. E., Headey, D., Ruel, M., Scott, N., ... & Heidkamp, R. (2021). The COVID-19 crisis will exacerbate maternal and child undernutrition and child mortality in low-and middle-income countries. *Nature Food*, *2*(7), 476-484

Owais, A., Merritt, C., Lee, C., & Bhutta, Z. A. (2021). Anaemia among Women of Reproductive Age: An Overview of Global Burden, Trends, Determinants, and Drivers of Progress in Low-and Middle-Income Countries. *Nutrients*, *13*(8), 2745.

Panchal, P. D., Ravalia, A., Rana, R., Puthussery, S., Vaze, G., Mavlankar, D., & Menon, K. (2022). Impact of Nutrition Interventions for Reduction of Anemia in Women of Reproductive Age in Low-and Middle-Income Countries: A Meta-Review. *Current Developments in Nutrition*, *6*(12), nzac134.

Parker, J. A., Barroso, F., Stanworth, S. J., Spiby, H., Hopewell, S., Doree, C. J., ... & Allard, S. (2012). Gaps in the evidence for prevention and treatment of maternal anaemia: a review of systematic reviews. *BMC pregnancy and childbirth*, *12*(1), 1-10

Pasricha, S.R., Drakesmith, H., Black, J., Hipgrave, D., & Biggs, B.A. (2013).Control of iron deficiency anemia in low and middle income countries. *Blood*,*121*, 2607–2617.

Rahman, M. M., Abe, S. K., Rahman, M. S., Kanda, M., Narita, S., Bilano, V., ... & Shibuya, K. (2016). Maternal anaemia and risk of adverse birth and health outcomes in low-and middle-income countries: systematic review and meta-analysis, 2. *The American journal of clinical nutrition*, *103*(2), 495-504

Reeves, B. C., Wells, G. A., & Waddington, H. (2017). Quasi-experimental study designs series—paper 5: A checklist for classifying studies evaluating the effects on health interventions—a taxonomy without labels. Journal of Clinical Epidemiology, 89, 30–42. https://doi.org/10.1016/j.jclinepi.2017.02.016

Ruel, M. T., Quisumbing, A. R., & Balagamwala, M. (2018). Nutrition-sensitive agriculture: what have we learned so far?. *Global Food Security*, *17*, 128-153

Sarkar, D., Murphy, H., Fisseha, T., Koroma, A. S., Hodges, M. H., Adero, N., ... & Namaste, S. M. (2018). Understanding the process of strengthening multi-sectoral efforts for anemia reduction: Qualitative findings from Sierra Leone and Uganda. *The International journal of health planning and management*, *33*(4), 1024-1044.

Scott, S., Lahiri, A., Sethi, V., de Wagt, A., Menon, P., Yadav, K., ... & Nguyen, P. H. (2022). Anaemia in Indians aged 10–19 years: Prevalence, burden and associated factors at national and regional levels. Maternal & Child Nutrition, 18(4), e13391.

Sedlander, E., Talegawkar, S., Ganjoo, R., Ladwa, C., DiPietro, L., Aluc, A., & Rimal, R. N. (2021). How gender norms affect anemia in select villages in rural Odisha, India: A qualitative study. *Nutrition*, *86*, 111159.

Shekar, M., Kakietek, J., Eberwein, J. D., & Walters, D. (2017). *An investment framework for nutrition: reaching the global targets for stunting, anemia, breastfeeding, and wasting*. World Bank Publications.

Shet, A. S., Zwarenstein, M., Mascarenhas, M., Risbud, A., Atkins, S., Klar, N., & Galanti, M. R. (2015). The Karnataka Anemia Project 2—design and evaluation of a community-based parental intervention to improve childhood anemia cure rates: study protocol for a cluster randomized controlled trial. *Trials*, *16*(1), 1-10.

Snilstveit, B., Bhatia, R., Rankin, K., & Leach, B. (2017). 3ie evidence gap maps: A starting point for strategic evidence production and use (3ie Working Paper 28). International Initiative for Impact Evaluation (3ie). https://www.3ieimpact.org/sites/default/files/2019-01/wp28-egm.pdf

Snilstveit, B., Vojtkova, M., Bhavsar, A., Stevenson, J., & Gaarder, M. (2016). Evidence & Gap Maps: A tool for promoting evidence informed policy and strategic research

SPRING (). *Anemia Research Working Groups. Activities*. Accessed on: 17th of January, 2023. https://www.spring-nutrition.org/about-us/activities/anemia-research-working-groups

SPRING. 2018. *Strengthening Partnerships, Results, and Innovations in Nutrition Globally (SPRING) Project: Life of Project Report 2011-2018.* Arlington, VA: Strengthening Partnerships, Results, and Innovations in Nutrition Globally (SPRING) project.

Stevens, G. A., Paciorek, C. J., Flores-Urrutia, M. C., Borghi, E., Namaste, S., Wirth, J. P., ... & Rogers, L. M. (2022). National, regional, and global estimates of anaemia by severity in women and children for 2000–19: a pooled analysis of population-representative data. *The Lancet Global Health*, *10*(5), e627-e639.

Sustainable Development Solutions Network (2012). Indicators and a Monitoring Framework. Available at: https://indicators.report/

Thomas, J., Graziosi, S., Brunton, J., Ghouze, Z., O'Driscoll, P., & Bond, M. (2020). EPPI-Reviewer: Advanced software for systematic reviews, maps and evidence synthesis. UCL Social Research Institute.

Thomas, J., Graziosi, S., O'Driscoll, P., Bond, M., Brunton, J. (2010). "Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)". EPPI – Reviewer. Available at: https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2936

The United Nations Children's Fund and the World Food Programme (2021). Nutrition Crisis Looms as More than 39 billion In-School Meals Missed Since Start of Pandemic. UNICEF. URL: https://www.unicef.org/press-releases/nutrition-crisis-looms-more-39-billion-school-meals-missed-start-pandemic-unicef-and

U.S. Agency for International Development (2017) USAID Multi-Sectoral Anemia Task Force Meeting. USAID Preventing Anemia Events. Available at: https://www.spring-nutrition.org/events/usaid-multi-sectoral-anemia-task-force-meeting-0

U.S. Agency for International Development (2022). *USAID Advancing Nutrition Anemia Task Force.* Accessed by: 19th of January 2023. Available at: https://www.advancingnutrition.org/what-we-do/monitoring-evaluation-and-learning/anemia-task-force

Waddington, H., White, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Davies, P., ... & Tugwell, P. (2012). How to do a good systematic review of effects in international development: a tool kit. *Journal of development effectiveness*, *4*(3), 359-387.

Williams, P. A., Poehlman, J., Moran, K., Siddiqui, M., Kataria, I., Rego, A. M., ... & Saldanha, N. (2020). Strategies to address anaemia among pregnant and lactating women in India: a formative research study. *Public health nutrition*, *23*(5), 795-805

Wirth, J. P., Woodruff, B. A., Engle-Stone, R., Namaste, S. M., Temple, V. J., Petry, N., ... & Aaron, G. J. (2017). Predictors of anemia in women of reproductive age: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *The American journal of clinical nutrition*, *106*(suppl_1), 416S-427S.

World Health Organization. (2023). *Publications.* Available at: https://www.who.int/publications/i?healthtopics=ff708c2a-1e12-4ea7-ae5c-cd7090045c41&healthtopics-hidden=true

World Health Organization. (2014). *Global nutrition targets 2025: Stunting policy brief* (No. WHO/NMH/NHD/14.3). World Health Organization.

World Health Organization. (2017). Nutritional anaemias: tools for effective prevention and control.

World Health Organization. (2020). Global anaemia reduction efforts among women of reproductive age: impact, achievement of targets and the way forward for optimizing efforts.

World Health Organization. (2021). *Anaemia in women and children*. Accessed on: 10 June 2022. Available at: https://www.who.int/data/gho/data/themes/topics/anaemia_in_women_and_children.

World Health Organization. (2021). Monitoring flour fortification to maximize health benefits: a manual for millers, regulators, and programme managers.

Wüstefeld, M., Marzara, S., & Korenromp, E. (2015). Nutrition Targets and Indicators for the Post-2015 Sustainable Development Goals. *SCN News*, *41*, 37-43

Yilma, H., Sedlander, E., Rimal, R. N., Pant, I., Munjral, A., & Mohanty, S. (2020). The reduction in anemia through normative innovations (RANI) project: study protocol for a cluster randomized controlled trial in Odisha, India. *BMC Public Health*, *20*(1), 1-13.

6. Appendices

Appendix A: EGM advisory group

The advisory group for this work is composed of:

- Lisa Rogers, WHO
- Maria Elena Jefferds, CDC
- Melissa Young, Emory University
- Denish Moorthy, Tufts University
- Kathryn Dewey, UC Davis

Terms of reference for an EGM advisory group

EGM advisory groups are a requirement for all 3ie EGMs. They help authors determine the parameters of their proposed map and provide inputs throughout the research process to help ensure that the final product is policy relevant and useful in informing decision-making.

Members of the advisory group should be diverse including policymakers, program managers, researchers and other key stakeholders (e.g. the funder, if appropriate). Members will be asked to provide inputs on various aspects of the EGM throughout the mapping process.

The details of member inputs will be finalized by the project manager or principal investigator prior to member recruitment. The total time commitment is not likely to exceed two days and may be less depending on members' availability. Indicative inputs are listed here (the examples are not exhaustive):

- Advise on key decisions regarding the EGM scope, including refining the objectives and definitions of key concepts;
- Determine important outcomes;
- Suggest relevant background literature and studies for inclusion;
- Participate in up to 2-3 teleconferences for the duration of the EGM (title/scoping stage, draft protocol, draft report);
- Provide written comments on the draft protocol and draft report;
- Help the team draw policy implications from the EGM findings. This can involve participating in a brainstorming session or focus group meeting to review the lessons and implications of the EGM in terms of policy and research investments;

- Assist the study team in implementing the communication plan developed for the project. This can involve advising on the project team's plan, identifying key audiences or hosting launch events for the report;
- Identify opportunities for policy influence to increase investments in evidence production and synthesis; and
- Act as a knowledge broker, providing a link between the author team and the end users by facilitating access to, interpretation and translation of the EGM findings for use locally.

Appendix B: Study Participants

Only studies which consider populations in low- and middle-income countries (as defined using the World Bank Country and Lending Groups classification in first year of intervention or if not available then Publication year) will be included. The exception to this is if a country held high-income status for only one year before reverting to L&MIC status. These are included even if the intervention began in the high-income year. As of the writing of this protocol, this applies to Argentina (2014, 2017), Venezuela (2014), Mauritius (2019), and Romania (2019). If the study is conducted in a high-income country but measures impacts on people, firms, or institutions in an L&MIC, it can be included. For example, we do not exclude a study that measures impacts of New Zealand's immigration visa lottery on residents of Tonga.

World Bank Country and Lending Groups classification

LOW-INCOME ECONOMIES ($1,085 OR LESS)

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

LOWER-MIDDLE INCOME ECONOMIES ($1,086 TO $4,255)

Angola	India	Philippines
Algeria	Indonesia	Samoa
Bangladesh	Iran, Islamic Rep	São Tomé and Principe

Benin	Kenya	Senegal
Bhutan	Kiribati	Solomon Islands
Bolivia	Kyrgyz Republic	Sri Lanka
Cabo Verde	Lao PDR	Tanzania
Cambodia	Lebanon	Tajikistan
Cameroon	Lesotho	Timor-Leste
Comoros	Mauritania	Tunisia
Congo, Rep.	Micronesia, Fed. Sts.	Ukraine
Côte d'Ivoire	Mongolia	Uzbekistan
Djibouti	Morocco	Vanuatu
Egypt, Arab Rep.	Myanmar	Vietnam
El Salvador	Nepal	West Bank and Gaza
Eswatini	Nicaragua	Zimbabwe
Ghana	Nigeria	
Haiti	Pakistan	
Honduras	Papua New Guinea	

UPPER-MIDDLE-INCOME ECONOMIES ($4,256 TO $13,205)

Albania	Fiji	Namibia
American Samoa	Gabon	North Macedonia
Argentina	Georgia	Palau
Armenia	Grenada	Paraguay
Azerbaijan	Guatemala	Peru
Belarus	Guyana	Russian Federation
Belize	Iraq	Serbia
Bosnia and Herzegovina	Jamaica	South Africa
Botswana	Jordan	St. Lucia
Brazil	Kazakhstan	St. Vincent & Grenadines
Bulgaria	Kosovo	Suriname
China	Libya	Thailand
Colombia	Malaysia	Tonga
Costa Rica	Maldives	Türkiye
Cuba	Marshall Islands	Turkmenistan
Dominica	Mauritius	Tuvalu
Dominican Republic	Mexico	
Equatorial Guinea	Moldova	
Ecuador	Montenegro	

*Source: World Bank Country and Lending Groups (2023).

Appendix C: Further information on included study designs

We will include studies that use randomized or non-randomized designs to measure a change in outcomes that is attributed to an intervention. This includes studies that apply one of the following approaches:

a. Randomized controlled trials (RCTs), with assignment at individual, household, community, or other cluster level, and quasi-RCTs using prospective methods of assignment (such as alternation).
b. Natural experiments with clearly defined intervention and comparison groups, which exploit natural randomness in implementation assignment by decision makers (e.g., public lottery) or random errors in implementation.
c. Regression discontinuity designs (RDD) or fuzzy-RDD
d. Instrumental variables (IV)
e. Endogenous treatment-effects models, endogenous switching regression, and other methods synonymous to the Heckman two step model.
f. Difference-in-differences (DID), two-way fixed-effects (TWFE), and two-way Mundlak regressions (TWM).
    a. DiD models will include an interaction term between a time and intervention variable in a regression model. They may also regress an intervention variable on a outcome variable measuring the changes in outcomes over time or present a *t*-test comparing changes in outcomes over time between the intervention and control group.
    b. TWFE regressions must include time fixed-effects and unit fixed-effects at the level of the intervention (or lower). For example, if the intervention varies at a village level, it must include either village fixed-effects or fixed-effects of a smaller unit, such as households.
    c. TWM models should be synonymous with the approach described by Wooldridge (2019, 2021). This includes correlated random-effects and pooled OLS regression models that control for unit-specific time averages and time-period specific cross-sectional averages.
g. Interrupted time series (ITS) models, with or without a contemporaneous comparison group. The ITS regression model must adjust for autocorrelation, or it can use autoregressive integrated moving-average (ARIMA) models. An ITS model should include pre-intervention outcome data for a minimum of three time periods.

h. Weighting and matching approaches which control for observable confounding, including non-parametric approaches (e.g., statistical matching, covariate matching, coarsened-exact matching, propensity score matching) and parametric approaches (e.g., propensity-weighted multiple regression analysis).
i. The synthetic control method

Appendix D. Provisional data extraction form

Code	Subcode
Study Information	Study ID
	Coder name
	Title name
	Foreign Title
	Short title
	Language
Author Information	Authors Name
	Authors Affiliation Institution
	Authors Affiliation Country
Publication Information	Publication Type
	DOI
	Study status
	Abstract
	Keywords
	Journal name
	Other journal name
	Journal volume
	Journal issue
	Pages
	Year of Publication
	URL
	Publisher location
	Open access
Sector Information	Sector name
	Sub-sector name
	DAC rank
	Primary DAC Code
	Secondary DAC Code
	CRS-Voluntary (tertiary) Code
	SDGs
	WB first theme
	WB first sub-theme

Code	Subcode
	WB second theme
	WB second sub-theme
	WB third theme
	WB third sub-theme
	Other topics
	Equity focus
	Equity dimension
	Equity description
Geographic Information	First year of intervention
	Continent name
	Country name
	Additional country
	Country income level
	Region name
	State/province name
	District name
	City/town name
	Location name
Target population and cost data	Age
	Sex
	Setting
	Sexual orientation
	Specific population group
	Cost data
	Type of cost data
Methodological information	Evaluation Design
	Evaluation Method
	Mixed Method
	Additional quantitative Methods
	Additional qualitative Methods
	Unit of Observation
Program, Funding and Implementation Information	Project Name
	Implementation Agency Category
	Implementation Agency Name
	Program Funding Agency Category
	Program Funding Agency Name
	Researching Funding Agency Category
	Researching Funding Agency Name
Intervention Information	Treatment group/Arm 1
	Treatment group/Arm 1 Description
	Treatment group/Arm 2
	Treatment group/Arm 2 Description

Code	Subcode
	(Create additional options as necessary)
Outcome Information	Outcome
	Outcome description
	(Create additional options as necessary)

Appendix E: Critical appraisal tool

Checklist for making judgements about how much confidence to place in a systematic review of effects. This checklist has been adapted from Supporting the Use of Research Evidence (SURE) Collaboration. SURE, checklist for making judgements about how much confidence to place in a systematic review. As noted in section 4.5, we will not critically appraise impact evaluations as this is beyond the realistic scope of the EGM.

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

Question	Criteria
A.3 Does the review cover an appropriate time period? *Is the search period comprehensive enough that relevant literature is unlikely to be omitted?*	Yes; can't tell (only use if no information about time period for search); no; unsure *Coding guide:* *YES: Generally, this means searching the literature at least back to 1990* *NO: Generally, if the search does not go back to 1990* *CAN'T TELL: No information about time period for search* *Note: With reference to the above – there may be important reasons for adopting different dates for the search, e.g. depending on the intervention. If you think there are limitations with the timeframe adopted for the search which have not been noted and justified by the authors, you should code this item as a NO and specify your reason for doing so in the comment box below. Older reviews should not be downgraded, but the fact that the search was conducted some time ago should be noted in the quality assessment. Always report the time period for the search in the comment box.*
A.4 Was bias in the selection of articles avoided? Did the authors specify: ▪ Independent screening of full text by at least 2 reviewers ▪ List of included studies provided ▪ List of excluded studies provided	Yes; partially; no *Coding guide:* *YES: All three should be yes, although reviews published in journals are unlikely to have a list of excluded studies (due to limits on word count) and the review should not be penalized for this.* *PARTIALLY: Independent screening and list of included studies provided are both reported* *NO: All other. If list of included studies provided, but the authors do not report whether or not the screening has been done by 2 reviewers review is downgraded to NO.*
A.5 Did the authors use appropriate criteria to assess the quality and risk of bias in analysing the studies that are included? ▪ The criteria used for assessing the quality/ risk of bias were reported ▪ A table or summary of the assessment of each included study for each criterion was reported ▪ Sensible criteria were used that focus on the quality/ risk of bias (and not other qualities of the studies, such as precision or applicability/external validity). "Sensible" is defined as a recognized quality appraisal tool/ checklist, or similar tool which assesses bias in included studies. Please see footnotes for details of the main types of bias such a tool should assess.	Yes; partially; no *Coding guide:* *YES: All three should be yes* *PARTIALLY: The first and third criteria should be reported. If the authors report the criteria for assessing risk of bias and report a summary of this assessment for each criterion, but the criteria may be only partially sensible (e.g. do not address all possible risks of bias, but do address some), we downgrade to PARTIALLY.* *NO: Any other*

Question	Criteria
A.6 Overall – how much confidence do you have in the methods used to identify, include and critically appraise studies? *Summary assessment score A relates to the 5 questions above.* *High confidence applicable when the answers to the questions in section A are all assessed as 'yes'* *Low confidence applicable when any of the following are assessed as 'NO' above: not reporting explicit selection criteria (A1), not conducting reasonably comprehensive search (A2), not avoiding bias in selection of articles (A4), not assessing the risk of bias in included studies (A5)* *Medium confidence applicable for any other – i.e. section A3 is assessed as 'NO' or can't tell and remaining sections are assessed as 'partially' or 'can't tell'*	**Low confidence** (limitations are important enough that the results of the review are not reliable) **Medium confidence** (limitations are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found) **High confidence** (only minor limitations)
Section B: Methods used to analyse the findings	
B.1 Were the characteristics and results of the included studies reliably reported? Was there: □ Independent data extraction by at least 2 reviewers □ A table or summary of the characteristics of the participants, interventions and outcomes for the included studies □ A table or summary of the results of all the included studies	Yes; no; partially; not applicable (e.g. no included studies) *Coding guide:* *YES: All three should be yes* *PARTIALLY: Criteria one and three are yes, but some information is lacking on second criteria.* *No: None of these are reported. If the review does not report whether data was independently extracted by 2 reviewers (possibly a reporting error), we downgrade to NO.* *NOT APPLICABLE: if no studies/no data*
B.2 Are the methods used by the review authors to analyse the findings of the included studies clear, including methods for calculating effect sizes if applicable?	Yes; partially; no; not applicable *Coding guide:* *YES: Methods used clearly reported. If it is clear that the authors use narrative synthesis, they don't need to say this explicitly.* *PARTIALLY: Some reporting on methods but lack of clarity* *NO: Nothing reported on methods* *NOT APPLICABLE: if no studies/no data*

Question	Criteria
B.3 Did the review describe the extent of heterogeneity? Did the review ensure that included studies were similar enough that it made sense to combine them, sensibly divide the included studies into homogeneous groups, or sensibly conclude that it did not make sense to combine or group the included studies? Did the review discuss the extent to which there were important differences in the results of the included studies? If a meta-analysis was done, was the I^2, chi square test for heterogeneity or other appropriate statistic reported? If no statistical test was reported, is a qualitative justification made for the use of random effects?	Yes; partially; no; not applicable *Coding guide:* *YES: First two should be yes, and third category should be yes if applicable should be yes* *PARTIALLY: The first category is yes* *NO: Any other* *NOT APPLICABLE: if no studies/no data*
B.4 Were the findings of the relevant studies combined (or not combined) appropriately relative to the primary question the review addresses and the available data? How was the data analysis done? ▪ Descriptive only ▪ Vote counting based on direction of effect ▪ Vote counting based on statistical significance ▪ Description of range of effect sizes ▪ Meta-analysis ▪ Meta-regression ▪ Other: specify ▪ Not applicable (e.g. no studies or no data) How were the studies weighted in the analysis? ▪ Equal weights (this is what is done when vote counting is used) ▪ By quality or study design (this is rarely done) ▪ Inverse variance (this is what is typically done in a meta-analysis) ▪ Number of participants (sample size) ▪ Other: specify ▪ Not clear ▪ Not applicable (e.g. no studies or no data) Did the review address unit of analysis errors? ▪ Yes - took clustering into account in the analysis (e.g. used intra-cluster correlation coefficient) ▪ No, but acknowledged problem of unit of analysis errors ▪ No mention of issue ▪ Not applicable - no clustered trials or studies included	Yes; partially; no; not applicable (e.g. no studies or no data); can't tell. *Coding guide:* *YES: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors addressed (if appropriate).* *PARTIALLY: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors not addressed (and should have been).* *NO: If narrative OR vote counting (where quantitative analyses would have been possible) OR inappropriate reporting of table, graph or meta-analyses.* *NOT APPLICABLE: if no studies/no data* *CAN'T TELL: if unsure (note reasons in comments below)*

Question	Criteria
B.5 Does the review report evidence appropriately? The review makes clear which evidence is subject to low risk of bias in assessing causality (attribution of outcomes to intervention), and which is likely to be biased, and does so appropriately Where studies of differing risk of bias are included, results are reported and analysed separately by risk of bias status	Yes; partially; no; not applicable *Coding guide:* *YES: Both criteria should be fulfilled (where applicable)* *NO: Criteria not fulfilled* *PARTIALLY: Only one criterion fulfilled, or when there is limited reporting of quality appraisal (the latter applies only when inclusion criteria for study design are appropriate)* *NOT APPLICABLE: No included studies* *Note on reporting evidence and risk of bias: For reviews of effects of 'large n' interventions, experimental and quasi-experimental designs should be included (if available). For reviews of effects of 'small n' interventions, designs appropriate to attribute changes to the intervention should be included (e.g. pre-post with assessment of confounders)*
B.6 Did the review examine the extent to which specific factors might explain differences in the results of the included studies? Were factors that the review authors considered as likely explanatory factors clearly described? Was a sensible method used to explore the extent to which key factors explained heterogeneity? ▪ Descriptive/textual ▪ Graphical ▪ Meta-analysis by sub-groups ▪ Meta-regression ▪ Other	Yes; partially; no; not applicable *Coding guide:* *YES: Explanatory factors clearly described and appropriate methods used to explore heterogeneity* *PARTIALLY: Explanatory factors described but for meta-analyses, sub-group analysis or meta-regression not reported (when they should have been)* *NO: No description or analysis of likely explanatory factors* *NOT APPLICABLE: e.g. too few studies, no important differences in the results of the included studies, or the included studies were so dissimilar that it would not make sense to explore heterogeneity of the results*
B.7 Overall - how much confidence do you have in the methods used to analyse the findings relative to the primary question addressed in the review? *Summary assessment score B relates to the 5 questions in this section, regarding the analysis.* *High confidence applicable when all the answers to the questions in section B are assessed as 'yes'.* *Low confidence applicable when any of the following are assessed as 'NO' above: critical characteristics of the included studies not reported (B1), not describing the extent of heterogeneity (B3), combining results inappropriately (B4), reporting evidence inappropriately (B5).* *Medium confidence applicable for any other: i.e. the "Partial" option is used for any of the 6*	**Low confidence** (limitations are important enough that the results of the review are not reliable) **Medium confidence** (limitations are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found) **High confidence** (only minor limitations)

Question	Criteria
preceding questions or questions and/or B.2 and/ or B.6 are assessed as 'no'.	
Section C: Overall assessment of the reliability of the review	
C.1 Are there any other aspects of the review not mentioned before which lead you to question the results?	▪ Additional methodological concerns – only one person reviewing ▪ Robustness ▪ Interpretation ▪ Conflicts of interest (of the review authors or for included studies) ▪ Other ▪ No other quality issues identified
C.2 Are there any mitigating factors which should be considered in determining the reviews reliability?	▪ Limitations acknowledged ▪ No strong policy conclusions drawn (including in abstract/ summary) ▪ Any other factors
C.3 Based on the above assessments of the methods how would you rate the reliability of the review? **Low confidence in conclusions about effects**: **Medium confidence in conclusions about effects**: The systematic review has the following limitations... **High confidence in conclusions about effects**: If applicable: The review has the following minor limitations... *Coding guide:* **High confidence in conclusions about effects**: high confidence noted overall for sections A and B, unless moderated by answer to C1. **Medium confidence in conclusions about effects**: medium confidence noted overall for sections A or B, unless moderated by answer to C1 or C2. **Low confidence in conclusions about effects**: low confidence noted overall for sections A or B, unless moderated by answer to C1 or C2. Limitations should be summarized above, based on what was noted in Sections A, B and C.	