

## Online appendix A: Additional methods detail

### A.1 PICOS

Specifying an EGM framework requires delineating key parameters, typically using the acronym PICO, which stands for population, intervention, comparison and outcome (Higgins and Green, 2013). PICO has been modified to also include an S for study design to limit number of articles generated in the systematic search (Methley et al., 2014). The following section presents the framework through the specification of PICOS.

#### ***Population (types of study participants)***

We included interventions that service any participant type providing they resided in a low- or middle-income country (L&MIC) when the intervention was delivered. We used the World Bank income status classification for defining L&MICs and studies were classified according to their status in the year the intervention was implemented.

Studies were included if they included at least one estimate of effectiveness for a population based in an LMIC. Studies that included both HIC and LMICs were included if the results are reported separately for at least one LMIC. Multi-country studies were included if results are provided for each country separately.

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

- Is the study primarily designed to determine to what extent a specific technique, technology, treatment, procedure or service works under ideal condition rather than attempt to answer a question relevant to the roll- out of a large programme?
- Is the intervention being carried out by the researchers themselves (e.g., by applying fertiliser in test plots to measure effects on plant growth), or by the people who would carry it out at scale (e.g., farmers applying fertiliser to their crops)?
- Does the study evaluate an intervention that is, or could easily be implemented as, a social policy or programme, or is it “basic science” research on biophysical mechanisms?

## Interventions

Table A1 provides an overview of all interventions eligible for inclusion. The domain column includes the three included domains, the intervention and within each intervention the subgroups of interventions. The last two columns provide definitions and examples of the interventions.

**Table A1: List of included interventions and their definitions**

Domain	Intervention	Intervention subgroup	Definitions	Examples
A. Food Supply Chain	AA. Production system	AAA. Provision of improved water access and management systems	Interventions that provide improved water access and management, through the drilling of wells / pumps, construction of dams or artificial lakes, irrigation, and other water management resources	Drought management, irrigation
A. Food Supply Chain	AA. Production system	AAB. Provision of free or reduced-cost access to agricultural inputs	Improved seed varieties may be resistant to drought or pests or to have higher nutritional content. Improved seed varieties include those created through traditional breeding programs and genetically modified seeds	Improved seed varieties, provision of biofortified seed varieties
A. Food Supply Chain	AA. Production system	AAB. Provision of free or reduced-cost access to agricultural inputs	Provision of free or reduced costs fertilizers	Fertilisers
A. Food Supply Chain	AA. Production system	AAB. Provision of free or reduced-cost access to agricultural inputs	Provision of free or reduced costs pesticides	Pesticides
A. Food Supply Chain	AA. Production system	AAB. Provision of free or reduced-cost access to agricultural inputs	Such as fencings, animal housing, manual tools, protective equipment, antibiotics	Other free or reduced-cost inputs/ livestock health/ feed inputs
A. Food Supply Chain	AA. Production system	AAC. Provision of mechanical equipment	Provision of mechanical equipment such as tractors, tillage machines, and more	Mechanical equipment
A. Food Supply Chain	AA. Production system	AAD. Education / information	Farmer field schools bring together a group of farmers to learn agricultural techniques. They meet regularly during a production cycle, setting up experimentation and engaging in hands-on learning to improve skills and knowledge that will help adapt practices to their specific context. Demonstration farms may be used in farmer field schools or separately to show the use of certain agricultural techniques	Farmer field schools and demonstration farms. Training provided to agricultural producers relating to entrepreneurship, sales and commercialisation
A. Food Supply Chain	AA. Production system	AAD. Education / information	Trained agents visit communities to teach current practices, organize cooperatives, and engage in other secondary activities	Agriculture extension programs coop formation only
A. Food Supply Chain	AA. Production system	AAD. Education / information	If information/reminders are provided over the phone, such as via voice messages or SMS	Information/reminders via mobile phones

Domain	Intervention	Intervention subgroup	Definitions	Examples
A. Food Supply Chain	AA. Production system	AAD. Education / information	Other educational programs supporting the adoption of new agricultural techniques. All mediums of education are included here so long as the information being exchanged is related to agricultural techniques / animal husbandry. Programs related to other educational topics (e.g. literacy) would not be included	Other educational programs, authentication scheme to help farmers identify which fertiliser products are real and which are counterfeiters
A. Food Supply Chain	AA. Production system	AAE. Other efforts to improve the production system	Other interventions that seek to improve the production system in some way	Insurance schemes for agricultural producers, land resettlement/distribution/property rights schemes, pest control management schemes, general credit scheme and conservation practices, contractual agreements to purchase agricultural output (demand side interventions), product quality analysis, e.g., milk quality analysis, GM crop production regulations, fishing regulations
A. Food Supply Chain	AB. Distribution and storage	ABA. Support for creating storage structures at farms	Efforts that provide the technical, logistical, and financial support to improve on farm storage practices	the use of sheds with off-the-ground storage
A. Food Supply Chain	AB. Distribution and storage	ABB. Trade regulations	Regulations that stipulate how much of specific goods can enter or leave a country. Largely restricted to international trade arrangements. Food quality standards in relation to quality with a focus on production; quality or Fairtrade certification if in relation to production	Import / export quotas or bans
A. Food Supply Chain	AB. Distribution and storage	ABB. Trade regulations	Tariffs are taxes / duties placed on foods as they enter or leave a country / region	Tariffs
A. Food Supply Chain	AB. Distribution and storage	ABB. Trade regulations	Any other trade agreements / regulations regarding which foods can enter or exit a country	Other trade agreements / regulations, sustainability/ quality certification
A. Food Supply Chain	AB. Distribution and storage	ABC. Implementation of distribution centres	Farmers cooperatives and other organizations that gather crops/animals from multiple farms to a central location before selling / distributing the goods	
A. Food Supply Chain	AB. Distribution and storage	ABD. Improved transportation from farms to markets	Initiatives that are specifically designed to facilitate the transport of food. This does not include initiatives to facilitate transport broadly, but which incidentally improve food transport as well	

Domain	Intervention	Intervention subgroup	Definitions	Examples
A. Food Supply Chain	AB. Distribution and storage	ABE. Education regarding improved storage and distribution techniques	Educational programs to support storage and distribution techniques. This reflects education itself, without necessarily providing additional services. Including for small food vendors storing foodstuffs, e.g., street food vendors	Training programs for vendors, pre- and post-harvest training
A. Food Supply Chain	AB. Distribution and storage	ABF. Cold chain initiatives	Interventions which support the use of refrigeration to keep food from spoiling between harvest and consumption	Provision of refrigerated technology
A. Food Supply Chain	AC. Processing and packaging	ACA. Fortification	Any intervention / regulation in which micronutrients are added to foods. The fortification includes: at the mill, between the mill and household, local/household, water and national regulations regarding fortification	Providing iron pots for cooking, sprinkles, provision of biofortified food, adding oils to food; fish protein isolate
A. Food Supply Chain	AC. Processing and packaging	ACB. Packaging	Packaging food, largely for preservation purposes both plastic and non-plastic	
A. Food Supply Chain	AC. Processing and packaging	ACC. On farm, post-harvest processing	Any processing of harvested goods that occurs before the foods leave the farm. This is often drying, but can include threshing, winnowing, pounding / milling, and more	Drying
A. Food Supply Chain	AC. Processing and packaging			Other on farm, post-harvest processing
A. Food Supply Chain	AC. Processing and packaging	ACD. Provision of good or services to support food processing and processes or business models	These provisions aim to support food processing and processes or business model. This relates to any processing activity that could be completed outside the farm, such as food preparation (e.g. cooking) in the home	Provision of goods to aid the cooking process, e.g., spoons
A. Food Supply Chain	AC. Processing and packaging	ACE: Education regarding improved processing and packaging techniques	Educational programs to support processing and packaging	Training sessions on packaging
A. Food Supply Chain	AD. Food loss and waste management	ADA. Private food donation	Programs that allow for the donation of foods which will not be consumed before they spoil	
A. Food Supply Chain	AD. Food loss and waste management	ADB. Use of and education regarding the use of spoiled, near spoiled, or traditionally uneaten food	Programs that support the use of spoiled, near spoiled, or traditionally uneaten foods / agricultural by-products. Includes interventions to re-purpose foods / agricultural by-products for non-food use. Agricultural by-products may be re-purposed outside of the food system, but, because the by-product was produced within the food system, this will be included	
A. Food Supply Chain	AD. Food loss and waste management	ADC. Composting	Programs that provide education and resources for composting	Compost making training

Domain	Intervention	Intervention subgroup	Definitions	Examples
<b>B. Food Environment</b>	BA. Availability	BAA. Designations of space and zoning laws	Studies where schools change the foods that are allowed to be sold on campus and request that vendors near campus sell healthier foods. Includes designations of areas as (farmers) markets and explicit exclusion of certain types of retail in some areas. Opening of canteen/restaurant in school, office	List of healthy foods provided to shop keepers, healthier options provided in tuck shops
<b>B. Food Environment</b>	BA. Affordability	BAB. Governmental price manipulations (excluding tariffs)	Tax incentives to retailers. May specifically incentivize certain foods in certain areas	Tax breaks to retailers
<b>B. Food Environment</b>	BA. Affordability	BAB. Governmental price manipulations (excluding tariffs)	Taxes on foods that are considered unhealthy. Most often, sugar sweetened beverages	Taxes for unhealthy foods
<b>B. Food Environment</b>	BA. Affordability	BAB. Governmental price manipulations (excluding tariffs)	Consumer subsidies are often paid by the government to retail locations so that the location can lower the prices on certain foods while remaining profitable	Consumer subsidies
<b>B. Food Environment</b>	BA. Affordability	BAB. Governmental price manipulations (excluding tariffs)	State controls on food prices	Price controls
<b>B. Food Environment</b>	BA. Availability	BAC. Direct provision of foods	State outlets that distribute food at free or reduced cost	Ration shops and other state outlets
<b>B. Food Environment</b>	BA. Availability	BAC. Direct provision of foods	Meals provided at free or reduced cost at school	School feeding programmes
<b>B. Food Environment</b>	BA. Availability	BAC. Direct provision of foods	Provision of food outside of state outlets and school meals. Often relates to the charitable distribution of food by religious or civil society groups	Other direct provision, ins-supplements, macronutrient supplementation
<b>B. Food Environment</b>	BA. Affordability	BAD. Cash-for-food programs	Programs that provide the poor cash for the specific use of household food purchases	Conditional cash transfers for the poor
<b>B. Food Environment</b>	BA. Affordability	BAD. Cash-for-food programs	Programs that provide mothers cash for the specific use of household food purchases	Conditional cash transfers for mothers
<b>B. Food Environment</b>	BA. Availability	BAE. Provision or use of supplements	Programs where the intervention population are provided supplements	Multiple micronutrient supplements
<b>B. Food Environment</b>	BB. Promotion and labelling	BBA. Advertising regulations	Regulations on the types, times, and venues through which food can be advertised	Banning advertisement for high sugar, salt and fat foods during TV shows for children
<b>B. Food Environment</b>	BB. Promotion and labelling	BBB. Innovative store design	Alterations to store design to support the purchasing of healthy foods	Includes menu design

Domain	Intervention	Intervention subgroup	Definitions	Examples
<b>B. Food Environment</b>	BB. Promotion and labelling	BBC. Labelling regulations	Regulations regarding the labelling of food including restrictions on how health claims can be made and the nutrient content of food. Quality or Fairtrade certification if in relation to consumer choice	Improved information and educational messages on outer packaging of micronutrient powders
<b>B. Food Environment</b>	BC. Quality and safety	BCA. Food safety regulations	Regulations regarding the safety of food, including its storage and the testing for contamination	
<b>C. Consumer Behaviour</b>	CA. Efforts to increase women's decision-making power	CAA. Efforts to increase women's decision-making power	Women's empowerment efforts targeted at increasing women's abilities to make decisions regarding the purchase and consumption of healthy foods	Credit for women, women SHG
<b>C. Consumer Behaviour</b>	CB. Information / behaviour change communication	CBA. Peer support / counsellors	The use of peer support or counsellors to increase healthy eating	Includes home visits and other work by community health workers
<b>C. Consumer Behaviour</b>	CB. Information / behaviour change communication	CBB. Professional services (dietitians / nurses)	The use of professional services such as dietitians or nurses to provide messages regarding healthy eating. Can be provided one-on-one or in a group. However, the service provided must be related to their direct training, and not simply general information. The provision of general information to a group is a class. (ex. they must be acting as a dietician and not a teacher, "group counselling" and not "group education"). If they are paid, it is professional. If the researcher provided a service, goes here. Includes the training of professionals (ex. if you train midwives in lactation counselling, are their mothers more likely to BF?)	Breastfeeding promotion interventions delivered by home visits or classes to mothers
<b>C. Consumer Behaviour</b>	CB. Information / behaviour change communication	CBC. Community meetings	The use of community meetings to provide messages regarding healthy eating. Education provided to a specific group would be a class not a community meeting. A community meeting must be a public engagement for discussion and mobilization, not simply education. Education within established groups (such as women's self-help group or microfinance groups) does not count as these are not open to the public	Village meetings organised to improve nutrition in village
<b>C. Consumer Behaviour</b>	CB. Information / behaviour change communication	CBD. Classes	The use of a classroom structure to provide messages regarding healthy eating. This includes classrooms outside of school	Ambiguous reference to "nutrition education" or "education sessions" should be assumed to be classes. Includes online classes

Domain	Intervention	Intervention subgroup	Definitions	Examples
<b>C. Consumer Behaviour</b>	CB. Information / behaviour change communication	CBE. Healthy food social marketing campaigns	Healthy food social marketing campaigns	Provision of media, e.g., magazines and newsletters to promote healthy eating. Including e and m-health, movies and multimedia, text messages, newsletters, posters, games
<b>C. Consumer Behaviour</b>	CB. Information / behaviour change communication	CBF. Door-to-door campaigns	The use of door-to-door campaigns to provide messages regarding healthy eating. Home visits are not door-to-door campaigns. Depending on who did a home visit, this would go into peer support or professional services. Door-to-door means that someone goes from one door to the next, not a specific door	Food hygiene behaviour change intervention through door-to-door household visits

Source: 3ie (2020).

## Outcomes

Table A2 shows all outcomes eligible for inclusion in the EGM. The first column shows whether the outcome is intermediate or final. The next column presents the outcome group and the following one the more specific outcome subgroup. The last two columns provide definitions and examples of for each subgroup.

**Table A2: List of final and intermediate outcomes and their definitions**

Stage in the theory of change	Outcome group	Outcome sub-group	Definitions	Examples indicators
H. Intermediate	HA. Economic	HAA. Income	The total monetary income earned from some activity by an individual, household or firm. For example, income earned from selling fish stock at a market or a salary earned by a labourer. This does not account for the costs incurred by the individual, household or firm	Farm, non-farm income, revenue, sales, costs, profits: Business costs, GM, profit measures, income diversification
H. Intermediate	HA. Economic	HAB. Assets	Assets refer to the property owned by individuals, households, or companies. There are many metrics to measure assets. This may be an indicator of whether or not households have things like TVs, refrigerators, and lights or more complex scoring systems. Livestock and landownership are included elsewhere (HBB and HBD)	Agricultural assets
H. Intermediate	HA. Economic	HAC. Output value	Some measure of the value of the output produced as a result of an intervention. A study might be looking at the change in market value for a crop as a result of changing a seed variety. Here is may be the case that stock has not yet been sold, so income has not been generated	Crop value, livestock value
H. Intermediate	HA. Economic	HAE. Prices received for goods	Measures of the price producers are able to sell their goods for	Sales prices for crops
H. Intermediate	HA. Economic	HAG. Other SES indicators	The amount of money collected through taxes	Wages, wealth, Measures of poverty (absolute, relative), labour productivity, measures of credit access and use, savings, measure of market access for producers, access to business service
H. Intermediate	HA. Economic	HAH. Tax revenue		
H. Intermediate	HA. Economic	HAI. Purchasing behaviour	May include any of the following: -Measures of what foods are bought and how much of the foods are bought. -Measures related re-allocation of budgets due to the intervention. For example, an intervention could increase food prices, and this could result in a smaller proportion of the household budget being used for education. -Measures related to the proportion of goods purchased which were intended to be purchased when the consumer entered the store	Food purchases, re-allocation of budgets, impulse purchases, measures of consumption or expenditure
H. Intermediate	HB. Agricultural	HBA. Water-related	Any measure of on-farm water quality and/or quantity	Amount of water that reach fields due to the intervention, Hours water pumped, mechanically;



Stage in the theory of change	Outcome group	Outcome sub-group	Definitions	Examples indicators
				Distance travelled / time to retrieve water; Animals / herders / farmers using the water source
H. Intermediate	HB. Agricultural	HBB. Animal husbandry	Any measure on animal well-being, health, growth (weight, height) or reproduction	Animals owned, Measures of animal size / health, livestock asset ownership, stress sales of livestock, output/quality for livestock products, livestock disease incidence, animal score, FAO
H. Intermediate	HB. Agricultural	HBC. Plant/crop production	Any measure of plant/crop productivity, growth, health, nutrient composition or measures of presence of pests, diseases or similar. Investment in agricultural inputs for production	Crop yield, crop health, crop loss. Vol of input sales, crop losses/bad harvest/ spoilage pre-farm gate, grain moisture (in context of quality arising from changes to storage), grain losses
H. Intermediate	HB. Agricultural	HBD. Land-related	Measured in hectares, acres, or other units of size Measures related to replacing one crop with another	Area farmed, changes in crops planted, land use, size of cultivated land or land used for agricultural production, area cultivated, proportion of different crops produced (changes thereof), intensity of land use, crop diversity, multiple cropping index, land investment, crop switching
H. Intermediate	HB. Agricultural	HBE. Quality of agricultural inputs	May include measures of soil nutrients or other soil quality indicators, quality and quantity of compost produced	Soil nutrients, compost measures. Quantity and quality of agricultural inputs. Investment in agricultural inputs, value of purchased agricultural inputs, investment in farm modernisation, input purchases, including electricity; compost/soil quality
H. Intermediate	HB. Agricultural	HBF. Agricultural cooperatives	Any measure related to agricultural cooperatives:  Could be related to the profitability of the cooperative, savings, and debt  Could be related to cooperation between producers, the reliability of fees being paid, and the ability of the cooperative to meet its obligations to the farmers	Number of cooperatives created, farms/cooperative, Financial viability of cooperatives, Measures of the functioning of cooperatives
H. Intermediate	HC. Bio nutritional outcomes	HCA. Food nutrient content	Measures of the nutrient content of food, including changes in nutrient content of food due to farming practices, product reformulation, and cooking habits	Breastmilk nutrient content
H. Intermediate	HC. Bio nutritional outcomes	HCB. Caloric requirements	Some interventions may change the caloric requirements of their target population. This outcome measures these changes	Mechanising otherwise labour-intensive tasks

Stage in the theory of change	Outcome group	Outcome sub-group	Definitions	Examples indicators
H. Intermediate	HC. Bio nutritional outcomes	HCC. Nutrient bioavailability	Sometimes, a nutrient is in foods, but for chemical reasons, we cannot absorb / use the nutrient. These studies measure the amount of the nutrient which can actually be used by the body	Rates of absorption, e.g., iron absorption
H. Intermediate	HD. Advertising and labelling	HDA. Exposure to advertisement	Measures of exposure to advertisement	Number of advertisements viewed, number per hour on TV / radio, recognition of advertising slogans, considers if people read labels, understand labels, and if their food choices are informed by the labels
H. Intermediate	HD. Advertising and labelling	HDB. Advertisement topics	The topics of the advertisement and who they are targeting	If they are health foods vs. fast foods and if targeting children
H. Intermediate	HD. Advertising and labelling	HDC. Accuracy of advertisement	Measures of the frequency and accuracy of health claims and disclaimers after advertisements. Considers how often health claims are made and if these claims are accurate	
H. Intermediate	HE. Food distribution	HEA. Import/export	Measures of how much food enters / leaves a region / country	Production and sale of export and non-export crops
H. Intermediate	HE. Food distribution	HEB. Movement of food	Includes measures of the country of origin of food, distance food travels, and time food spent in transport	
H. Intermediate	HE. Food distribution	HEC. Location of foods in stores	Where food is located within stores, schools, canteens, offices	Healthy foods are highlighted at the front of the store
H. Intermediate	HE. Food distribution	HED. Food distribution centres	Measures of the amount of time in line or waiting for rations to become available, Measures of how often foods are not available in retail / distribution locations	Wait lines, stock outs, amount of food distributed, number of individuals who received food
H. Intermediate	HF. Environmental impacts of the food system	HFA. Climate impact	CO2, methane, and other environmental impacts of food production and transport. Measures of the environmental impacts of the food value chain	Can be measured with environmental impact quotient (EIQ)
H. Intermediate	HF. Environmental impacts of the food system	HFB. Non-food waste produced	Waste produced through the production of food that is not the food / crop. Could be related to packaging and processing of foods	
H. Intermediate	HG. Food loss	HGA. Time food remains unspoiled	Measures of how long food remains unspoiled	

Stage in the theory of change	Outcome group	Outcome sub-group	Definitions	Examples indicators
H. Intermediate	HG. Food loss	HGB. Food spoilage	Amount of spoiled, near spoiled, or traditionally uneaten food / agricultural byproducts used, Amount (mass, weight, etc.) of re-purposed food / agricultural by-products	Amount of fruit damaged by fruit flies of total production
H. Intermediate	HG. Food loss	HGC. Food loss	Loss of food after it leaves the farm	Including food wastes
H. Intermediate	HH. Intrinsic motivators	HHA. Consumer preferences	Stated preferences regarding foods and purchases. Includes indicators of how often purchases are informed by perceived "healthiness"	Acceptability measures, demand, sensory assessments
H. Intermediate	HH. Intrinsic motivators	HHB. Perceptions	An individual's perception of how able they are to adopt behaviours which have been targeted by the intervention	Self-efficacy
H. Intermediate	HH. Intrinsic motivators	HHB. Perceptions	What one expects one will gain from adopting the behaviour promoted in an intervention	Benefits of behaviour adoption
H. Intermediate	HH. Intrinsic motivators	HHB. Perceptions	Reasons that one believes one cannot adopt the behaviours promoted by an intervention	Barriers to behaviour adoption
H. Intermediate	HH. Intrinsic motivators	HHB. Perceptions	Consequences one believes could occur if the promoted behaviours are not adopted	Susceptibility
H. Intermediate	HH. Intrinsic motivators	HHB. Perceptions	An individual's perception of how others tend to behave	Perceived norms, self-efficacy, attitudes, perception of other traits and emotions, perceived benefits and costs of behaviours, Pender's HPM constructs, HBM constructs, reported motivations
H. Intermediate	HH. Intrinsic motivators	HHC. Knowledge	Knowledge gained through the intervention, measured by pre-post-tests or similar means	Nutrition knowledge, IYCF knowledge
H. Intermediate	HI. Women's empowerment	HIA. Decision making	Measures of women' participation in household decision making, may include financial decisions or intra-household food allocation	Women's bargaining power
H. Intermediate	HI. Women's empowerment	HIB. Ownership	Outcomes with regards to women's' ownership of land or assets	Livestock ownership, agricultural assets
H. Intermediate	HI. Women's empowerment	HIC. Control of resources	Measures of women's control of resources and the ability to make decisions on how the household resources are used without having to consult other members of the household	Control of money to purchase food
H. Intermediate	HI. Women's empowerment	HID. Self-esteem	Any measure of self-esteem	Empowerment of having paid work
H. Intermediate	HI. Women's empowerment	HIE. Time use	Assessments of women's time use	

Stage in the theory of change	Outcome group	Outcome sub-group	Definitions	Examples indicators
H. Intermediate	HI. Women's empowerment	HIF. Other women's empowerment outcome	Other measures or composite measures of the above	WEIA, PRO-WEAI, measures of IPV
H. Intermediate	HJ. Regulations	HJA. Violations	Measures of how often regulations are violated	
H. Intermediate	HJ. Regulations	HJB. Fines	Measures of the fines levied due to regulatory violations	
H. Intermediate	HJ. Regulations	HJC. Other regulation outcome	Other steps taken due to non-compliance: Measures reflecting other steps taken due to regulatory violations	
H. Intermediate	HK. Economic, social, and political stability	HKA. Economic, social, and political stability	Measures of economic, social, and political stability	Border dispute risk
H. Intermediate	HL. Time use	HLA. Time use	Changes in time use due to the interventions. Includes changes in time spent in manual labour, stores, cooking, and eating while watching TV. Women's time use is not included here, but in HIE	Hours or days of labour
H. Intermediate	HM. Behaviour change	HMA. Behaviour change	Many of the relevant interventions support the adoption of new behaviours. This measures the number / proportion of individuals who adopted the behaviours. In relation to the food system. Education, health and physical activity behaviour change is excluded as they are not a part of the food system	Adoption of agricultural methods, technology etc. Change in dietary habits
H. Intermediate	HN. Other steps taken due to non-compliance	HNA. Other steps taken due to non-compliance	Measures reflecting other steps taken due to regulatory violations	
I. Final	IA. Anthropometric	IAA. Linear growth	Any measure of length or height. This includes units in centimetres/inches, in standardized units such as length-for-age and height-for-age z-scores, and binary outcomes reflecting length/height thresholds such as stunting. Can be measured as average/prevalence at a national or sub-national level. Birth outcomes are included elsewhere (IAE)	Stunting prevalence, cm/m, LAZ, HAZ
I. Final	IA. Anthropometric	IAB. Weight	Any measure of weight. This includes units in kilograms/pound, in standardized units such as weight-for-age z-scores, and binary outcomes reflecting weight thresholds such as underweight. Can be measured as average/prevalence at a national or sub-national level. Birth outcomes included elsewhere (IAE)	Kg, lbs

Stage in the theory of change	Outcome group	Outcome sub-group	Definitions	Examples indicators
I. Final	IA. Anthropometric	IAC. Relative weight	Any measure of the weight of an individual relative to that person's height or age. This includes weight-for-height z-scores (WHZ), weight-for-age z-scores (WAZ), body mass index (BMI), and BMI z-scores. Also includes Severe Acute Malnutrition (SAM) and Moderate Acute Malnutrition (MAM) prevalence, when assessed with WHZ. Can be measured as average/prevalence at a national or sub-national level. Birth outcomes included elsewhere (IAE)	WAZ, WHZ, BMI, wasting prevalence, overweight/obesity prevalence, underweight prevalence, SAM/MAM prevalence, fat-free mass indices (FFMI)
I. Final	IA. Anthropometric	IAD. MUAC	Any measure of the mid-upper arm circumference (MUAC) of an individual, including centimetres/inches, z-scores, and binary outcomes reflecting thresholds. Also includes Severe Acute Malnutrition (SAM) and Moderate Acute Malnutrition (MAM) prevalence, when assessed by MUAC. Can be measured as average/prevalence at a national or sub-national level. Birth outcomes included elsewhere (IAE)	SAM or MAM prevalence, cm, inches
I. Final	IA. Anthropometric	IAE. Birth outcomes	Any measure of anthropometry at birth. Could be measures other than weight, but these are less common. Can be measured as average/prevalence at a national or sub-national level	Low birth weight (LBW), small/large for gestational age (z-score)(SGA), macrosomia, large for gestational age (LGA), birth weight, birth length, birth head circumference, LAZ
I. Final	IA. Anthropometric	IAF. Anthropometric other	This includes all other anthropometric measures	Head circumference, Hip-to-waist ratio, skinfold thickness (subscapular, tricipital) in mm, waist circumference, bone length, body fat
I. Final	IB. Developmental outcomes	IBA. Physical	Any measure of child physical development and growth not captured in anthropometric outcomes, this may include bone or organ development or other	Tooth eruption and bone closure
I. Final	IB. Developmental outcomes	IBB. Other developmental outcome	All other outcome measures of child development, such as cognitive and socioemotional or other	Bayley Psychmotor scales, Raven's matrices, ASQ, Denber II-Jimma screener, Griffiths' developmental scores; measures of IQ, RCPM, Universal Nonverbal Intelligence Test31 (UNIT); Stroop (numbers) test, Backward Digit Span test, Movement Assessment Battery for Children 32 (MABC), finger-tapping test
I. Final	IC. Micronutrient status	ICA. Iron	Any measure of micronutrient content within the body (as opposed to intake). Measures can include markers in the blood, urine, hair, ect. Also included are clinical signs of deficiency or toxicity	Iron deficiency anaemia, serum ferritin level, transferrin saturation, haemoglobin level
I. Final	IC. Micronutrient status	ICB. Iodine	Any measure of micronutrient content within the body (as opposed to intake). Measures can include markers in the blood, urine, hair, ect. Also included are clinical signs of deficiency or toxicity	Iodine deficiency disorder (IDD), goitre, hyperthyroidism, cretinism, thyroid-stimulating hormone (TSH) level

Stage in the theory of change	Outcome group	Outcome sub-group	Definitions	Examples indicators
I. Final	IC. Micronutrient status	ICC. Vitamin A	Any measure of micronutrient content within the body (as opposed to intake). Measures can include markers in the blood, urine, hair, ect. Also included are clinical signs of deficiency or toxicity	Night blindness or xerophthalmia prevalence, plasma/serum retinol
I. Final	IC. Micronutrient status	ICD. Zinc	Any measure of micronutrient content within the body (as opposed to intake). Measures can include markers in the blood, urine, hair, ect. Also included are clinical signs of deficiency or toxicity	Serum zinc, plasma zinc
I. Final	IC. Micronutrient status	ICE. Other micronutrient status outcome	Any measurement of any micronutrient or trace mineral not listed above. Specify the micronutrient in 'other'	Includes salt, vitamin B6, C, D (serum 25 (OH)D; 25-hydroxyvitamin D in nmol/L), E, B-12, calcium, copper, magnesium, thiamine (vit B1), folate (Vit B9), riboflavin (B2); potassium excretion, Beta carotenoids
I. Final	ID. Diet quality and adequacy	IDA. Breastfeeding	Any measurement related to breastfeeding this may include: the proportions of children born in the last 24 months who were put to the breast within one hour of birth, infants 0–5 months of age who are fed exclusively with breast milk, children 12–15 months of age who are fed breast milk or infants 6–8 months of age who receive solid, semi-solid or soft foods. The indicators are largely based on the WHO Infant and Young Child Feeding (IYCF) guidelines	Breastfeeding initiation, EBF, Continued breastfeeding, introduction of complementary foods, ICFI, BF status, CF introduced, BF within given timescales, early initiation, IYCF scores WHO- CF feed, CF practices, paternal perceptions and attitudes toward breastfeeding
I. Final	ID. Diet quality and adequacy	IDB. Dietary diversity	Any quantification of types of food or food groups consumed. Dietary diversity indicates the amounts of different types of foods eaten. Changes quantities or frequencies of specific types does not reflect diversity unless you are going from none to some	Food consumption score, dietary diversity score, Shannon metric / modified functional attribute diversity, minimum dietary diversity, food variety score. Individual and household dietary diversity score. Example: the effect of CTs on food security was examined in seven programs while CT's effect on dietary diversity was examined in nine programs. W-MDD, FCS, DDS
I. Final	ID. Diet quality and adequacy	IDD. Insufficient diet	Measures of the inadequate diets in terms of frequency or quality of food, this may include the amount of processed or fast foods consumed. Indicates that something is sufficient/insufficient relative to an established standard. Minimum dietary diversity and IYCF are not included in this outcome	Alternative Healthy Eating Index (AHEI), 24-hour recall assessment, minimum meal frequency, micronutrient density adequacy, probability of adequacy
I. Final	ID. Diet quality and adequacy	IDE. Micronutrient intake	Measures of the intake of specific micronutrients as a result of the intervention, this may include supplements	% intake of recommended daily intake value (RDI)
I. Final	ID. Diet quality and adequacy	IDF. Other diet quality and adequacy	Any other measure not captured above	Measures the intake of specific food groups (e.g. fish, animal source foods, leafy greens, fast food), frequency of intake, consumption, DQI, intake of macronutrients, calories, consistency of food (infants)

Stage in the theory of change	Outcome group	Outcome sub-group	Definitions	Examples indicators
I. Final	IE. Food safety	IEA. Food toxins	Measures of the content of toxins in food, this may include	Aflatoxin, pesticides, arsenic measures
I. Final	IE. Food safety	IEB. Food borne illness	The frequency at which food born illnesses are reported	Diarrhoea incidence
I. Final	IE. Food safety	IEC. Other food safety outcome	Any other measure not captured above	Milk hygiene (total bacteria count)
I. Final	IF. Food affordability and availability	IFA. Food access	Access to markets as assess by the distance to a market or similar (Note: definitions of "market" may vary), number of different types of food-retail options in a given area. Trade regulations are not included here	Distribution or prevalence of markets or other food outlets, number of months per year with sufficient access
I. Final	IF. Food affordability and availability	IFB. Food availability and supply	Measures of the types of foods available in a given area, may include retail outlets in a given area or measures of national or domestic food supply	Import/export, national food availability tables, national food balance sheets
I. Final	IF. Food affordability and availability	IFC. Affordability	Measure of the cost of a healthy diet. Can be measured as the expected cost of a healthy diet for a family of five for one week. However, units and definition of healthy / sustainable will vary. Any definition used by the authors should be included. Food consumer price index (FCPI) for example reflects the average change in food prices over time	Food consumer price index (FCPI)
I. Final	IF. Food affordability and availability	IFD. Food insecurity measures	Composite measures of food insecurity, typically reflecting a household's reported food security	Self-reports of food insufficiency, Food insecurity experience scale, Food insecurity experience scale, Household food insecurity access scale, food security, household food security level/score
I. Final	IF. Food affordability and availability	IFE. Food stressed households	Reflects the proportion of food stressed households in an area. Any definition of food stress adopted by the authors will be accepted	

Source: 3ie (2020).

## **Study design and comparator**

We included impact evaluations where the comparison/control group receive no intervention, a different food system intervention, a placebo intervention or the study employs a pipeline (wait-list) approach.

We included both impact evaluations and systematic reviews:

- **Impact evaluation:** An impact evaluation is a study that uses rigorous methods to provide a quantitative estimate of the impact of an intervention. This is accomplished by constructing a counterfactual, which provides evidence about what would have happened in the absence of the intervention. In an impact evaluation, the outcomes of those who receive the intervention are compared with those of a comparison group that does not receive the intervention. The comparison group may be a specific population in the study area that does not receive the treatment (as in an RCT), or may be constructed by researchers (as in propensity score matching or interrupted time series). For an impact evaluation to be valid, there must be a sound statistical basis for claiming that the comparison group represents what would have happened to the treatment group had they not received the intervention.
- **Systematic review:** A systematic review is a synthesis of the research evidence on a particular topic, such as the effectiveness of water supply and sanitation, obtained through an exhaustive systematic literature search for all relevant studies using widely accepted scientific strategies to minimize error associated with appraising the design and results of studies. A systematic review uses methods of internal and external quality assurance that make it more similar to a primary study (e.g. double-coding of data, calculation of effect sizes from data reported, synthesis of finding).

We only included studies that implemented at least one of the following study designs that are widely used to evaluate intervention effectiveness (Reeves et al., 2017; Aloe et al., 2017):

- A) Prospective studies that allocate participants to and control groups using randomised or quasi-randomised mechanisms at individual or cluster levels:
  1. Randomised control trial (RCT) with assignment at the individual or cluster level.
  2. Quasi-RCT using a quasi-random method of prospective assignment (e.g. alternation of clusters)
- B) Non-randomised designs with either a known assignment variable(s) or a seemingly random assignment process:
  3. Regression discontinuity designs, where assignment is done on a threshold measured before intervention, and the study uses prospective or retrospective approaches of analysis to control for unobservable confounding.
  4. Natural experiments with clearly defined intervention and comparison groups, which exploit apparently random natural variation in assignment (e.g. lottery) or random errors in implementation, etc.
- C) Non-randomised studies with pre-intervention and post-intervention outcome data:
  5. Panel studies or pseudo-panels with analysis to account for time-invariant unobservables (e.g. difference-in-difference, DID, or fixed-effects models)
  6. Cross-sectional or panel (controlled before and after) studies with an intervention and comparison group using methods to match individuals and groups statistically (e.g. PSM) or control for observable confounding in adjusted regression.
  7. Interrupted time series (ITS)
- C) Other non-randomised designs:
  8. Studies that build a counterfactual through synthetic control approaches.
  9. Cross-sectional studies using multi-stage or multivariate approaches to account for unobservables, including instrumental variable (IV) approaches such as two- or three-stage least squares procedures, or Heckman two-step estimation approaches



D) Systematic effectiveness reviews were included if they described the search, data collection and synthesis methods according to the 3ie database of systematic reviews protocols (Snijlsteit et al. 2016). Any evidence reviews, such as literature reviews, that did not adopt these methods was excluded. We excluded systematic reviews that were not effectiveness reviews (i.e. those which do not aim to synthesise the evidence of the effects of a relevant intervention on priority outcomes of interest), such as systematic reviews of driving factors of nutrition sensitive methods. Where reviews included a mixture of evidence from both high-income and LMICs, we included them if they presented disaggregated evidence for LMICs, or if more than 50 per cent of the evidence of non-disaggregated results were from LMICs. Where there were no disaggregated results for LMICs and more than 50 per cent of the evidence for consolidated findings in a systematic review came from high-income countries, or where it was impossible to ascertain the composition of evidence by income level, the studies were excluded. If reviews included multiple research methods, we included these if over at least 50 percent of studies include at least one impact evaluation design specified above.

We excluded before-after studies or cross-sectional studies that did not attempt to control for selection bias or confounding in anyway.

Case-control studies were also excluded.

Studies that present the results of randomised block designs, where the farm fields or field sections are the blocking unit only were excluded.

Studies that only examine willingness-to-pay for goods, services, process and business models were excluded.

### ***Other criteria***

We also applied the following criteria when selecting studies for inclusion.

- **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 their publication date was 2000 or after to make the search results manageable for the study team to screen in the timeframe that was available.
- **Status of studies:** We will include ongoing and completed impact evaluations and systematic reviews. For on-going studies, we will include prospective study records, protocols and trial registries. Providing an indication of the prevalence and characteristics of on-going evaluation evidence is expected to enrich the analysis of current evidence gaps and support decision making in relation to evidence generation.

## A2. Systematic screening process to identify relevant studies for inclusion

### *Screening process*

Screening of studies was managed using EPPI-Reviewer 4 software (EPPI)<sup>1</sup> (Thomas et al. 2010) and was completed by implementing the following steps:

- **Prepare study records:** All output files of the implemented search strategy were imported into EPPI. Studies that were identified through the additional means specified were added to EPPI manually. An automated process within EPPI was used to remove known duplicate files.
- **Title and abstract screening:** The title and abstracts of all imported and de-duplicated studies were single screened. The screener assigned one code which indicated that either the study should be included for full-text screening, that the study should be excluded, or that they were unsure. If a study was excluded, the reviewer coded the rationale for exclusion. Where screeners had any difficulty in applying the eligibility criteria, a study was screened by a second reviewer, also known as a safety-first approach. We also held periodic meetings to discuss and resolve screening decisions for studies that screeners had coded as 'unsure'.
- **Full-text screening:** We then retrieved a full text for each study that met all the title/abstract inclusion criteria. Initially, two reviewers examined each full text in detail against the protocol and applied a code to indicate whether the study was included or why the study was excluded. Disagreements were reconciled periodically. From August 2020, full texts were single screened with safety first due to time constraints. Overall, 30 percent of studies included at the title and abstract stage were double screened. This process identified a set of studies coded as meeting eligibility criteria. In some cases, there were studies that were ongoing from prior to 2018. If we could not find the final articles, we contacted the authors and if we did not get an answer, we assumed the study was withdrawn and excluded the study.
- **Checks for linked publications:** We grouped publications that were not independent attempts to examine impacts, that is the publications shared sample participants. This typically occurred when authors followed a group of participants over time, published multiple versions of the same study in different formats (for example, a working paper later published as a journal article), or updated a systematic review. Descriptive information was only included once for each group of linked publications, so that each study was independent.

### *Example boundary decisions taken during screening*

This section provides an illustration of decisions made during the screening process with respect to boundary cases, i.e. cases where the eligibility criteria provided in the protocol did not always enable the application of a clear screening decision. The core team meet to discuss these issues on a case-by-case basis to produce and disseminate guidance to the screening team:

- If the outcome were health problems that are not the direct cause of poor nutrition the studies were excluded. Health problems that were included were the following:
  - Iron – anaemia, haemoglobin, haematocrit
  - Iodine – goitre
  - Vitamin A – night blindness, xerophthalmia, retinol or carotenoids; not retinol binding protein)

---

<sup>1</sup> More information available at: <http://eppi.ioe.ac.uk/cms/>. Date accessed: 28/10/20.

- Vitamin D – rickets
  - Vitamin C – scurvy
  - Folate – spina bifida or neural tube defects
  - Calcium – osteoporosis or osteopenia
- We excluded clinical studies unless there was a clear link to the food system as we have defined it. As such, we included studies that evaluate interventions that seek to introduce nutrition-related changes to prevent worsening health conditions for a population in general, and we excluded studies that seek to treat, or prevent the harmful consequences of, specific medical conditions.
  - Agricultural studies that did not mention an attempt to modify farmer practices in some way, were excluded. Studies were also excluded if they were ambiguous on whether a target population was engaged, assuming that if this engagement was a key part of the study it would have been mentioned.
  - Conditional cash transfers were included providing they were related to the food system in some way. For example, if the condition related to food or if parts of the cash transfer was meant to be spent on food. Unconditional cash transfers were all excluded.
  - Cash crops, interventions concerning crops that are not produced for food for humans or animals, were considered outside the food system and therefore excluded.
  - Interventions targeting drivers of the food system without an explicit food system focus were excluded as they do not actually function within the food system. For example, interventions to support conflict resolution in areas experiencing active conflict were not included despite the fact that these interventions would likely support the functioning of the food system.
  - Interventions that seek to enhance the performance of a specific niche populations, such as athletes, the military, astronauts or actors/models were excluded. Interventions that provide nutrition intravenously (i.e. parenteral nutrition) were excluded on LMIC.
  - We excluded studies where specifically migrants form the population of interest – for example, migrants from high-income countries based in LMICs, or vice versa.

### **A3. EGM advisory group**

3ie contacted several organisations to join an impartial advisory group for the project. Currently the following organisations have confirmed their involvement in advising the project's development:

- AECD – Benin
- CARES International
- Global Alliance for Improved Nutrition  
International Food Policy Research Institute
- UNICEF – West and Central Africa
- USAID
- World Bank

#### ***Synthesis project advisory group - Terms of Reference***

Authors of synthesis projects establish stakeholder advisory review groups to help them determine the parameters of their proposed project and to provide inputs throughout the project process to help ensure that the final analytical outputs are policy relevant and have an audience of policy and practice actors that understand the results and that are interested in using the findings. Members of the advisory group can be policymakers, practitioners, influencers, researchers, and other stakeholders with an interest in the project. Members of the advisory group will be asked to provide inputs on various aspects of the project throughout the process. The total time commitment is not likely to exceed two days. The tasks of the advisory group members may include:

- Advice on key decisions regarding the scope of the project, including refining the research question(s) and definitions of key concepts.
- Determine and group important outcomes.
- Suggest relevant background literature and studies for inclusion.
- Provide written comments on draft protocol and draft outputs.
- Help the team draw implications from the findings. This could involve participating in meeting or workshop to review the implications of the project in terms of policy and practice.
- Assist the study team with policy engagement. This can involve advising the team on key stakeholders with whom to communicate to build interest in and understanding of the project results, contribute to developing a communication and uptake plan, facilitate engagement with key audiences and communicate findings.