Systematic review

A systematic review of agricultural interventions that aim to improve nutritional status of children

by Edoardo Masset
Lawrence Haddad
Alex Cornelius
Jairo Isaza-Castro

May 2011

UKaid from the Department for International Development
IDS Institute of Development Studies
3ie
EPPI CENTRE
IOE LONDON Leading education and social research Institute of Education University of London

The EPPI-Centre reference number for this report is 1904.

This report should be cited as: Masset E, Haddad L, Cornelius A and Isaza-Castro J (2011), A systematic review of agricultural interventions that aim to improve nutritional status of children. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

ISBN: 978-1-907345-09-8

Contents

List of abbreviations

3ie	International Initiative on Impact Evaluation
Agris	Agricultural Science and Technology Database
BLDS	British Library of Development Studies
DFID	Department for International Development
Eldis	Electronic Development and Environment Information System
EPPI-Centre	Evidence for Policy and Practice Information and Coordinating Centre
IBSS	International Bibliography of the Social Sciences
IDS	Institute of Development Studies
IEG	Independent Evaluation Group
IFPRI	International Food Policy Research Institute
IOE	Institute of Education
WHO	World Health Organisation

Abstract

This report is a systematic review of the impact of potential ‘win-win’ agricultural interventions that aim to improve children’s nutritional status by improving the incomes and the diet of the rural poor. Previous reviews on the same subject found mixed results or no impact of agricultural interventions on nutritional status. The differences in results across these reviews are the result of the different timeframes and methodologies adopted, and of the different types of agricultural interventions reviewed. We build on and expand previous reviews by covering the period 1990-2010 and we find results similar to those of previous reviews, but we attribute the lack of impact of agricultural interventions on child nutrition to methodological weaknesses of the studies reviewed rather than to specific characteristics of these interventions.

The review is based on a systematic search of the published and unpublished literature. The search was broken down by interventions of the following types: bio-fortification interventions; home gardens; aquaculture and small fisheries; dairy development; and animal source food promotion. During the search we found more than 7,000 studies, but only 23 qualified for final inclusion based on the exclusion criteria set.

We outlined a programme theory of the interventions and we assessed the efficacy of the interventions on five outcome indicators: programme participation; income; diet composition; micronutrients intake; and children’s nutritional status. Of all the studies we reviewed, 23 met our criteria for establishing a credible counterfactual. We found little information on participation rates and characteristics of programme participants. We found that the agricultural interventions considered have a positive impact on the production of agricultural goods promoted by the interventions, but poor evidence of impact on total households’ income. We found only one study that tested for impact on total household income. This study found a positive effect of the intervention.

We found that the interventions were successful in promoting consumption of specific foods but very little evidence was available on changes in the diet of the poor. We found no evidence of impact on the absorption of iron and some evidence of the impact on absorption of vitamin A. Nineteen studies attempted to assess the impact of the interventions on diet composition. Two of these studies undertook no statistical test on diet change, four found no statistically significant impact and 13 found a significant and positive impact on the consumption of food targeted by the intervention. None of the studies assessed whether the interventions improved the quality of the whole diet. Five studies undertook tests for impacts on iron intake. Four tests showed no statistically significant difference at the 5% level and one showed a positive impact at the 5% level. Nine studies tested for programme impact on vitamin A intake, but only four reported data to be able to verify whether there was indeed an impact. The summary effect, assessed by meta-analysis, of these four studies reveals a positive difference in vitamin A intake between project and control groups.

We found no evidence of impact on prevalence rates of stunting, wasting and underweight among children under five. Eight studies examined the impact on childrens’ nutritional status. Of these, only one found a positive and significant impact on stunting prevalence, three found a positive and significant impact on underweight and two found a positive and significant impact on wasting. Five of the eight studies showed no impact on any of the three indicators.

We performed ex-post calculations of the statistical power of the selected studies in detecting differences in nutritional status between the programme and the control groups. We found that none of the studies reviewed would have been able to detect a 'small' impact on prevalence rates of malnutrition; few would have detected a 'medium' impact; and only 50% of the studies would have detected a 'large impact'. Based on this analysis, we concluded that the absence of any reported statistically significant impact of agricultural interventions on childrens' nutritional status found by this review, as well as by other reviews that preceded this one, should not be attributed to the inefficacy of these interventions. Rather it is the lack of power of the studies reviewed that could have prevented the identification of such impact, if any.

We also conducted a validity assessment of the methodologies adopted by the studies reviewed. Few studies performed a rigorous counterfactual analysis of the impact of the interventions. Most studies neglected the analysis of the characteristics of programme participants. Sample sizes were often inadequate and power calculations for determining sample size were rarely performed or presented. Most studies were based on good conceptual framework and analysed intermediate outcomes, but often relied on inappropriate outcome indicators. Finally, all studies neglected the analysis of heterogeneity of impact and were unable to extrapolate results outside the area of the interventions considered.

We also conducted a separate assessment of the existing evidence on the impact of bio-fortification interventions. We built a programme theory of these interventions and reviewed the existing evidence along five intermediate outcomes: successful plant breeding; farmers' response; consumers' response; bioavailability; and nutritional status. We found that consumers' acceptance of bio-fortified staple food is good and that micronutrients in staple food are successfully absorbed by the body. Seven of the studies reviewed found that consumers are willing to pay a premium for food with higher micronutrient contents and only one study found that a discount was needed for consumers to accept bio-fortified staple food. We found little evidence of farmers' acceptance of bio-fortified crops, with only one study reporting farmers' adoption rates. We found little evidence of any impact of these interventions on nutritional status. Only two studies assessed the impact of bio-fortification interventions on nutritional status and found a positive impact.

Acknowledgments

This review was financially supported by DFID. Thanks to Imran Choudury (DFID), Max Gasteen (DFID) and Todd Benson (IFPRI) for their guidance in the process of designing the protocol. Thanks to Jeff Brunton (IOE) for providing and supporting the use of the EPPI-Reviewer software. Thanks to Jeremy Lynd (IDS) and Christophe Bené (IDS) for their help in refining the search of the pastoralist and fisheries literatures. Finally, thanks to Anna Taylor (DFID), Tim Wheeler (DFID), Hugh Waddington (3ie) and particularly Todd Benson (IFPRI) for their comments on an earlier draft. The views and the conclusions expressed in this review are those of the authors and do not necessarily reflect views of DFID or of the aforementioned people.

1 Background

1.1 Aim and rationale of the review

Agricultural interventions aimed at improving nutrition have been implemented by governments and development agencies since the 1960s. However, the focus and types of interventions have changed considerably over the years (World Bank 2007). Changes have followed theoretical developments in the understanding of the causes of hunger and its remedies. Early agricultural interventions focused on increasing food production and agricultural productivity because undernourishment was seen as the result of lack of food. Agricultural extension and irrigation projects are examples of this type of intervention.

It was soon realised that increasing food production alone, whilst ignoring distributional issues, was not sufficient to eradicate malnutrition, unless the poorest were given access to food. Data from the 1960s and the 1970s showed that poor nutritional status coexisted with adequate food supplies (Reutlinger and Pellekaan 1986). From the late 1970s, and particularly after the seminal work on famines by Sen (1981), malnutrition was linked to food security rather than food availability. Food may be available to poor people but they may not have access to it because of lack of means or other constraints. As a result, projects focused on increasing incomes and livelihoods rather than food production. In the fight against hunger, agricultural interventions were replaced by interventions that promoted food security, such as income and food transfers. Interventions increasing agricultural productivity were still implemented but only if targeted to the poorest sections of the rural population.

A shift in thinking about agricultural interventions occurred after research on budget studies in developing countries documented the low calorie elasticity of income (Bouis and Haddad 1992, Strauss and Thomas 1995). Empirical studies using microdata showed that changes in incomes did not immediately translate into changes in the consumption of calories, and highlighted the limitations of increasing income alone in order to improve nutritional status. This research inspired a new wave of agricultural interventions in the 1990s that aimed at simultaneously increasing income and the intake of nutritious food. Undernourishment was to be reduced not only by increasing the incomes of the poor, but also by shifting their diets towards the consumption of more nutritious food in terms of their caloric, protein or micronutrients contents. Unlike their predecessors, these agricultural interventions had the explicit objective of reducing undernourishment. Projects of this type include production diversification projects (such as dairy development, and the promotion of vegetable gardens, fisheries and livestock), and bio-fortification projects (projects increasing the nutritional content of staple foods). This review focuses on this type of intervention.

1.2 Research background

This systematic review builds on previous reviews of the nutritional impact of agricultural interventions. Some of these reviews looked at the nutritional impact of various kinds of agricultural interventions while others focused on a particular subset of agricultural interventions, such as animal food interventions or aquaculture. These reviews have been identified through an extensive search of the available literature and their results are summarised in this section.

Berti et al. (2003) conducted a systematic review to assess whether agricultural interventions of any type have an impact on nutritional outcomes. Their review covered the period from 1985 to 2001 and identified 30 studies. Nutritional impact was assessed using any of the following indicators: dietary intake, biomarkers, anthropometry and morbidity. They found that 19 studies out of 30 had a positive impact on nutritional status measured by any of the indicators above. The authors concluded that 'agriculture interventions had mixed results in terms of improving nutritional status in participating households.'

A review by the World Bank (2007) expanded the work by Berti et al. (2003) by covering the period from 2001 to 2007. This review appears less systematic in design as the search methods and results were not fully described and many of the conclusions are drawn from previous reviews of Ruel (2001), Berti et al. (2003) and Leroy and Frongillo (2007). The objective was to assess the contribution of any agricultural intervention to nutritional outcomes. A total of 52 studies were identified and analysed in relation to three outcomes: food expenditure, caloric intake and anthropometry. The authors found that 'impacts on child nutritional status were limited and mixed' in the case of interventions involving staple foods; interventions involving animal source food 'showed mixed results'; and home gardening interventions 'that did not include a nutrition education component failed to achieve significant impacts on nutritional outcomes'. The authors concluded that 'agricultural interventions have not always been successful in improving nutritional outcomes.'

Ruel (2001a) reviewed the evidence of the impact of food-based interventions designed to increase the intake of micronutrients on four outcomes: production and income; knowledge, attitudes and practices; dietary intake; and nutritional status. The review had a strong focus on home gardening interventions and on the intake of vitamin A and iron. The literature search was restricted to the period 1995 to 1999, and 14 studies were identified. The author concluded that there is some evidence of a positive impact of home gardening interventions on production and knowledge and practices, but that 'the evidence is still scant' on their impact on nutritional status.

Leroy and Frongillo (2007) reviewed the evidence on the nutritional impact of interventions promoting animal production (including aquaculture, dairy development and poultry). A total of 14 studies were identified and analysed across seven indicators: production; income and expenditure; caregiver income; caregiver time and workload; zoonosis; dietary intake and nutritional status (measured by any indicator, including anthropometric measurements, iron deficiency, serum retinol concentration and haemoglobin levels). The authors found a clear positive impact on production and income but not on other indicators. Only four of the studies reviewed reported nutritional outcomes, and the results were not always positive. The authors concluded that 'the evidence is insufficient to answer whether the promotion of animal production is an effective means to alleviate undernutrition'.

Kawarazuka (2010) conducted a systematic review of the impact of aquaculture and small-fisheries interventions on dietary intake and nutritional status. The search identified 23 studies that were analysed across the following indicators: fish consumption, income, women's income and nutritional status. The author found that many interventions increased fish consumption and income and actively involved women. Very few studies analysed impact on nutritional status. The author concluded that 'there is little evidence of the positive changes in nutritional status among households taking up aquaculture' and that the nutritional outcomes of small-scale fisheries 'were not clearly demonstrated'.

Table 1.1 Summary of reviews of the impact of agricultural interventions on nutrition

Review	Period covered	Studies	Interventions	Nutritional impact
Berti et al. (2003)	1985-2001	30	All agriculture: home gardens, animal husbandry, cash cropping, irrigation, land reforms, credit and extension, duck-fish promotion	Mixed results
World Bank (2007)	1985-2007	52	All agriculture: agricultural commercialisation, horticulture, animal source food, and mixed interventions	Mixed results
Ruel (2001)	1995-1999	14	Interventions designed to increase production and intake of micronutrient-rich food through: home gardens, animal husbandry, aquaculture and nutrition education	Some evidence of impact on vitamin A intake but evidence is scant and studies are poorly designed
Leroy and Frongillo (2007)	Not specified (but oldest study is 1987 and most recent is 2003)	14	Animal interventions: aquaculture, dairy production and poultry production	Some evidence of impact but few studies available and often poorly conducted
Kawarazuka (2010)	Not specified (but oldest study is from 2000 and most recent is in 2009)	23	Aquaculture and small-fisheries	Few studies available and very little evidence of impact

Table 1.1 summarises the characteristics of previous systematic reviews and their conclusions regarding the nutritional impact of agricultural interventions. These reviews covered different periods, employed different criteria for inclusion, and covered different types of agricultural interventions. Some regularities are however discernible. First, the reviews found very few studies that assessed the nutritional impact of agricultural interventions. Second, the nutritional outcomes observed were not always positive. Third, all reviews stressed several methodological weaknesses of the studies reviewed.

1.3 Interventions and objectives of the review

The present review synthesises the evidence on the effectiveness of potential 'win-win' agricultural interventions that promote the adoption of new technologies to improve income and the composition of the diet of the poor. Not all agricultural interventions are reviewed, but only those that have the explicit goal of improving the nutritional status of children via an increase in income and a change in diet. In order to clarify, Table 1.2 below presents a list of agricultural interventions that have been included and excluded.

Table 1.2 Types of agricultural interventions included and excluded by the review

Included	Excluded
Bio-fortification	Irrigation
Home gardening	Watershed development
Aquaculture	Credit and microfinance
Small scale fisheries	Land reforms
Poultry development	Marketing
Animal husbandry	Agricultural extension
Dairy development	Food processing and storage

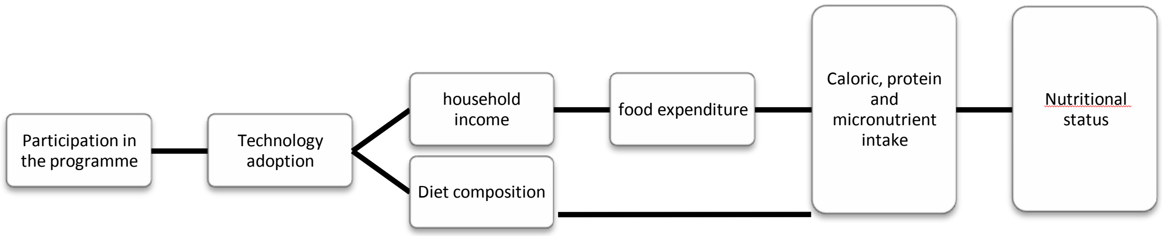
We adopt an economics definition of technology (Varian 1992), whereby a technology is any combination of household resources or other available resources as inputs in the production of an agricultural output. Thus the adoption of improved seeds, the setting up of fisheries, and the introduction of vegetable gardens are all examples of agricultural technologies. Income is the monetary value of the flow of household production net of its cost. A diet is the composition of the food basket consumed by the household, and an improvement in diet can be the result of a diversification of the diet (the addition to the diet of food items that are rich in caloric and protein contents such as milk or fish) or of the enrichment of food items currently consumed (such as the addition of micronutrients to staple foods).

Projects promoting the adoption of new technologies for higher incomes and better diets fall in two main categories:

- Production diversification projects: in particular those promoting dairy production, fisheries, vegetable gardens and livestock.
- Bio-fortification projects: by conventional crop breeding or genetic engineering that increase the content of iron, zinc and vitamins in crops such as rice, wheat and sweet potato.

Figure 1.1 sketches how these interventions are expected to impact on children's nutritional status. a) The first step in the causal chain is farmers' participation in the programme. Not all targeted population are able to join programmes, and vulnerable groups can be missed by the interventions. b) Participation in the programme implies the adoption of technology which may consist of new production activities (such as livestock and fisheries) or new varieties of existing production activities (such as the production of fortified food). c) The effect of technology adoption is twofold: first, the new technology increases household income; second, the new technology changes the food basket consumed by households or the nutritional contents of existing baskets. d) Both effects - higher incomes and better diets - lead to improvements in nutritional status. A more detailed exposition of the 'programme theory' can be found in Appendix 1.2.

Figure 1.1 Pathways of impact of agricultural interventions on nutrition



This review assesses agricultural interventions along the logical pathway shown in Figure 1.1 using five indicators: programme participation; household income; diet diversity; micronutrient intake; and nutritional status. The choice of the specific metric for each indicator is not only based on a judgment of what we believe is the best metric available for each indicator, but also on the availability of data in the reviewed studies. Table 1.3 lists and briefly describes the indicators considered. Programme participation is assessed through programme participation rates, particularly among vulnerable groups such as the poor, women and children. Household income is total household income rather than income from agriculture because substitutions effects are likely (as farm income increases households may employ less labour in other activities). Diet diversity is assessed in terms of changes in the consumption of specific food items such as fish, milk, vegetables and fruit. Ideally we would use diet diversity indices, but these are rarely employed in the literature. Micronutrient intake is measured in terms of vitamin A because this has been the preferred indicator of analysis in the literature.[1] Nutritional status is measured by the prevalence rates of stunting, wasting and underweight that are obtained through anthropometric measurement of children under five. Ideally we would use continuous indicators of undernutrition (such as Z-scores) rather than rates; focus on long term indicators of malnutrition (such as stunting) rather than on short term ones (such as underweight or wasting); and assess nutritional status of children above the age of four. However, research efforts have focused on the three anthropometric indicators above.

Table 1.3 Description of outcome indicators employed in the review

Indicator	Description
Programme participation	Characteristics of targeted population and participation rates
Income	Total household income
Diet diversity	Consumption of calorie, protein and micronutrient rich food
Micronutrient intake	Vitamin A intake
Nutritional status	Prevalence rates of stunting, underweight and wasting among children under five

[1] The international research effort has focused on three micronutrients: iron, zinc and vitamin A Mason JB, Lofti M, Dalmiya N, Sethuraman K, Deitcher M (2001) *The Micronutrient Report: Current Progress and Trends in the Control of Vitamin A, Iodine, and Iron Deficiencies*. Ottawa: The Micronutrients Initiative.. Vitamin A is relatively easy to measure by the concentration of serum retinol from blood samples, prevalence of night blindness and Bitot's spot at the clinical level, and by estimation from the consumption of food items from budget studies. Iron intake can be measured by the concentration of serum ferreting from blood samples or by the prevalence of anaemia at the clinical level, while estimation from intake of specific foods from budget studies is imprecise. There are currently no standard methods to assess the concentration of zinc in the human body.

2 Methods

2.1 Defining relevant studies

Based on the conceptual framework developed in section 1, we set a number of exclusion criteria in order to screen the studies relevant to this review. A study was excluded if:

- Date: it was produced before 1990.
- Language: it was not written in English.
- Geographic location: it was conducted in a high income country as classified by the World Bank.
- Intervention: it did not investigate the impact of agricultural interventions on any of the key outcome indicators identified by the review (see Table 1.3).
- Study design: it did not employ a credible methodology for assessing programme impact.

While the first three screening criteria used (date, language, and location) are rather uncontroversial, the last two (type of intervention and study design) need some qualifications. The number of development interventions that could be classified as 'agricultural interventions' is very large. This review focuses on those interventions aimed at improving childrens' nutritional status via an increase in income and a change in the diet of programme participants. Whilst this definition clarifies the scope of the review, it does not provide sufficient guidance for identifying studies to be included in the analysis. Therefore we specified a set of agricultural interventions that fall within this category. These interventions are: bio-fortification (both conventional breeding and genetic engineering); vegetables home gardens; dairy development; fisheries (aquaculture and small scale fisheries); and animal husbandry (including poultry). This review is confined to these five types of interventions. While it might be argued that some types of interventions were incorrectly excluded from our review (e.g. natural resources management), selecting specific programme types enabled us to set clear boundaries for this review.

A study methodology was considered credible when the impact of the programme on the outcome indicators (participation; income; diet diversity; and nutritional status) was based on a counterfactual analysis whereby the programme group was compared to a control group and a meaningful effort had been made to minimise selection bias. Agricultural interventions are often targeted at specific population groups or attract individuals with specific characteristics. Typically, the poor or the very poor participate in these programmes. Because these groups have pre-intervention characteristics, including nutritional status, which differ from those of the general population, they cannot be compared to the general population. Randomisation, longitudinal studies and matching methods deal efficiently with this issue. In order to facilitate the screening task, the studies reviewed were coded under one the following categories:

- No control group.
- Before-after comparison of participants.
- Comparison of participants and non-participants.
- Cross-sectional project-control comparison.
- Cross-sectional project-control comparison over time.
- Panel project-control comparison over time.
- Randomised field trials.

Studies not employing a control group and studies based on comparisons of outcomes before and after the interventions were excluded. Studies comparing participants and non-participants over a cross section were included if they attempted to address selection bias by either matching project and control observations on pre-intervention individual, households or community characteristics; or by testing differences in characteristics between the project and control observations before the interventions. Studies based on longitudinal observations and randomised field trials were accepted.

2.2 Search strategy

Searching the literature to build a database of studies requires finding a balance between general key words that result in the retrieval of a large number of papers, many of which are irrelevant, and more specific key words that may miss relevant studies. In order to find a satisfactory set of key words for our search, we conducted a number of tests. First, we searched using very inclusive terms such as 'agriculture', 'projects', 'nutrition'. This search retrieved papers in the order of tens of thousands, which was beyond the scope of our work. We then searched using keywords for the five types of interventions selected, namely: bio-fortification; home gardens; dairy development; fisheries; and animal husbandry. The advantage of this research strategy is that it enables other researchers to extend the review to these other types of programmes whilst maintaining clear boundaries regarding what is included.

Next we identified multiple key words for the five intervention types in three different ways. First, using a thesaurus we identified synonymous terms for each intervention. Second, we browsed some of the studies initially retrieved in order to find commonly used synonyms. Third, we used a 'text-mining' function of EPPI-Reviewer to list the most commonly used words in selected documents to help us find further synonyms. Keywords for intervention types were then paired with keywords for programme outcomes subject to review. We tested the use of synonyms of programme outcomes in the search, but this resulted in the retrieval of an unmanageable number of studies. We therefore decided to opt for a listing of programme outcomes with a limited number of synonyms. A detailed list of the keywords used, searches run and tables of results is in Appendix 2.

The literature search was conducted in ten databases that cover comprehensively the available literature on the following broad thematic areas: nutrition; agriculture; rural development; and social sciences. These ten databases also enabled us to include both the published literature (Econnlit, IBSS, PubMed and Web of Science) and unpublished literature (Agris, Eldis, IDEAS, IFRPI, Jolis and World Bank).

In order to minimise the loss of relevant papers, we included in our search strategy the following additional steps. First, we searched the reference lists of key studies. Second, we applied the 'forwards citation' utility in Google scholar in order to find papers that cited these studies. Third, we reviewed the full reference lists of the systematic reviews previously conducted on the same subject (see section 1.2). Finally, we contacted a number of experts in relevant field who provided further references particularly in the area of aquaculture and small fisheries.[2]

[2] Howarth Bouis (IFPRI) was contacted about bio-fortification projects. Christophe Bené (IDS) provided extensive and useful inputs in the search of aquaculture related research. Jeremy Lind (IDS) provided input in the search of the animal husbandry literature.

The search returned a total of 10,885 studies which were reduced to 7,239 after the elimination of duplicates. The titles and abstracts of these studies were uploaded to the EPPI-Reviewer database. More than 75% of studies identified were from the published literature. Studies on fisheries and dairy represented respectively 26% and 23% of the total number of studies returned. The other three types of interventions (bio-fortification, home gardens and animal husbandry) each represented about 17% of the total number of studies.[3]

2.3 Screening studies

The studies identified by our search were screened in two stages. In the first, one reviewer read titles and abstracts and applied all selection criteria outlined in section 2.1, with the exception of study design. This process led to the exclusion of 6,932 studies from the 7,239 originally selected and to the inclusion of 307 studies for further screening. The large majority of studies excluded were written before 1990, concerned a high income country or failed to mention in the title and abstract any of the outcome indicators specified in section 2.1. Only a few studies were excluded because they were written in a language other than English. Agris, which accounts for 12% of total search results, is the only database that returned a large number of non-English studies (around a quarter of total). Therefore less than 5% of the total number of studies originally searched was excluded based on the language of publication. The full text of each study that passed the first stage of screening was uploaded into EPPI-Reviewer.

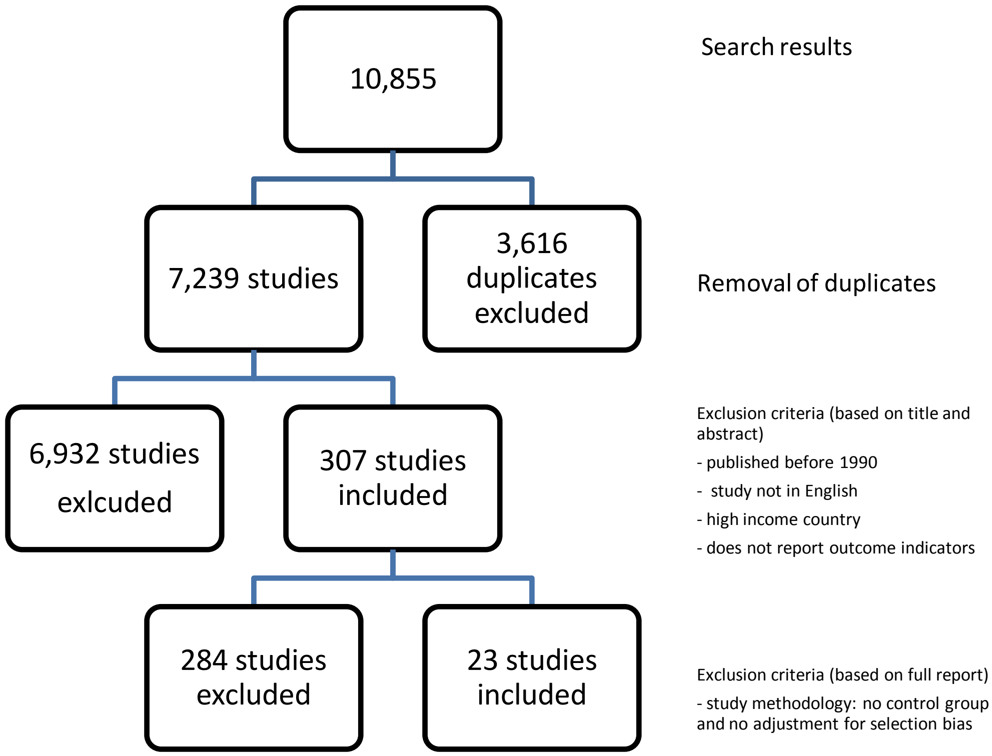
In the second stage of the screening process two other reviewers read the full articles of the studies and coded them based on the methodology employed to assess programme impact (see section 2.1). Studies that did not meet the methodological standards set by the review were excluded. An in-depth assessment of the methodology employed by the studies included was conducted by the lead reviewer in order to assess the internal and external validity of the studies in great detail. This exercise served two purposes. First, it validated the screening exercise already performed.[4] Second, it enabled us to provide recommendations for further research on this subject. This process led to the exclusion of 284 studies and the inclusion of 23 studies for the final review. The two stages of the screening process are illustrated in Figure 2.1.

In the course of the second stage selection process, the lead reviewer performed a thorough assessment of the methodology employed by the included studies. The internal and external validity of the studies was analysed in great detail.

[3] See table in Appendix 2.

[4] Based on the coding of the study methodology, 30 studies had been included in the second stage of the screening process. The in-depth analysis of the methodology employed by these 30 studies resulted in further exclusion of 7, thus leading to the final inclusion of 23 studies. A careful scrutiny of the methodology employed by these 7 studies revealed a series of flaws compromising their internal validity that could not be picked up through the coding exercise.

Figure 2.1 Stages of the screening process



2.4 Methods for synthesis

The number of studies included in the review was too small to conduct a full meta-analysis of their results, despite the outcome indicators selected being amenable to this type of analysis. Nutritional outcomes in particular are objectively measurable with considerable precision. The measurement of micronutrient intake can be complex if based on blood samples or on detailed expenditure surveys, but anthropometric measurements of children under five require a very simple technology and only moderately skilled staff to perform the measurements. Measurement errors are possible due to misreporting of the age of the children, errors in the use of the measurement tools and in data reporting. However, large samples obviate measurement error and population estimates of the height and weight of children can be fairly accurate (WHO Expert Committee 1995). It is therefore surprising that the measurement of childrens' height and weight is so rarely performed by researchers, particularly in the evaluation of projects aimed at reducing undernutrition among children. Yet less than half of the studies included in the review had collected anthropometric measurements of children in the study area.

An additional impediment to performing a full meta-analysis is that the 23 studies included in the review were not identified based on the reporting of a single outcome indicator. Rather, five different indicators were considered for the review and a study was included provided it reported data on any of these indicators. As a result, the summary of results may count on a varying number of studies and this number is invariably small. A consequence of the paucity of the data available is that we could not conduct any meaningful analysis of impact disaggregated by type of intervention, geographic area of intervention or other grouping. This is very unfortunate because the variability in the outcomes of such complex interventions is likely to depend on the type of interventions, the characteristics of implementation (including the staff capacity) and the concomitant health and socio-economic factors that affect nutritional status.

Further, even when the impact on one indicator is assessed by several studies it is often the case that these use different metrics that are not strictly comparable. For example, some researchers collect income data while other collect expenditure data. Changes in the diet are sometimes measured by indices of diet diversity or simply collecting data on the intake of specific food items. Nutritional status can be measured by prevalence rates of undernutrition or by standardised Z-scores, and so forth. This further reduces the scope for aggregating the results from different studies in a single metric.

Because of these difficulties, in some cases we report results from the included studies using summary tables of impacts on income, diet composition and nutritional status. These summaries report the number of statistically significant effects against non-statistically significant ones, thus providing a very general indication of the direction of impact based on the available evidence. This procedure, known as 'vote-counting', can be misleading and a word of caution in the interpretation of the results is necessary (Borenstein et al. 2009). First, authors often do not report results that are statistically non-significant in such a way that we might be left with the pool of statistically significant effects found. This can lead to an overestimation of programme impact. Second, lack of significance should not be confused with lack of impact. Studies may show no statistical significance because of poor study design (small sample size) rather than because of absence of programme impact. This can lead to an underestimation of programme impact. In other words, the summary results by vote counting come with an implicit level of incertitude and should not be taken at face value.

In the analysis we also make use of forest plots in reporting the programme impact on intake of vitamin A and on undernutrion rates. A forest plot is a graphical representation of a meta-analysis. In a meta-analysis all effect sizes from the different studies are pooled and a single effect size for the pooled data, with its standard error and confidence interval, is calculated. A forest plot reports separately for each study, the effect size and the confidence interval, plus the relative weight in determining the summary result.

When conducting a meta-analysis, a choice needs to be made between a fixed-effect model or a random-effect model (Borenstein et al. 2009). The fixed-effect model assumes that the effect size of an intervention varies across studies only because of a random error in each study (for example resulting from measurement error). The random-effect model on the other hand assumes that variation in effect size depends not only on random error but also on the specific characteristics of the population and of the intervention considered in each study. When effects are collected from studies performed by different researchers for different programmes and different populations, a random-effect model is more appropriate. However, when the number of studies is small, as in our case, the estimates of a random-effect model are imprecise. In this case a fixed-effect model is preferable, but it should be noted that, under these conditions, this approach provides a simple descriptive analysis of the studies reviewed and that the summary effect reported has very little generalisability to a wider population.

2.5 Methods for assessing validity

As part of the second stage of the screening process we conducted an assessment of the methodology employed by the studies reviewed. The goal of the assessment was to judge the internal and external validity of the studies in order to obtain a snapshot of the current state of research and to provide guidance for future research. Internal and external validity have been defined in many ways and can take different meanings (Shadish et al. 2002). We define internal validity as the ability of a study to establish causality. We adopted two criteria of internal validity: the use of a valid control group in the analysis and the use of a sample size sufficiently large to detect a difference in the outcome of interest. Studies conduct counterfactual analysis with varying degrees of precision. For example, some studies compare nutritional outcomes among participants before and after an intervention while others adopt a randomised design. Similarly, sampling of the units of observations can be performed with varying degrees of accuracy. At one end of the spectrum, studies use small sample sizes and do not conduct power analysis, while at the other end the sample size is defined through power analysis.

We define external validity as the ability of a study to establish causal relationships that can be extrapolated to areas or countries different from those in which the study was originally conducted. Internal validity is often privileged over external validity with the consequence that rigorous research ends up providing little guidance for policy (Pelletier 2002). In order to offer guidance for policy, the causal effect should be analysed over varying characteristics of the population (heterogeneity of impact). This type of analysis increases the predictive ability of the findings when the same intervention is introduced to a new area with different characteristics. Studies may or may not conduct an analysis of the heterogeneity of impact. For example, some studies report the average intervention effect on the population, while others analyse how the impact varies depending on the socio-economic conditions of the households. In addition, a study would offer better guidance for policy if the intervention is analysed not just over its final outcomes but also over its intermediate outcomes. The understanding of the operation of a project over its intermediate steps, and of the factors that prevent an efficient operation, helps the adaptation of the same intervention to other contexts. Some studies simply ignore the programme theory of an intervention, while others proceed to test the success of all intermediate steps of a project.

Based on the considerations above, we built a scale to score each study along two dimensions of internal validity (counterfactual analysis and power) and two dimensions of external validity (assessment of intermediate outcomes and of heterogeneity of impact). We defined simple criteria to score the quality of a study as low, medium or high, along each of these dimensions. These criteria and the scores are listed in Table 2.1. The studies selected for the review were assessed and scored based on these criteria for internal and external validity. A visual summary of this assessment is in Figure 3.2, which shows the proportion of studies scoring low, medium and high for each of the four dimensions assessed.[5]

[5] Scores and justification of the assigned scores are presented for each study in Appendix 6.

Table 2.1 Criteria and scores employed in the validity assessment

	Criteria	High score	Medium score	Low score
Internal validity	Counterfactual analysis	• Randomised experiment • Sound matching technique • Difference in difference analysis • Credible selection correction	• RCT , matching or double difference analysis poorly performed • Selection correction not credible	• Comparison of participants to unmatched non-participants • Before-after comparisons • No control group
	Power	• Power calculations over variables of interest	• Weak power calculations	• No power calculations • Sample from a specific area or population
External validity	Programme theory	• Intermediate outcomes estimated	• Intermediate outcomes considered but not analysed	• No intermediate outcomes considered
	Heterogeneity	• Heterogeneity of impact analysed	• Heterogeneity considered but not analysed	• No heterogeneity considered

3 Identifying and describing studies: results

This section presents the results of the review for each outcome of the programme theory outlined in section 1.3: programme participation; impact on incomes; impact on diet composition; impact on micronutrient intake; and nutritional status of children. Details of the outcomes for each included study by programme type are described in Appendix 5. As only two studies on the impact of bio-fortification interventions were found, a separate review of the outcomes of the programme theory of bio-fortification interventions was produced and is reported in Appendix 3.

3.1 Characteristics of included studies

The largest number of studies returned by the initial literature search was on fisheries and dairy interventions (see Table 3.1). However, these types of interventions are poorly represented among the 23 studies selected for the final review where home gardens interventions figure predominantly. The discrepancy between the frequency distribution of search results and final inclusions across studies is to be explained by the methodological approaches adopted by researchers evaluating different programmes. Impact evaluations of dairy development, animal husbandry and fisheries projects are extremely rare. Studies in these areas are often conducted by researchers who have little familiarity with statistical methods and principles of counterfactual analysis. Conversely, a large number of home gardens interventions have been evaluated by researchers competent in the use of quantitative evaluation methods. In the case of bio-fortification programmes the lack of evidence is largely due to the novelty of the interventions. Few bio-fortification programmes have been in operation for a sufficiently long time to be rigorously evaluated. There is however a large literature testing the necessary preconditions for the effectiveness of these project. This literature was thoroughly reviewed and our analysis is reported in Appendix 3.

Table 3.1 Screening of studies by type of agricultural intervention

Programme type	Search results	1st stage screening	2nd stage screening
Bio-fortification	833	87	2
Home gardens	1,347	65	16
Fisheries	2,088	81	3
Dairy development	1,709	38	1
Animal husbandry	1,262	36	1
TOTAL	7,239	307	23

3.2 Programme participation

Most studies described the population targeted by the interventions only in very general terms. This normally consisted of one of the following categories: poor geographic areas; women; poor households; remote communities. Data were collected on children under five and occasionally on mothers or the entire household. No study reported participation rates or the ability of the programmes to reach the targeted population. No study described the socio-economic characteristics of the programme participants or estimated a model identifying the determinants of participation.

3.3 Income

Only five studies reported programme impact on total household income. One of the studies (Low et al. 2007) collected income data from project and control areas but did not present the results. All studies that collected data on total household income found a large impact of the interventions, but only in one case was the statistical difference between project and control groups tested (see Table 3.2). Absence of a test is not proof of lack of statistically significant difference, but it is difficult to draw any conclusion from the results reported by these studies.

Some studies reported project impact on incomes from a particular source, such as income from home gardening (Bushamuka et al. 2005, Olney et al. 2009, Talukder et al. 2010), or on cash income from sales of the food item promoted by the intervention, for example from market sales of home garden produce (Chakravarty 2000). This is however a very imprecise measure of income because substitution effects in production are possible. As income from a source increases and family labour supply also increases for that source, income and labour supply from another income source may well decrease. As a result, the overall impact on household income, and therefore on food expenditure, remains unclear. For this reason we decided to exclude from the analysis data on market sales or income from a particular source. These data may provide evidence of the programme success in promoting production of a specific good but do not represent evidence of an overall increase in household resources.

Table 3.2 Impact of the interventions on total household income

Study	Impact	Description
Hoorweg et al. (2000), Kenya	Positive (statistically significant at 0.001)	Income among dairy farmers is 40% higher than among non dairy farmers
Marsh et al. (1998), Bangladesh	Positive (no statistical test)	Households with home garden have slightly higher incomes
Murshed-e-Jahan (2010), Bangladesh	Positive (no statistical test)	Income increases more rapidly among farmers in aquaculture programme and at the endline is 40% higher in the project group
Nielsen et al. (2003), Bangladesh	Positive (no statistical test)	Income is 15% higher among households in a poultry promotion programme
Schipani et al. (2002), Thailand	Positive (no statistical test)	Income is 60% higher among families with home gardens

3.4 Diet composition

Of the 23 studies selected for the review, 19 reported programme impact on the diet composition of the beneficiary population. Impacts on diet composition across studies could not be standardised in a meaningful way because different studies reported consumption of different food items and did so in different ways. First, studies reported consumption of food items that were specifically targeted by the intervention. Second, some studies collected data on household monetary expenditures while others used the physical quantities consumed. Third, studies reported consumption using different recall periods - 24 hours, 7 days or 1 month – and frequency was reported in different ways: as yes/no consumption over the recall period (such as whether a child consumed any fish over the last week); number of times a food item was consumed over some time period (such as the number of times fish was eaten over a week); consumption frequencies compared to a cut-off frequency (such as eating fish more or less often than once a week); or in terms of a diet diversity index.

Due to the difficulty of standardising programme effect on different food items that are measured and reported in different ways, we decided to produce a narrative summary of the impact of the interventions. Table 3.3 details for each study the food item whose consumption increased, the statistical significance of the observed change, and the population on which the measurement was taken (whether, for example, they were households or children).

The large majority of studies found a positive impact of the interventions on the consumption of specific foods. As expected, home garden programmes increased the consumption of fruit and vegetables; aquaculture and small fisheries interventions increased the consumption of fish, while dairy development projects increased the consumption of milk. There are, however, a number of exceptions where no changes were observed.

One difficulty in interpreting the results of these studies is that the majority focused on the impact on the consumption of the food item that was targeted by the intervention (for example fish, eggs or sweet potato) thus ignoring substitution effects in consumption. For example, Bushamuka et al. (2005) found that while consumption of vegetables, rice and fish increased after the intervention, consumption of pulses decreased. In another example, de Pee et al. (1998) found that home gardening increased the consumption of plant food but at the same time reduced the consumption of vitamin A-rich animal food. This is a familiar phenomenon to economists analysing consumer behaviour. The economic analysis of consumer behaviour often assumes that households make consumption decisions regarding broad categories of food items before deciding about specific food items within each category (Deaton and Muellbauer 1980). Thus it is perfectly plausible for a family to reduce the consumption of equally protein-rich food, such as meat or milk, when the consumption of fish increases. This suggests that consumption of any single category of food is not an appropriate indicator of impact and that studies should rather focus on indicators of diet diversity by looking at the composition of the whole food basket consumed by the household. The only studies reviewed that did so are Low et al. (2007) and Olney et al. (2009), while the studies by Bushamuka et al. (2005) and de Pee et al. (1998) quoted above reported consumption figures on each item of the food basket separately.

Changes in the diet appear to be independent of the unit of observations within the family. Some studies collected data on food consumption of children, other studies collected data on consumption by mothers, while others studies collected consumption data by the whole household. There is no indication that children and women did not benefit from the interventions in terms of diet improvements. A full understanding of the rules defining intra-household food allocation would require data on changes in consumption of all household members separately. However, the data at hand do not offer strong support to the hypothesis that women and children are discriminated against in the allocation of food under the selected interventions.

Table 3.3 Impact of the interventions on diet composition

Study	Statistical significance of the difference observed	Food items	Observed population
Bushamuka et al. (2005)	**	Higher consumption of vegetables, rice, fish, oils; stable consumption of meat; reduction in consumption of pulses	Households
de Pee et al. (1998)	**	Higher consumption of plant food but lower consumption of animal food	Women
Faber et al. (2002)	**	Higher consumption of carrots, butternut squash and spinach	Children 2-5 years
Greiner and Mitra (1995)	**	Higher consumption of oil rich foods, green leafy vegetables but not of yellow fruit	Children 1-6 years
Hoorweg et al. (2000)	***	40% higher milk intake	Households
Jones et al. (2005)	***	Higher consumption of green leafy vegetables and orange-fleshed potatoes and fruit, but no difference in the consumption of animal source food	Households
Kidala et al. (2000)	***	Higher consumption of vitamin A rich food	Children 6-71 months
Laurie and Faber (2008)	**	Higher consumption of butternut, sweet potato, but not of carrot and pumpkin	Children 1-5 years
Low et al. (2007)	***	Higher dietary diversity and higher consumption of papaya, orange-fleshed potato and dark green leaves	Children 6-59 months
Marsh (1998)	(no statistical test)	Higher consumption of vegetables	Children under five
Murshee-e-Jahan et al. (2010)	(no statistical test)	Higher consumption of fish but lower consumption of staple cereals	Households
Nielsen et al. (2003)	n.s.	No difference found in any element of the diet composition	Women and 6-12 year old daughters
Olney et al. (2009)	**	Higher dietary diversity and higher consumption of liver, meat and eggs among children but not mothers	Women and children under five
Roos et al. (2003)	n.s.	No differences in fish intake	Households
Schipani et al. (2002)	n.s.	No difference in any component of the diet	Household and children under five
Shmidt and Vorster (1995)	n.s.	No difference in the consumption of vegetables	Children 6 -13 years
Smitasiri et al. (1999)	***	Increase in the consumption of vitamin A rich food and fats	Preschool girls, schoolgirls 10-13 years, lactating and pregnant women
Talukder et al. (2010)	*	Higher consumption of chicken liver and animal food but not of vegetables	Household and children 6-59 months
Vijayaraghavan et al. (1997)	**	Higher consumption of vitamin A rich food	Preschool children

Note: in the second column, n.s. is not statistically significant, * is statistically significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level

3.5 Micronutrient intake

The international research effort on micronutrient deficiencies has focused on three micronutrients: vitamin A, iron and zinc. There is currently no standard methodology to assess levels of zinc in the human body and the studies included in this review have investigated the impact of interventions on iron or vitamin A. Iron and vitamin A deficiencies are a major concern in the developing world. The average prevalence of iron deficiency in developing countries among children under five is estimated to be 45% (Micronutrient Initiative 2009), with little sign of progress over the period from 1970 to 2000 (Mason et al. 2001). There is evidence that iron deficiency increases maternal mortality risk (via haemorrhage) and that it impairs the cognitive development of children (Micronutrient Initiative 2009). Adequate levels of vitamin A in the body prevent blindness in childhood and reduce mortality among children under five (Micronutrient Initiative 2009). Despite a clear decrease in the prevalence of Vitamin A deficiency globally (Mason et al. 2001), it is believed that vitamin A deficiency is compromising the immune system of about 40% of children under five in developing countries and leading to the death of approximately one million children every year (Micronutrient Initiative 2004).

We found only four studies that assessed the impact of the interventions on iron intake. The findings of these studies cannot be aggregated in a summary figure of impact because iron intake measurement were taken and reported in different ways. One study (Talukder et al. 2010) reported a statistically significant reduction in anaemia prevalence among non-pregnant women in project areas and no change in non-intervention areas. The other three studies found no statistically significant impact of the interventions on iron intake (see Table 3.4).

Table 3.4 Impact of interventions on iron intake

Study	Iron measurement	Statistical significance of the difference	Impact
Olney et al. (2009)	Haemoglobin	n.s.	No differences in levels among women and children under five
Schipani et al. (2002)	Serum ferritin and haemoglobin	n.s.	No differences in levels among women and children under five
Roos et al. (2003)	Food consumption	n.s.	No difference in household iron intake
Talukder et al. (2010)	Anaemia prevalence	** (in Nepal) * (in Bangladesh)	Difference found among non-pregnant women in Bangladesh and Nepal but not in Cambodia

Note: in the second column, n.s. is not statistically significant, * is statistically significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level

Vitamin A levels in the body can be assessed in three ways. First, vitamin A deficiency can be detected at the clinical level by the prevalence of night blindness and Bitot's spot. Second, it can be measured by the levels of serum retinol concentration in blood samples. Third, it can be deducted from food consumption data by employing micronutrients conversion factors. The second method (serum retinol concentration) is the most accurate because, unlike consumption, it can be assessed at the individual level and because, unlike night blindness, it can be measured on a continuous scale. In addition, vitamin A is stored in fat and therefore it is not linearly correlated with intake of vitamin A rich food making estimates from food consumption imprecise.

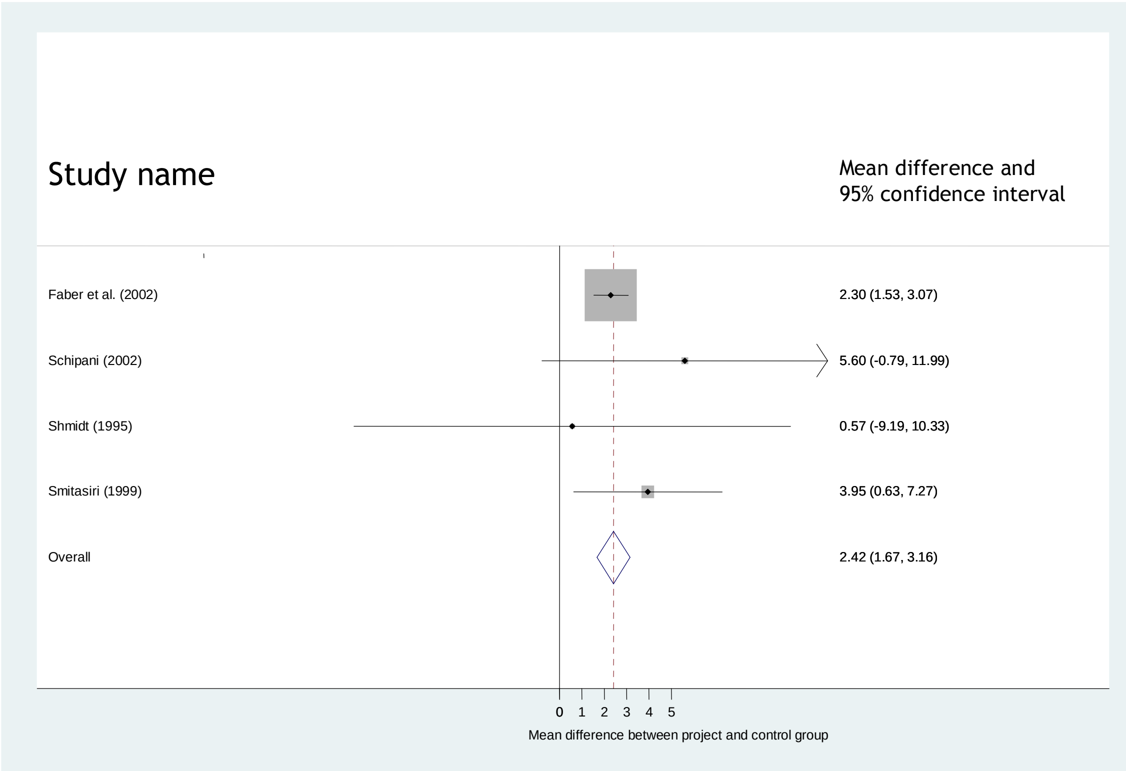
Nine studies reported mean concentration of serum retinol in blood samples from project and control areas. [6] However, only four of these studies reported means and standard deviations of children in project and control areas. Attig et al. (1993) and Vijayaraghavan et al. (1997) reported results only for programme participants and not

[6] Serum retinol concentration is sometimes reported in molar measures of vitamin A (based on the molecular weight of retinol), sometimes in micrograms. All measures were converted to micrograms/dl. using the following equivalence: 1 micromole of retinol is equal to 286.46 mg.

for control groups. Low et al. (2006) reported a double difference estimate of 2.2 microgram/dl. between project and control areas (statistically significant at the 1% level), but using a regression model. de Pee et al. (1998) reported a difference of 1.7 microgram/dl. (statistically significant at the 5% level), but among women and not children. Kidala et al. (1997) found a statistically significant lower concentration of serum retinol in project areas, but attributed this result to an extraordinary parasitic infestation in the intervention areas that compromised the validity of the experiment.

The results of the four studies reporting mean differences and relative standard deviations are summarised in the forest plot of Figure 3.1. The difference between the mean serum retinol concentration in the project and control group is reported for each study with a 95% confidence interval. The size of the square represents the weight of each study in the calculation of the summary effects. Since we used a fixed effect model, the weight is the inverse of the study variance and is higher for studies with larger sample size that provide more precise estimates. The summary impact of the interventions is an overall difference of 2.4 microgram/dl. serum retinol between project and control areas (Z test of significance is 6.35 and P value<0.001). This summary is simply the weighted mean of the effects found by the individual studies and is represented by the position of the diamond in Figure 3.1. The width of the diamond is the confidence interval of the summary effect and represents its level of precision. This meta-analysis provides some support to the hypothesis that agricultural interventions improve intake of vitamin A among children under five.

Figure 3.1 Forest plot of differences between project and control areas in serum retinol concentrations (micrograms/dl.) among children under five



3.6 Nutritional status

Children's nutritional status is normally measured by collecting anthropometric data (WHO Expert Committee 1995). Heights and weights of children under five are compared to height and weight of a healthy 'reference' population of children of the same age, and a score is calculated (Z-score – the number of standard deviations from the median of the reference population). Children with scores below -2 and -3 are classified as undernourished and severely undernourished respectively. This procedure is performed separately for three different indicators: stunting (low height-for-age), underweight (low weight-for-age) and wasting (low weight-for-height). Stunting is a long term indicator of health and hunger, while wasting is an indicator of acute undernutrition, and underweight is an indicator of more difficult interpretation because an underweight child could be either stunted or wasted or both (WHO Expert Committee 1995).

Anthropometric data were collected by 13 of the 23 studies included in the review. However, three of these studies (de Pee et al. 1998, Kidala et al. 2000, Talukder et al. 2010) did not report prevalence rates of undernutrition, while the study by Marsh (2002) reported only changes in underweight rates but without a statistical test of significance. The remaining studies employed the anthropometric data to calculate prevalence rates of stunting, underweight or wasting and performed statistical tests of significance. The only exception is the study by Gunaratna et al. (2010) which employed rates of growth in height and weight, rather than prevalence rates, and found positive and statistically significant impact on nutrition by the interventions considered.

The results of these studies are listed in Table 3.5. We decided not to summarise the results by meta-analysis because this is likely to be a biased sample of studies. Some of the studies collected anthropometric data but did not report the differences observed or the results of statistical tests of significance. If the lack of reporting is due to the absence of statistical significance of programme impact, the meta-analysis of only studies that reported prevalence rates and statistical tests would be biased, and the summary outcome would be overestimated.

Only one study found a statistically significant impact on prevalence of stunting, while three studies (out of eight) found an impact on prevalence rates of underweight and two found an impact on wasting. The relatively greater success of agricultural interventions in reducing the prevalence of underweight and wasting compared to stunting can be explained at least in two ways. The interventions considered may be better suited to addressing short-term undernutrition rather than chronic undernutrition. An alternative explanation is that the studies assessed impact shortly after the interventions had taken place and could therefore not capture long-term impact such as chronic undernutrition.

Table 3.5 Impact of interventions on nutritional status of children

Study	Stunting (height-for-age)	Underweight (weight-for-age)	Wasting (weight-for-height)
Aiga et al. (2002)	n.s	**	n.s
Faber et al. (2002)	n.s	n.s	n.s.
Hoorweg et al. (2000)	**	**	**
Makhotla and Hendriks (2004)	n.s.	n.s.	n.s.
Low et al. (2007)	n.s.	**	**
Olney et al. (2009)	n.s.	n.s.	n.s
Schipani et al. (2002)	n.s.	n.s.	n.s.
Shmidt and Vorster (1995)	n.s.	n.s.	n.s.

Note: in the second column, n.s. is not statistically significant, * is statistically significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level

Overall these results provide little support to the hypothesis that agricultural interventions help reduce undernutrition. However, they should not be interpreted as evidence of the absence of impact. Lack of significance can be the result of absence of impact as well of absence of statistical power (Borenstein et al. 2009), and many of the studies reviewed were conducted over small samples of children.

In empirical research it is standard practice to set a null hypothesis stating, for example, the equality of the means of two populations. If a difference between the means is found and if the probability that this difference is the result of chance is less than a critical value (normally 5% - the 'alpha'), the null hypothesis is rejected and the difference between the means is accepted as 'true'. If the statistical test finds a probability higher than 5 or 10% that the difference is the result of chance, then the null hypothesis is accepted and the difference is considered 'false'.

Non-rejection of the null hypothesis should not be confused with its acceptance. Accepting the null when it is false is a type II error. The likelihood of committing a type II error depends on power, which is the probability of finding a difference when there is indeed a difference. Power is normally used to determine sample size. Any of the following variables - difference between populations, sample size, alpha and power - can be determined given the other three. Hence, given a desired probability of detecting a difference when there is a difference (power), an acceptable probability of detecting a difference when there is no difference (alpha), and an expected difference between the two populations, the sample size can be calculated. A common number for alpha is 5% which tells us that in 5% of the experiments of a given sample size we will find an impact when there is none. A common number for power is 80% which tells us that, in 80% of the experiments of a given sample size conducted in a given population we will be able to identify an effect when there is an effect.

Power can also be calculated ex-post in order to assess, for example, the probability of an experiment of finding an effect. This is rarely done and the lessons from doing it are not always learned (Sedlmeier and Gigerenzer 1989). We calculated power for the studies (Table 3.6) to assess the probability of finding differences in the prevalence rate of undernutrition. Since true project effects are not known we calculate power for three hypothetical project effects: small, medium and large.

Target 1c of the Millennium Development Goals sets the objective of halving the population suffering from hunger by 50% over the period from 1990 to 2015 as measured by the prevalence of underweight. It has been estimated that over the same period a reduction by 36% is more realistic (de Onis et al 2004), which implies a reduction rate of 1.78% per year. In another article, de Onis et al. (2008) estimated a 31% reduction in stunting rates from 1980 to 2000, which is equal to a reduction rate of 1.84% per year. We use these figures to set upper and lower bounds of project effects. If a project achieves a reduction in undernutrition that would take 20-25 years, according to the estimates above, then the project effect is 'large' (30%). If a project achieves a reduction in undernutrition that would take one year, according to the estimates above, then its effect is 'small' (2%). To set a 'medium' effect we used an average of effects of nutrition interventions that have been rigorously evaluated and summarised in a report by the World Bank (IEG 2010) which gave us a value of 10%.[7]

[7] This figure was obtained by calculating a simple average of effects reported in the IEG review (average of 5.5% for stunting and of 11% for underweight).

Table 3.6 Ex-post power calculations assessing the ability of the studies to detect impact on nutritional status

Study	Sample size	Samples ratio (control/ project)	Change in stunting prevalence			Change in underweight prevalence		
			Small (2%)	Medium (10%)	Large (30%)	Small (2%)	Medium (10%)	Large (30%)
Shmidt et al. (1995)	36	1.00	0.05	0.04	0.08	0.05	0.04	0.09
Schipani et al. (2002)	60	1.00	0.04	0.04	0.06	0.05	0.04	0.04
Aiga et al. (2002)	66	1.00	0.03	0.06	0.35	0.03	0.03	0.12
Hoorweg et al. (2000)	102	1.00	0.03	0.03	0.14	0.03	0.03	0.13
Faber et al. (2002)	165	0.35	0.03	0.05	0.29	0.03	0.04	0.20
Olney et al. (2009)	445	0.44	0.03	0.14	0.80	0.03	0.09	0.54
Low et al. (2007)	741	0.33	0.04	0.49	0.99	0.03	0.16	0.89
Marsh (1998)	1,200	0.17	0.03	0.17	0.90	0.03	0.12	0.75
Makhotla and Hendriks (2004)	2,688	0.25	0.05	0.52	1.00	0.04	0.25	0.98
Average power			0.04	0.15	0.51	0.04	0.09	0.42

Note: the study conducted by Marsh (1998) is not reported in Table 3.5 because did not perform statistical test of the observed differences in prevalence rates.

We then calculated for each study the probability of finding a 'small', 'medium' and 'large' project effect.[8] We also calculated an average power across all studies for each project effect. Table 3.6 shows the striking results of these calculations. None of the studies reviewed would be able to detect a 'small' impact. The probability of detecting a small effect for these studies is less than 5%. Even a 'large' effect, i.e. a reduction in malnutrition of 30%, would only be detected by 50% of the studies. Finally, a more realistic medium project effect of 10% would be detected only by 15% of the studies. These numbers are very unsatisfactory. Based on this analysis, we conclude that the absence of any reported statistically significant impact of agricultural interventions on children nutritional status found by this review, as well as by other reviews that preceded this one, should not be attributed to the inefficacy of these interventions. Rather it is the lack of power of the studies reviewed that could have prevented the identification of such impact, if any.

Finally we calculated the sample size that would be required to detect small, medium and large project effects. To do so, we assumed that the sample would be allocated to project and control group in equal proportions and we set a value of power of 80% and alpha of 5%. The results are reported in Table 3.7. A minimum of 500 observations is required even to detect a large impact on prevalence rates. A medium effect requires samples of 5,000 observations and above. Small effects, as defined by this review, cannot be detected as the number of observations required exceeds the hundreds of thousands.

[8] We performed these calculations using the SAMPSI command in the Stata software for the calculation of power and sample sizes in two-sample comparison of proportions. The Stata command applies a formula for power calculations that can be found in Fleiss et al. Fleiss JL, Levin B, Paik MC (2003) *Statistical Methods for Rates and Proportions 3rd ed.* New York: Wiley.. Notice that these calculations do not take into account intra-cluster correlations, and therefore tend to overestimate power.

Table 3.7 Sample sizes required to detect impacts on prevalence rates of stunting and underweight

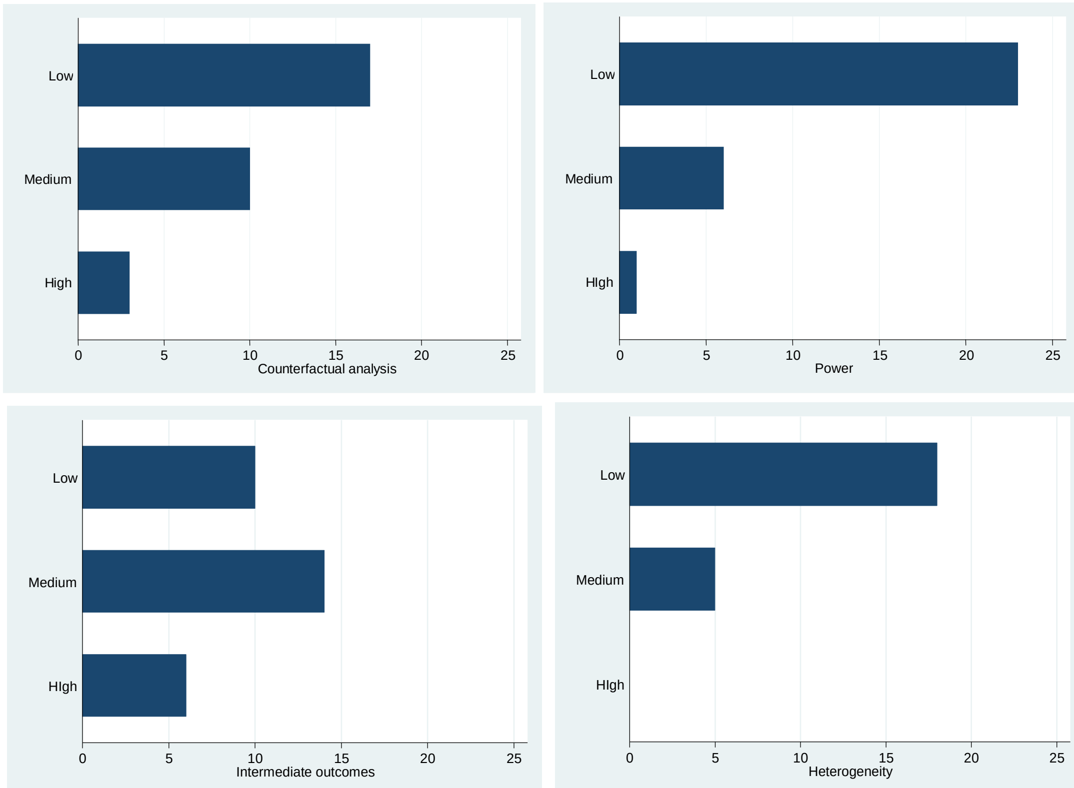
Study	Power	Alpha	Samples ratio (control/ project)	Sample required to detect a change in stunting prevalence		
				Small (2%)	Medium (10%)	Large (30%)
Stunting	.80	.05	1	117,832	4,722	520
Underweight	.80	.05	1	182,752	7,240	778

3.7 Validity of studies

Few studies were found to be of high quality according to the standards specified for assessing validity. Rigorous counterfactual analysis has not been a major concern of this literature. Several studies were excluded because based on before-after comparisons or because comparing outcomes of participants and non-participants from the same project areas without the minimum attempt to address selection bias. Only one study exploited a random allocation of the intervention (Kidala et al. 2000) but the number of clusters surveyed was so small that the observed impact on nutrition was affected by a parasitic infestation in one of the project villages. Another study consisted of a meta-analysis of a series of unpublished randomised interventions in bio-fortification (Gunaratna et al. 2010). Most studies consisted of project-control comparisons where the controls, either households or villages, were selected based on similarity of characteristics that were either very few or not made explicit.

In addition to ignoring selection bias issues, the studies reviewed neglected the analysis of the determinants of participation in the programmes. Data on participation rates and characteristics of project participants were rarely collected or reported. This is striking given that these interventions were often targeted to specific groups, such as marginalised households, women or the poor. Without data on characteristics of participants and non-participants it is impossible to judge, and correct for, the equality of the samples of project and control groups. It is also impossible to understand the factors that induce people to join or not join the programme, thus limiting our ability to understand programme outcomes.

Figure 3.2 Average quality scores of the selected studies by indicators of internal and external validity



Power calculations for selecting sample size were rarely performed or presented. Samples were often small in terms of clusters covered and no power calculations were found that included intra-cluster correlation. The few power calculations we found were performed over a single indicator even though the quantitative analysis was conducted over a large number of outcomes. For example, a sample size was defined using power formulae aimed at detecting the impact of the intervention on vitamin A, but then the same sample size was also used to assess prevalence of undernutrition for which the sample size was inadequate. Lack of statistical power is of great concern because the absence of statistical significance is often identified with absence of impact as it was discuss at length in section 3.6.

The studies were reasonably good in exposing the programme theory of the interventions. Intermediate outcomes were often described and analysed in detail and many studies explained the pathways of change implicit in the interventions. Many studies tested programme impact on intermediate outcomes such as income generation and diet diversification. However, there were problems with the specific metrics employed. Studies often assessed impact on income from a specific activity (such as dairy) or from market sales, rather than impact on total household income. Given substitution effects discussed in section 3.3, the latter is a much better indicator than the former. Similarly, there are substitution effects in consumption. People may decrease consumption of some foods while increasing consumption of food promoted by the project. However, the majority of the studies focused on the impact of the interventions on fostering consumption of specific healthy foods rather than their overall impact on the composition of a healthy diet.

We found no study that performed a high quality impact analysis differentiated over households with different characteristics such as wealth, gender of head of household or location of residence. Heterogeneity of impact has not been a major concern of this literature. This is an important limitation of these studies, because even the few that employed a rigorous methodology for assessing impact cannot extrapolate their findings to areas with different characteristics.

4 Conclusions

This review assessed the impact of agricultural interventions aimed at improving nutritional status of children on five outcome indicators: participation in the programme; household income; diet composition; vitamin A intake; and nutritional status of children. Of all the studies we reviewed, 23 met our criteria for establishing a credible counterfactual. These 23 studies broke down into the following intervention types: bio-fortification (2); home gardens (16); fisheries (3); dairy development (1); and animal husbandry (1). No study reported participation rates in the programme. Only one study tested for impact on household income and found a positive effect. Nineteen studies attempted to assess the impact of the interventions on diet composition. Two of these studies undertook no statistical test on diet change, four found no statistically significant impact and 13 found a significant and positive impact on the consumption of food targeted by the intervention. None of the studies assessed whether the interventions improved the quality of the whole diet. Five studies undertook tests for impacts on iron intake. Four tests showed no statistically significant difference at the 5% level and one showed a positive impact at the 5% level. Nine studies tested for programme impact on vitamin A intake, but only four reported data to be able to verify whether there was indeed an impact. The summary effect, assessed by meta-analysis, of these four studies reveals a positive difference in vitamin A intake between project and control groups. Eight studies examined the impact on children nutritional status. Of these, only one found a positive and significant impact on stunting, three found a positive and significant impact on underweight and two found a positive and significant impact on wasting. Five of the eight studies showed no impact on any of the three indicators.

The results of the review can be summarised as follows:

- No data is available on participation rates or characteristics of participants in agricultural interventions. As a result, little is known about: the impact of these interventions on specific vulnerable groups; the targeting efficiency of the interventions; the characteristic of programme participants.
- Agricultural interventions appear to have a positive impact on the production of the food item promoted by the intervention. However, it is less clear whether these interventions have a positive impact on total household income. The evidence available is very weak. Given the generally low response of food consumption and particularly of calories consumption to income changes, it is unlikely that the interventions considered had an impact on nutritional status via a simple income effect.
- Agricultural interventions change the diet of the beneficiary households in the expected way. There is considerable evidence that the interventions analysed are successful in promoting the consumption of specific food items such as vegetables, fish or milk. However, consumers can, for example, compensate for an increase in the consumption of fish with a reduction in the consumption of other protein rich food such as meat. The overall impact of the interventions on the diet of the poor remains unexplored.
- The impact of agricultural interventions on micronutrients is unclear. There is no evidence of an impact of the interventions on iron intake. There is some evidence of a positive impact on vitamin A intake, but the number of studies available is too small to generate robust results as the summary results are

very sensitive to the inclusion of one or two studies. Evaluations of bio-fortification interventions are positive, but again, the number of these evaluations is too small to provide conclusive answers.

- The studies reviewed report little or no impact of agricultural interventions on the nutritional status of children. This result confirms the results of previous systematic reviews on the same topic. However, unlike previous reviews, we attribute this result to the lack of statistical power of the studies reviewed rather than to the lack of efficacy of these interventions.
- The studies reviewed found a greater impact of the intervention on the prevalence of short term indicators of hunger (wasting and underweight) versus long-term indicators (stunting). However, this result could be a consequence of the short time frame adopted by the evaluations, which is not well suited to detect long term effects.

This review also conducted an assessment of the methodological validity of the studies reviewed. This assessment found that studies are often based on solid conceptual frameworks and that intermediate indicators of impact are often assessed. This is extremely helpful in understanding reasons of failure or success of the interventions and allowed us to structure the assessment of these interventions along the steps of the causal chain. However, much less attention was given to how impact varies across beneficiaries with different socio-economic characteristics or vulnerable groups. This limits the possibility of extrapolating the results of these studies to other populations.

We also found that the absence of adequate control groups, the use of before-after comparisons and the neglect of selection bias are common features of many evaluations of agricultural interventions. We found a general disregard for understanding the determinants of participation in the programmes and the characteristics of participants. As a result, the ability of the interventions to reach the poorest sections of the population and to do so effectively remains unexplored. Finally, we found that most studies suffer from poor statistical power and are not able to detect a positive impact of the interventions on the nutritional status of children.

We conclude with a number of recommendations for primary research. First, studies investigating the impact of agricultural interventions on nutrition should be better designed. We found few studies based on a randomised design (all randomised trials found had been reviewed by a meta-analysis of studies on quality protein maize) and the sample sizes adopted by studies adopting a project-control design were often inadequate to detect an impact on undernutrition rates. More randomised trials and longitudinal studies should be conducted, and power calculations should form the basis for the choice of sample size.

Second, studies should make an effort to collect data on both participants and non-participants in the interventions, both across and within communities. Data on the characteristics of participants and non-participants are not only needed to address selection bias, but also to investigate the determinants of participation and the impact of the interventions across vulnerable groups.

Third, studies should refine the metrics used to assess the impact of agricultural projects. Total household income data should be collected rather than income from particular sources. Data on consumption should be disaggregated by category or summarised by indices of diet diversity. Measurement of height and weight of children should be always performed.

Finally, we found several studies that did not report observed differences in nutritional outcomes when the data had been collected, while in some cases statistical test of the observed difference were not performed or reported. Observed differences and statistical tests should always be performed and reported regardless of the results obtained.

To conclude, the very important question of whether agricultural interventions have a positive impact on nutritional status of children does not currently have an answer. We recommend that more rigorous impact evaluation studies of agricultural interventions be conducted in the future.

5 References

Alderman H, Bouis H, Haddad L (1997) Aggregation, flexible forms, and estimation of food consumption parameters: Comment. American Journal of Agricultural Economics 79: 267-268.

Aluru M, Xu Y, Guo R, Wang Z, Li S, White W, Wang K, Rodermel S (2008) Generation of transgenic maize with enhanced provitamin A content. Journal Of Experimental Botany 59: 3551-3562.

Berti P, Krasevec J, Fitzgerald S (2003) A review of the effectiveness of agricultural interventions in improving nutrition outcomes. Public Health Nutrition 7: 599-609.

Bhaskaram P (2002) MIcronutrient Malnutrition, Infection, and Immunity: An Overview. Nutriton Reviews 60: S(40)-S(45).

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009) Introduction to Meta-Analysis. Chippenham: Wiley.

Bouis HE, Haddad LJ (1992) Are Estimates of Calorie Income Elasticities Too High - a Recalibration of the Plausible Range. Journal of Development Economics 39: 333-364.

Bouis HE, Graham RD, Welch RM (1999) The CGIAR Micronutrients Project: Justification, HIstory, Objectives, and Summary of Findings. Improving Human Nutrition through Agriculture: The Role of International Agricultural Research Workshop October 5-7, 1999.

Brooks S (2010) Rice Biofortification: Lessons for Global Science and Development. London: Earthscan.

Bushamuka V, de Pee S, Talukder A, Kiess L, Panagides D, Taher A, Bloem M (2005) Impact of a homestead gardening program on household food security and empowerment of women in Bangladesh. Food and Nutrition Bulletin 26: 17-25.

Chowdhury S, Meenakshi JV, Tomlins K, Owori C (2009) Are Consumers Willing to Pay More for Biofortified Foods? Evidence from a Field Experiment in Uganda. HarvestPlus Working Paper 3.

Copenaghen Consesus Center (2008) Copenaghen Consesus 2008 - Results. Copenhagen Busines School, Frederiksberg.

Darnton-Hill IB, de Benoist MW, Brown LR (2002) Micronutrient restoration and fortification: Communicating change, benefits and risks. Asia Pacific Journal Of Clinical Nutrition 11(S6): S184-S196.

De Groote H, Kimenju CS (2008) Comparing consumer preferences for color and nutritional quality in maize: Application of a semi-double-bound logistic model on urban consumers in Kenya. Food Policy 33(4): 362-370.

de Pee S, Bloem MWW, Gorstein J, Sari M, Satoto, Yip R, Shrimpton R, Muhilal (1998) Reappraisal of the role of vegetables in the vitamin A status of mothers in Central Java, Indonesia. The American journal of clinical nutrition 68: 1068-1074.

Deaton A, Muellbauer J (1980) Economics and consumer behaviour. Cambridge: Cambridge University Press.

Deaton A (1997) The Analysis of Household Surveys. Washington DC: The World Bank.
Depositario DPT, Nayga RM, Wu X, Laude TP (2009) Effects of Information on Consumers' Willingness to Pay for Golden Rice. Asian Economic Journal 23: 457-476.
Donangelo CM, Woodhouse LR, King SM, Toffolo G, Shames DM (2003) Iron and zinc absorption from two bean (Phaseolus vulgaris L.) genotypes in young women. Journal of Agricultural And Food Chemistry 51: 5137-5143.
Ezedinma CI, Nkan NM (2008) The effect of quality on gari prices in Nigeria: A hedonic analysis. Journal Of Food Agriculture & Environment 6(1): 18-23.
Faber M, Phungula AS, Venter SL, Dhansay MA, Spinnler Benadé AJ (2002) Home gardens focusing on the production of yellow and dark-green leafy vegetables increase the serum retinol concentrations of 2-5-y-old children in South Africa. The American journal of clinical nutrition 76: 1048-1054.
Fleiss JL, Levin B, Paik MC (2003) Statistical Methods for Rates and Proportions 3rd ed. New York: Wiley.
Gibson RSS, Hotz C (2001) Dietary diversification/modification strategies to enhance micronutrient content and bioavailability of diets in developing countries. The British journal of nutrition 85 Suppl 2: S159-166.
Gonzalez C, Johnson N, Qaim M (2009) Consumer Acceptance of Second-Generation GM Foods: The Case of Biofortified Cassava in the North-east of Brazil. Journal Of Agricultural Economics 60(3): 604-624.
Gunaratna NS, De Groote H, Nestel P, Pixley KV, McCabe GP (2010) A Meta-analysis of Community-Based Studies on Quality Protein Maize. Food Policy: 202-210.
Haas DJ, Beard JL, Murray-Kolb LE, del Mundo AM, Felix A, Gregorio GB (2005) Iron-biofortified rice improves the iron stores of nonanemic Filipino women. The Journal of nutrition 135: 2823-2830.
Haddad L (2000) A conceptual framework for assessing agriculture-nutrition linkages. Food and Nutrition Bulletin 21: 367-373.
Hotz C, McClafferty B (2007) From harvest to health: challenges for developing biofortified staple foods and determining their impact on micronutrient status. Food And Nutrition Bulletin 28: S271-279.
IEG (2010) What Can We Learn from Nutrition Impact Evaluations: Lessons from a Review of Interventions to Reduce Child Malnurition in Developing Countries. Washington DC: World Bank.
Kawarazuka N (2010) The contribution of fish intake, aquaculture, and small-scale fisheries to improving food and nutrition security: A literature review. Penang, Malaysia: The World Fish Center.
Khush GS (2002) The promise of biotechnology in addressing current nutritional problems in developing countries. Food And Nutrition Bulletin 23: 354-357.
Kidala D, Greiner T, Gebre-Medhin M (2000) Five-year follow-up of a food-based vitamin A interventionin Tanzania. Public Health Nutrition 3: 425-431.

King JC (2002) Evaluating the impact of plant biofortification on human nutrition. The Journal of nutrition 132: 511S-513S.
Leroy JL, Frongillo EA (2007) Can Interventions to Promote Animal Production Ameliorate Undernutrition? The Journal of Nutrition 137: 2311-2316.
Li S, Nugroho A, Rocheford T, White WS (2010) Vitamin A equivalence of the {beta}-carotene in {beta}-carotene-biofortified maize porridge consumed by women. The American journal of clinical nutrition 92(5):1105-12.
Lönnerdal B (2007) The importance and bioavailability of phytoferritin-bound iron in cereals and legume foods. International journal for vitamin and nutrition research Internationale Zeitschrift für Vitamin- und Ernährungsforschung Journal international de vitaminologie et de nutrition 77: 152-157.
Low JW, Arimond M, Osman N, Cunguara B, Zano F, Tschirley D (2006) A Food-Based Approach Introducing Orange-Fleshed Sweet Potatoes Increased Vitamin A Intake and Serum Retinol Concentrations in Young Children in Rural Mozambique. Journal of Nutrition 137:1320-1327.
Low JW, Arimond M, Osman N, Cunguara B, T. ZF (2007) Ensuring the supply of and creating demand for a biofortified crop with a visible trait: Lessons learned from the introduction of orange-fleshed sweet potato in drought-prone areas of Mozambique. Food And Nutrition Bulletin 28: S258-S270.
Mason JB, Lofti M, Dalmiya N, Sethuraman K, Deitcher M (2001) The Micronutrient Report: Current Progress and Trends in the Control of Vitamin A, Iodine, and Iron Deficiencies. Ottawa: The Micronutrients Initiative.
Meenakshi JV, Banerji A, Manyong VM, Tomlins K, Hamukwala P, Rodah Z, Mungoma C (2010a) Consumer Acceptance of Provitamin A Orange Maize in Rural Zambia. HarvestPlus Working Paper.
Meenakshi JV, Johnson JV, Johnson NL, Manyong VM, DeGroote H, Javelosa J, Yanggen DR, Naher F, Gonzalez C, Garcia J, Meng E (2010b) How Cost-Effective is Biofortification in Combating Micronutrient Malnutrition? An Ex ante Assessment. World Development 38IS 1: 64-75.
Micronutrient Initiative (2004) Vitamin & mineral deficiency: a global progress report. Ottawa: The Micronutrient Initiative.
Micronutrient Initiative. (2009) Global Report 2009. Investing in the future: A united call to action on vitamin and mineral deficiencies. ISBN: 987-1-894217-31-6
Naico Ata Lusk JL (2010) The Value of a Nutritionally Enhanced Staple Crop: Results from a Choice Experiment Conducted with Orange-fleshed Sweet Potatoes in Mozambique. Journal Of African Economies 19: 536-558.
Olney DK, Talukder A, Iannotti LL, Ruel MT, Quinn V (2009) Assessing impact and impact pathways of a homestead food production program on household and child nutrition in Cambodia. Food and Nutrition Bulletin 30: 355-369.

Pelletier DL (2002) Toward a Common Understanding of Malnutrition. Assessing the Contributions of the UNICEF Framework. Background Papers World Bank/UNICEF Nutrition Assessment.
Poletti S, Gruissem W, Sautter C (2004) The nutritional fortification of cereals. Current Opinion in Biotechnology 15 (2): 162-165.
Qaim M, Stein AJ (2009) Biofortification of Staple Crops How well does it work and what does it cost? Ernahrungs-Umschau 56 (5): 274-280.
Quaim M, Stein AJ, Meenakshi JV (2007) Economics of biofortification. Agricultural Economics 37: 119-133.
Reutlinger S, Pellekaan H (1986) Poverty and Hunger: Issues and Options for Food Security in Developing Countries. Washington DC: The World Bank.
Rosado JL, Hambidge KM, Miller L, Garcia O, Westcott J, Gonzalez K, Conde J, Hotz C, Pfeiffer W, Ortiz-Monasterio I, Krebs NF (2009) The Quantity of Zinc Absorbed from Wheat in Adult Women Is Enhanced by Biofortification. Journal Of Nutrition 139IS 10: 1920-1925.
Ruel MT, Bouis HE (1998) Plant breeding: a long-term strategy for the control of zinc deficiency in vulnerable populations. American Journal Of Clinical Nutrition 68(suppl): 488S-494S.
Ruel MT (2001a) Can Food-Based Strategies Help Reduce Vitamin A and Iron Deficiencies? A Review of Recent Evidence. Washington DC: IFPRI.
Ruel MT (2001b) Can Food-Based Strategies Help Reduce Vitamin A and Iron Deficiencies? A Review of Recent Evidence. Food Policy Review 5.
Sedlmeier P, Gigerenzer G (1989) Do Studies of Statistical Power Have an Effect on the Power of Studies. Psychological Bulletin 105: 309-316.
Sen A (1981) Poverty and Famines. Oxford: Clarendon Press.
Smith L. C., Haddad L (2000) Overcoming Child Malnutrition in Developing Countries: Past Achievements and Future Choices. Washington DC: IFPRI.
Stevens R, Winter-Nelson A (2008) Consumer acceptance of provitamin A-biofortified maize in Maputo, Mozambique. Food Policy 33 (4): 341-351.
Strauss J, Thomas D (1995) Human resources: empirical modeling of household and family decisions. In: Behrman J, Srinivasan TN (eds) Handbook of Development Economics 3a. Amsterdam: Elsevier, pages 1883-2023.
Svedberg P (2000) Poverty and Undernutrition. New York: Oxford University Press.
Talukder A, Haselow NJ, Osei AK, Villate E, Reario D, Kroeun H, SokHoing L, Uddin A, Dhungel S, Quinn V (2010) Homestead food production model contributes to improved household food security and nutrition status of young children and women in poor populations - lessons learned from scaling-up programs in Asia (Bangladesh, Cambodia, Nepal and Philippines). Field Actions Science Report.
Tang G, Qin J, Dolnikowski GG, Russell RM, Grusak MA (2009) Golden Rice is an effective source of vitamin A. The American journal of clinical nutrition 89: 1776-1783.
van Jaarsveld PJ, Faber M, Tanumihardjo SA, Nestel P, Lombard CJ, Spinnler Benadé AJ (2005) ß-Carotene–rich orange-fleshed sweet potato improves

the vitamin A status of primary school children assessed with the modified-relative-dose-response test. American Journal of Clinical Nutrition 81 (5):1080-7..

Varian HR (1992) Microeconomic Analysis. New York: W.W. Norton and Company.

Vijayaraghavan K, Nayak MU, Bamji MS, Ramana GNV, Reddy V (1997) Home gardening for combating vitamin A deficiency. Food and Nutrition Bulletin 18: 337-343.

Welch MR, Graham RD (2004) Breeding for micronutrients in staple food crops from a human nutrition perspective. Journal of Experimental Botany 55: 353-364.

WHO Expert Committee (1995) Physical Status: The Use and Interpretation of Anthropometry. WHO Technical Report Series No 854, Geneva.

World Bank (2007) From Agriculture to Nutrition: Pathways, Synergies and Outcomes. Washington DC: The World Bank Agriculture and Rural Development Department.

Appendix 1.1: Authorship of this report

Authors
Edoardo Masset, Institute of Development Studies
Lawrence Haddad, Institute of Development Studies
Alex Cornelius, Institute of Development Studies
Jairo Isaza-Castro, University of Sussex

Contact details
Edoardo Masset
Research Fellow
Institute of Development Studies (IDS) at the University of Sussex
Brighton BN1 9RE UK
+44(0)1273 915807
e.masset@ids.ac.uk

Review group
International Initiative for Impact evaluation (3ie) and Department for International Development (DFID)

Advisory group
The Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre)

Conflict of interest
The authors declare they have no conflict of interest related to the publication of this systematic review.

Acknowledgements
This review was financially supported by DFID. Thanks to Imran Choudury (DFID), Max Gasteen (DFID) and Todd Benson, International Food Policy Research Institute (IFPRI) for their guidance in the process of designing the protocol. Thanks to Jeff Brunton (IOE) for providing and supporting the use of the EPPI-Reviewer software. Thanks to Jeremy Lynd (IDS) and Christophe Bené (IDS) for their help in refining the search of the pastoralist and fisheries literatures. Finally, thanks to Anna Taylor (DFID), Tim Wheeler (DFID), Hugh Waddington (3ie) and particularly Todd Benson (IFPRI) for their comments on an earlier draft. The views and the conclusions expressed in this review are those of the authors and do not necessarily reflect views of DFID or of the aforementioned people.

Appendix 1.2: Programme theory of agricultural interventions

Undernutrition is a complex phenomenon determined by multiple causes. The popular UNICEF conceptual framework originally distinguished immediate causes of undernutrition (food security, health and caring practices), and underlying basic causes of undernutrition (the existing economic, social and political structure) (Pelletier 2002). This framework has been further elaborated over the years to include immediate, underlying and basic determinants which are linked via a series of feed-back mechanisms (Smith L. C. and Haddad 2000).

Within this broad framework, agricultural interventions can impact on nutrition by increasing household food security. Following the framework developed by Haddad (2000), technological innovations in agriculture may improve food security through a generic increase in income or through a specific increase in the consumption of food and particular types of food. Table A1.1 illustrates the pathways leading from technological innovations to a reduction in malnutrition including a pathway through income stabilisation that was not included in the original Haddad's framework (2000).

The effect of an income increase on nutritional status is well known and normally analysed through the estimation of elasticities: the proportionate change in calories consumption for a unit change in income. The increase in consumption of calories following an increase in income is often found to be small in size (Deaton 1997), because households tend to spend their additional income in tastier and less nutritious food. Further, higher caloric intake does not immediately translate into improved nutritional status, as food intake is only one of the main determinants of nutritional status together with caring practices and an healthy environment (Smith L. C. and Haddad 2000). Finally, additional caloric intake can be spent in additional activity for work and play, thus not resulting in changes in nutritional status (Svedberg 2000).

Impact of agricultural interventions on the composition of the diet and the nutritional content of food has been less investigated. This effect follows from the production and consumption of nutritious foods obtained through the undertaking, for example, of dairy production or vegetable home gardens. This effect is achieved through a change in relative prices (making production of specific food less costly), a modification of the nutritional contents of the current diet (via bio-fortification), or by a change in consumer preferences (for example home gardens interventions are often accompanied by nutrition education classes illustrating the virtues of a balanced food diet). The addition of nutritious food items to an existing diet should have an immediate and direct effect on the intake of calories and micronutrients. This in turn should positively affect nutritional status, though the effect is moderated by the impact of other determinants of malnutrition and by the balance of energy intake and outtake of the body.

Table A1.1 Programme theory of the impact of agricultural interventions on nutrition

Effect of intervention	**Effect on nutrition**
Generic effects: Increase in production and/or productivity	
Increase in real income via increase in the quantities produced and sold (for sellers) or via a reduction in retail prices (for buyers)	The change in income improves nutrition through the calorie income elasticity
Higher agricultural productivity reduces household labour assigned to production	Nutrition improves because more time is available for child care in the family
Higher women's income or women's assets (for example land and livestock) increase women's decisional power within the household	Spending decisions made by women are more favourably oriented to the consumption of food and health care
Specific effects: increase in the production of food and particular food items	
Increase in food production or in the production of specific food commodities	Food consumption or the consumption of calories and micronutrients increases
Changes in the processing and storage of food	Improved technology for processing and storing food preserve the nutritious content of aliments
Genetic modification of crops containing higher quantities of nutrients	Food produced for own consumption becomes more nutritious
Stabilisation effects: income and food production are more stable over the seasons and over the years	
Reduction in the variability of income or food production over the seasons and over the years	Negative impact of seasonal and other short term shocks on child growth is averted

Note that this programme theory abstracts from behavioural changes induced by the programme which may affect programme impact in two ways. First, the increase in the nutritional contents of food consumed through dietary improvements may induce households to reduce overall food expenditure. Individuals may have a target in terms of energy intake in such a way that, for example, an increase in the consumption of a protein rich food (such as fish) might replace the previous consumption of another protein rich food (such as meat) in such a way that the net effect on nutrient intake is uncertain.

Second, food can be differently allocated to household members. For example, children may obtain a large portion of the family improved diet thus resulting in larger than expected improvement in nutritional status. Conversely, children or female household members may be denied food due to prevailing cultural norms. Sharing rules of income and food within the household are largely unknown and very context-specific.

This programme theory also abstracts from general equilibrium effects of the interventions. The introduction of new technologies on a massive scale can have powerful effects on prices, thus affecting both producers and consumers of agricultural commodities that were not originally targeted by the interventions. If price effects are sufficiently strong the comparison between programme participants and non-participants can be contaminated by the benefit accruing to the latter.

Appendix 2: Search methodology

Bio-fortification

1. (bio-fortif* OR bio-fortif* OR “conventional selective breeding” OR “golden rice” OR “harvest plus” OR “plant breeding”) AND (nutrition* OR “diet* diversification” OR “diet* change” OR expenditure OR income OR anthropom* OR micronutrient* OR particip* OR livelihood*)

2. (bio-fortif* OR bio-fortif* OR conventional-selective-breeding OR golden-rice OR harvest-plus OR plant-breeding) AND (nutrition* OR diet-diversification OR expenditure OR income OR anthropom* OR micronutrient* OR particip* OR livelihood*)

Home Gardens

1. (“home garden” OR “home gardening” OR “home gardens” OR “home-garden” OR “home-gardens” OR “home-gardening” OR “vegetable garden” OR “homestead food production” OR “household garden” OR “household gardening” OR “household gardens” OR “garden based nutrition program” OR “kitchen garden” OR “kitchen gardens” OR “kitchen gardening” OR “project garden” OR “project gardens” OR “project garden” OR “homestead plot” OR “food garden” OR “food gardens” OR “food gardening” OR HFP OR HFPP OR home based food OR home-based food OR home based garden OR home-based garden) AND (nutrition* OR diet* OR expenditure OR income OR malnutrition* OR anthropom* OR micronutrient)

2. (home garden OR home gardening OR home gardens OR home-garden OR home-gardens OR home-gardening OR vegetable garden OR homestead food production OR household garden OR household gardening OR household gardens OR garden based nutrition program OR kitchen garden OR kitchen gardens OR kitchen gardening OR project garden OR project gardens OR project garden OR homestead plot OR food garden OR food gardens OR food gardening OR HFP OR HFPP OR home based food OR home-based food OR home based garden OR home-based garden) AND (nutrition* OR diet* OR expenditure OR income OR malnutrition* OR anthropom* OR micronutrient)

3. home garden* OR vegetable garden* AND nutrition

Dairy Development

1. (“dairy development” OR “dairy farming” OR “dairy program” OR “dairy programme” OR “smallholder dairy development” OR “dairy development” OR “dairy cooperative” OR “dairy extension”) AND (nutrition* OR “diet diversification” OR income OR malnutrition* OR anthropom* OR micronutrient)

2. (dairy development OR dairy farming OR dairy program OR dairy programme OR smallholder dairy development OR dairy development OR dairy cooperative OR dairy extension) AND (nutrition* OR diet* OR expenditure OR income OR malnutrition* OR anthropom* OR micronutrient)

3. (dairy-development OR dairy-farming OR dairy-program OR dairy-programme OR smallholder-dairy-development OR dairy-cooperative OR dairy-extension) AND (nutrition* OR diet diversification OR income OR malnutrition* OR anthropom* OR micronutrient)

Fishponds and Aquaculture

1. (fish-pond OR fishpond OR fisheries OR fishery OR aquaculture OR aqua-culture OR aquafarm* OR aqua-farm* OR fishfarm* OR fish-farm OR “capture fisheries” OR “pond polyculture” OR mariculture OR mari-culture OR “small-scale fisheries” OR “small-scale fishery” OR “small fish species” OR “fish consumption”) AND (nutrition* OR “diet diversification” OR diversi* OR expenditure OR income OR anthropom* OR particip*)

2. (fish-pond OR fishpond OR fisheries OR fishery OR aquaculture OR aqua-culture OR aquafarm* OR aqua-farm* OR fishfarm* OR fish-farm OR fishfarm OR capture fisheries OR pond polyculture OR mariculture OR mari-culture OR small-scale fisheries OR small-scale fishery OR small fish species OR fish consumption) AND (nutrition* OR diet diversification OR expenditure OR income OR anthropom* OR micronutrient* OR particip*)

3. (fish-pond OR fishpond OR fisheries OR fishery OR aquaculture OR aqua-culture OR aquafarm* OR aqua-farm* OR fishfarm* OR fish-farm OR capture-fisheries OR pond-polyculture OR mariculture OR mari-culture OR small-scale-fisheries OR small-scale-fishery OR small-fish-species OR fish-consumption) AND (nutrition* OR diet-diversification OR diversif* OR expenditure OR income OR anthropom* OR particip*)

Animal Husbandry

1. 1 ("animal source foods" OR "animal-source foods" OR "animal production" OR "livestock promotion" OR "livestock production" OR "poultry promotion" OR "poultry production" OR "chicken promotion" OR "chicken production" OR pastoralism OR "pastoral farming" OR pastoralist OR "agro-pastoralist" OR "cattle production" OR "camel production" OR "goat production" OR "sheep production" OR "small ruminates") AND (nutrition* OR diet diversification OR expenditure OR income OR anthropom* OR micronutrient* OR particip*)

2. 2 (animal husbandry OR animal source foods OR animal-source foods OR animal production OR livestock promotion OR livestock production OR poultry promotion OR poultry production OR chicken promotion OR chicken production OR pastoralism OR pastoral farming OR pastoralist OR agro-pastoralist OR cattle production OR camel production OR goat production OR sheep production OR small ruminates) AND (nutrition* OR diet diversification OR expenditure OR income OR anthropom* OR micronutrient* OR particip*)

3. 3 (animal-husbandry OR animal-source-foods OR animal-production OR livestock-promotion OR livestock-production OR poultry-promotion OR poultry-production OR chicken-promotion OR chicken-production OR pastoralism OR pastoral-farming OR pastoralist OR agro-pastoralist OR cattle-production OR camel-production OR goat-production OR sheep-production OR small-ruminates) AND (nutrition* OR diet diversification OR expenditure OR income OR anthropom* OR micronutrient* OR particip*)

Table A2.1 Key searches employed

Database	Bio-fortification	Home gardens	Fisheries	Dairy development	Animal husbandry
Econlit	2	1	1	1	2
Web of Science	2	2	3	3	3
Agris	2	3	Aquaculture AND Nutrition	2	Pastoral* AND nutrition
PubMed	1	1	1	1	1
World Bank	Plant-breeding AND nutrition	3	Aquaculture AND Nutrition	1	pastoral*OR animal husbandry AND nutrition
IFPRI	Bio-fortif OR bio-fortif AND nutrition	3	Aquaculture	Dairy	Pastoral*
Eldis	bio-fortification	3	Aquaculture	Dairy development	Pastoral* AND nutrition
IBSS	2	home garden* OR vegetable garden*	Aquaculture AND Nutrition	dairy and nutrition	Pastoral* AND nutrition
Ideas	bio-fortification and nutrition	1	1	1	2
Jolis	bio-fortification	3	Aquaculture AND Nutrition	1	Pastoral* AND nutrition

Table A2.2 Studies found by the key searches (including duplicates)

Database	Bio-fortification	Home gardens	Fisheries	Dairy	Animal husbandry	TOTAL
Econlit	40	12	457	20	109	638
Web of Science	526	981	1049	993	1007	4556
Agris	845	70	26	198	141	1280
PubMed	281	658	600	1125	345	3009
World Bank	150	5	313	36	148	652
IFPRI	23	2	30	56	12	123
Eldis	4	6	81	49	30	170
IBSS	6	6	14	32	0	58
Ideas	3	10	246	10	118	387
Jolis	0	1	5	4	2	12
TOTAL	1878	1751	2821	2523	1912	

Appendix 3: Bio-fortification interventions

Bio-fortification is the use of traditional crop breeding practices or modern biotechnology to produce micronutrient-dense staple crops to reduce micronutrient deficiencies (Micronutrient Initiative 2009). Humans require 49 micronutrients to meet their metabolic needs and required intakes have been established (Welch and Graham 2004). The international research effort on bio-fortification has focused on three micronutrients in particular: iron, zinc and vitamin A.

It is estimated that 45% of children under the age of five in developing countries are iron deficient, 40% of children are lacking appropriate amount of vitamin A, and that a third of the world's population live in countries at high risk of zinc deficiency (Micronutrient Initiative 2009). Mason et al. (2001) use WHO data from several developing countries to show that prevalence rates deficiency of vitamin A decreased between 1980 and 2000, while no reduction of anaemia prevalence was observed between 1970 and 2000. There are currently no agreed indicators for the measurement of zinc deficiency and prevalence rates are not available.

The evidence reported by the Micronutrient Initiative (2009) suggests that a reduction in micronutrient deficiencies would be beneficial in several ways. Provision of vitamin A reduces mortality of children under five, prevents blindness in childhood and reduces chances of xerophtalmia (night blindness). The provision of iron during pregnancy lowers maternal mortality due to haemorrhage, the risk of premature birth and low birth weight. In addition, iron supplementation helps mental development in children. Zinc can reduce the incidence of diarrhoea, the incidence of lower respiratory tract infections, and the prevalence of stunting.

The Micronutrient Initiative identified three factors determining micronutrient deficiencies in developing countries (2009). First, poor people obtain the largest share of energy intake from cereals but a varied diet including meat, eggs, fish, milk, legumes, fruits and vegetables is the basis for micronutrients adequacy. Second, infections and intestinal worms reduce micronutrients in the body, and micronutrients deficiencies and infections interact in a complex way leading to a vicious circle of malnutrition and infections (Bhaskaram 2002). Third, low prevalence of exclusive breastfeeding in the first six months of life increases deficiencies because breast milk is an important source of micronutrients.

There are several ways to address micronutrient deficiencies (Micronutrient Initiative 2009). Supplementation of iron, zinc and vitamin A provides vulnerable groups (children and women of child-bearing age) with micronutrients in the form of low-cost tablets, capsules or syrups. Fortification of foods adds vitamins and minerals to food regularly consumed by the population. Nutrition education informs the public about the need to consume supplements or micronutrients-rich food. Livelihood promotion programmes promote production activities that enhance the diet's micronutrients content. Disease control programmes reduce the incidence of malaria, measles, diarrhoea and parasitic infections that impair micronutrient absorption.

Bio-fortification is yet another strategy to reduce micronutrient deficiencies. Research has shown that breeding for enhanced nutritional value does not result in lower yields and that bio-fortified crops can be very productive (Ruel and Bouis 1998). Bio-fortification is now considered a 'win-win' intervention (Bouis et al. 1999). First, increasing micronutrients store in seeds increases seedlings vigour in poor soils.

Seedlings of micronutrient-dense varieties have more and longer roots under micronutrient-poor soils, thus allowing seedlings to scavenge more for micronutrients which can increase yields. Second, seedlings from micronutrient-dense seeds are more disease resistant and more stress tolerant (Welch and Graham 2004).

There are two additional advantages of bio-fortification. First, bio-fortification can reach poor populations in remote areas that rely on consumption of staple foods and have no access to fortified food (Darnton-Hill et al. 2002). Second, bio-fortification is considered a cost-effective intervention (Meenakshi et al. 2010b), and was ranked fifth among cost-effective interventions to tackle malnutrition and hunger by the panel of experts of the Copenhagen consensus of 2008 (Copenaghen Consesus Center 2008).

Theory of change of bio-fortification programmes

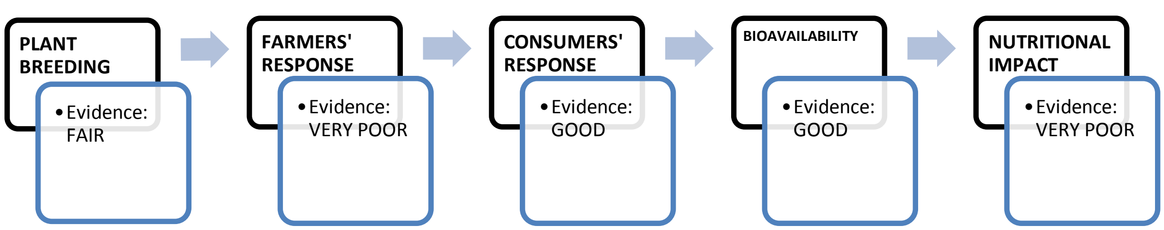
In order to build a programme theory of bio-fortification interventions, we reviewed work by Bouis et al. (1999), Ruel (2001b), King (2002), Quaim et al. (2007), Quaim et al. (2009), Hotz and McClafferty (2007), and Brooks (2010). Bio-fortification interventions will have an impact on nutritional status provided the following steps are met (see Figure A3.1): a) plants are developed that retain a large amount of micronutrients in their edible parts; b) farmers adopt bio-fortification technologies; c) consumers (often the producers themselves) buy and consume staples from bio-fortified crops; d) the micronutrient content of bio-fortified foods are absorbed by the human body (bioavailability); e) finally consumption of bio-fortified food results in improved nutritional status.

Figure A3.1 Programme theory bio-fortification interventions

Results of bio-fortification interventions

We reviewed the available evidence on the success of bio-fortification interventions along each step of the causal chain. Figure A3.2 summarises the results of this review.

Figure A3.2 Evidence on impact of bio-fortification interventions along the programme theory



Successful plant breeding

The first step of a bio-fortification intervention is the production, either by traditional selective breeding or by bioengineering, of staple food varieties that are micronutrient-dense (Bouis et al. 1999). However, the nutritional fortification of cereals is still in its infancy, and little is known about the physiological and biochemical mechanisms that control micronutrient accumulation in plants (Poletti et al. 2004). There are a number of difficulties in obtaining plants that accumulate more micronutrients. Current knowledge of all these processes is limited and little is known about their variability across different soil characteristics and climatic conditions (Welch and Graham 2004). Nonetheless, the production of plants accumulating desired contents of micronutrients is likely to be within reach (Hotz and McClafferty 2007). For example, promising results have been obtained in the production of rice with beta-carotene (Khush 2002), and maize with vitamin A (Aluru et al. 2008).

Farmers' response

Farmers' adoption of bio-fortified varieties critically depends on yields and farmers' expectations about yields (Bouis et al. 1999, Ruel and Bouis 1998). Farmers need to know that bio-fortified crops are at least as productive (both in terms of yields and in terms of their economic value) as traditional ones. We could only find one study addressing this issue (Low et al. 2007). A research project in a drought-prone area of rural Mozambique promoted the production of an orange-fleshed variety of sweet potatoes rich in vitamin A. The study found that only 38% of households manifested interest in the project in the intervention areas. Of the households that joined the intervention, 90% produced the potato and 30% sold the produce on the market. The average size of potato farms in project areas was ten times the size of that in control areas.

Consumers' response

Bio-fortification may affect flavour, appearance and other characteristics of food which could affect consumers' acceptance. In general, consumers' acceptance will depend on costs, cultural preferences and organoleptic properties (taste, sight, smell and touch) (Ruel and Bouis 1998). Other factors such as time required for food preparation, shelf-life and wastage may also matter (King 2002). Several recent studies have assessed, through experiments or econometric demand analysis, consumers' willingness to pay for bio-fortified varieties of staple foods. In general, consumers are willing to pay a premium for food with higher micronutrient contents. These results were found for bio-fortified orange maize in rural Zambia (Meenakshi et al. 2010a), provitamin A GM cassava in north-east Brazil (Gonzalez et al. 2009), orange-fleshed sweet potato in rural Mozambique (Naico Ata Lusk 2010), golden rice (Depositario et al. 2009), vitamin A bio-fortified gari (a staple processed from cassava) in rural Nigeria (Ezedinma and Nkan 2008), provitamin A-bio-fortified maize in urban Mozambique (Stevens and Winter-Nelson 2008), and orange-fleshed sweet potato in rural Uganda (Chowdhury et al. 2009). Only one study (De Groote and Kimenju 2008) in urban Kenya (Nairobi) found that a discount was needed for consumers to accept provitamin A bio-fortified yellow maize.

Bioavailability

Bioavailability is the fraction of ingested nutrient that is utilised by the human body for normal physiological functions or storage and is affected by three main factors (King 2002, Ruel and Bouis 1998, Welch and Graham 2004). First, the characteristics of the individual determine the rate of absorption. For example, individuals with poor iron status absorb iron more rapidly. Second, other components of the meal affect the rate of absorption (Gibson and Hotz 2001) because most foods contain both promoter and inhibitor substances. Third, food preparation may alter bioavailability. For example, fermentation of cereals before cooking improves bioavailability of iron and zinc. Because of high cost of running bioavailabilty tests on humans, studies often perform availability tests in vitro (see for example Lonnerdal (2007)) or on rats. However, we only report here the results of tests performed on humans.

An experiment on 12 women fed with two varieties of beans (one common and the other with high zinc and iron concentration) concluded that high-zinc bean genotypes may improve zinc status, but high-iron genotypes have little effect on iron status (Donangelo et al. 2003). Bioavailability of vitamin A from beta-carotene bio-fortified rice was found to be high in a sample of five healthy adults of the Boston area (Tang et al. 2009). An experiment involving six women consuming standard porridge and porridge prepared with bio-fortified wheat found good bioavailability of beta-carotene (Li et al. 2010). An experiment among adult women used common tortillas and tortillas prepared from zinc bio-fortified wheat and found a potential increases in absorption of zinc (Rosado et al. 2009). An experiment on 192 adult women of metro Manila compared the effects of standard rice and bio-fortified rice and found increases in iron stores (Haas et al. 2005). Another experiment tested the efficacy of orange-fleshed potato among 180 children in South Africa and found a good level of absorption of vitamin A (van Jaarsveld et al. 2005).

<u>Nutritional status</u>

We could only find two studies that assessed the impact of bio-fortification interventions on nutritional status. Low et al. (2007) conducted a study of an intervention promoting orange-fleshed potato in three districts of a drought-prone area of Mozambique. Children from project and control areas were followed over a two year period in order to test differences in vitamin A intake. After two years children in intervention areas showed vitamin A intakes which were nearly eight times those of control children. While at baseline anthropometric Z-scores did not differ between the two populations, after two years wasting and underweight prevalence were lower in project areas, though no difference was found in prevalence of stunting. Gunaratna et al.(2010) conducted a meta-analysis of nine unpublished results of randomised trials in Africa and Latin America. The interventions consisted of the introduction of modern varieties of quality protein maize. Though bio-fortification is normally intended as a process increasing micronutrients level in the body, it can also include plant breeding, enhancing the bioavailability of proteins. The meta-analysis concluded that consumption of quality protein maize versus conventional maize resulted on average in an increase in the rates of growth of children by 12% (weight) and by 9% (height).

Conclusions on impact of bio-fortification interventions

- There is insufficient evidence to assess the programme theory of bio-fortification interventions and only few of the causal links have been successfully explored.
- Little is known about farmers’ acceptance of bio-fortified crops and no evidence is available on impact on yields and farm profits.
- Only two studies report evidence on the nutritional impact of bio-fortification programmes and the results are positive.

- Impact evaluations of bio-fortification programmes under different climatic and socio-economic conditions are needed in order to assess their acceptance by farmers and their effectiveness.

Appendix 4: List of included and excluded studies

Studies excluded after the validity assessment are marked with an asterisk (*)

*Ahmed MM, Jabbar M, Ehui S. (2000) Household-level economic and nutritional impacts of market-oriented dairy production in the Ethiopian highlands. Food and nutrition bulletin 21 (4):460-465.

Aiga H, Matsuoka S, Kuroiwa C, Yamamoto S. (2009) Malnutrition among children in rural Malawian fish-farming households. Transactions of the Royal Society of Tropical Medicine and Hygiene 103 (8):827-33.

Attig GA, Smitasiri S, Ittikom K, Dhanamitta S. (1993) Promoting home gardening to control vitamin A deficiency in northeasternThailand. Food, Nutrition and Agriculture 7:18-25.

*Ayele Z, Peacock C. (2003) Improving access to and consumption of animal source foods in rural households: the experiences of a women-focused goat development program in the highlands of Ethiopia. The Journal of Nutrition 133 (11 Suppl 2):3981S-3986S.

*Begum JM. (1994) The impact of dairy development on protein and calorie intake of pre-school children. Indian Journal of Medical Sciences 48 (3):61-4.

*Bloem MW, Huq N, Gorstein J, Burger S, Kahn T, Islam N, Baker S, Davidson F. (1996) Production of fruits and vegetables at the homestead is an important source of vitamin A among women in rural Bangladesh. European Journal of Clinical Nutrition 50 Suppl. 3:S62-S67.

Bushamuka VN, de Pee S, Talukder A, Kiess L, Panagides D, Taher A, Bloem M. (2005) Impact of a homestead gardening program on household food security and empowerment of women in Bangladesh. Food and Nutrition Bulletin 26 (1):17-25.

*Chakravarty I. (2000) Food-based strategies to control vitamin A deficiency. Food and Nutrition Bulletin 21 (1):135-143.

de Pee S, Bloem MW, Gorstein J, Sari M, Satoto YR, Shrimpton R, Muhilal. (1998) Reappraisal of the role of vegetables in the vitamin A status of mothers in Central Java, Indonesia. The American Journal of Clinical Nutrition 68 (5):1068-74.

Faber M, Phungula MSA, Venter SL, Dhansay MA, Benadé A, Spinnler J. (2002) Home gardens focusing on the production of yellow and dark-green leafy vegetables increase the serum retinol concentrations of 2-5-y-old children in South Africa. The American Journal of Clinical nutrition 76 (5):1048-54.

Greiner T, Mitra SN. (1995) Evaluation of the impact of a food-based approach to solving vitamin A deficiency in Bangladesh. Food Nutrition Bulletin 16 (3):193-205.

Gunaratna NS, De Groote H, Nestel P, Pixley KV, McCabe GP. (2010) A Meta-analysis of Community-Based Studies on Quality Protein Maize. Food Policy (3):202-10.

Hoorweg J, Leegwater PH, Veerman W. (2000) Nutrition in Agricultural Development: Intensive Dairy Farming by Rural Smallholders. Ecology of Food and Nutrition 39:395-416.
Jones KM, Specio SE, Shrestha PK, Brown KH, Allen LH. (2005) Nutrition knowledge and practices, and consumption of vitamin A rich plants by rural Nepali participants and nonparticipants in a kitchen-garden program. Food and Nutrition Bulletin 26 (2):198-208.
Kidala D, Greiner T, Gebre-Medhin M. (2000) Five-year follow-up of a food-based vitamin A intervention in Tanzania. Public Health Nutrition 3 (4):425-431.
Laurie SM, Faber M. (2008) Integrated community-based growth monitoring and vegetable gardens focusing on crops rich in beta-carotene: Project evaluation in a rural community in the Eastern Cape, South Africa. Journal of the Science of Food and Agriculture 88 (12):2093-2101.
Low JW, Arimond M, Osman N, Cunguara B, Zano F, Tschirley D. (2007) Ensuring the supply of and creating demand for a bio-fortified crop with a visible trait: Lessons learned from the introduction of orange-fleshed sweet potato in drought-prone areas of Mozambique. Food And Nutrition Bulletin 28 (2 Suppl.): S258-S270.
Makhotla L, Hendriks S. (2004) Do home gardens improve the nutrition of rural pre-schoolers in Lesotho? Development Southern Africa 21 (3): 575-581.
Marsh R. (1998) Building on traditional gardening to improve household food security. Food, Nutrition and Agriculture 22: 4-14.
Murshed-e-Jahan K, Ahmed M, Belton B. (2010) The impacts of aquaculture development on food security: lessons from Bangladesh. Aquaculture Research 41 (4):481-495.
Nielsen H, Roos N, Thilsted SH. (2003) The impact of semi-scavenging poultry production on the consumption of animal source foods by women and girls in Bangladesh. The Journal of Nutrition 133 (11 Suppl 2):4027S-4030S.
Olney DK, Talukder A, Iannotti LL, Ruel MT, Quinn V. (2009) Assessing impact and impact pathways of a homestead food production program on household and child nutrition in Cambodia. Food and Nutrition Bulletin 30 (4):355-69.
Roos N, Islam MM, Thilsted SH. (2003) Small indigenous fish species in bangladesh: contribution to vitamin A, calcium and iron intakes. The Journal of Nutrition 133 (11 Suppl 2):4021S-4026S.
Schipani S, van der Haar F, Sinawat S, Maleevong K. (2002) Dietary intake and nutritional status of young children in families practicing mixed home gardening in northeast Thailand. Food and Nutrition Bulletin 23 (2):175-80.
Schmidt MI, Vorster HH. (1995) The effect of communal vegetable gardens on nutritional status. Development Southern Africa 12 (5):713-724.
Smitasiri S, Sa-ngobwarchar K, Kongpunya P, Subsuwan C, Banjong O, Chitchumroonechokchai C, Rusami-Sopaporn W, Veeravong S, Dhanamitta S. (1999) Sustaining behavioural change to enhance

micronutrient status through community- and women-based interventions in north-east Thailand: Vitamin A. Food and Nutrition Bulletin 20 (2):243-251.
*Taher A, Talukder A, Sarkar NR, Bushamuka VN, Hall A. (2002) Homestead gardening for combating vitamin A deficiency: The Helen Keller International, Bangladesh, experience. In Roos N, Bouis HE, Hassan N, Kabir KA. (Eds.) (2004) Alleviating Malnutrition Through Agriculture in Bangladesh Bio-fortification & Diversification as Sustainable Solutions, Proceedings of the Workshop on Alleviating Micronutrient Malnutrition Through Agriculture in Bangladesh: Biofortification And Diversification As Long-Term, Sustainable Solutions, Gazipur And Dhaka, Bangladesh April 22–24, 2002. pp 68-74.
Talukder A, Haselow NJ, Osei AK, Villate E, Reario D, Kroeun H, SokHoing L, Uddin A, Dhungel S, Quinn V. (2010) Homestead food production model contributes to improved household food security and nutrition status of young children and women in poor populations - lessons learned from scaling-up programs in Asia (Bangladesh, Cambodia, Nepal and Philippines). Field Actions Science Report.
*Thompson PM, Sultana P, Nuruzzaman Md, Firoz Khan AKM. (2000) Impacts of Aquaculture Extension on Pond Operators and the Rural Community. Dakha: International Center for Living Aquatic Resources Management.
Vijayaraghavan K, Uma Nayak M, Bamji MS, Ramana GNV, Reddy V. (1997) Home gardening for combating vitamin A deficiency. Food and Nutrition Bulletin 18:337-343.

Appendix 5: Characteristics of selected studies

Bio-fortification studies

	Study & country	Intervention	Nutritional impact	Micronutrients impact	Dietary impact	Income effect
1	Low et al. (2007 Mozambique	An integrated package of agriculture, nutrition education and market intervention focused on introduction and promotion of orange-fleshed sweet potato	Reduction in prevalence rates of children's wasting and underweight but not stunting	Vitamin A intake eight times higher among project children after the intervention compared to control children	Project children more likely to consume orange-fleshed sweet potato, dark-green leaves and papaya and show higher dietary diversity	Not discussed or estimated
2	Gunaratna et al. (2010), Ghana, Ethiopia, India, Mexico and Nicaragua	Farmers are given seeds to produce protein-fortified maize, and fortified food	An increase in the rate of growth of weight among children of families cropping fortified maize versus conventional maize			Income was estimated but the results of the follow up interviews are not reported

Home gardening studies

	Study & country	Intervention	Nutritional impact	Micronutrients impact	Dietary impact	Income effect
3	Attig et al. (1993), Thailand	A social marketing programme of vitamin A-rich food including nutrition education and home and school gardening	Not assessed	Larger improvement in vitamin A among pregnant and lactating mothers	Not assessed	Not assessed
4	Bushamuka et al. (2005), Bangladesh	A gardening and nutritional surveillance project encourages poor households to produce vegetables rich in vitamin A all year round	Not assessed	Not assessed	Consumption of vegetables and fruit was higher in the control group	Income difference are not assessed, but there is some discussion of how income from gardening is spent (on food, education and other expenditures

	Study & country	Intervention	Nutritional impact	Micronutrients impact	Dietary impact	Income effect
5	de Pee et al. (1998), Indonesia	A social marketing campaign encouraging production and consumption of vitamin A-rich food	Measured but not reported	Women with home gardens have higher serum retinol concentration and are more likely to have serum retinol concentration above the median	Women with a home garden had a higher vitamin A intake from plant foods and a lower vitamin A intake from animal foods than women without a home garden	Not assessed
6	Faber et al. (2002), South Africa	A home gardening programme promoting production of yellow and dark-green leafy vegetables alongside growth monitoring and promotion in a rural village	No differences in stunting, wasting or underweight found either at baseline or follow-up	Serum retinol concentration increased over time in the project village while it decreased in the control village	At follow-up children from the project village consumed yellow and green leafy vegetables more often than children in the control village	Not assessed
7	Greiner and Mitra (1995), Bangladesh	Home gardening and nutrition education	Not assessed	Not assessed	Considerable difference in the consumption of green vegetables, but no difference in the consumption of yellow fruit	Not assessed
8	Jones et al. (2005), Nepal	The programme supported the production of high-value crops through technical assistance, together with nutrition education and home gardens	Not assessed	Not assessed	Consumption of vegetables and fruit was significantly higher in the kitchen-garden group	Not assessed
9	Kidala et al. (2000), Tanzania	Home gardening and nutrition education	Not reported	Vitamin A intake was considerable higher in the control group (explained by helminths infestation)	Considerably higher consumption of vitamin A-rich food in the project area	Not assessed

	Study & country	Intervention	Nutritional impact	Micronutrients impact	Dietary impact	Income effect
10	Laurie and Faber (2008), South Africa	A home gardening intervention combined with nutrition education, and community based GMP	Not assessed	Not assessed	Higher consumption of vitamin A rich vegetables (carrot, orange potato, butternut and spinach) among project participants	Not assessed
11	Makhotla and Hendriks (2004), Lesotho	Home gardens programmes run by the government and NGOs nationwide	No difference in stunting, wasting and underweight between children from families with and without gardens	Not assessed	Not assessed	Not assessed
12	Marsh (1998), Bangladesh	Home gardens and nutrition education for landless or near landless women	A larger decrease in the prevalence of severe underweight in the control group	Not reported	A larger increase in the consumption of vegetables in the control group	The target group reported incomes slightly higher than those of the control group, but statistical tests are not performed
13	Olney et al. (2009), Cambodia	Women are given seeds and technical assistance to grow vegetables that are rich in vitamin A	No difference found in stunting, wasting and underweight of children under five	No difference found in haemoglobin (anaemia) of children and mothers	A larger fraction of households consumes vegetables rich in vitamin A in the intervention group	Not assessed, though it is observed that project households earn a larger share of income from gardens
14	Schipani et al. (2002), Thailand	Government promotion of establishment of mixed home gardens to produce fish, small animals, and vegetables	No differences were found in stunting, wasting or underweight	No difference found in the concentration of haemoglobin, ferritin and retinol among young children of the two groups	No differences in the dietary intake of calories, protein, vitamin A, iron, vitamin C and fats	Annual income was considerably higher among gardening families compared to non-gardening families
15	Shmidt and Vorster (1995), South Africa	NGOs operated a communal vegetable garden in an extremely poor village	No differences found in stunting, wasting and underweight	No differences found in concentrations of vitamin A, beta-carotene and vitamin E	No significant differences in the intake of vegetables between the two	Not assessed

	Study & country	Intervention	Nutritional impact	Micronutrients impact	Dietary impact	Income effect
					groups	
16	Smitasiri et al. (1999), Thailand	A social marketing programme of vitamin A-rich food including nutrition education and home and school gardening	Not assessed	Larger increase in vitamin A concentrations in the project group	Similar changes in intakes of vitamin A of children and mothers in project and control groups	Not assessed
17	Talukder et al. (2010), Bangladesh, Cambodia, Nepal and Philippines	Homestead food production programme coupled with nutrition education	Not reported	No change in anaemia prevalence among mothers. Reduction of anaemia among children in Bangladesh and Cambodia. Reduction in the prevalence of night-blindness.	Higher consumption of vegetables by children in the intervention groups	Total income is not assessed but income from gardens is higher in intervention villages and is mostly spent in purchasing other food
18	Vijayaraghavan et al. (1997), India (AP)	Seeds and seedlings of carotene rich food distributed to households with pre-school children.	Not assessed	No significant change in the prevalence of Bitot's spot	The frequency of consumption of carotene-rich foods increased by about 50% over that observed at baseline	Not assessed

Fisheries studies

	Study & country	Intervention	Nutritional impact	Micronutrients impact	Dietary impact	Income effect
19	Aiga et al. (2002), Malawi	Household fish farming	Reduction in prevalence of underweight and severe underweight, but not differences in stunting and wasting	Not assessed	Not assessed	Incomes of the two groups are not reported though it is found that undernutrition is correlated with the share of income obtained from farming
20	Murshed-e-Jahan et al. (2010), Bangladesh	Training support to farmers on aquaculture at the farm level by the diffusion of low-cost aquaculture technologies	Not assessed	Not assessed	Consumption of fish increased more by project farmers than by non-project farmers	Income increases for the participating farmers over time while it is stable for the non-participating households

21	Roos et al. (2003), Bangladesh	Integration of mola and other small indigenous fish in carp polyculture in small seasonal ponds among poor communities	Not assessed	Not reported	No differences in fish consumption between the fish-producing and non-fish producing households	Not assessed

Dairy development

	Study & country	**Intervention**	**Nutritional impact**	**Micronutrients impact**	**Dietary impact**	**Income effect**
22	Hoorweg et al. (2000), Kenya	Promotion of intensive dairy farming among rural smallholders	Lower prevalence of stunting, wasting and underweight in the participant group	Not assessed	Higher consumption of milk in the participant group	Income is found to be considerably higher among dairy farmers

Animal husbandry

	Study and country	Intervention	Nutritional impact	Micronutrients impact	Dietary impact	Income effect
23	Nielsen et al. (2003), Bangladesh	Promotion of semi-scavenging poultry production	Not assessed	Not assessed	No difference found in the consumption of chicken, eggs and other food items	Monthly household income was higher among adopting compared to non-adopting households, but no significance tests are reported

Appendix 6: Validity assessment

In the tables below, each study is scored as LOW, MEDIUM or HIGH along four dimensions (counterfactual analysis, power, intermediate outcomes and heterogeneity), where LOW means the methodological quality of the study is weak in that particular dimension, and HIGH is a high quality methodology.

Bio-fortification

	Study & country	Counterfactual analysis	Power	Intermediate outcomes	Heterogeneity
1	Gunaratna et al. (2010), Ghana, Ethiopia, India, Mexico and Nicaragua	This is a meta-analysis of RCTs only HIGH internal validity	There are nine studies included in the meta-analysis and the quality of the sampling is not clear. Some samples are very small MEDIUM	They are discussed but not addressed in the analysis. MEDIUM internal validity	No disaggregation is carried out LOW external validity
2	Low et al. (2007a and 2007b), Mozambique	The study is based on a longitudinal sample. However, participants households self-selected in intervention areas. Though control villages are matched on agroecological conditions, households are not. The paper claims addressing selection bias by running a child-level fixed effects model and household control variables MEDIUM	It is not clear how many villages were sampled. Power was calculated over serum retinol concentration to observe a 0.05 difference with alpha=0.005 and beta=0.95. Not clear that intracluster correlation was taken into account MEDIUM	Main factors are discussed. Knowledge, production and consumption are discussed and estimated HIGH	No disaggregation by wealth, education or other LOW

Home gardens

	Study & country	Counterfactual analysis	Power	Intermediate outcomes	Heterogeneity
3	Attig et al. (1993), Thailand	Methodology for the selection of the control group both at the village and household level is not reported LOW	Sampling methodology and power are not discussed LOW	Intermediate outcomes (knowledge, attitudes and practices) are discussed and assessed HIGH	No disaggregated analysis LOW
4	Bloem et al. (1996), Bangladesh	There is no control group. Effects of three different garden types are compared with regression and correlation analysis without controlling for selection in each of the treatments EXCLUDED	No power calculations are reported and criteria for the definition of the sample size are not reported LOW	No pathways of impact or additional determinants of vitamin A status are discussed LOW	No attempt to differentiate impact by any category LOW
5	Bushamuka et al. (2005), Bangladesh	The sample comprises participants, former participants and a control group. The control group is composed of households with similar socio-economic characteristics, though the methodology is not explained MEDIUM	No power calculations are reported and criteria for the definition of the sample size are not reported LOW	Several intermediate outcomes are discussed and analysed, including use of garden income, use of garden produce and changes in women decision-making power HIGH	No attempt to differentiate impact by any category LOW
6	Chakravarty (2000), India (West Bengal)	No control group, the impact before and after the intervention on participants is assessed EXCLUDED	No power calculations are reported and criteria for the definition of the sample size are not reported LOW	The programme theory is not fully discussed, but some intermediate indicators are analysed as household expenditure and production MEDIUM	No attempt to differentiate impact by any category LOW
7	de Pee et al. (1998), Indonesia	The study compares participants and non-participants from the same villages LOW	Power calculations are not reported but the sample size seems to be sufficiently large to provide reliable estimates MEDIUM	The programme theory is not discussed and no intermediate outcomes are analysed LOW	Impact of education, breastfeeding status and availability of latrine is assessed, but not comparatively for the two groups MEDIUM

	Study & country	Counterfactual analysis	Power	Intermediate outcomes	Heterogeneity
8	Faber et al. (2002), South Africa	The study compares participants and non-participants over 1 year, but does not control for self-selection in the group of participants MEDIUM	Power calculations are reported, but there are only two villages in the sample, one project and one control LOW	The programme theory is not discussed, differential consumption patterns are reported LOW	No attempt to differentiate impact by any category LOW
9	Greiner and Mitra (1995), Bangladesh	The study employs a control group, but it is not clear that the characteristics of the area selected are similar, that there is no contamination of the intervention and that self-selection in the programme is controlled for MEDIUM	Power calculations are reported and the sample looks sufficiently large, but the number of clusters and the sampling strategy are not reported MEDIUM	A programme theory is discussed and intermediate outcomes such as household nutritional knowledge, breastfeeding practices and consumption patterns are analysed MEDIUM	No attempt to differentiate impact by any category LOW
10	Jones et al. (2005), Nepal	The study compares participants and non-participants from the same areas without attempt to control for selection bias MEDIUM	Power calculations are not reported and the sampling strategy is not entirely clear LOW	The programme theory is briefly discussed and intermediate outcomes such as knowledge and consumption patterns are analysed MEDIUM	No attempt to differentiate impact by any category LOW
11	Kidala et al. (2000), Tanzania	This is an ex-post comparison of a randomly allocated intervention. HIGH	Power calculations are not reported but the sampling strategy is fully described MEDIUM	There is some discussion of the programme theory and the study analysed changes in knowledge MEDIUM	No attempt to differentiate impact by any category LOW
12	Laurie and Faber (2008), South Africa	Neighbouring non-participating households are the control group, resulting in large selection bias LOW	Power calculations are not reported and the sampling strategy is unclear LOW	The programme theory is partially discussed and outcomes in terms of cultivation, acceptability and consumption of vegetables are reported MEDIUM	No attempt to differentiate impact by any category LOW

	Study & country	Counterfactual analysis	Power	Intermediate outcomes	Heterogeneity
13	Makhotla and Hendriks (2004), Lesotho	Households with and without vegetable gardens from the same areas are compared without controlling for selection bias LOW	Power calculations are not reported and sampling strategy is unclear LOW	No intermediate outcomes are reported and no programme theory is discussed LOW	Variation of malnutrition rates across districts is analysed MEDIUM
14	Marsh (1998), Bangladesh	Participant households in project areas are compared to households in non-interventions areas. No attempt to match villages and households on characteristics MEDIUM	Power calculations are not reported and the sampling strategy is not entirely clear LOW	A number of intermediate outcomes including: time allocation, production, gender effects and constraints MEDIUM	No attempt to differentiate impact by any category LOW
15	Olney et al. (2009), Cambodia	The control group consists of households living in neighbouring villages selected by NGO staff using the same criteria used for the selection of the target households MEDIUM	Power calculations are not reported and the sampling strategy is not entirely clear LOW	Pathways of impact are fully discussed and some intermediate outcomes are analysed, such as production and consumption MEDIUM	Dietary diversity is analysed across household size, education, land ownership and gender of head of household MEDIUM
16	Schipani et al. (2002), Thailand	Children from households with gardens are matched to children from households without gardens and similar socio-economic characteristics MEDIUM	Power calculations are not reported and the sample size (60 children) is very small LOW	Programme theory is not discussed and intermediate indicators are not analysed LOW	No attempt to differentiate impact by any category LOW
17	Schmidt and Vorster (1995), South Africa	Children from families participating in communal gardens are matched to an identical number of children from non-participating families from the same village based on few socioeconomic characteristics MEDIUM	Power calculations are not reported and the sample size (36 children) is very small LOW	The reasons for the failure of the intervention are discussed and some intermediate indicators (production and consumption) are analysed MEDIUM	No attempt to differentiate impact by any category LOW

	Study & country	Counterfactual analysis	Power	Intermediate outcomes	Heterogeneity
18	Smitasiri et al. (1999), Thailand	Difference in difference analysis. One concern is that localities were found comparables based on government information and team observation HIGH/MEDIUM	Power calculations are reported with details and properly conducted HIGH	Knowledge, attitudes and practices with respect to consumption of vitamin A rich food is the only intermediate outcome considered MEDIUM	No attempt to differentiate impact by any category LOW
19	Taher et al. (2002), Bangladesh	Before after comparison without a control group EXLCUDED	Power calculations are not reported and the sampling strategy is unclear LOW	No discussion of the programme theory and analysis of intermediate indicators LOW	No attempt to differentiate impact by any category LOW
20	Talukder et al. (2010), Bangladesh, Cambodia, Nepal and Philippines	Control group is composed of households randomly selected from villages where programme was not implemented. The two village types are similar but it is not explained how MEDIUM	Power calculations are not reported and the sampling strategy is unclear LOW	The programme theory is poorly described and some intermediate outcomes result on production and consumption are reported MEDIUM	No attempt to differentiate impact by any category LOW
21	Vijayaraghavan et al. (1997), India (AP)	Before after comparison of participants and non-participants, but the methodology is unclear MEDIUM	Power calculations are not reported and the sampling strategy is unclear LOW	The programme theory is not discussed, but there are some results on knowledge, attitudes and practices MEDIUM	No attempt to differentiate impact by any category LOW

Fisheries studies

	Study & country	Counterfactual analysis	Power	Intermediate outcomes	Heterogeneity
22	Aiga et al. (2002), Malawi	Children of fishing and non-fishing households are compared. Non-fishing children are selected from the same age and gender groups LOW	Power calculations are reported, though intracorrelation coefficients are ignored and samples are small (77 children per group) MEDIUM	There is a clear programme theory described and the following intermediate indicators are used: share of income from fish farming and frequency of intake of oils and fats HIGH	Correlation with breast-feeding practices is reported MEDIUM
23	Murshed-e-Jahan et al. (2010), Bangladesh	The control group is composed of participating farmers before project starts and of non-participating farmers from the same areas LOW	Power calculations are not reported and the rationale for the sampling procedure is unclear LOW	A conceptual framework is presented and the following intermediate outcomes: labour employment and consumption HIGH	Correlation of aquaculture with land ownership and gender of head of household is analysed, but no subgroup analysis of nutritional impact is performed MEDIUM
24	Roos et al. (2003), Bangladesh	The study compares households participating in the trial to a sample of neighbouring households non-participating but having similar socio-economic status. The matching process is not fully explained LOW	Power calculations are not performed, the sampling strategy is unclear and the sample size is rather small (59 participants against 25 non-participants) LOW	No conceptual framework is presented but some intermediate outcomes in terms of production and consumption of fish are reported MEDIUM	No attempt to differentiate impact by any category LOW
25	Thompson et al. (2000), Bangladesh	Participants' households are compared to non-participants' households without attempts to control for selection bias by matching villages or households EXLCUDED	Power calculations are not reported, the sampling strategy is not entirely clear and the sample size is rather small (100 project households against 60 in the control) LOW	The programme theory is not spelled out, but the impact on production and consumption is discussed MEDIUM	Some intermediate outcomes are analysed for households of different landholding size MEDIUM

Food animal source studies

	Study & country	Counterfactual analysis	Power	Intermediate outcomes	Heterogeneity
26	Ahmed et al. (2000), Ethiopia	Outcomes of adopters and non-adopters are compared in regression framework without controlling for selection bias EXLCUDED	Power, significance and sampling strategy are not described LOW	Programme theory is underdeveloped and no intermediate outcomes are assessed LOW	No attempt to differentiate impact by any category LOW
27	Ayele and Peacock (2003), Ethiopia	Outcomes of project participants are compared over time with no attempt to build a control group EXCLUDED	Power calculations are not reported, the sample is taken from one district and the sampling strategy is unclear LOW	No programme theory is developed and intermediate outcomes are not analysed LOW	No attempt to differentiate impact by any category LOW
28	Begum (1994), India (Karnataka)	Outcomes among children of cooperative members and non-members are compared with no attempt to control for selection bias EXLCUDED	Power calculations are not reported LOW	The programme theory is not discussed and intermediate outcomes are not analysed LOW	Outcomes are assessed separately for boys and girls and for milk producers of different operational size MEDIUM
29	Hoorweg et al. (2000), Kenya	Project participants are compared to farmers from a control group LOW	Power calculations are not reported, sampling strategy unclear and small sample size (54 project households against 90 in the control group) LOW	The programme theory is not discussed and intermediate outcomes are not analysed LOW	No attempt to differentiate impact by any category LOW
30	Nielsen et al. (2003), Bangladesh	Participants are compared to households from non-intervention areas LOW	Power calculations are not shown and sample size is very small (35 project households against 35 control households) LOW	Programme theory not developed and production and sales are the only intermediate outcomes considered MEDIUM	No attempt to differentiate impact by any category LOW

The authors of this report were supported by the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) and the International Initiative for Impact Evaluation (3ie).

The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) is part of the Social Science Research Unit (SSRU), Institute of Education, University of London.

Since 1993, we have been at the forefront of carrying out systematic reviews and developing review methods in social science and public policy. We are dedicated to making reliable research findings accessible to the people who need them, whether they are making policy, practice or personal decisions. We engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Founded in 1990, the Social Science Research Unit (SSRU) is based at the Institute of Education, University of London. Our mission is to engage in and otherwise promote rigorous, ethical and participative social research as well as to support evidence-informed public policy and practice across a range of domains including education, health and welfare, guided by a concern for human rights, social justice and the development of human potential.

This material has been funded by the Department for International Development. However the views expressed do not necessarily reflect the department's official policies. The report was first published in 2011 by:

Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)
Social Science Research Unit
Institute of Education, University of London
18 Woburn Square
London WC1H 0NR

Tel: +44 (0)20 7612 6397
http://eppi.ioe.ac.uk
http://www.ioe.ac.uk/ssru

ISBN: 978-1-907345-09-8

The International Initiative for Impact Evaluation (3ie) works to improve the lives of people in the developing world by supporting the production and use of evidence on what works, when, why and for how much. 3ie is a new initiative that responds to demands for better evidence, and will enhance development effectiveness by promoting better informed policies. 3ie finances high-quality impact evaluations and campaigns to inform better program and policy design in developing countries. 3ie Systematic Reviews examine the range of available evidence regarding a particular intervention.

International Initiative for Impact Evaluations (3ie)
Tel: + 44(0)20 7958 8351/8350
http://www.3ieimpact.org

This document is available in a range of accessible formats including large print. Please contact the Institute of Education for assistance:
telephone: +44 (0)20 7947 9556 email: info@ioe.ac.uk