

Protocol:

The Effects of Training, Innovation and New Technology on African Smallholder Farmers' Wealth and Food Security: A Systematic Review

Ruth Stewart, Yvonne Erasmus, Hazel Zaranyika, Natalie Rebelo Da Silva, Evans Muchiri, Marcel Korth, Laurenz Langer, Nolizwe Madinga, Nicola Randall, Thea de Wet

Submitted to the Coordinating Group of:

<input type="checkbox"/>	Crime and Justice
<input type="checkbox"/>	Education
<input type="checkbox"/>	Disability
<input checked="" type="checkbox"/>	International Development
<input type="checkbox"/>	Nutrition
<input type="checkbox"/>	Social Welfare
<input type="checkbox"/>	Other:

Plans to co-register:

<input checked="" type="checkbox"/>	No		
<input type="checkbox"/>	Yes	<input type="checkbox"/> Cochrane	<input type="checkbox"/> Other
<input type="checkbox"/>	Maybe		

Date Submitted: 18th October 2013

Date Final Revision Submitted: 22th May 2014

Approval Date:

Publication Date: 01 September 2014

Note: Campbell Collaboration Systematic Review Protocol Template version date: 24 February 2013

BACKGROUND

The Issue: Food security and poverty reduction for smallholder farmers in Africa

A large proportion of the world's poor live in rural areas, dependent on subsistence farming for their survival (FAO, 2011). Smallholder farmers have been credited with providing up to 80 per cent of food in developing countries (IFAD, 2012) and have the potential to feed themselves and also supply urban markets. Vietnam's smallholder farmers are often credited, for example, with transforming the country from a net-importer of food, to a major exporter (ibid).

Whilst definitions of smallholder farming vary, the concept usually incorporates a number of key elements (Morton, 2007): farms on which labour is predominantly family ('family farms') (IFAD, 2009); farmers and farms that are resource poor (Nagayets 2005; Dixon et al., 2003); farms of a particular size, most commonly two hectares (Nagayets, 2005; Hazell et al. 2010; Wiggins et al., 2010; World Bank 2003; IFAD, 2011a); and farms which are predominantly subsistence, but might also include a mix of commercial and subsistence activities (Narayan and Gulati, 2002).

Smallholder farming is of particular significance to Africa for a number of reasons. Africa's economy is dominated by agriculture (Massett et al., 2011) and the vast majority of farmers in Africa are smallholders (World Bank, 2007). Smallholder farmers contribute significantly to food security on the continent, for example, in sub-Saharan Africa smallholder farmers contribute up to 80 per cent of the food supply (IFAD, 2011b). Smallholder farmers also include many of the continent's poorest and most marginalised people (World Bank, 2007). In southern Africa in particular, large numbers of women and girls rely on smallholder farming, and it provides a survival strategy for many of the continent's young people, many of whom are orphans and head of their households. Furthermore, supporting Africa's development is a priority within the G8, and working towards increasing food security in the region is high on the agendas of the majority of international donors, including the Canadian Foreign Affairs, Trade and Development agency (DFATD), who have commissioned this review.

Given both the importance of smallholder farming in Africa and its potential to contribute to the food security of so many, it is not surprising that considerable efforts are being invested in its success. Both national and international agencies are investing in improving the productivity of smallholder farming, including the International Fund for Agricultural Development (IFAD) and the Canadian Foreign Affairs, Trade and Development agency (DFATD). Additionally, in 2009, the G8's L'Aquila initiative pledged \$22 billion USD for agriculture in developing countries (G8, 2009). In 2012, IFAD launched the Adaptation for Small Holder Agriculture Programme (www.ifad.org/climate/asap/). On a national level, heads of state in Africa are increasingly stressing the need for support for smallholder

farmers. For instance, South African President Jacob Zuma emphasized the need for support of smallholder farmers in his 2013 State of the Nation Address (RSA, 2013). Parallel efforts are being invested in agricultural research, such as impact evaluations and systematic reviews, to assess the effectiveness of these agricultural programmes.

Agricultural productivity of particularly starchy cereals is important since this category of crop accounts for two thirds of the region's energy intake as well as 70 per cent of the income of the extremely poor population living in Africa (AGRA, 2013). The 2013 Africa agriculture report issued by AGRA (ibid) identifies the general production trend for Africa as somewhat erratic, but with most countries reporting a steady increase in production. According to the report, Kenya, Ghana, Mali, Tanzania, Uganda, Zambia, and possibly Mozambique have all reported steady increases in agricultural production (AGRA, 2013, 21). Areas that have experienced civil unrest, political instability, or mismanagement of the macroeconomics of the country in the last decade have seen a decrease in agricultural productivity; included among these are Sierra Leone and Liberia (AGRA, 2013, 21). Technologies typically used to increase agricultural productivity in the region include the "increased use of agricultural inputs, modern farming techniques, and reduced market inefficiencies" (AGRA, 2013, 20). However, a much larger array of factors come to bear on agricultural productivity. Political, technological, physical environmental, and micro- and macroeconomic factors related to each country play a pivotal role in shaping the region's agricultural productivity. World prices of inputs and outputs, and international trade policies also influence the agricultural productivity within countries (AGRA, 2013). As such, any technology addressing any one of these aspects may be expected to have an influence on the food security or income of smallholder farmers in Africa. Examples of specific technologies include treadle pump irrigation technology (Adeoti et al., 2009); biofortification and health information (de Brauw et al., 2013); and adopting an export crop and marketing techniques (Ashraf et al., 2008).

In the context of this considerable and growing emphasis on smallholder farming, there is a need to understand the relative effectiveness of the different interventions targeting smallholder agriculture in achieving various outcomes.

Proposed solutions considered in this review and how they might work

Today, there are a multitude of agricultural interventions in place across Africa (Sapa, 2009). The focus of these interventions has shifted as understanding of the relationship between agriculture and poverty has developed (Masseti et al., 2011). Early interventions focused on increasing productivity to meet a perceived lack of food. With the realisation that undernourishment persists alongside high levels of production (Reutlinger & Pellekaan, 1986), structural issues came to the fore and the concept of food security was introduced (Sen, 1981). Interventions shifted towards income generation, access to markets and the production of more nutritious and calorific foods. Two groups of interventions have

specifically sought to increase food security and reduce poverty by training farmers and / or encouraging them to adopt agricultural innovations and new technologies.

Interventions that are categorised as innovations emphasise the introduction of a ‘new’ farming method, product, or service. An example of this kind of intervention is the introduction of home gardens to increase the intake of vitamin A. A new technology intervention places emphasis on the introduction of a previously unfamiliar agricultural input. This could be a different piece of equipment or genetically modified seeds. Training interventions would place emphasis on providing some kind of training to farmers. The content of such training may not necessarily be new to farmers, but rather previously unemplyed. However, we acknowledge that some training interventions will centre on the introduction of new technology and/or innovation. For instance, an example of this kind of study is de Brauw et al.’s evaluation of biofortification and a health information intervention on the food security of smallholder farmers (2013).

A recent systematic map of the evidence of interventions targeting smallholder farmers (Stewart et al., 2013) found that there were gaps in the African evidence base, including: 1) a lack of systematic reviews addressing various interventions’ impacts on the financial wealth of smallholder farmers, and 2) a lack of assessments of the impact of interventions on smallholders’ food security. The scope of the present review has been influenced by these gaps as well as by consultation with our advisory group and funders.

Agricultural Training / Knowledge

Training interventions for farmers vary considerably. Some interventions focus directly on teaching farmers, using top-down ‘train and visit’ approaches (Hume, 1991). Such training interventions are also often packaged as ‘extension services’, a broad term for programmes which aim to “support and facilitate people engaged in agricultural production to solve problems and to obtain information, skills and technologies” (Anderson, 2007:6). Although traditionally considered as a top-down approach to training, extension services have over time become more participatory in nature (Waddington et al., *in press*). Specifically ‘farmer field schools’, which may be one component of broader agricultural extension services, use a more bottom-up approach to training and knowledge transfer (Waddington et al., 2009). Farmer field schools are participatory, empowering and experiential in nature and draw on problems and priorities identified by farmers themselves, rather than those determined by outsiders (Waddington et al., *in press*). Initially developed to tackle an over-reliance on pesticides, field schools have now been implemented across over 80 countries (van den Berg, 2004).

Another important aspect of these interventions relates to the training objectives: there is a clear distinction in the literature between courses that are directed to improve agricultural practices and increase yields (for example, training on natural resource management; integrated pest management; conservation agriculture), and those which focus on aspects of

farm management (for example, social organisation; management; institutional development).

Perhaps the most straightforward way of considering the range of training interventions available is to consider three facets of the interventions: how experiential or participatory the training is; the duration of the training; and the content of the training – see Table 1.

Table 1: Dimensions of training interventions

Was the training experiential or participatory?	Fully participatory designed to empower farmers and provide experiential learning Partly participatory with limited experience provided Limited participation by farmers with didactic teaching approaches
How long did the training last?	Less than one day 1-7 days Longer than one week
What was the content of the training?	A new technology or innovation (see Table 2 below) Other

An example of an evaluation of this kind of intervention is Anyango et al. (2010), who evaluated a five-year project in Kenya aimed at improving the income and food security of smallholder farmers through a number of training interventions, including introducing new cultivars to the farmers and training them through farmer field schools, as well as providing training on marketing skills, amongst other things. Ashraf et al. (2008), on the other hand, evaluated agricultural interventions that worked with pre-existing farmer self-help groups and provided farmers with information and short orientation-sessions about switching to export crops and offered in-kind loans and facilitated transactions with exporters.

New Technology and Innovation

Agricultural innovation interventions aim to facilitate adoption of new technologies including: fertilisers; new crops (including genetic modification, Hall, 2010); more nutritious crops; and new industries (Ton et al., 2013); and incorporate these technical developments with new systems (Adjei-Nsiah et al., 2008). Sunding and Zilberman (2001) provide a useful framework of these interventions, in terms of mechanical, biological, chemical, agronomic, biotechnological, process and product innovations – see Table 2.

Table 2: A framework of innovation and new technologies

Component	Example
Mechanical innovations	Tractors and combines
Biological innovations	New seed varieties

Chemical innovations	Fertilisers and pesticides
Agronomic innovations	New management practices
Biotechnological innovations	Computer technologies
Process innovations	How to plant a crop
Product innovation	New varieties of vegetables or potatoes

The literature includes a number of examples of evaluations of these innovations and new technologies. For example, Bennett et al. (2003) evaluated the impact of a biological innovation - the introduction of insect-tolerant Bt cotton in South Africa. Panin (1995), on the other hand, assessed a mechanical innovation, evaluating the effectiveness of mechanisation (the introduction of tractor farm technology) on factors such as smallholder farming yield, income and resource utilisation in Botswana.

How the interventions might work

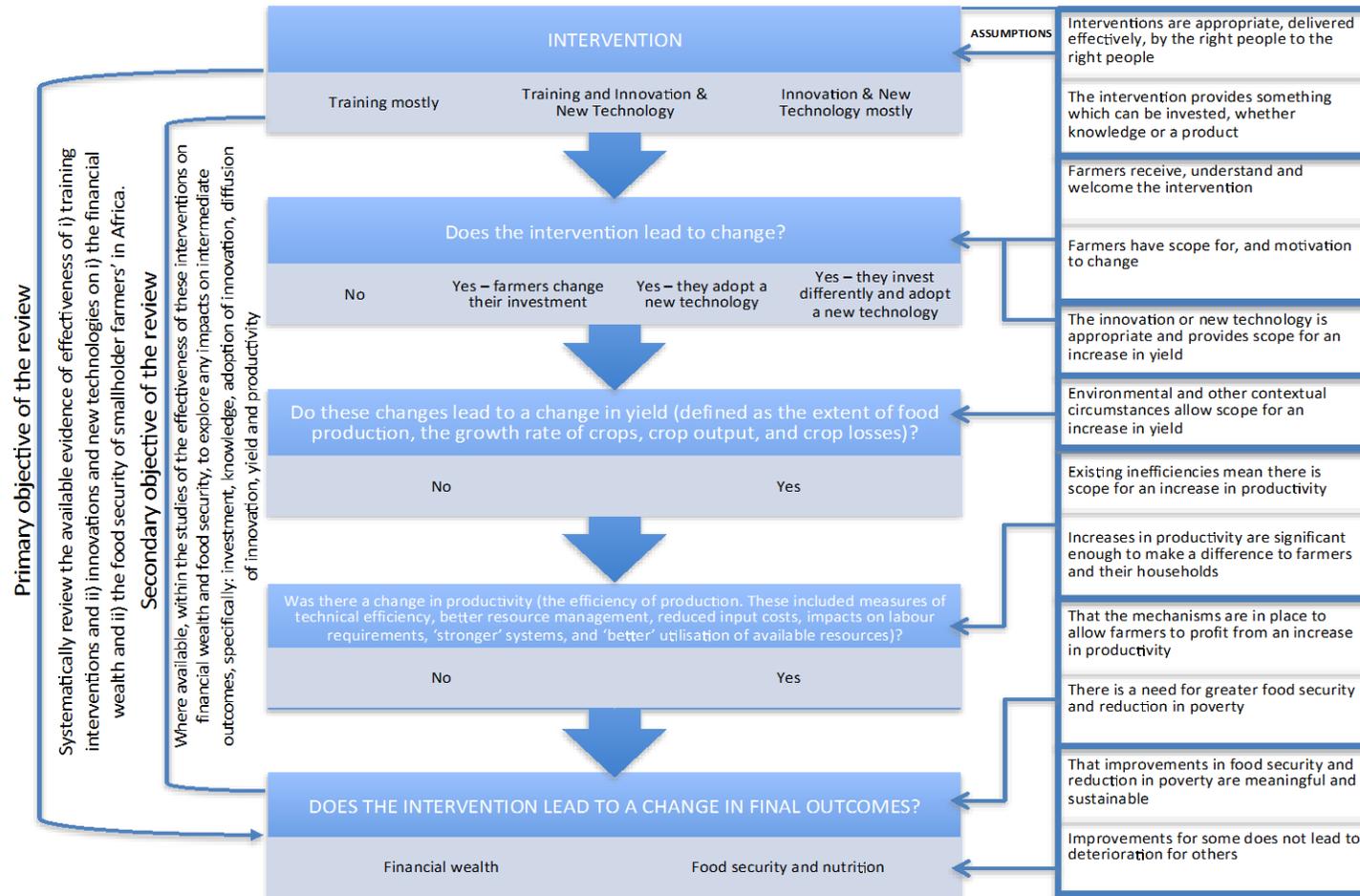
The intended outcomes of these innovations are wide-ranging: from investment (in seed, land, livestock, or labour), to increased yields, productivity, income generation, health, nutrition, food security, and poverty reduction (World Bank, 2007). In particular, there is increasing emphasis amongst international donors on the ‘end-point’ outcomes of food security and poverty reduction. Smallholder farming has long been credited with the potential to end food insecurity (Sen, 1981; Reutlinger & Pellekaan, 1986). The argument is that it is both an effective subsistence strategy and a potential income-generating activity enabling poor farmers to purchase additional food (IFAD, 2012). Furthermore, it is thought to benefit those segments of the population that are most vulnerable to the effects of poverty, namely women, children, and youth (World Bank, 2007).

Whilst there is demand for evidence of the effectiveness of training and innovation and new technology on the financial wealth and food security of smallholder farmers, the mechanisms by which these interventions work involve several intermediate steps. These are multi-faceted, and are dependent on factors such as the environmental context, political stability and economic climate, as well as more direct elements such as farmers’ scope to change their practice and increase their productivity. As Figure 1 illustrates, there are key intermediate outcomes on the pathway to increased food security and financial wealth, specifically: investment, knowledge transfer, adoption of innovation, diffusion of innovation, and increase in yield and productivity.

We will interpret yield broadly. We will include studies that refer to the extent of food production, the growth rate of crops, crop output, as well as crop losses (Stewart et al., 2013). These categories may include studies of the impact of interventions on: the improvement or conservation of soil fertility, of the quality of output, disease resistance or reduction, and food storage conditions (Stewart et al., 2013). Productivity is defined by Stewart and colleagues (2013) in the systematic map of African evidence as “the efficiency of production”.

This includes various aspects, such as “measures of technical efficiency, better resource management, reduced input costs, impacts on labour requirements, ‘stronger’ systems, and ‘better’ utilisation of available resources” (Stewart et al., 2013).

Figure 1: An initial causal pathway



Why it is important to do this review

The importance of smallholder farming in Africa, and the multitude of interventions to increase the wealth and food security of smallholder farmers has been outlined already. We initiated discussions with government agencies and non-governmental organisations supporting these farmers to identify their priorities for evidence to inform their programmes. Having consulted widely on the range of interventions implemented and their intended outcomes, we identified the need for clear evidence on the effectiveness of innovation and / or training interventions, and their impacts on both poverty reduction and food security. An initial scoping review to ascertain the extent to which published reviews had already answered these questions highlighted how more focussed reviews provided evidence on one intervention, but did not answer the question that donors and NGOs raised around which intervention to invest in and why. (See Box 1 and Appendix 1 for more on this preliminary scoping work.)

Box 1: An overview of our ‘review of reviews’ (reproduced from Stewart et al. 2013, with authors’ permission)

A total of 21 systematic reviews of relevance to smallholder farming in Africa were found. Of these, 18 reviews were complete, two protocols were published (Loevinsohn & Sumbug 2012; Knox et al., 2013) and a third protocol is currently under peer review (Dorward et al. 2013). The protocols both focus on agricultural infrastructure (Loevinsohn & Sumbug, 2012; Knox et al., 2013), whilst Dorward and colleagues’ review will focus on agricultural finance. The scopes of the 18 completed reviews were categorised into four broad intervention categories: training, innovation and new technology, infrastructure and finance. Only one of the 18 focused on training, specifically farmer field schools (Waddington et al. 2013). Reflecting the search for new and better ways of farming, we found nine systematic reviews that evaluated the impacts of innovation and new technology (Bayala et al., 2012; Bennet & Franzel, 2009; Berti et al., 2004; Hall et al., 2012; IOB 2011; Girad et al., 2012; Gunaratna et al., 2010; Masset et al., 2011; Rusinamhodzi et al., 2011). These included evaluations of the effectiveness of conservation agriculture in general (Bayala et al., 2012, Bennet & Franzel, 2009, Rusinamhodzi et al., 2011), as well as specific conservation agriculture interventions, including: parkland trees associated with crops (Bayala et al., 2012), coppicing trees (Bayala et al., 2012), green manure (Bayala et al., 2012), mulching (Bayala et al., 2012), crop rotation and intercropping (Bayala et al., 2012; Rusinamhodzi et al., 2011), traditional soil and water conservation (Bayala et al., 2012), tillage management (Rusinamhodzi et al., 2011), and residue retention (Rusinamhodzi et al., 2011). These systematic reviews also considered the impacts of organic agriculture (Bennet & Franzel, 2009) and genetically modified crops (Hall et al. 2012), as well as specific interventions aimed at increasing nutritional status of households, such as home gardening (Berti et al., 2004; Girad et al. 2012; Masset et al., 2011), cash cropping (Berti et al., 2004), irrigation (Berti et al. 2004), and biofortification (Masset et al., 2011; Gunaratna et al. 2010). The impact of interventions to increase food production have been reviewed (IOB, 2011), including particular forms of agriculture,

specifically livestock (Berti et al., 2004), and in particular poultry development (Masset et al., 2011), animal husbandry (Masset et al. 2011) and dairy development (Masset et al., 2011); fish ponds (Masset et al., 2011), aqua culture (Masset et al. 2011), and mixed garden and livestock (Berti et al., 2004). Five completed reviews have considered finance for farmers, specifically: index insurance (Cole et al., 2012), micro-credit (Duvendack et al., 2011; Stewart et al., 2010, 2012), micro-savings (Stewart et al., 2010, 2012), micro-leasing (Stewart et al., 2012), and agricultural investment grants (Ton et al., 2013). Lastly, three systematic reviews focused on the impact of agricultural infrastructure interventions, specifically agricultural interventions and food security (IOB, 2011); infrastructural investments in roads, electricity and irrigation (Knox et al., 2013); and land property rights (Hall et al., 2012).

Despite the somewhat extensive literature base outlined in Box 1, Stewart and colleagues (2013) found that there were three gaps in the African evidence base, two of which will be addressed by this review, namely the lack of systematic reviews addressing various interventions' impacts on 1) the financial wealth of smallholder farmers, and 2) on their food security.

Given the potential for African smallholder farmers to contribute to the food security across the region, coupled with the increasing investment in the industry, there is a need for evidence as to which interventions are most effective. The wide range of options facing policy-makers and practitioners and the need to focus the limited resources available increases the importance of this review.

OBJECTIVES

Our objectives in conducting this Campbell systematic review are to:

1. Systematically review the available evidence of effectiveness of i) training interventions and ii) innovations and new technologies on i) the financial wealth and ii) the food security of smallholder farmers' in Africa.
2. Where available within the studies of the effectiveness of these interventions on financial wealth and food security, to explore any impacts on intermediate outcomes, specifically: investment, knowledge transfer, adoption of innovation, diffusion of innovation, yield and productivity.

In doing so, we will provide valuable information to decision-makers, not in the least being DFATD who have commissioned this work.

METHODOLOGY

I. Criteria for including studies in the review:

To be included in this review, a study must use an experimental or quasi-experimental design. Eligible designs include those in which the authors use a control or comparison group and in which *one* of the following is true:

- Participants are randomly assigned (using a process of random allocation, such as a random number generation);
- A pseudo-random method of assignment has been used and pre-treatment equivalence information is available regarding the nature of the group differences (and groups generated are essentially equivalent);
- Participants are non-randomly assigned but matched on pre-tests and/or relevant demographic characteristics (using observables, or propensity scores) and/or according to a cut-off on an ordinal or continuous variable (regression discontinuity design); or, participants are non-randomly assigned, but statistical methods have been used to control for differences between groups (for example, using multiple regression analysis, including difference-in-difference, cross-sectional (single differences), or instrumental variables regression).

To be included a study must have:

- pre-test data
- post-test data
- an intervention group
- a control group

For this review, the control or comparison conditions in these studies may include farmers receiving no treatment, treatment as usual, or an alternative treatment. No restriction will be placed on duration of follow up. Studies for which the impacts within Africa cannot be isolated, will be excluded from the review.

Given our focus on the end impacts of the interventions (that is, financial wealth and food security), studies for the review must include a minimum follow-up period of at least 6 months between receipt of intervention and measurement of these end impacts. Shorter follow-up may produce misleading results. For example, an intervention that introduces a new breed of cattle, could lead to increased access to meat in the diet in the immediate term,

but it would be misleading to label the consumption of these cattle as an increase in food security.¹

I a) Types of participants

To be included, a study must include African farmers of smallholder farms.

Farmers include both men and women who either own their farms or farm land owned by others. We will not limit by age as we acknowledge that there are large numbers of child-headed households in Africa, and it is feasible that smallholder farmers will be very young. For the purposes of sub-group analysis later in our review, we will define young farmers as those under the age of 20.

Smallholder farms can be defined in a number of ways. Whilst size of farm is often cited – most commonly less than 2 hectares – the productivity of the land can mean that in some countries much larger farms are considered to be ‘smallholdings’. In Tanzania for example, farms of up to 50 hectares have been classified as smallholder farms. The nature of the land, the crops grown and the types of livestock kept all shape the resource-level of farms. Farmers may own their land, although this is often not the case. Similarly smallholder farms are usually assumed to be rural, yet peri-urban farms can also be included. This review will employ a definition of smallholder farms as ‘resource-poor’, where the “resources of land, water, labour and capital do not currently permit a decent and secure family livelihood” (Chalmers, 1985). Table 3 provides a framework for how we will operationalise our definition of smallholder farms.

Women farmers, young farmers and landless labourers have been highlighted by our advisory group as key populations of interest within this review. All three groups will be included within the review, and study populations coded accordingly.

Table 3: Defining smallholder farming for this review

Of the range of dimensions that are relevant for the definition of smallholder farms, we have selected four, at least two of which should be met to qualify for inclusion in our systematic map:

1. Limited size of farm (reported as below two hectares or as compared to other farms in the sector)
2. Mostly dependent on family labour, but also incorporating landless labourers
3. Subsistence farming or mix of subsistence and market-oriented farming, often with limited market access

¹ The danger of measuring end-impacts such as financial wealth using short-term measures is discussed in more detail in Korth et al. (2012).

-
4. Reportedly limited resources in terms of land, technical and technological support, and/or capital for maintenance and investment.

Studies that use the term 'smallholder farm' but do not define it will be included in our review.

I b) Types of interventions

This review focuses on two broad intervention types. Studies will be included in the review if they meet at least one of the following criteria:

- Their main focus is the transfer of knowledge and / or experience to smallholder farmers
- They seek to train farmers in the use of one of the following types of innovation or new technology: mechanical, biological, chemical, agronomic, biotechnological, process or product innovations
- They introduce, or otherwise promote, a new technology or innovation to smallholder farmers
- They introduce, or otherwise promote, a technology to smallholder farmers which is new to the farmers, even if it may already be used by others.
- The means of introducing a new technology, that is, a previously unfamiliar agricultural input or innovation, such as the introduction of a 'new' farming method, product, or service, would primarily be through some form of what we are labelling broadly as 'training'.

Interventions that do not target smallholder farmers specifically will be excluded.

Studies will not be excluded by duration or frequency of programme delivery.

I c) Types of outcome measures

Primary outcomes

This review focuses on two primary outcomes areas: financial wealth and food security.

Financial wealth

We define financial wealth as any form of finance or asset that a household generates, for example, income from selling food products or savings from not having to buy food products.

Specifically we will extract data on the following outcome measures for financial wealth:

- Household income (including intra-household distribution of income if available)
- Household accumulation of financial assets

- Household accumulation of non-financial assets
- Household food expenditure

Bennet et al. (2006) is an example of a study in which the economic impact of genetically modified cotton on South African smallholders' profits is assessed.

Studies that do not consider one of these primary outcomes will be excluded from the review.

Food security

According to the 2009 Declaration of the World Summit, food security exists when all people, at all times, have physical and economic access to sufficient, safe, nutritious food to meet their dietary needs and food preferences for an active life (FAO, 2013). Therefore, food security is essentially the availability of food and one's access to it. We consider the above definition of food security in our review. Based on this definition, our review takes into account improved access, availability and nutrition.

We will extract data on the following specific outcome measure for food security:

- Household food consumption by weight. Low et al. (2007) is an example in which the introduction of OFSP was assessed in an integrated agriculture and nutrition intervention, which aimed to increase vitamin A intake and serum retinol concentrations in young children.
- Per capita calorific intake (Low et al., 2007)
- Household perceptions of food security (for example, Ayalew et al. (1999) Reducing vitamin A deficiency in Ethiopia)

We have also included an 'other category' which will consider any measures which do not fall under those mentioned above.

Secondary outcomes

Secondary outcomes considered in this review will include:

- Investment in capital
- Knowledge transfer
- Adoption of innovation
- Diffusion of innovation
- Yield
- Productivity

I d) Other criteria for including or excluding studies

Studies will not be excluded from the review on the basis of language. Searches will be conducted in English and translations obtained for foreign language papers where possible.

We will include only studies conducted since 1990. Both the methodologies for assessing impact of these interventions, and the nature of the interventions have developed

significantly since 1990 (Romani et al., 2003; Sapa, 2009) making it highly unlikely that we will identify any relevant literature prior to this date. We will search only for papers published since 1990 and screen on study dates. Where any data in a study published after 1990 was collected prior to 1990 (e.g. baseline), the study will be excluded.

II. Search Strategy

In order to identify the literature for this review as comprehensively as possible, we have designed our search strategy to include both general and specialist sources, with both broad search terms and more specialist ones. We have taken advice from two search specialists, from the Campbell Collaboration and the EPPI-Centre, in the design of these searches.

Electronic Sources

- CAB Abstracts
www.cabdirect.org
- Web of Science – specifically, the Social Science Citation Index and Science Citation Index
- 3ie impact evaluations database
<http://www.3ieimpact.org/evidence/impact-evaluations/>
- IDEAS
<http://ideas.repec.org/>
- Africa bib. databases (specifically, African periodical literature/ African Women's Bibliographic database).
<http://www.africabib.org/>
- AGRIS, the research database of the FAO
<http://agris.fao.org/>
- BLDS
<http://bldscat.ids.ac.uk/>
- Agricola
<http://www.ebscohost.com/academic/agricola>
- AgEcon
<http://ageconsearch.umn.edu>
- Africa Wide
<http://www.ebscohost.com/academic/africa-wide-information>

Other sources (including websites and grey literature)

- IFAD evaluation reports
http://www.ifad.org/evaluation/public_html/eksyst/doc/index.htm
- IFPRI publications
<http://www.ifpri.org/publications>
- JPAL evaluations
<http://www.povertyactionlab.org>

- The Millennium Challenge Corporation
<http://www.mcc.gov>
- USAID
<http://www.usaid.gov>
- Bill and Melinda Gates Foundation
<http://www.gatesfoundation.org>
- CGIAR
<http://www.cgiar.org>

Search terms

The key concepts in our review are summarised below.

- i) Smallholder farm
- ii) Impact evaluation
- iii) Africa
- iv) Intervention (specifically training and innovation / new technology)

We have some concerns that combining four concepts in our searches may be too narrow and exclude some relevant studies. From test searches, the 'Africa' concept is challenging to search for (many search engines won't accept the large numbers of search terms required), and is also relatively easy to screen for (the country where a study is conducted is usually reported clearly in the abstract). We will therefore search for only three concepts in these new databases, combining the concepts for smallholder farms, impact evaluation and the interventions of interest in the following way:

((smallholder farm AND impact evaluation AND (training OR innovation)))

We have developed detailed search strings in order to ensure we capture all possible search terms for the concepts we seek (see Appendix 2). However, some databases use relatively simple search functions making long strings of terms difficult to employ. The proposed strings will therefore be adapted to suit each of the databases as appropriate. Where available, we will search within the title and abstract fields. Where this option is not included, we will search the full record. We will also seek appropriate controlled terms where available.

Additional searches

- i. Citation searches will be conducted using Google Scholar, Web of Knowledge, and Scopus for related systematic reviews and key impact evaluations, as listed in Appendix 2.
- ii. Both the included and excluded lists of the identified overlapping systematic reviews will be screened for relevance to this review (see Appendix 1 for list of these SRs).

- iii. We will search the reference lists of all potentially relevant impact evaluations. In addition, we will check the reference lists of a recently published scoping map of agricultural innovation in sub-Saharan Africa:

Percy R, Tsui J and Sutherland A (May 2013) *Agricultural Innovation in Sub-Saharan Africa and South Asia: A Scoping Study*.

http://www.3ieimpact.org/media/finder/2013/06/28/3ie_scoping_study_report_1.pdf

- iv. Requesting relevant studies from key contacts:

We will write to key contacts requesting relevant studies. These will include members of our project advisory group and first authors of relevant reviews, as listed in Appendix 2.

Selection of studies

Two reviewers will independently assess the full text papers against the inclusion criteria, and each author will extract data from included studies. Discrepancies will be resolved by consensus, and a third reviewer will be available to resolve any disagreements.

III. Description of methods used in primary research

Methods used in the primary research considered relevant to this review include randomised controlled trials, cluster randomised controlled trials, and a range of designs that employ non-randomised allocation approaches. These include regression discontinuity designs, natural experiments where external factors determine allocation, and self-selected assignment (by the research team, or the research participants) (Waddington et al., 2012). Of those study designs in which allocation to intervention or control is self-selected, we will only include studies that have a comparison group with before and after data, assessing impact using double difference approaches.

Only studies that have a well-defined intervention-control group will be included in this review. Included studies must assign participants at the individual, group, cluster, districts, or provincial levels.

Ashraf and colleagues (2008) study is an example of a randomised controlled trial which we anticipate including in this review. They collect baseline data and assess impacts at one year, across two treatment groups (both of which receive the 'DrumNet' intervention, and one of which also receive microcredit), and a control group.

In another example, Smale and colleagues (2012) assess the impact of three linked programmes using baseline and endline data, comparing beneficiaries and non-beneficiaries

of the programmes using difference-in-difference analyses. They used propensity score matching to try to retrospectively match their intervention and control groups.

IV. Data extraction of study information

We will use a detailed coding sheet with screening information that determines whether a study is to be included or excluded for this review (see Appendix 3). We will then extract data from the included studies as outlined in the coding sheet, including details on the target population, the type of intervention, scale of intervention, outcomes and how they were measured, and funding agencies. We will use the coding sheet included in Appendix 4 to extract data to allow us to calculate effect sizes, including sample sizes, means, standard deviations, confidence intervals, and rates of dropouts for both control and intervention at each time of follow-up.

Initial coding and screening will be done on EPPI-Reviewer, and additional data extraction for included studies will be done in Microsoft Excel. This will facilitate standardisation of effect measures for outcomes in included studies that will be used in meta-analysis.

A randomly selected sample of the coded studies (10% of full texts) will be double-coded by an independent member of the review team, and inter-rater reliability scores (percentage matches) will be calculated and Cohen's kappa applied (Higgins et al. 2011). Disagreements will be discussed and resolved and a consensus code adopted.

V. Assessing risk of bias

We will assess the methodological quality of the included studies using the risk of bias tool developed by the Cochrane Methods group (Higgins et al., 2011) and adopted for non-randomised studies (Stern et al., 2013). Specifically, these will include screening questions to determine whether particular bias is controllable in a given study, and guidance for the reviewer to rely on while scoring the risk of bias for the outcome, and the justification for making a judgment for every domain and outcome reported. The six domains are listed below, with full details of the tool included in Appendix 5.

1. *Bias due to baseline confounding*
2. *Bias due to selection of participants into the study*
3. *Bias due to departure from intended interventions*
4. *Bias due to missing data*
5. *Bias due to measurement of outcomes*
6. *Bias due to selection of results*

Risk of bias assessments will be done for every relevant outcome in all the six domains. The risk of bias will be deemed as low, moderate, high, or critical, and where sufficient details to make judgment are unavailable, the risk will be deemed as unclear. The findings of those studies judged to be at critical risk of bias will be reported, but not considered for synthesis. See Appendix 5 for more details on each of these judgements.

VI. Effect sizes

We anticipate a variety of scales measuring effect sizes and we will standardise these measures in order to obtain a uniform scale to facilitate synthesis in a meta-analysis (Borenstein, 2009). For continuous outcomes we will use the Hedges' *g* statistic to calculate standardised effect sizes of included studies (Hedges & Olkin, 1985). For dichotomous outcomes we will calculate odds ratio and risk ratio as the measure of effect size. Effect sizes, standard errors and their 95% confidence intervals will be computed as recommended by Borenstein (2009).

Where information is missing, and can be calculated from other variables, we will do so, as per Higgins et al. (2011). Where crucial statistics are not reported, authors will be contacted, but if information is still unavailable manipulations will be conducted to derive desired statistics in order to perform meta-analysis, as specified by Lipsey and Wilson (2001).

For those studies where no standardised effect size can be extracted for meta-analysis, effect sizes and confidence intervals will still be reported but not included in the forest plot or meta-analysis. Furthermore we will seek to indicate on all forest plots how many studies reported relevant findings but could not be included due to an inability to extract the effect size. This is to avoid any over-simplification of the evidence base within our forest plots, without an awareness of the wider evidence.

VII. Selecting interventions in multi-armed studies

In cases where studies have used a single control group and multiple intervention groups:

- If only one of the intervention groups is closely matched to our intervention group of interest, we will select one pair of intervention/control group and exclude others (Deeks et al. 2011; Borenstein et al. 2009; Higgins, 2008). The selected pair of groups will be closely matched to our intervention of interest using our pre-specified criteria.
- If more than one intervention group is closely related to our intervention of interest when compared to our pre-specified inclusion criteria, and if possible, all relevant experimental intervention groups will be combined to create a single experimental group to compare with the control group (Higgins, 2008). Similarly, all relevant control groups will be combined to create a single control group for comparison with relevant experimental groups (as per Borenstein, 2009). For dichotomous outcomes, both the sample sizes and the numbers of people with events can be summed across groups. For

continuous outcomes, means and standard deviations will be combined as described in The Cochrane Collaboration handbook (Higgins, 2008). Where correlation measures are reported, multiple-treatment meta-analysis may be adopted, where findings from several arms of the same study will be included in the meta-analysis after adjusting for correlation.

VIII. Dependent effect sizes

We will follow Campbell Guidance (Becker et al., 2007) so that each meta-analysis pools only findings that are statistically independent. Dependent effect sizes may occur where more than one study is reported in a single paper, where several papers report the findings of one study, when studies include several intervention arms with only one control or measure outcomes at more than one point in time. Multiple measures of the same outcome within one study will not be synthesised. We will take steps to ensure that only independent findings are included in any one synthesis as outlined below.

Where more than one paper or report is identified on a single study, we will choose one as the 'main' paper (the one with the most relevant data) and the others will be considered 'secondary reports' which we draw on only for additional information about that one study, including outcome measures not reported in the main paper.

Where a single report describes more than one study, these will be separated into two or more 'studies', which will be coded and analysed separately. To help us identify papers which might be linked or split in these ways, we will collect information on funding bodies, and intervention programme names in our preliminary coding questions. Our aim will be to identify all affiliations between studies/reports before detailed coding takes place.

Where individual studies report multiple outcome measures for the same outcome construct we will select the outcome that is most commonly reported across included studies, and if not available, then the outcome that is most accurately measured will be used.

For studies reporting follow-up effects at multiple points in time, we will take the final follow-up measure, assuming more follow-up time has been accrued, thus increasing the statistical power of detecting an effect of the intervention.

Finally, some studies will report results for two or more instruments to assess the same participants, which provides basis for comparing the two instruments. Including both instruments into a statistical analysis violates the assumption of independent findings and moderator analysis will be performed to assess the potential bias of using different instrumentation. The specification with the lowest level of bias will be included in the synthesis.

IX. Unit of analysis and accounting for clustering

Unit of analysis errors occur when interventions are allocated at a different level than the

unit of analysis. Clustering has the effect of narrowing the confidence intervals from the true confidence intervals because observations from individuals in the same cluster are likely to be more similar than individuals across the cluster (Higgins, 2011). Therefore, observations from the same cluster cannot be considered to be independent from one another. In clustered designs, reviewers will observe if the assignment units are the same units reported in analysis, and if this is not the case, whether appropriate statistical procedures of analysis including multi-level analysis have been applied. For clustered designs with unit of analysis errors, corrected standard errors and confidence intervals will be used to adjust for clustering

We will adjust standard errors or sample sizes from cluster randomised trials using the methods described in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins, 2011). The adjustment method requires the intra-class correlation co-efficient (ICC) where available to estimate the design effect. The 'design effect' is given as

$$deff=1+ICC(m-1) \text{ where } m \text{ is the mean cluster size}$$

$$\text{The 'effective sample size' } ESS = \text{total sample size}/\text{design effect}$$

Standard errors associated with the SMD in cluster studies may be inflated by the square root of the design effect unless where authors have implicitly adjusted for the clustering.

X. Approach to Meta-analysis

Quantitative effect estimates will be synthesised using inverse-variance random effects model meta-analysis in STATA 13.0 (StataCorp, 2013). Using the random effects model allows the true effects in each study to vary according to some distribution.

We will separately analyse results from experimental studies and quasi-experimental studies as per The Campbell Collaboration guidelines. Only quasi-experimental studies that have adjusted for similar baseline covariates, or matched on related covariates, or applied similar statistical models to estimate effect sizes will be combined as these are deemed more comparable. If the outcomes are deemed to be too dissimilar to combine in a meta-analysis, we will display effect sizes in a forest plot, but without the final pooled effect size.

Sensitivity analysis will be conducted, where outlying studies that exhibit a high or low effect size will be excluded, as well as those with wide confidence intervals. Meta-regression will be performed to investigate sources of heterogeneity if sufficient studies are included for statistical synthesis.

We acknowledge that meta-analysis may not be always sensible if studies are not addressing the same question, measuring the effect of a similar intervention, or if the populations are different. This might imply that the outcomes are different and combining these for a mean

estimate would not be sensible.

XI. Assessment of heterogeneity

Heterogeneity will be assessed using inspection of the forest plots for lack of overlap of confidence intervals, and using statistical tests of heterogeneity using the Q statistic (Hedges & Olkin, 1985). We will also calculate and report the I^2 , and Tau^2 to provide estimates of the magnitude of the variability that is due to heterogeneity (Higgins, 2002; Higgins, 2003).

XII. Moderator analyses

It is possible that the effect sizes of intervention on outcomes vary according to levels of some moderating covariates. If feasible, we plan to investigate three types of moderators as laid out by Lipsey (2009): extrinsic, methodological, and substantive. The specific moderators are listed below.

Extrinsic moderators (unchangeable characteristics of the study): year of publication (to investigate whether treatment effects have changed over time), funder, type of publication.

Methodological (aspects of the study designs): study design, interventions evaluated, outcomes reported, and risk of bias status (Higgins et al., 2011)

Substantive moderators: type of implementer, settings, type of training received, age of participants, gender of participants, socio-economic status of participants, geographical location, type of crop or livestock, climate zone

Moderator analysis will be conducted as a form of sensitivity analysis according to the values taken by the moderator covariates, as specified by Borenstein (2009). The relationship between two categorical moderators will be examined by creating a two-way table where a chi-square test will be determined according to the number of studies with these particular moderator variables. The Pearson correlation will be used to examine the effect of continuous moderators (Cooper & Hedges, 1994).

XIII. Publication bias

We acknowledge that studies that may have evaluated similar interventions and outcomes may be missing. It is important to include all studies where possible and to get additional information, but there is also the possibility that the missing studies are different from the published studies. We will address publication bias in two ways. Firstly, we will conduct an extensive search of the published and unpublished literature and will include studies regardless of publication status. Secondly, we will use statistical methods to test and control for publication bias (Egger et al. 1997; Palmer et al., 2008). The presence of publication bias will be evaluated empirically in studies included in our meta-analysis using funnel plots. The

funnel plot is based on the fact that precision in estimating the underlying treatment effect will increase as sample size increases. Results from small studies will scatter widely at the bottom of the plot, with the spread narrowing among larger studies. In the absence of bias the plot will resemble a symmetrical inverted funnel. However, if there is publication bias, funnel plots will often be skewed and asymmetrical.

XIV. Interpreting findings of syntheses

Our findings and judgments will be integrated into summary of findings tables to ease the readers' understanding of our how we have reached our review conclusions. We will apply the GRADE tool to our synthesis to enable transparent and structured interpretation of results (Guyatt et al., 2011).

XV. Treatment of qualitative data

We will not include qualitative studies, but we will extract information about intervention implementation and context from qualitative research or process evidence incorporated within the included studies, as outlined in the coding framework. We will use this information to inform our grouping and analysis of included studies.

REFERENCES

- AGRA (2013). Africa agriculture status report: focus on staple crops. Nairobi, Kenya.
- Anyango, J.J., Wambugu, F.M., Nkanya, J., Kyalo, G., & Onyango, C. (2010). Introduction of tissue culture banana technologies and their impact on producer wellbeing. *Acta Horticulturae*, 879, 767-772
- Ashraf, N., Gine, X., & Karlan, D. (2008) Finding missing markets (and a disturbing epilogue): evidence from an export crop adoption and marketing intervention in Kenya. Washington DC: Word Bank
- Ayalew, W., Gebriel, Z. W., & Kassa, H. (1999). Reducing Vitamin A deficiency in Ethiopia: linkages with a women-focused dairy goat farming project. Research Report Series 4. International Center for Research on Women.
- Bayala, J., Sileshi, G.W., Coe, R., Kalinganire, A., Tchoundjeu, Z., Sinclair, F. & Garrity, D. (2012). Cereal yield response to conservation agriculture practices in drylands of West Africa: A quantitative synthesis. *Journal of Arid Environments*, 78, 13-25
- Becker, W.B., Hedges, L.V., & Pigott, T. D. (2007). Statistical analysis policy brief. Available from http://www.campbellcollaboration.org/artman2/uploads/1/C2_Statistical_Analysis_Policy_Brief-2.pdf
- Bennet, R., Morse, S., & Ismael, Y. (2006). The economic impact of genetically modified cotton on south African smallholders: Yield, Profit and Health Effects. *Journal of Development Studies*, 42(4), 662–677
- Bennett, M. and Franzel, S. (2009). Can organic and resource-conserving agriculture improve livelihoods? A meta-analysis and conceptual framework for site-specific evaluation. ICRAF Occasional Paper No. 11. Nairobi: World Agroforestry Centre.
- Bennett, R., Buthelezi, T.J., Ismael, Y., & Morse, S. (2003). Bt cotton, pesticides, labour and health: a case study of smallholder farmers in the Makhathini Flats, Republic of South Africa. *Outlook on Agriculture*, 32, 123-128
- Berti, R. P., Krasevec, J., & FitzGerald, F. (2004). A review of the effectiveness of agriculture interventions in improving nutrition outcomes. *Public Health Nutrition*, 7 (5), 599-609.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009) Introduction to meta-analysis. Chichester: John Wiley & Sons.
- Cole, S., Bastian, G., Vyas, S., Wendel, C. & Stein, D. (2012) The effectiveness of index based micro-insurance in helping smallholders manage weather-related risks. London: EPPI-

Centre, Social Science Research Unit, Institute of Education, University of London.

- Cooper, H., & Hedges, L. (1994). *The Handbook of Research Synthesis*. New York: Russel Sage Foundation
- De La Rue, L., Polanin, J. R., Espelage, D. L., & Pigott, T. (2013) Protocol: School-based interventions to reduce dating and sexual violence: A systematic review. *Campbell Library*. <http://www.campbellcollaboration.org/lib/project/268/>
- Deeks, J. J., Higgins, J. P. T., & Altman, D.G. Eds. (2011). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Dixon, J., Taniguchi, K., & Wattenbach, H., Eds. (2003). Approaches to assessing the impact of globalization on African smallholders: Household and village economy modeling. Proceedings of a working session on Globalization and the African Smallholder Study. FAO (Agricultural Support Systems Division [AGS] and Agricultural and Development Economics Division [ESA]) and the World Bank. Rome: FAO
- Dorward, A., Roberts, P.D., Finegold. C., Hemming, D. J., Chirwa, E., Wright, H. J., Osborn, J., Lamontagne-Godwin, J., Harman, L., Parr, M. & Barbosa, V. (in press). What are the impacts of agricultural input subsidies on productivity, farm incomes, consumer welfare and wider growth in low- and middle- income countries: A systematic review.
- Duvendack, M., Palmer-Jones, R., Copestake, J. G., Hooper, L., Loke, Y. & Rao, N. (2011). What is the evidence of the impact of microfinance on the well-being of poor people? London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Egger, M., Smith, G.D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 315, 629-634.
- FAO, IFAD & WFP. (2013). *The State of Food Insecurity in the World 2013. The multiple dimensions of food security*. Rome: FAO
- Girard, A.W., Self, J.L., McAuliffe, C. & Oludea, O., (2012). The effects of household food production strategies on the health and nutrition outcomes of women and young children: A Systematic Review. *Paediatric and Perinatal Epidemiology*, 26 (Suppl. 1), 205–222.
- Gunaratna, N. S., De Groote, H., Nestel, P., Pixley, K. V. & McCabe, G. P., (2010). A meta-analysis of community-based studies on quality protein maize. *Food Policy*, 35, 202–210.

- Guyatt, G., Oxman A., Akl, E. A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., deBeer H., Jaeschke R., Rind D., Meerpohl J., Dahm, P., & Schünemann, H. J. (2011) GRADE guidelines: 1. Introduction to GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology*, *64*, 383-394.
- Hall, C., Knight, B., Ringrose, S. & Knox, O., (2012). What have been the farm-level economic impacts of the global cultivation of GM crops? Systematic Review, No.CEE 11-002.
- Hazell, P., Poulton, C., Wiggins, S. & Dorward, S. (2010). The future of small farms: trajectories and policy priorities. *World Development*, *38* (10), 1349-1361
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL.: Academic Press
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*; *21*: 1539-1558.
- Higgins J. P. T., Thompson, S. G., & Deeks, J. J., Altman, D.G. (2003) Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*, 557-560.
- Higgins, J. et al. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal*, *343*, d5928 doi: 10.1136/bmj.d5928
- Higgins, J. P. T. & Green, S. (2008). Chapter 9: Analysing data and undertaking meta-analyses. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration. Available from: www.cochrane-handbook.org.
- IFAD [International Fund for Agricultural Development] (2009). *Smallholder farming in transforming economies of Asia and the Pacific: Challenges and opportunities*. Discussion Paper, Thirty Third session of Governing Council. IFAD
- IFAD [International Fund for Agricultural Development] (2011a). Latin America: The state of smallholders in agriculture. Conference on New Directions for Smallholder Agriculture 24-25 January 2011. Rome: IFAD
- IFAD [International Fund for Agricultural Development], (2011b). Smallholders can feed the world. Rome: IFAD
- IOB. (2011). Improving food security. A systematic review of the impact of interventions in agricultural production, value chains, market regulation, and land security. IOB Study No 363
- Knox, J., Daccache, A., & Hess, T. (2013). What is the impact of infrastructural investments in roads, electricity and irrigation on agricultural productivity? CEE review 11-007.

Collaboration for Environmental Evidence:
www.environmentalevidence.org/SR11007.html.

Korth, M., Stewart, R., van Rooyen, C., & De Wet, T. (2012). Microfinance: Development Intervention or Just Another Bank? *Journal of Agrarian Change*, 12(4), 575-586.
DOI: 10.1111/j.1471-0366.2012.00375.x

Lipsey, M. W. (2009). Identifying interesting variables and analysis opportunities. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.) *The handbook of research synthesis and meta-analysis*, (pp. 147-158), New York: Sage Publication

Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis (Applied Social Research Methods Series, Vol. 49). Thousand Oaks, CA: Sage.

Loevinsohn, M. & Sumberg, J. (2012). Under what circumstances and conditions does adoption of technology result in increased agricultural productivity? – Protocol. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London

Low, J., Arimond, M., Osman, N., Cunguara B., Zano, F., & Tschirley, D., (2007). A food-based approach introducing orange-fleshed sweet potatoes increased vitamin A intake and serum retinol concentrations in young children in rural Mozambique. The United Nations University. Food and Nutrition Bulletin, 28 (2).

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

Morton, J.F. (2007) *The Impact of Climate Change on Smallholder and Subsistence agriculture*. University of Greenwich, Natural Resources Institute

Nagayets, O. (2005). Small farms: current status and key trends. In: The future of small farms. Proceedings of a Research Workshop, Wye, UK, June 26-29. Washington, D.C.: International Food Policy Research Institute

Narayanan, S., & Gulati, A. (2002). Globalization and the smallholders: A review of issues, approaches and implications. Markets and Structural Studies Division Discussion Paper No.50. Washington D.C.: International Food Policy Research Institute and the World Bank

Panin, A. (1995). Empirical evidence of mechanization effects on smallholder crop production systems in Botswana. *Agricultural Systems*, 47, 199 – 210.

- Palmer, T., Peters, J., Sutton, A.J., & Moreno, S. (2008). Contour enhanced funnel plots for meta-analysis. *The Stata Journal*, 8; 242-254.
- Reutlinger, S. & Pellekaan, H. (1986). Poverty and hunger: Issues and options for food security in developing countries. Washington DC: The World Bank.
- Rusinamhodzi, L., Corbeels, M., van Wijk, M.T., Rufino, M.C., Nyamangara, J. & Giller, K.E. (2011). A meta-analysis of long-term effects of conservation agriculture on maize grain yield under rain-fed conditions. *Agronomy Sustainable Development*, 31, 657–673.
- Sen, A. (1981). *Poverty and Famine*. Oxford: Clarendon Press
- Sunding, D. & Zilberman, D. (2001). The agricultural innovation process: research and technology adoption in a changing agricultural sector. *Handbook of Agricultural Economics*, 1(A), 207-261.
- StataCorp. (2013). *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
- Stern, J., Higgins, J., & Reeves, B. (in development). Extending the Cochrane Risk of Bias tool to assess risk of bias in non-randomized studies.
- Stewart, R., van Rooyen, C., Dickson, K., Majoro, M. & de Wet, T. (2010). What is the impact of microfinance on poor people? A systematic review of evidence from sub-Saharan Africa. Technical report. London: EPPI-Centre, Social Science Research Unit, University of London.
- Stewart, R., van Rooyen, C., Korth, M., Chereni, A., Rebelo Da Silva, N. & de Wet, T. (2012). Do micro-credit, micro-savings and micro-leasing serve as effective financial inclusion interventions enabling poor people, and especially women, to engage in meaningful economic opportunities in low- and middle-income countries. A systematic review of the evidence. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Stewart, R., Erasmus, E., Zaranyika, H., Rebelo Da Silva, N., Korth, M., Langer, L., Randall, N., Madinga, N. & de Wet, T. (2013). The size and nature of the evidence-base for smallholder farming in Africa: a systematic map. *Journal of Development Effectiveness*, 6(1), 58-68.
- Ton, G., de Grip, K., Klerkx, L., Rau, M.L., Douma, M., Friis-Hansen, E., Triomphe, B., Waters-Bayer, A. & Wongtschowski, M. (2013). Effectiveness of innovation grants to smallholder agricultural producers: an explorative systematic review. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Available at: <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=4soRhylhV4A%3D&tabid=3401>

Tripney, J., Hombrados, J., Newman, M., Hovish, K., Brown, C., Steinka-Fry, K., & Wilkey, E. (2013). Technical and Vocational Education and Training (TVET) interventions to improve the employability and employment of young people in low- and middle-income countries: A systematic review. *Campbell Systematic Reviews*. 9 DOI: 10.4073/csr.2013.9

Waddington, H., White, H., Snilstveit, B., Hombrados, J.G., Vojtkova, M., Davies, P., Bhavsar, A., Evers, J., Koehlmoos, T. P., Petticrew, M., Valentine, J. C., & Tugwell, P. (2012) How to do a good systematic review of effects in international development: a tool kit, *Journal of Development Effectiveness*, 4(3), 359-387, DOI: 10.1080/19439342.2012.711765

Waddington, H., Snilstveit, B., Hombrados, J., Vojtkova, M., && White, H. (2013 – *in press*) Farmer Field Schools for improving farming practices and farmer outcomes in low and middle-income countries: a systematic review. *Campbell Systematic Reviews*.

Wiggins, S., Kirsten, J., & Llambi, L. (2010). The future of small farms. *World Development* 38(10): 1341-1348.

World Bank (2003). *Reaching the rural poor: A renewed strategy for rural development*. Washington D.C.: World Bank

World Bank (2007). *World Development Report 2008: Agriculture for development*. Washington, D.C.: World Bank
http://siteresources.worldbank.org/INTWDR2008/Resources/WDR_00_book.pdf

SOURCES OF SUPPORT

This review has been made possible thanks to the generosity of many individuals and organisations. It is funded by Foreign Affairs, Trade and Development, Canada (DFATD, formerly CIDA), managed through the International Initiative for Impact Evaluation (3ie), who have also provided some financial support. As such, some extent of the review scope was predetermined. However, after a detailed conversation with funders following two pieces of additional work – a review of reviews and a systematic map – as well as consulting with our advisory group, the scope of the review was refined to its present form.

In addition, the Centre for Anthropological Research at the University of Johannesburg has generously allocated additional staffing to the review. With thanks too to our international advisory group, and peer reviewers for their input to our protocol.

DECLARATIONS OF INTEREST

None. No authors have any involvement in any of the primary studies included in this review. The same review team are also conducting a related review on the impacts of urban agriculture on food security and nutrition.

Stewart, R., Korth, M., Langer, L., Rafferty S., Rebelo Da Silva, N., & van Rooyen, C. (2013) What are the impacts of urban agriculture programs on food security and nutrition in low and middle income countries? A systematic review protocol. *Environmental Evidence*, 2(7): 1-13.

REVIEW AUTHORS

Lead review author:

The lead author is the person who develops and co-ordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base and takes responsibility for the on-going updates of the review.

Name:	Ruth Stewart
Title:	Dr
Affiliation:	EPPI-Centre (London) and Centre for Anthropological Research, University of Johannesburg
Address:	Social Science Research Unit, 18 Woburn Square
City, State, Province or County:	London
Postal Code:	WC1H 0NR
Country:	United Kingdom
Email:	r.stewart@ioe.ac.uk ; ruths@uj.ac.za

Co-author(s): (In alphabetical order)

Name:	Thea de Wet
Title:	Prof
Affiliation:	Centre for Anthropological Research
Address:	House 10, Research Village, Bunting Road Campus, University of Johannesburg

City, State, Province or County: Johannesburg

Country: South Africa

Name: **Yvonne Erasmus**

Title: Dr

Affiliation: Centre for Anthropological Research

Address: House 10, Research Village, Bunting Road
Campus, University of Johannesburg

City, State, Province or County: Johannesburg

Country: South Africa

Name: **Marcel Korth**

Title: Mr

Affiliation: Centre for Anthropological Research

Address: House 10, Research Village, Bunting Road
Campus, University of Johannesburg

City, State, Province or County: Johannesburg

Country: South Africa

Name: **Laurenz Langer**

Title: Mr

Affiliation: Centre for Anthropological Research

Address: House 10, Research Village, Bunting Road
Campus, University of Johannesburg

City, State, Province or County: Johannesburg

Country: South Africa

Name:	Nicola Randall
Title:	Dr
Affiliation:	Harper Adams Agricultural University
Address:	Crop and Environment Sciences
City, State, Province or County:	Newport, Shropshire, TF10 8NB
Country:	United Kingdom

Name:	Natalie Rebelo Da Silva
Title:	Ms
Affiliation:	Centre for Anthropological Research
Address:	House 10, Research Village, Bunting Road Campus, University of Johannesburg
City, State, Province or County:	Johannesburg
Country:	South Africa

Name:	Hazel Zaranyika
Title:	Ms
Affiliation:	Centre for Anthropological Research
Address:	House 10, Research Village, Bunting Road Campus, University of Johannesburg
City, State, Province or County:	Johannesburg
Country:	South Africa

ROLES AND RESPONSIBILITIES

We have a large team, deliberately formulated to enable us to build review experience within our Centre in Johannesburg. Some members will have only small roles on the review, whilst others will take the lead on specific elements as outlined below.

- *Content:*

RS, YE, HZ and NRDS will be conducting much of the work required for this review, with occasional support from MK, LL and NM. RS and MK have both conducted systematic reviews of relevant literature (on microfinance including agricultural finance). The team is currently undertaking a parallel review on urban agriculture.

The team will be supported by the other PIs on the review, who have been chosen for their specific expertise: in international development in Africa (MK, YE and TW), in agricultural research (NR), and a new member of the team who specialises in biostatistics and meta-analysis.

- *Systematic review methods:*

RS will be responsible for leading on methods, designing the study and taking responsibility for all stages of the review. Her depth of systematic review experience and previous reviews using a wide range of approaches (including reviews of reviews, systematic maps, traditional effectiveness reviews and rapid evidence assessments) and working with a number of organisations (Cochrane, Campbell, EPPI-Centre, 3ie and Collaboration for Environmental Evidence) means she is well equipped to lead this complex review.

TW, YE, MK, NR, LL and NRDS have all attended training in systematic review methods and have experience of working on reviews. They will draw on this experience in their varying roles in this review:

TW, and NR will be commenting on draft products as the review progresses, as well as systematic review tools (such as the coding framework), and leading our dissemination activities;

HZ and NRDS will be working day-to-day on the review conducting much of the many body of work, collecting, screening, coding, appraising and synthesizing studies alongside RS; An additional statistician is joining the team and will lead the statistical elements of the day to day tasks;

NRDS, supported by LL, will be responsible for the administration of the review, the management of our specialist database, and assisting with collecting, screening and coding the literature.

- *Statistical analysis:*

RS will take the lead in the data analysis for this review, supported by our new statistician, a medical statistician with experience of conducting meta-analyses for systematic reviews.

In addition, an experienced Cochrane-trained bio-statistician, Alfred Musikewa, with expertise in conducting meta-analysis and in providing training to others has offered his advice to the team. Alfred has already provided training to the team and is available to advise on the statistical meta-analysis for the review as necessary.

MK and RS are both trained in statistics for social research.

RS and NR have advanced training in statistical meta-analysis.

RS, NR and MK are all experienced in narrative synthesis in systematic reviews.

- Information retrieval:

Additional technical input on systematic searching is being provided by the EPPI-Centre's Information Scientist, Claire Stansfield. Claire is a specialist in developing and implementing search strategies for systematic reviews in development. She will be supported in this role by RS, who also has experience of these systematic review tasks.

NRDS is experienced in collecting publications for inclusion in systematic reviews and is doing most of our 'collecting' supported by the rest of the team. We will benefit from having 3 centres included in this review (EPPI-Centre, Harper Adams University and University of Johannesburg), all of which have different access to publications.

N.B. While the following team are already committed to contributing to this review, we have been approached by a number of stakeholders from international agencies that specialise in agriculture. As well as contributing their expertise via our advisory group, they have expressed an interest in helping with the practical work of the review. It is therefore possible that additional agriculture specialists will provide input to the review. Should they be involved in conducting review tasks they will be trained by RS and will be 'partnered' by an experienced systematic reviewer to support their learning whilst assuring the rigour of the review.

PRELIMINARY TIMEFRAME

Approximate date for submission of the systematic review: April 2014

- Scope agreed with funders: August 2013
- Protocol submitted for peer review: October 2013
- Searches and screening: until December 2013
- Collection of full texts: September - December 2013²
- Detailed coding including quality assessment and data extraction: November - April 2014
- Synthesis: May 2014
- Preliminary findings available for discussion with our advisory group: early 2014
- Writing and submission of full review: end May 2014

² Note that many full texts have already been identified, screened, and collected as part of our preliminary scoping work

PLANS FOR UPDATING THE REVIEW

This review is reliant on external funding. Updates will similarly depend on the availability of funds, which is ultimately dependent on the importance of the subject to international agencies. There has been considerable interest in the review from not only our funders DFATD, but also other international agencies such as IFAD.

Our plans are therefore to approach possible funders for backing to update this review in 2016/2017. Ruth Stewart will take responsibility for exploring the potential for funding and liaising with the Campbell International Development Coordinating Group about updates.

AUTHORS' RESPONSIBILITIES

By completing this form, you accept responsibility for preparing, maintaining and updating the review in accordance with Campbell Collaboration policy. The Campbell Collaboration will provide as much support as possible to assist with the preparation of the review.

A draft review must be submitted to the relevant Coordinating Group within two years of protocol publication. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every five years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

PUBLICATION IN THE CAMPBELL LIBRARY

The support of the Campbell Collaboration and the relevant Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review and subsequent updates in the Campbell Library. Concurrent publication in other journals is encouraged. However, a Campbell systematic review should be published either before, or at the same time as, its publication in other journals. Authors should not publish Campbell reviews in journals before they are ready for publication in the Campbell Library. Authors should remember to include a statement mentioning the published Campbell review in any non-Campbell publications of the review.

I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

A handwritten signature in blue ink that reads "R Stewart". The signature is written in a cursive style with a large initial "R" and a stylized "S".

Form completed by: Dr Ruth Stewart

Date: October 2013

Appendix 1: An initial review of the overlapping reviews (completed in March 2013)

Systematic review	Interventions that overlap with our review (Our category, and the review-specific interventions)	Outcomes that overlap with our review (Our category, and the review-specific outcomes)	Population	Region
Bayala et al. (2012) Cereal yield response to conservation agriculture practices in drylands of West Africa: A quantitative synthesis	INNOVATION & NEW TECHNOLOGY Conservation Agriculture including: Parkland trees associated with crops Coppicing trees Green manure Mulching Crop rotation and intercropping Traditional soil/water conservation	YIELD & PRODUCTIVITY Crop yield Yield response	Famers and pastoralists living in dryland areas	4 sub-Saharan African countries: Burkina Faso, Mali, Niger and Senegal
Bennet et al. (2009) Can organic and resource-conserving agriculture improve livelihoods?	INNOVATION & NEW TECHNOLOGY Organic Agriculture Conservation Agriculture	YIELD & PRODUCTIVITY Crop yield Reduced input costs INCOME / FINANCE Income	Smallholder farmers	Africa
Berti et al. (2004) A review of the effectiveness of agriculture interventions in improving nutrition outcomes	INNOVATION & NEW TECHNOLOGY Agricultural interventions that aim to increase nutritional status of households: Home gardening Livestock Mixed garden and livestock Cash cropping Irrigation	FOOD SECURITY & NUTRITION Nutritional status of household members	Unspecified	LMICs
Cole et al. (2012) The effectiveness of index-based micro-insurance in helping smallholders manage weather-related risks	FINANCE Index Insurance [Weather & Yield]	YIELD / PRODUCTIVITY Production Yields INCOME / FINANCE Income	Smallholder farmers / developing country farmers	LMICs
Duvendack et al. (2011) What is the evidence of the impact of microfinance on the well-being of poor people?	FINANCE Microcredit	INCOME / FINANCE Income	Poor/ disadvantage people, households, enterprises	LMICs
Girad et al. (2012) The effects of household food production strategies on the health and nutrition outcomes of women and young children: a systematic review	INNOVATION & NEW TECHNOLOGY Home gardens Crop production	FOOD SECURITY & NUTRITION Maternal, neonatal and young child nutrition Maternal, neonatal and young child health	Children 0-59 months of age Women of reproductive age	LMICs

Systematic review	Interventions	Outcomes	Population	Region
Gunaratna et al. (2010) A meta-analysis of community-based studies on quality protein maize	INNOVATION & NEW TECHNOLOGY Biofortification [QP Maize]	FOOD SECURITY & NUTRITION Child Growth Nutritional Status	Children under five with mild to moderate undernutrition	Global review
Hall et al. (2012 – final draft) What have been the farm-level economic impacts of the global cultivation of GM crops	INNOVATION & NEW TECHNOLOGY GM crops	INCOME / FINANCE Gross profit	Unspecified	Global review
IOB. (2011) Improving food security. A systematic review of the impact of interventions in agricultural production, value chains, market regulation, and land security.	INFRASTRUCTURE & INNOVATION AND NEW TECHNOLOGY Increasing agricultural production Developing the value chain Reforming market regulations Enhancing land tenure security	YIELD / PRODUCTIVITY & FOOD SECURITY & NUTRITION Level of food access Level of stability in food access	Food insecure people	Developing countries
Masset et al. (2011) A systematic review of agricultural interventions that aim to improve nutritional status of children	INNOVATION & NEW TECHNOLOGY Bio-fortification Home gardens Aqua culture Fish ponds Poultry development Animal husbandry Dairy development	FOOD SECURITY & NUTRITION Improvement in children's diets	Unspecified	LMICs
Rusinamhodzi et al. (2011) A meta-analysis of long-term effects of conservation agriculture on maize grain yield under rain-fed conditions	INNOVATION & NEW TECHNOLOGY Conservation agriculture, including: Crop rotation and intercropping Tillage management Residue retention	YIELD & PRODUCTIVITY Increased yields	Unspecified	Global review
Stewart et al. (2010) What is the impact of microfinance on poor people?	FINANCE Micro-credit Micro savings	INCOME / FINANCE Income Savings Expenditure Accumulation of assets	Poor people	Sub-Saharan Africa

Systematic review	Interventions	Outcomes	Population	Region
Stewart et al. (2012) Do micro-credit, micro-savings and micro-leasing as effective inclusion interventions, enabling poor people, especially women to engage in meaningful economic activities in low and middle income countries?	FINANCE Micro-credit Micro savings Micro-leasing	INCOME / FINANCE Engagement in economic opportunities Income Savings Expenditure Accumulation of assets	Poor people, especially women	LMICs
Waddington et al. (in press) Farmer field schools for improving farming practices and farmer outcomes in low and middle income countries	TRAINING Farmer field schools	YIELDS & PRODUCTIVITY Increased yields INCOME / FINANCE Increased profits	Smallholder farmers	Global review

Appendix 2: Detailed search strategy

Below we present as much detail as possible on our search strategy, including how we have adapted our generic search strings (in 1.1.1 below) for application into our specific sources.

2.1.1. Search strings to be adapted for each database

Smallholder farming:

(Poor NEAR/3 farm*) OR
(Poor NEAR/3 agricultur*) OR
(socioeconomic NEAR/3 farm*) OR
(socioeconomic NEAR/3 agricultur*) OR
("Low income" NEAR/3 farm*) OR
("Low income" NEAR/3 agricultur*) OR
(Subsistence NEAR/3 farm*) OR
(Subsistence NEAR/3 agricultur*) OR
("low fertilizer" NEAR/3 farm*) OR ("low fertiliser" NEAR/3 farm*) OR
("low fertilizer" NEAR/3 farm*) OR ("low fertiliser" NEAR/3 farm*) OR
("Small scale" NEAR/3 (agricultur*OR farm*) OR ("Small-scale" NEAR/3 agricultur* OR farm*) OR (Smallscale NEAR/3 agricultur* OR farm*) OR
(Impoverished NEAR/3 farm*) OR (Disadvantaged NEAR/3 farm*) OR ("food insecure" NEAR/3 farm*) OR TS=("small plot" NEAR/3 farm*) OR
(Impoverished NEAR/3 agricultur*) OR (Disadvantaged NEAR/3 agricultur*) OR ("food insecure" NEAR/3 agricultur*) OR TS=("small plot" NEAR/3 agricultur*) OR
("low input" NEAR/3 farm*) OR ("low labor" NEAR/3 farm*) OR ("low labour" NEAR/3 farm*) OR
("low input" NEAR/3 agricultur*) OR ("low labor" NEAR/3 agricultur*) OR ("low labour" NEAR/3 agricultur*) OR

(Small-hold* NEAR/3 farm*) OR (smallhold* NEAR/3 farm*) OR("smallhold*" NEAR/3 farm*) OR
(Small-hold* OR smallhold* OR"smallhold*") OR
(Peasant* near/3 farm*) OR
(Peasant* near/3 agricultur*) OR
("Small-holder agricultur*" OR "Smallholder agricultur*" OR "Small holder agricultur*") OR
("Smallscaleagricultur*" OR "Small scale agricultur*" OR "small-scale agricultur*") OR
("Subsistence agriculture") OR (subsistence NEAR/3 agricultur*) OR (subsistence NEAR/3 farm*) OR
("Low input agriculture")OR ("Low-input agriculture") OR
(agro-pastoral* OR agro pastoral* OR pastoral*ORagropastoral*)

Impact evaluation:

(impact OR outcome OR evaluation OR trial*OR comparison study OR non-comparison study OR social performance NEAR/3 assess* OR Imp-Act OR randomi*ed controlled trial OR controlled clinical trial OR placebo OR clinical trials OR random* OR controlled OR control group OR comparison group OR control group* OR comparison groups OR Intervention OR RCT OR experiment* OR program* evaluation OR "controls (experimental)" OR pilot scheme(s) OR Pilot study/ies OR pilot program* OR effectiveness NEAR/3 intervention* OR performance assessment OR time series OR before NEAR/2 after study OR comparative analysis OR Quasi-experiment* OR post-test* OR posttest* OR posttest* OR pre-test OR pretest OR pre test OR "participat* rural apprais*" OR performance apprais* OR project apprais*³ OR (random* NEAR/3 allocat*))

Training

("practical education" OR "extension education" OR "education program*" OR "community education" OR "agricultural education" OR "inservice training" OR "vocational training" OR "innovation adoption" OR "participatory extension" OR "agricultural advisory" OR "agricultural extension" OR "rural extension" OR course* OR class* OR lesson* OR teach* OR taught OR train* OR skill* OR adult w/5 educat* OR "adult learning" OR community w/5 educat* OR "Community learning" OR farmskills OR educating OR capacity building OR participatory learning OR "education* material*" OR "extension program*" OR "education* program*" OR "agricultural knowledge" OR "extension education" OR "technical knowledge" OR "technology transfer" OR "field school*" OR "farmer field school*")

Innovation and new technology

(innovation OR adoption OR "technological innovation*" OR "innovation technique*" OR "technical innovation*" OR "farming innovation" OR "agricultural technolog*" OR "agricultural biotechnolog*" OR "biotechnological innovation*" OR "new technolog*" OR "environmental technolog*" OR "agricultural innovation*" OR "agronomic innovation*" OR "social innovation*" OR "economic innovation*" OR "organizational innovation*" OR "management innovation*" OR "mechanical innovation*" OR "biological innovation*" OR "chemical innovation*" OR "process innovation*" OR "product innovation*" OR "local innovation*" OR "traditional innovation*" OR "Breeding technolog*" OR "innovative crop technolog*" OR "crop production technolog*" OR "plant and livestock breed*" OR "weed management" OR "storage technolog*" OR "post harvest management" OR "agro forestry" OR "cropping patterns" OR "soil conservation" OR "water harvest*" OR "soil and crop improvement" OR

³ Adapted from Stewart et al (2012)

“conservation agriculture” OR “conservation farm*” OR “pest management” OR “disease management” OR “farm machinery” OR “organic farming innovation” OR “crop management” OR “pest control technologies” OR “crop improvement” OR “crop production” OR “crop diversification” OR “crop protection” OR “water management” OR “livestock and fisheries management” OR “post harvest technolog* and value addition” OR irrigation OR fertilizer OR manure OR “water management” OR “water conservation” OR “water harvesting” OR “maize storage” OR “seed storage” OR “contract farming” OR “organic farming” OR “organic certification” OR “land certification” OR “household gardens” OR “urban agriculture” OR “soil fertility” OR “soil conservation” OR “tillage practices” OR “cropping patterns” OR “pest control” OR “weed control” OR “disease control” OR “export horticulture” OR biofortification OR “genetically modified crops” OR “seed varieties” OR “improved seeds” OR “improved agriculture” OR “improved technology”)

2.1.2 Africana Periodical Database search strategy

All searches will have the date limit applied of 1990 onwards. The database is coded on subject and some records, but not most, have an abstract. Therefore, the search will consist of searching the subject terms and supplementing these with title searches where applicable. Some searching of abstracts was applied in test searches, but these did not seem fruitful.

#1: Subject: Small farms

#2: Subject: Agricultural Projects

#3: Subject: Agriculture AND title: evaluation

#4: Subject: Agriculture AND title: performance

#5: Subject: Agriculture AND title: intervention

#6: Subject: Agriculture AND title: small-scale

2.1.3 African Women Bibliographic Database search strategy

The database is coded on subject and some records, but not most, have an abstract. Therefore, the search will consist of searching the subject terms and supplementing these with title searches where applicable. Some searching of abstracts was applied in test searches, and will be included where they seemed fruitful.

#1 Subject: Small farms

#2 Subject: Agricultural Projects

#3 Subject: Small enterprises

#4 Abstract: smallholders

#5 Abstract: small-scale

#6 Title: smallholders

#7 Title: trial AND subject: Agriculture

#8 Title: impact AND subject: Agriculture

#9 Title: evaluation AND subject: Agriculture

#10 Title: intervention AND subject: Agriculture

#11 Title: performance AND subject: Agriculture

#12 Title: Impact AND subject: Agriculture

2.1.4 Web of Science search strategy

We will search the following databases together:

Social Science Citation Index, Science Citation Index; Social Science Citation Index; Science Citation Index Expanded; Conference Proceedings Citation Index- Social Science & Humanities; Conference Proceedings Citation Index- Science.

The following search strings have been developed in test searches and will be employed combining the concepts of 'small-holder', 'intervention' and 'Africa'.

TS=((Poor NEAR/3 farm*) OR (Poor NEAR/3 agricultur*) OR (socioeconomic NEAR/3 farm*) OR (socioeconomic NEAR/3 agricultur*) OR ("Low income" NEAR/3 farm*) OR ("Low income" NEAR/3 agricultur*) OR (Subsistence NEAR/3 farm*) OR (Subsistence NEAR/3 agricultur*) OR ("low fertilizer" NEAR/3 farm*) OR ("low fertiliser" NEAR/3 farm*) OR ("low fertilizer" NEAR/3 farm*) OR ("low fertiliser" NEAR/3 farm*) OR ("Small scale" NEAR/3 (agricultur* OR farm*)) OR ("Small-scale" NEAR/3 (agricultur* OR farm*)) OR (Smallscale NEAR/3 (agricultur* OR farm*)) OR (Impoverished NEAR/3 farm*) OR (Disadvantaged NEAR/3 farm*) OR ("food insecure" NEAR/3 farm*) OR ("small plot" NEAR/3 farm*) OR (Impoverished NEAR/3 agricultur*) OR (Disadvantaged NEAR/3 agricultur*) OR ("food insecur*" NEAR/3 agricultur*) OR ("small plot" NEAR/3 agricultur*) OR ("low input" NEAR/3 farm*) OR ("low labor" NEAR/3 farm*) OR ("low labour" NEAR/3 farm*) OR ("low input" NEAR/3 agricultur*) OR ("low labor" NEAR/3 agricultur*) OR ("low

labour" NEAR/3 agricultur*) OR ("Small-hold*" OR smallhold* OR "small hold*") OR (Peasant* near/3 farm*) OR (Peasant* near/3 agricultur*) OR ("Small-holder agricultur*" OR "Smallholder agricultur*" OR "Small holder agricultur*") OR ("Smallscale agricultur*" OR "Small scale agricultur*" OR "small-scale agricultur*") OR ("Subsistence agriculture") OR (subsistence NEAR/3 agricultur*) OR (subsistence NEAR/3 farm*) OR ("Low input agricultur") OR ("Low-input agricultur") OR ("agro-pastoral*" OR "agro pastoral*" OR agropastoral* OR pastoral*))

TS=(impact OR outcome OR evaluation* OR trial* OR "comparison study" OR "non-comparison study" OR ("social performance" NEAR/3 assess*) OR "Imp-Act" OR placebo OR random* OR controlled OR "control group" OR "comparison group" OR "control group*" OR "comparison groups" OR Intervention* OR RCT OR RCTs OR experiment* OR (program* NEAR/1 evaluation) OR "controls (experimental)" OR "pilot scheme*" OR "pilot study" OR "pilot studies" OR "pilot program*" OR (effectiveness NEAR/3 intervention*) OR "performance assessment*" OR "time series" OR (before NEAR/2 "after study") OR "comparative analysis" OR "Quasi-experiment*" OR "post-test*" OR posttest* OR "post test*" OR "pre-test" OR pretest OR "pre test" OR (participat* NEAR/1 "rural apprais*") OR "performance apprais*" OR "project apprais*")

TS=(Africa* OR sahara* OR sub-sahara* OR Maghreb* OR sahel* OR rift valley OR Swahili* OR Fula* OR Mandinka* OR Balanta* OR Papel* OR Manjaco* OR Mancanha* OR Bantu* or centrafricaine or Baya* or Banda* or Ovimbundu* OR ambundu* OR Hutu* OR Tutsi* OR Kikuyu* OR Luhya* OR Dahomey OR batswana* OR Somali* OR Benadiris* OR Burkina* OR Mosotho* OR Basotho* OR Nubia* OR Swazi* OR Ngwane OR Swatini OR Nyasaland OR Kongo* OR Malian* OR Tchad OR Chadian* OR M?uritani* OR togo OR Togolese OR ivo?rian OR Mo?ambi* OR Erythree OR caboverde* OR KabuVerd* OR Kabuverdianu OR Madagas* OR Malagas* OR Algeria* OR Angola* OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR Camero* OR "Canary Island" OR "Canary Islands" OR Cape Verde* OR Chad OR Comor* OR Congo* OR "Democratic Republic of Congo" OR DRC OR Djibouti* OR "Equatorial Guinea" OR Eritrea* OR Ethiopia* OR Egypt* OR Gabon* OR Gambia* OR Ghana* OR Guinea OR "Guinea Bissau" OR Bissau* OR "Ivory Coast" OR "Cote d'Ivoire" OR Kenya* OR Lesotho OR Liberia* OR Libya* OR Madagasca* OR Malawi* OR Mali OR Mauritania* OR Mauriti* OR Mayot* OR Morocc* OR Mozambiq* OR Mocambiq* OR Namibi* OR Niger* OR Nigeria* OR Principe* OR Reunion* OR Rwanda* OR "Sao Tome" OR Senegal* OR Seychelles OR "Sierra Leone" OR Somali* OR "St Helena" OR "saint Helena" OR Sudan* OR Swazi* OR Tanzania* OR Togo OR Uganda* OR "Western Sahara" OR Zaire OR Zambia* OR Zimbabwe)

TS=((developing or "under developed" or underdeveloped or "middle income" or underserved or "under served" or deprived or poor*) NEAR/1 (countr* or nation OR nations or population or populations or world))

TS=((developing or "under developed" or underdeveloped or "middle income") NEAR/1 (economy or economies))

TS=(lmic or lmics or "third world" or "lami countr*" OR "transitional countr*")

TS=((("less developed" NEAR/1 (countr* or nation OR nations or population or populations or world or economy or economies)) OR ("lesser developed" NEAR/1 (countr* or nation OR nations or population or populations or world or economy or economies)))

TS=("low* income" NEAR/1 (countr* or nation OR nations or population or populations or world or economy or economies))

SO=(Africa* OR sahara* OR sub-sahara* OR Maghreb* OR sahel* OR rift valley OR Swahili* OR Fula* OR Mandinka* OR Balanta* OR Papel* OR Manjaco* OR Mancanha* OR Bantu* or centrafricaine or Baya* or Banda* or Ovimbundu* OR ambundu* OR Hutu* OR Tutsi* OR Kikuyu* OR Luhya* OR Dahomey OR batswana* OR Somali* OR Benadiris* OR Burkina* OR Mosotho* OR Basotho* OR Nubia* OR Swazi* OR Ngwane OR Swatini OR Nyasaland OR Kongo* OR Malian* OR Tchad OR Chadian* OR M?uritani* OR togo OR Togolese OR ivo?rian OR Mo?ambi* OR Erythree OR caboverde* OR KabuVerd* OR Kabuverdianu OR Madagas* OR Malagas* OR Algeria* OR Angola* OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR Camero* OR "Canary Island" OR "Canary Islands" OR Cape Verde* OR Chad OR Comor* OR Congo* OR "Democratic Republic of Congo" OR DRC OR Djibouti* OR "Equatorial Guinea" OR Eritrea* OR Ethiopia* OR Egypt* OR Gabon* OR Gambia* OR Ghana* OR Guinea OR "Guinea Bissau" OR Bissau* OR "Ivory Coast" OR "Cote d'Ivoire" OR Kenya* OR Lesotho OR Liberia* OR Libya* OR Madagasca* OR Malawi* OR Mali OR Mauritania* OR Mauriti* OR Mayot* OR Morocc* OR Mozambiq* OR Mocambiq* OR Namibi* OR Niger* OR Nigeria* OR Principe* OR Reunion* OR Rwanda* OR "Sao Tome" OR Senegal* OR Seychelles OR "Sierra Leone" OR Somali* OR "St Helena" OR "saint Helena" OR Sudan* OR Swazi* OR Tanzania* OR Togo OR Uganda* OR "Western Sahara" OR Zaire OR Zambia* OR Zimbabwe)

TS=((("integrated control" OR "integrated production" OR "integrated management" OR "integrated pest" OR "integrated nutrient" OR "crop management") OR ("practical education" OR "extension education" OR "education program*" OR "community education" OR "agricultural education" OR "inservice training" OR "vocational training" OR "innovation adoption" OR "participatory extension" OR "agricultural advisory" OR "agricultural extension" OR "rural extension" OR course* OR class* OR lesson* OR teach* OR taught OR train* OR skill* OR (Adult NEAR/5 educat*) OR "Adult learning" OR (community NEAR/5 educat*) OR "Community learning"))

TS=(microfinance OR micro-finance OR micro finance OR micro-lease OR micro-insurance OR micro insurance OR micro-savings OR micro savings OR micro lease OR microlease OR microinsurance OR microsavings OR microfranchise OR microfranchis* OR micro-franchise OR micro-franchis* OR micro-enterprise OR microenterprise OR microleasing OR

micro-leasing OR micro-banking OR micro-banks OR micro-business* OR microinsurance OR micro-insurance OR (banking NEAR/10 development) OR (banks NEAR/10 development) OR (bank NEAR/10 development) OR (savings NEAR/10 development) OR (lease NEAR/10 development) OR (finance NEAR/10 development) OR (banking NEAR/10 poverty) OR (banks NEAR/10 poverty) OR (bank NEAR/10 poverty) OR (savings NEAR/10 poverty) OR (lease NEAR/10 poverty) OR (banking NEAR/10 "the poor") OR (banks NEAR/10 "the poor") OR (bank NEAR/10 "the poor") OR (savings NEAR/10 "the poor") OR (lease NEAR/10 "the poor") OR (finance NEAR/10 "the poor") OR (finance NEAR/10 poverty) OR ((finance OR insurance OR savings) AND poverty))

TS=("seed bank*" OR "colostrum bank*" OR "river bank*" OR "blood bank*" OR "fishes" OR genetic)

TS=((insurance AND (crop OR weather OR rain* OR index* OR climat* OR precipitation)) OR (risk AND crop) OR (risk AND weather) OR (risk AND rain*) OR (risk AND index*) OR (risk AND climat*) OR (risk AND precipitation))

2.1.5 CAB Abstracts search strategy

Notes

Searches will be limited to those published on or after 1st January 1990.

The mode 'Apply related words' will be used for all searches.

Terms to be searched

#1: AB (agricultur* OR farm*) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#2: AB (poor OR socioeconomic OR low income OR subsistence) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#3: AB (low fertiliser OR low fertilizer OR small scale OR smallscale OR small-scale) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#4: AB (impoverished OR disadvantaged OR food insecure OR small plot OR low input OR low labor OR low labour) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#5: AB (small hold* OR smallhold* OR small-hold* OR small scale OR smallscale OR small-scale OR subsistence OR low-input) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#6: AB (agropastoral OR agro-postoral OR agro pastoral OR pastoral) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#7: #2 OR #3 OR #4 OR #5 OR #6

#8: #1 AND #7

#9: AB (comparison study OR non-comparison study OR social performance OR impact OR outcome OR evaluation OR trial* OR Imp-act OR randomised control trial OR clinical trial* OR random* OR controlled OR controlled group* OR control group OR comparison group OR intervention OR rct) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#10: AB (assess OR program* evaluation OR pilot scheme* OR pilot stud* OR pilot program* OR performance assess* OR time series OR comparative analys* OR quasi-experiment OR posttest OR post test OR pre test OR pretest OR participant* rural apprais* OR performance apprais* OR project apprais*) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#11: AB (effectiveness AND intervent*) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#12: AB (before AND after) NOT AB America NOT AB China NOT AB India NOT AB Brazil NOT AB Bangladesh NOT AB Pakistan

#13: #9 OR #10 OR #11 OR #12

#14: #8 AND #13

2.1.6 IFAD database search strategy

Terms to be searched

Poor AND farm\$; poor AND agriculture\$; socioeconomic AND farm\$; socioeconomic AND agriculture\$; "low income" AND farm\$; "low income" AND agriculture\$; subsistence AND farm\$; subsistence AND agriculture\$; "low fertilizer" AND farm\$; "low fertiliser" AND farm\$; "low fertilizer" AND farm\$; "low fertiliser" AND farm\$; "small scale" AND agricultur\$; "small-scale" AND farm\$; smallscale AND agricultur4; smallscale AND farm\$; impoverished AND farm\$; disadvantaged AND farm\$; "food insecure" AND farm\$; "small plot" AND farm\$; impoverished AND agriculture\$; disadvantaged AND agriculture\$; "food insecure" AND agriculture\$; "small plot" AND agriculture\$; "low input" AND farm\$; "low labor" AND farm\$; "low labour" AND farm\$; "low input" AND agriculture\$; "low labor" AND agriculture\$; "low labour" AND agriculture\$; small-hold\$ AND farm\$; smallhold\$

AND farm\$; "smallhold\$" AND farm\$; small-hold\$; smallhold\$; "smallhold\$"; peasant\$ AND farm\$; peasant\$ AND agriculture\$; "small-holder agriculture\$"; "smallholder agriculture\$"; "small holder agriculture\$"; "smallscaleagriculture\$"; "small scale agriculture\$"; "small-scale agriculture\$"; "subsistence agriculture\$"; subsistence AND agriculture\$; subsistence AND farm\$; "low input agricultur\$"; "low-input agriculture\$"; agro-pastoral\$; agro pastoral\$; pastoral\$; agropastoral\$.

2.1.7 AGRIS database search strategy

Search string

+(+agrovoc:"Africa" +("small farms"~3 "small agriculture"~3 subsistence "small scale"~2 "low input" "low labour" "low labor" "peasant farming") +(intervention* impact* effectiveness pilot trial* random* apprais* performance "controlled study" comparison assessment experiment* program*)) +date:[1990 TO 2013]Social Science Citation Index and Science Citation Index

2.1.8 3ie impact database search strategy

Given the limited scope for sophisticated searches within this database, we will screen all 789 entries.

2.1.9 Citation searches will be conducted for the following reviews and impact evaluations

Systematic reviews

- Berti RP, Krasevec J, FitzGerald F (2004) A review of the effectiveness of agriculture interventions in improving nutrition outcomes. *Public Health Nutrition*,7 (5): 599-609.
- Girard AW, Self JL, McAuliffe C, Oludea O (2012) The Effects of Household Food Production Strategies on the Health and Nutrition Outcomes of Women and Young Children: A Systematic Review. *Paediatric and Perinatal Epidemiology* 26(Suppl. 1), 205–222.
- Gunaratna NS, De Groot H, Nestel P, Pixley KV, McCabe GP (2010) A meta-analysis of community-based studies on quality protein maize, *Food Policy*, (35): 202–210

- Hall C, Knight B, Ringrose S, Knox O (2012) *What have been the farm-level economic impacts of the global cultivation of GM crops?* Systematic Review No.CEE 11-002.
- IOB (2011) *Improving food security. A systematic review of the impact of interventions in agricultural production, value chains, market regulation, and land security.* IOB Study No 363.
- 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.
- Waddington H, Snilstveit B, Hombrados J, Vojtkova M, White H (2013) *Farmer Field Schools for improving farming practices and farmer outcomes in low and middle-income countries: a systematic review.* Campbell Systematic Reviews.

Impact evaluations

- Abonesh Tesfaye, Ayalneh Bogale, Namara R E; Dereje Bacha (2008) The impact of small-scale irrigation on household food security: the case of Filtino and Godino irrigation schemes in Ethiopia. *Irrigation and Drainage Systems.* 22: 145-158.
- Asfaw S, Shiferaw B, Simtowe F, Lipper L (2012) Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia. *Food Policy.* 37(3): 283-295.
- Ashraf N, Gine X, Karlan D (2008) Finding missing markets (and a disturbing epilogue): evidence from an export crop adoption and marketing intervention in Kenya.. Washington DC: World Bank.
- Bennett R, Buthelezi T J; Ismael Y, Morse S (2003) Bt cotton, pesticides, labour and health: a case study of smallholder farmers in the Makhathini Flats, Republic of South Africa. *Outlook on Agriculture.* 32: 123-128.
- Dercon S, Gilligan D O; Hoddinott J, Woldehanna T (2008) The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages. : International Food Policy Research Institute (IFPRI).
- Katrak H (2006) Better Health, More Wealth: The Impacts of Farmer Training in Developing Countries.. *Pesticide News.* 73: 18-20. Searching reference lists of included studies.
- Friis-Hansen E (2008) Impact assessment of farmer institutional development and agricultural change: Soroti district, Uganda. *Development in Practice.* 18: 506-523.
- Low JW, Arimond M, Osman N, Cunguara B, Zano F, Tschirley D (2007) 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.
- Smale M, Mathenge M K; Jayne T S; Magalhaes E, Olwande J, Kiriimi L, Kamau M, Githuku J (2012) Income and poverty impacts of USAID-funded programs to

promote maize, horticulture, and dairy enterprises in Kenya, 2004-2010. East Lansing; USA: Michigan State University, Department of Agricultural Economics.

- Todo Y, Takahashi R (2011) Impact of Farmer Field Schools on Agricultural Income and Skills: Evidence from an Aid-Funded Project in Rural Ethiopia. JICA-RI Working Paper, No. 30, May 2011.
- World-Bank (2007). Project performance assessment report Ethiopia. Seed Systems Development Project. National Fertilizer Sector Project. Independent Evaluation Group, Washington DC.

2.1.10 Requesting relevant studies from key contacts

We will write to key contacts requesting relevant studies. These will include members of our project advisory group and first authors of relevant reviews, as listed below.

Advisory group members

Adelina Mensah	University of Ghana
Birte Snilsveit	3ie
Constanza Di Nucci	IFAD
David Rohrbach	World Bank
Karen Nortje	CSIR
Marzia Perilli	IFAD
Nikita Eriksen-Hamel	DFATD
Phiko Kavinya	Ministry of Agriculture and Food Security (Malawi)
Samuel Amanquah	AGRA

First authors of overlapping reviews identified in our Review of Reviews

Peter Berti, lead author of Berti et al. 2004. A review of the effectiveness of agriculture interventions in improving nutrition outcomes

Amy Girard, lead author of Girard et al. 2012. The Effects of Household Food Production Strategies on the Health and Nutrition Outcomes of Women and Young Children: A Systematic Review.

Nilupa Gunaratna, lead author of Gunaratna NS et al. 2010 A meta-analysis of community-based studies on quality protein maize.

Clare Hall, lead author of Hall et al. 2012 What have been the farm-level economic impacts of the global cultivation of GM crops?Systematic Review.

Dr Ferko Bodnár and Dr Bart de Steenhuijsen Piters (Royal Tropical Institute, KIT - Amsterdam), who led the IOB review, 2011. Improving food security.A systematic review of the impact of interventions in agricultural production, value chains, market regulation, and land security.

Edoardo Masset, lead author of Masset et al 2011. A systematic review of agricultural interventions that aim to improve nutritional status of children.

Hugh Waddington, lead author of Waddington et al. 2013. Farmer Field Schools for improving farming practices and farmer outcomes in low and middle-income countries: a systematic review).

Additional contacts recommended in previous feedback on this project

Ken Giller , Theoretical Production Group at Wageningen University

Professor Milla Mclachlan, The Food Security Project, Stellenbosch University

Circulation of requests for reports of impact evaluation through our advisory group members at IFAD.

Appendix 3: Coding Sheet

1. General information

- 1.1. Coded by:
- 1.2. Date coded:
- 1.3. Checked by:
- 1.4. Study title:
- 1.5. Author/s name:
- 1.6. Date study published:
- 1.7. Country / countries of intervention:

2. Confirming study as included or excluded

2.1. Was the data collected on or after 1990?

- Yes: include
- No: exclude (list date)

2.2. Study design

Are before-intervention / after-intervention dates reported?

- Yes
- No; exclude
- Not sure, please check

Are intervention AND control groups identified?

- Yes
- No; exclude
- Not sure; please check

2.3. Interventions

- New technology / innovation
- Training
- Not relevant intervention; exclude.

2.4. Outcomes

- Food security
- Financial wealth
- Not relevant outcome; exclude.

Include
 Exclude
(provide reason):

Please check
(provide reason):

References to
collect

2.5. Target population

Does the target population include smallholder farmers?

- No; exclude.
- Uses the term 'smallholder farm' but is not defined
- Yes; describe AND tick all that apply (if information is available):
 - Farmers who have a limited size of farm; specify size:
 - Farmers who are mostly dependent on family labour
 - Farmers who practice subsistence farming or mix of subsistence and market-oriented farming
 - Farmers who have limited resources in terms of land, technical and technological support, and/or capital for maintenance and investment.
 - Farmers who are young farmers (under the age of 20)
 - Farmers who are female farmers
 - Farmers who are landless labourers
 - Other; describe:

3. Describing the study

3.1. Interventions

3.1.1. New innovation & technology

a) How do they describe the intervention (provide as much detail as possible):

b) Describe the type of innovation (tick all that apply):

- Process innovation (a way to modify a gene in a plant)

- Product innovation (new varieties of vegetables or potatoes)**

- Mechanical (tractors, roads, water, irrigation)**

- Biological and chemical (new seed varieties, chemical, fertilisers and pesticides)**

- Agronomic (new management practices)**

- Organisation of inputs (seed, fertiliser, pesticides)**
- Organisation of output markets (diversification, processing, trade)**

- Biotechnological (computer technologies)**

- Other; describe:**

3.1.2. Training & knowledge

a) Type of training:

- Farmer Field Schools
- Other; describe:

b) Was the training experiential or participatory?

- Fully participatory designed to empower farmers and provide experiential learning
- Partly participatory with limited experience provided
- Limited participation by farmers with didactic teaching approaches

c) How long did the training last?

d) What was the content of the training?

- A new technology or innovation
- Other; describe:

3.1.3. Were the interventions delivered alongside other interventions?

- No
- Yes (specify):

3.1.4. Intervention components

a) What was the main element of intervention (tick one):

- Genetic improvements
- Environmental improvements
- Improved management methods
- Knowledge transfer
- Access to assets
- Other; describe

b) List any secondary elements (tick all that apply)

- None
- Genetic improvements
- Environmental improvements
- Improved management methods
- Knowledge transfer
- Access to assets
- Other; describe

3.1.5. Number of participants:

a) Unit of measurement of participants (describe):

b) List numbers in:

- Control group:
- Intervention group:
- Total number of people evaluated:
- Other; describe:

3.1.6. Scale of intervention

- Regional level (intervention delivered across a group of countries)
- National level (intervention delivered within a country)
- Intermediate level (intervention delivered within provinces within a country)
- Local level (intervention delivered within a group of villages)
- Individual level

3.1.7. What was the duration of intervention? (Describe):

3.1.7. Who delivered the intervention?

- Limited reporting

3.1.8. Who funded the study?

- Not sure

3.2. Outcomes

3.2.1. Food security and nutrition:

- Household food consumption by weight
- Per capita calorific intake
- Household perceptions of food security
- Other; describe:

3.2.2. Financial wealth (tick all that apply)

- Household income
- Household accumulation of financial assets
- Household accumulation of non-financial assets
- Household food expenditure
- Other; describe:

3.2.3. Intermediate outcomes

- Investment in capital
- Knowledge transfer
- Adoption of innovation
- Diffusion of innovation
- Yield
- Productivity
- Other; describe:

3.2.4. How was each of the outcomes measured (what did they measure and how did they measure it?)

Outcome 1 (name):

- Not reported
- Reported; describe:

Outcome 2 (name):

- Not reported
- Reported; describe:

Outcome 3 (name):

- Not reported
- Reported; describe:

Outcome 4 (name):

- Not reported
- Reported; describe:

Outcome 5 (name):

- Not reported
- Reported; describe:

3.2.5. When were the outcomes measured?

- Not reported
- Reported; describe:

4. Reference list checked?

- Yes, no additional references.
- Yes, additional references (number of studies identified):
 - Reference list printed and study details written on list
 - In pile 'to find'

Appendix 4: Data extraction for Effect Size Calculation

Effect size data		
	Which page(s) contain the effect size data?	Open answer
	Sample size unit of analysis	1= Individual Farmers 2= Farmer Groups 3= Clusters (e.g. Villages, Districts) 4= Other 5= Not clear
	Initial sample size treatment group	#
	Initial sample size control group	#
	Number of drop-outs	#
	Number of drop-outs	#
	Number of treatment observations after attrition (individuals)	#
	Number of control observations after attrition (individuals)	#
	What treatment effect is estimated?	1=ITT 2=ATE 4=LATE
Outcomes - continuous	Does the study give a precise definition of outcome X?	1=Yes 2=No 3=Partially
	What definition of outcome x given	Open answer
	State result of baseline outcome for treatment group	#
	State SD of baseline outcome measure for treatment group	#
	State sample size at baseline	#
	State result of baseline outcome for control group	#
	State SD of baseline outcome measure for contol group	#

	State sample size at baseline	#
	State result of post intervention outcome for treatment group	#
	State SD of post intervention outcome measure for treatment group	#
	State sample size post intervention	#
	State result of post intervention outcome for control group	#
	State SD of post intervention outcome measure for control group	#
	State sample size post intervention	#
	State result of 1st follow up outcome measure for treatment group	#
	State SD 1st follow up outcome measure for treatment group	#
	State sample size first follow up	#
	State result of 1st follow up outcome measure for treatment group	#
	State SD 1st follow up outcome measure for treatment group	#
	State sample size first follow up	#
	Repeat the above for any additional follow up measures	
Outcomes - dichotomous	Does the study give a precise definition of outcome X?	1=Yes 2=No 3=Partially
	What definition of outcome x given	Open answer
	State result of baseline outcome for treatment group	#
	State sample size at baseline	#
	State proportion with outcome at baseline in treatment	#
	State result of baseline outcome for treatment group	#
	State sample size at baseline	#
	State proportion with outcome at baseline in control	#
	State number with outcome post intervention for treatment group	#
	State sample size for treatment group post intervention	#

	State proportion with outcome post intervention in control group	#
	State number with outcome post interventionfor control group	#
	State sample size for control group post intervention	#
	State proportion with outcome post intervention in control group	#
	State number with outcome at 1st follow up for treatment group	#
	State sample size at 1st follow up for treatment group	#
	State proportion with outcome at 1st follow up in treatment group	#
	State number with outcome at 1st follow up for control group	#
	State sample size at for control group at 1st follow up	#
	State proportion with outcome at 1st follow up in control group	#
	Repeat the above for any additional follow up measures	
Sub groups	Does the study conduct sub group analysis	1=Yes 2=No
	State any sub-groups for which the study includes outcome measures	
	Extract data necessary to calculate effect sizes for each outcome where sub-group analysis is conducted	

Appendix 5: Risk of Bias Tool

Tool for assessing risk of bias (ROB)

This tool is closely based on the Cochrane Collaboration's new tool for assessing ROB in non-randomised studies, currently being piloted by their ROB Methods Group. For more information, contact Julian Higgins. (Julian Higgins julian.higgins@bristol.ac.uk)

An overview 'form' is provided, with more detailed guidance provided in Section 2.

Study details	
Study title	
Authors, Year Published	
Population targeted	
Main Intervention	
Other Interventions	1.
	2.
	3.
Outcomes	1.

	2.
	3.
	4.
Range of dates the study was conducted	

Section 1: Risk of Bias (RoB) Tool

Bias due to confounding		
Outcome 1	<i>Screening question: Is confounding potentially controllable in the context of this study?</i>	Description Rationale for RoB judgment
	1.1.1 Did the authors conduct an appropriate analysis that controlled for all the critically important confounding domains?	
	1.1.2 If yes to 1.1.1: Were all of the confounding domains measured validly and reliably by the variables adjusted for in this study?	
	1.1.3. Did the authors avoid adjusting for post-intervention variables?	
Risk of bias judgment		
Outcome 2	<i>Screening question: Is confounding potentially controllable in the context of this study?</i>	Description Rationale for RoB judgment
	1.2.1. Did the authors conduct an appropriate analysis that controlled for all the critically important confounding domains?	
	1.2.2. If yes to 1.2.1: Were all of the confounding domains measured validly and reliably by the variables adjusted for in this study?	
	1.2.3. Did the authors avoid adjusting for post-intervention variables?	

Risk of bias judgment		
Outcome 3	<i>Screening question: Is confounding potentially controllable in the context of this study?</i>	Description Rationale for RoB judgment
	1.3.1. Did the authors conduct an appropriate analysis that controlled for all the critically important confounding domains?	
	1.3.2. If yes to 1.3.1: Were all of the confounding domains measured validly and reliably by the variables adjusted for in this study?	
	1.3.3. Did the authors avoid adjusting for post-intervention variables?	
Risk of bias judgment		

Bias in selection of participants into the study		
Outcome 1		Description Rationale for RoB judgment
	2.1.1 Do start of follow-up and start of intervention coincide?	
	2.1.2. If no to 2.1.1: Were adjustment techniques used that are likely to correct for the presence of selection biases?	
Risk of bias judgment		
Outcome 2		Description Rationale for RoB judgment
	2.2.1 Do start of follow-up and start of intervention coincide?	
	2.2.2. If no to 2.2.1: Were adjustment techniques used that are likely to correct for the presence of selection biases?	
Risk of bias judgment		
Outcome 3		Description Rationale for RoB judgment
	2.3.1 Do start of follow-up and start of intervention coincide?	
	2.3.2. If no to 2.3.1: Were adjustment techniques used that are likely to correct for the presence of selection biases?	
Risk of bias judgment		
Bias due to departures from intended interventions		
Outcome 1	<i>Screening question: Were the intended interventions sufficiently clearly defined and implemented such that a reasonable comparison of them can be made?</i>	Description Rationale for RoB judgment
	3.1.1 Were the critical co-interventions balanced across intervention groups?	

	3.1.2. Were treatment switches low enough that they do not threaten the validity of the estimated effect of intervention?	
	3.1.3. Was implementation failure minor and unlikely to threaten the validity of the outcome estimate?	
	3.1.4. If no to 3.1.1, 3.1.2 or 3.1.3: Were adjustments techniques used that are likely to correct for switches, unbalanced co-intervention and implementation failure?	
Risk of bias judgment		
Outcome 2	<i>Screening question: Were the intended interventions sufficiently clearly defined and implemented such that a reasonable comparison of them can be made?</i>	Description Rationale for RoB judgment
	3.2.1 Were the critical co-interventions balanced across intervention groups?	
	3.2.2. Were treatment switches low enough that they do not threaten the validity of the estimated effect of intervention?	
	3.2.3. Was implementation failure minor and unlikely to threaten the validity of the outcome estimate?	
	3.2.4. If no to 3.2.1, 3.2.2 or 3.2.3: Were adjustments techniques used that are likely to correct for switches, unbalanced co-intervention and implementation failure?	
Risk of bias judgment		

Outcome 3	<i>Screening question: Were the intended interventions sufficiently clearly defined and implemented such that a reasonable comparison of them can be made?</i>	Description Rationale for RoB judgment
	3.3.1 Were the critical co-interventions balanced across intervention groups?	
	3.3.2. Were treatment switches low enough that they do not threaten the validity of the estimated effect of intervention?	
	3.3.3. Was implementation failure minor and unlikely to threaten the validity of the outcome estimate?	
	3.3.4. If no to 3.3.1, 3.3.2 or 3.3.3: Were adjustments techniques used that are likely to correct for switches, unbalanced co-intervention and implementation failure?	
Risk of bias judgment		

Bias due to missing data		
Outcome 1	<i>Screening question: Are the intervention groups free of critical differences in participants with missing data?</i>	Description Rationale for RoB judgment
	4.1.1. Are outcome data reasonably complete?	
	4.1.2. Was intervention status reasonably complete for those in whom it was sought?	
	4.1.3. Are data reasonably complete for other variables in the analysis?	
	4.1.4. If no to 4.1.1, 4.1.2 or 4.1.3: Are proportion of participants and reasons for missing data similar across interventions?	
	4.1.5. If no to 4.1.1, 4.1.2 or 4.1.3: Were appropriate statistical methods used to account for missing data?	
Risk of bias judgment		
Outcome 2	<i>Screening question: Are the intervention groups free of critical differences in participants with missing data?</i>	Description Rationale for RoB judgment
	4.2.1. Are outcome data reasonably complete?	
	4.2.2. Was intervention status reasonably complete for those in whom it was sought?	
	4.2.3. Are data reasonably complete for other variables in the analysis?	
	4.2.4. If no to 4.2.1, 4.2.2 or 4.2.3: Are the proportion of participants and reasons for missing data similar across interventions?	
	4.2.5. If no to 4.2.1, 4.2.2 or 4.2.3: Were appropriate statistical methods used to account for missing data?	
Risk of bias judgment		

Outcome 3	<i>Screening question: Are the intervention groups free of critical differences in participants with missing data?</i>	Description Rationale for RoB judgment
	4.3.1. Are outcome data reasonably complete?	
	4.3.2. Was intervention status reasonably complete for those in whom it was sought?	
	4.3.3 Are data reasonably complete for other variables in the analysis?	
	4.3.4. If no to 4.3.1, 4.3.2 or 4.3.3: Are the proportion of participants and reasons for missing data similar across interventions?	
	4.3.5. If no to 4.3.1, 4.3.2 or 4.3.3: Were appropriate statistical methods used to account for missing data?	
Risk of bias judgment		
Bias in measurement of outcomes or interventions		
Outcome 1		Description Rationale for RoB judgment
	5.1.1. Were outcome assessors unaware of the intervention received by study participants?	
	5.1.2. Was the outcome measure objective?	
	5.1.3. Were the methods of outcome assessment comparable across intervention groups?	
	Risk of bias judgment	
Outcome 2		Description Rationale for RoB judgment
	5.2.1. Were outcome assessors unaware of the intervention received by study participants?	
	5.2.1. Was the outcome measure objective?	

	5.2.3. Were the methods of outcome assessment comparable across intervention groups?	
Risk of bias judgment		
Outcome 3		Description Rationale for RoB judgment
	5.3.1. Were outcome assessors unaware of the intervention received by study participants?	
	5.3.2. Was the outcome measure objective?	
	5.3.3. Were the methods of outcome assessment comparable across intervention groups?	
	Risk of bias judgment	

Bias in selection of result reported		
Outcome 1		Description Rationale for RoB judgment
	6.1.1. Is it unlikely that the reported effect estimate is available primarily because it was a notable finding among numerous exploratory analyses?	
	6.1.2. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among multiple outcome measurements within the outcome domain?	
	6.1.3. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among multiple analyses of the outcome measurements?	
	6.1.4. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among different subgroups?	
Risk of bias judgment		
Outcome 2		Description Rationale for RoB judgment
	6.2.1. Is it unlikely that the reported effect estimate is available primarily because it was a notable finding among numerous exploratory analyses?	
	6.2.2. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among multiple outcome measurements within the outcome domain?	
	6.2.3. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among multiple analyses of the outcome measurements?	
	6.2.4. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among different subgroups?	
Risk of bias judgment		

Outcome 3		Description Rationale for RoB judgment
	6.1. Is it unlikely that the reported effect estimate is available primarily because it was a notable finding among numerous exploratory analyses?	
	6.2. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among multiple outcome measurements within the outcome domain?	
	6.3. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among multiple analyses of the outcome measurements?	
	6.4. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among different subgroups?	
Risk of bias judgment		
Overall risk of bias judgment		Description Rationale for RoB judgment

SECTION 2

1. Bias due to baseline confounding

Screening question

Is confounding potentially controllable in the context of this study?

(If 'No', go straight to judgment of 'Critical risk of bias')

Preliminary considerations

- a. Within each confounding domain listed in the review protocol, list the relevant variables, if any, measured in this study.

--

- b. List additional confounding domains, if any, specific to the setting of this particular study. Within each domain, list the relevant variables, if any, measured in this study

--

- c. List additional domains and corresponding measured variables, if any, that the study authors identified as potential confounders that are not included in the above domains

--

Relationship between confounding domains and potential confounders.

In the table below, “critically important” confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. “Validity” refers to whether the confounding variable or variables fully measure the domain, while “reliability” refers to the precision of the measurement (more measurement error means less reliability).

Confounding Domain	Is the domain critically important	Measured Variables	Did the authors demonstrate that controlling for this variable was unnecessary?*	Is the domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is adjusting for this variable (alone) expected to move the effect estimate up or down? **

Signaling Question	Rationale/Remark	
--------------------	------------------	--

1.1. Did the authors conduct an appropriate analysis that controlled for all the critically important confounding domains?	Appropriate analyses to adjust for measured confounders include stratification, regression, matching, standardization, and inverse probability weighting. They may adjust for individual variables or for the estimated propensity score. Inverse probability weighting is based on a function of the propensity score. Each method depends assuming that there is no unmeasured or residual confounding.	
1.2. If yes to 1.1, Were all of the confounding domains measured validly and reliably by the variables adjusted for in this study?	Appropriate control of confounding requires that the variables used are valid and reliable measures of the confounding domains. For some topics, a list of valid and reliable measures of confounding domains will be specified in the review protocol but for others such a list may not be available. Study authors may cite references to support the use of a particular measure. If authors control for confounding variables with no indication of their validity or reliability pay attention to the subjectivity of the measure. Subjective measures (e.g. based on self-report) may have lower validity and reliability than objective measures such as lab findings.	
1.3. Did the authors avoid adjusting for post-intervention variables?	Adjusting for post-intervention variables is not appropriate. Adjusting for mediating variables (those on the causal pathway from intervention to outcome) restricts attention to the effect of intervention that does not go via the mediator (the "direct effect") and may introduce confounding, even for RCTs. Adjusting for common effects of intervention and outcome causes bias.	

Risk of Bias Judgment		
Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	No confounding expected.	
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):	Confounding expected, all known critically important domains appropriately measured and adjusted for. <i>and</i> Reliability and validity of measurement of a critically important domain were sufficient that we do not expect serious residual confounding.	
Serious risk of bias (the study has some important problems);	At least one known critically important domain not appropriately	

	<p>measured, or not adjusted for.</p> <p><i>or</i></p> <p>Reliability or validity of measurement of a critically important domain was low enough that we expect serious residual confounding.</p>	
Critical risk of bias (the study is too problematic to provide any useful evidence);	Confounding inherently not controllable, or use of negative controls strongly suggests unmeasured confounding.	
No information on which to base a judgment about risk of bias for this domain.	Confounding expected, but no information on how or whether it is addressed in the reported result.	

2. Bias in selection of participants into the study

Signaling Question	Rationale/Remark	Key Variations
2.1. Do start of follow-up and start of intervention coincide?	If subjects are not followed from the start of the intervention then a period of follow up has been excluded, and individuals who experienced the outcome soon after intervention will be missing from analyses. This may occur when prevalent, rather than new (incident), users of the intervention are included in analyses.	
2.2. If no to 2.1, Were adjustments techniques used that are likely to correct for the presence of selection biases?	It is in principle possible to correct for selection biases, for example by using inverse probability weights to create a pseudo-population in which the selection bias has been removed, or by modeling the distributions of the missing follow up times and outcome events and including them using missing data methodology. However such methods are rarely used and the answer to this question will usually be "No".	
	The answer 'yes' corresponds to lack of selection bias, for example when controls were sampled from a defined population through random digit dialing, or from the patient register of the family doctor from which the case was recruited, or through a "nearest neighbor" procedure. Examples of (usually) "no" would be when controls are sampled from a hospital ward.	<p><i>For case-control studies:</i></p> <p>2.3 Were the controls sampled from the population that gave rise to the cases?</p>

Risk of Bias Judgment		
Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	Start of follow up and start of intervention coincide	
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):	Start of follow up and start of intervention do not coincide, but the authors used appropriate methods to adjust for the selection bias.	
Serious risk of bias (the study has some important problems);	Start of follow up and start of intervention do not coincide. A potentially important number of outcomes or potentially important amount of follow-up are likely to be missing from analyses	
Critical risk of bias (the study is too problematic to provide any useful evidence);	A substantial number of outcomes or substantial amount of follow-up is likely to be missing from analyses.	
No information on which to base a judgment about risk of bias for this domain.	There is no statement that the intervention group was restricted to incident users of the intervention, but no evidence that prevalent users were included	

3. Bias due to departures from intended interventions

Are the (pre-specified) co-interventions likely to be administered in the context of this study?

Screening questions

Were the intended interventions sufficiently clearly defined and implemented such that a reasonable comparison of them can be made?

(If 'No', go straight to judgment of 'Critical risk of bias')

Rationale/remark:

Signaling Question	Rationale/Remark	
<p>3.1. Were the critical co-interventions balanced across intervention groups?</p>	<p>Make a list of possible co-interventions that could differ between intervention groups and could have an impact on study outcomes. Are they likely to be administered in the context of this study?</p> <p>From this list of possible co-interventions select the critical one or more that is/are most likely to affect the impact of the intended intervention within each group. The selection may be based on the available literature or other plausible rationales.</p> <p>Consider if the co-interventions are balanced or unbalanced between intervention groups. In either case, a judgment must be made if the co-intervention has the potential to significantly impact the intended treatment effect.</p>	
<p>3.2. Were numbers of treatment switches low enough that they do not threaten the validity of the estimated effect of intervention?</p>	<p>Intervention switches (crossovers or contamination) introduce bias if the comparison of interest is analogous to the per-protocol effect in the target randomized trial.</p> <p>However they may not introduce bias if the comparison of interest is of initiation of treatment.</p> <p>When considering the interventions being compared, assess the following:</p> <ul style="list-style-type: none"> a) Is there a potential for people receiving one intervention to switch to the other? b) Are multiple switches possible or likely? c) Was the extent of switching sufficient to impact the study outcomes? d) Does the study design minimize the impact of switches? 	
<p>3.3. Was implementation failure minor and unlikely to threaten the validity of the outcome estimate?</p>	<p>Consider implementation fidelity in the context of complexity of the intervention</p> <ul style="list-style-type: none"> a) adherence of intervention administrators b) adherence of study participants c) context of study 	
<p>3.4. If no to 3,1, 3.2 or 3.3: Were adjustments techniques used that are likely to correct for switches, unbalanced co-intervention and</p>	<p>Such adjustment techniques will rarely be reported, and may need to address the potential for time-varying confounding. Specialist advice may be needed.</p>	

implementation failure?		
-------------------------	--	--

Risk of Bias Judgment		
Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	No bias due to departure from the intended intervention is expected, for example if both the intervention and comparator are implemented over a short time period, and subsequent interventions are part of routine medical care, or if the specified comparison relates to initiation of intervention regardless of whether it is continued.	
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):	Bias due to departure from the intended intervention is expected, and switches, co-interventions, and some problems with intervention fidelity are appropriately measured and adjusted for in the analyses. Alternatively, most (but not all) departures from intended intervention reflect the natural course of events after initiation of intervention.	
Serious risk of bias (the study has some important problems);	Switches in treatment, co-interventions, or problems with implementation fidelity are apparent and are not adjusted for in the analyses.	
Critical risk of bias (the study is too problematic to provide any useful evidence);	Substantial departures from the intended intervention are present and are not adjusted for in the analysis.	
No information on which to base a judgment about risk of bias for this domain.	Bias due to departure from the intended intervention is expected, but there is no or limited information on how or whether it is addressed in the reported results.	

4. Bias due to missing data

Screening question

Are the intervention groups free of critical differences in participants with missing data?

(If 'No', go straight to judgment of 'Critical risk of bias')

Rationale/remark: The question intends to address a combination of the numbers, the differential between intervention groups and the reasons for missingness. For example, if the number of participants with missing data is similar by group, but the reasons for them having missing data are likely to be very different (e.g. lack of efficacy vs. side effects) then the study might be considered to be at critical risk of bias.

Signaling Question	Rationale/Remark	Key Variations
4.1 Are outcome data reasonably complete?	This aims to elicit whether the proportion of missing observations is likely to result in missing information that could substantially impact on our ability to answer the question being addressed. Guidance will be needed on what is meant by 'reasonably complete'. One aspect of this is that review authors would ideally try and locate an analysis plan for the study. If there is no mention of missing data, the risk of bias judgment is likely to be 'Unclear'.	<i>For case-control study, alter wording:</i> Is exposure data reasonably complete?"
4.2 Was intervention status reasonably complete for those in whom it was sought?	Missing 'exposure' status (i.e. which intervention the participants received) may be a problem. This requires that the <i>intended</i> study sample is clear, which it may not be in practice. A special issue in case-control studies is that some investigators might simply replace individuals if exposure data cannot be collected (and we may or may not know about this from the report).	<i>For case-control study:</i> OMIT
4.3 Are data reasonably complete for other variables in the analysis?	This relates particularly to missing covariate values when attempts were made to adjust for them in the analysis.	
4.4 <i>If no to 4.1, 4.2 or 4.3,</i> Are the proportion of participants and reasons for missing data similar across interventions?	This aims to elicit whether either (i) differential proportion of missing observations or (ii) differences in reasons for missing observations could substantially impact on our ability to answer the question being addressed. When looking at unintended effects, an important consideration is whether the review authors are satisfied that follow-up has not systematically excluded nontrivial proportions of individuals in whom	

	<p>adverse effects may be prevalent (for example, if older people drop out more, and also have more adverse events).</p> <p>For case-control studies, an important consideration is whether the risk of missing exposure data differs systematically between cases and controls.</p>	
<p>4.4 If no to 4.1, 4.2 or 4.3, Were appropriate statistical methods used to account for missing data?</p>	<p>It is important to assess whether assumptions employed in analyses are clear and plausible. Both content knowledge and statistical expertise will often be required for this. For instance, use of a statistical method such as multiple imputations does not guarantee an appropriate answer. Review authors should seek naive (complete-case) analyses for comparison, and clear differences between complete-case and multiple imputation-based findings should lead to careful assessment of the validity of the adjustment method.</p>	

Risk of Bias Judgment		
Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	All data were available.	
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):	<p>Data were reasonably complete.</p> <p>or</p> <p>Proportions and reasons of missing participants were similar across intervention groups.</p> <p>or</p> <p>Missing data were addressed appropriately in the analysis.</p>	
Serious risk of bias (the study has some important problems);	<p>Proportions of missing participants differ substantially.</p> <p>or</p> <p>Reasons for missingness differ across interventions.</p> <p>or</p>	

	Missing data were not addressed appropriately in the analysis.	
Critical risk of bias (the study is too problematic to provide any useful evidence);	There were serious differences between interventions in participants with missing data.	
No information on which to base a judgment about risk of bias for this domain.	No information about missing data or the potential for data to be missing.	

5. Bias in measurement of outcomes or interventions

Signaling Question	Rationale/Remark	Key Variations
5.1 Were outcome assessors unaware of the intervention received by study participants?	In some studies, blinding of outcome assessors may lead to an answer "yes" to this question. In other studies, outcome assessors may be unaware of the interventions being received by participants despite there being no active blinding by the study investigators. In studies where participants report their outcomes themselves, for example in a questionnaire, the outcome assessor is the study participant.	<i>For case-control study, alter wording:</i> Were assessors of intervention received blinded to participants' status (case or control)?
5.2 Was the outcome measure objective?	An objective outcome measure involves negligible assessor judgment, e.g. all-cause mortality or non-repeatable automated laboratory assessments	<i>For case-control study:</i> OMIT
5.3 Were the methods of outcome assessment comparable across intervention groups?	Comparable assessment methods (i.e. data collection) would involve the same outcome detection methods and thresholds, same time point, same definition, same measurements.	<i>For case-control study, alter wording:</i> Were the methods of assessment of intervention received comparable for the case and control groups?

Risk of Bias Judgment

If an answer to either of the first two questions is “yes” then it would be low risk of bias. A: depends on whether or not blinding was likely to be broken? We also note that the low risk of bias when there is no blinding, but the outcome is objective then this is subject to how the outcome was collected and recorded.

Risk of Bias Judgment		
Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	Identical assessment methods were used for an objective outcome measure, OR assessors (including participants if important outcomes were self-reported) were adequately blinded to (or otherwise aware of) the intervention received by study participants.	
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):		
Serious risk of bias (the study has some important problems);		
Critical risk of bias (the study is too problematic to provide any useful evidence);	Assessment methods were not comparable across intervention groups.	
No information on which to base a judgment about risk of bias for this domain.		

6. Bias in selected of reported result

Signaling Question	Rationale/Remark	Key Variations
<p>6.1 Is it unlikely that the reported effect estimate is available primarily because it was a notable finding among numerous exploratory analyses?</p>	<p>Exploratory studies may be entirely justifiable at an early stage of knowledge about associations between an intervention and outcomes. However, they are not comparable to a randomized trial. A randomized trial will almost always be confirmatory, in that it will pre-specify one or more PICO research questions and a sample size that will allow the trial to detect an important target difference between intervention and control groups in a defined primary outcome. In an exploratory NRS there is likely to be a serious risk of selective reporting if the researchers are likely to have tested many associations and reported only the ones that were statistically significant (or that they selected in some other way).</p>	
<p>6.2. Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among multiple outcome <i>measurements</i> within the outcome domain?</p>	<p>For a specified outcome domain, it is possible to generate multiple effect estimates for different measurements. If multiple measurements were made, but only one or subsets are reported, there is a risk of selective reporting on the basis of results.</p>	
<p>6.3 Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results) from among multiple <i>analyses</i> of the outcome measurements?</p>	<p>Because of the limitations of using data from non-randomized studies for analyses of effectiveness (imbalance in prognostic factors, substantial missing data, etc), analysts may implement different analytic methods to address the limitations. Examples include unadjusted and adjusted models; use of final value vs change from baseline vs analysis of covariance; different transformations of variables; different sets of covariates used for adjustment; different analytic strategies for dealing with missing data. Application of such methods generates multiple effect estimates for a specific outcome metric. If the analyst does not pre-specify the methods to be applied, and multiple estimates are generated but only one or a subset are reported, there is a risk of selective reporting on the basis of results.</p>	<p><i>For case-control study, alter wording:</i> Were the methods of assessment of intervention received comparable for the case and control groups?</p>
<p>6.4 Is the reported effect estimate unlikely to be prone to selective reporting (on the basis of the results)</p>	<p>Particularly with large cohorts often available from routine data sources, it is possible to generate multiple effect estimates for different subgroups. If multiple estimates are generated but only one or subsets are reported, there is a risk of selective reporting on the basis of</p>	

from among different <i>subgroups</i> ?	results.	
---	----------	--

Risk of Bias Judgment	
Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	There is clear evidence (usually through examination of a pre-registered protocol or statistical analysis plan) that all reported results correspond to all intended outcomes, analyses and sub-cohorts.
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial):	<p>EITHER</p> <p>1. Outcome measurements are consistent with: a protocol or statistical analysis plan; or a registration record for the study; or IRB/REC approval;</p> <p>OR</p> <p>2. Outcome measurements are clearly defined in the paper and the paper is internally consistent regarding methods and results;</p> <p>AND</p> <p>The paper is externally consistent regarding measurements and analyses (e.g. there are no other papers with different measurements or analyses of the same measurements); AND</p> <p>There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.</p>
Serious risk of bias (the study has some important problems);	<p>Outcome measurements are internally inconsistent between methods and results; or</p> <p>The paper is externally inconsistent regarding measurements and analyses (e.g. other papers with different measurements or analyses of the same measurements); or</p> <p>The cohort or subgroup is selected from a larger study for analysis and reported on the basis of the results.</p>
Critical risk of bias (the study is too problematic to provide any useful evidence);	There is evidence or strong suspicion of selective reporting of results, and the unreported results are likely to be substantially different from the reported results. This is likely to arise from some explicit statement about selective reporting. It is impossible to specify a comprehensive list of what such statements might say but these could include a statement: (a) that results for outcomes relevant to the systematic review outcome domain were not reported because they were not significant; (b) that various cut-off criteria for dichotomizing/classifying a continuous variable were "tried out". The specific text provoking a judgment of critical bias must be recorded in the free text box.

No information on which to base a judgment about risk of bias for this domain.

There is too little information to make a judgment, for example if only an abstract is available for the study.