

Evaluating farm-level livestock interventions in low-income countries: a scoping review of what works, how, and why

Johanna F. Lindahl^{1,2,3} , Florence Mutua^{1,4} and Delia Grace¹

Review

Cite this article: Lindahl JF, Mutua F, Grace D (2020). Evaluating farm-level livestock interventions in low-income countries: a scoping review of what works, how, and why. *Animal Health Research Reviews* **21**, 108–121. <https://doi.org/10.1017/S1466252320000146>

Received: 30 August 2018
Revised: 4 May 2020
Accepted: 21 May 2020
First published online: 2 December 2020

Key words:

Animal health; animal-source food; food security; livestock production; nutrition security; smallholder farming

Author for correspondence:

Johanna F. Lindahl, Department of Biosciences, International Livestock Research Institute, P. O. Box 30709, 00100 Nairobi, Kenya.
E-mail: J.Lindahl@cgiar.org

¹Department of Biosciences, International Livestock Research Institute, P. O. Box 30709, 00100 Nairobi, Kenya; ²Department of Medical Biochemistry and Microbiology, Uppsala University, P. O. Box 70790 SE 750 07, Uppsala, Sweden; ³Department of Clinical Sciences, Swedish University of Agricultural Sciences, P. O. Box 70790 SE 750 07, Uppsala, Sweden and ⁴Department of Public Health, University of Nairobi, P.O.Box 29053-00625, Nairobi, Kenya

Abstract

Livestock interventions can improve nutrition, health, and economic well-being of communities. The objectives of this review were to identify and characterize livestock interventions in developing countries and to assess their effectiveness in achieving development outcomes. A scoping review, guided by a search strategy, was conducted. Papers needed to be written in English, published in peer-reviewed journals, and describe interventions in animal health and production. Out of 2739 publications systematically screened at the title, abstract, and full publication levels, 70 met our inclusion criteria and were considered in the study. Eight relatively high-quality papers were identified and added, resulting in 78 reviewed publications. Only 15 studies used randomized controlled trial designs making it possible to confidently link interventions with the resulting outcomes. Eight studies had human nutrition or health as outcomes, 11 focused on disease control, and four were on livestock production. Eight interventions were considered successful, but only four were scalable. We found good evidence that livestock-transfer programs, leveraging livestock products for nutrition, and helping farmers manage priority diseases, can improve human well-being. Our report highlights challenges in garnering evidence for livestock interventions in developing countries and provides suggestions on how to improve the quantity and quality of future evaluations.

Introduction

Although declining, poverty and hunger are persistent problems in many developing countries. In most of these, many poor people are involved in livestock: around one billion keep livestock and two billion are estimated to be involved in livestock value chains (HLPE, 2016). Livestock are a pathway out of poverty for poor producers; animal products are of high value and their demand is rapidly growing, driven by urbanization and increasing middle-income classes (Lindahl *et al.*, 2018). These factors make livestock-based interventions attractive to development agencies. Several logic models, or pathways, have been developed to understand how interventions could benefit smallholder livestock keepers, livestock value chain actors, and consumers of livestock products (Randolph *et al.*, 2007; Mayne and Johnson, 2015). These pathways include income generation from marketing of livestock and livestock-based products (including products used for fuel and building), increasing assets and resilience, direct provision of food, financial services (such as guaranteeing loans), generation of power for ploughing and transport, and providing manure for crops and aquaculture. The negative effects of livestock at community level include reducing the time available for mothers to take care of young children, transmission of zoonotic infections, and creating social discord when animals stray or are stolen.

A broad literature exists on the direct benefits of agricultural interventions on increasing productivity and production, but links to more distal, yet important, outcomes such as resilience, nutrition, health, and well-being are challenging to unravel and often difficult to establish. For example, previous reviews have found some evidence on improved production, improved livelihoods, and increased consumption but little on improved nutritional status in terms of height for age, weight for age, or micro-nutrient adequacy (Leroy and Frongillo, 2007; Masset *et al.*, 2011; Ruel *et al.*, 2018).

Given the potential for livestock to contribute to development objectives, it is important to understand which interventions have been implemented and what their effectiveness has been, and where possible analyze their potential for adoption and scalability, including identifying adverse effects potentially associated with their application, and how much could be mitigated. It is equally important to consider the risks associated with livestock keeping and the numerous concerns over the externalities linked to their production (especially that of climate change

© International Livestock Research Institute, 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. Key words used to search for publications evaluating livestock interventions

Geographical restriction	Evaluation	Intervention restriction	Outcomes and impacts
Africa	Experiment intervention	Livestock	Yield
Asia	Impact evaluation	Cattle	Food security
South America	Impact assessment	Sheep	Food safety
Latin America	Trial	Goat	Nutrition
'Country list' (a list of country names from the relevant regions)		Swine	Income
		Pig	Produc*
		Fish	Meat produc*
		Small ruminants	Milk produc*
		Meat	Dairy prod*
		Milk	Livelihood
		Egg	
		Dairy	
		Animal source food	
		Animal product	
		Foods of animal origin	

and disease). Currently, increasing attention is being given to the gathering and use of best evidence to support decision-making processes. However, this is constrained by a number of challenges including poor access to evidence and only a few studies, which are sometimes poorly designed, poorly conducted, or poorly reported (Alonso *et al.*, 2016a).

A broad category of livestock interventions exists (including technology transfers, policy change, infrastructure provision, training, and provision of information). Moreover, interventions can be implemented by a variety of different actors, for example (e.g., actors from disciplines of epidemiology, agriculture, economics, sociology, and development). Scoping reviews can be used to answer broad research questions and are recommended where findings from heterogeneous sources (either in methods or discipline) need to be summarized (Tricco *et al.*, 2016). Hence, a study was designed to identify and characterize livestock interventions and assess their effectiveness in achieving development outcomes in a data-poor context. The results were used to determine where the gaps are and to make recommendations to guide the future evaluation of livestock-based interventions. They would also inform the direction of future research (e.g. to determine what methodologies should be used to ensure a non-biased assessment of impact).

Material and methods

Study approach

A scoping review was conducted in a systematic way following the PRISMA guidelines (Moher *et al.*, 2009; Tricco *et al.*, 2016). Briefly, the methodology provides a framework to identify, review, and analyze evidence to support some specified research questions. A *a priori* study protocol was developed, based on what had successfully been used to answer similar questions (Alonso *et al.*, 2016a, 2016b) (Annex 1). The protocol was used to systematically find publications that could answer the research question. It was not pre-published.

Our research question was 'what is the evidence that farm-based livestock interventions in smallholder systems in developing countries have benefited people'.

The PICO elements were as described below:

- Population: people involved in smallholder livestock production in developing countries

- Intervention: any planned action that, when implemented, could impact on smallholder livestock production and has the potential to benefit people
- Comparison: the comparison for the intervention, it could be a control group or a 'before-and-after' comparison. Results are less meaningful in studies that lack a comparison group.
- Outcome: any documented outcome which describes or is plausibly related to human well-being (includes development outcome, primary outcome, secondary outcome, surrogate outcome, intermediate outcome, or end outcome) (see Table 1 for detailed listing).

Papers published before 23 May 2017, with no set start time, were considered in the study. Studies had to be in English, to describe interventions in animal production and health, and be published in peer-reviewed journals to be included in the review. Screening was applied at the title, abstract, and full paper review stages.

Eligibility criteria and definitions

We defined 'farm-based interventions' as any actions planned to bring change in smallholder livestock systems, and those with the potential to improve human well-being in developing countries. There is no unique or universal definition of a smallholder. They may be defined on the basis of either area of farm, number of animals owned, farming system used, farming purpose, farm income, farm capital and labor, or combinations of these (Grace *et al.*, 2008a, 2008b). For each publication, we relied on the information provided to determine if an intervention was a smallholder-based one or not. We considered developing countries to be low- or middle-income countries as assigned by the World Bank at the time this review was conducted (World Bank, 2019). Outcomes affecting human well-being were those with the potential to improve health, nutrition, and income (or livelihoods).

Studies that included the evaluation of farm-level interventions (in animal production and health) and had a comparison group (either a control group or before-and-after data) were eligible for consideration. Interventions implemented as experiments and in restricted environments such as research centers or on research stations (i.e. agricultural premises operated by a research organization) were excluded. Challenge trials were also excluded.

Such studies are systematically different from field studies (and over-optimistic of impact) and have limited external validity (Levitt and List, 2009; Wisener *et al.*, 2014). We did not consider trials with relevance only to large-scale intensive livestock production, even where there was evidence that the projects had been implemented in developing countries (Wallach *et al.*, 2008), largely because of their limited relevance to smallholder farming systems.

We captured development outcomes, development impact, epidemiological outcomes, surrogate outcomes, intermediate outcomes, and end outcomes using the following definitions. It is important to note that epidemiological research, economic research, and development initiatives often have different definitions for these commonly used words.

- Development outcomes and impacts: The OECD Development Assistant Committee (OECD, 2012) defines outcomes as *'the likely or achieved short-term and medium-term effects of an intervention's outputs'* and impacts as *'positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended'*.
- In epidemiological studies, outcomes can either be primary or secondary. Primary outcomes are those that are most relevant to the research question being addressed by the project, while secondary outcomes are the additional ones that are monitored to help interpret the primary ones (Sibbald and Roland, 1998). A primary outcome might be a reduction in mastitis while an associated secondary outcome could be an improvement in farmer knowledge. Outcomes should be stated before the research is implemented (Sibbald and Roland, 1998).
- Surrogate outcomes are measures that are not of direct practical importance but are believed to reflect the outcome of interest. For example, antibody response to a vaccine is not of real-world significance but may indicate that animals are being protected from the disease. Though easy to measure, surrogate outcomes may lead to a false interpretation of the efficacy of the intervention if the surrogate used is not a very good predictor of the primary outcome (Sibbald and Roland, 1998).
- Intermediate outcomes apply to project management and refer to changes on the way to the 'end outcome' or a higher-level change that an intervention hopes to bring about. For example, increased knowledge of mastitis by farmers has no direct benefit, but may lead to, or support, better mastitis control which may lead to higher milk production, more milk sales, and increased consumption within the household. In this case, knowledge about mastitis is an intermediate outcome. The end outcome may be better control of mastitis, or improved income and nutrition for farmers depending on the objectives of the intervention.

Information sources

Searches were done in PubMed and CAB Direct databases.

Search terms and study selection strategies

Keywords were chosen (Table 1) and search strings to capture relevant studies defined (Annex 1), based on the region where they were implemented, the type, livestock species involved, and the outcome measured. The terms were developed by the research team, tested, and judged for their suitability on the basis of the publications captured (and modified where appropriate). We kept the terms unspecific at the start to make it possible to capture as many intervention studies as possible, and any type of farm-

level intervention, either on animal production or on health, was considered at this early stage. Screening was done at the title, abstract, and full publication stage of the review. All titles and abstracts were reviewed by two reviewers, who also, independently, made decisions on whether to accept or reject. An article entered the next reviewing stage if it was accepted by at least one of the reviewers. First, the publication title was considered, and where judged eligible (based on the inclusion criteria specified in the protocol), its abstract was sought and reviewed to determine its eligibility. We checked for duplicate entries and removed them whenever present. Full publications linked to accepted abstracts were retrieved, reviewed to determine their eligibility, and data extracted from the eligible ones. We also extracted data from abstracts considered eligible but whose full papers could not be found. After having reviewed the results from the systematic search, the authors also included studies they considered relevant to the study question (i.e. those they were aware of but had not been captured by the search process) for the more in-depth review of interventions.

Data collection process

Data were extracted from full papers and abstracts meeting the inclusion criteria. The abstracts referred to are those that were judged acceptable but for which full papers were not available for review. Data extraction for each paper was carried out by one reviewer; however, to monitor the process and ensure accuracy, both reviewers extracted data in approximately 6% of the publications. An Excel® database was designed into which all the extracted data were entered. Codes were provided for most of the variables and an additional spacing provided to expand on selected options. In a few of the entries, no restrictions were applied as responses were expected to vary across publications. For example, the question 'which outcome variable was studied?' revealed a variety of responses, even within the same study. An observation was treated as 'missing' in the case where a response was difficult to find during the article review process. Table 2 is a summary of key data extracted from the eligible publications. Factors likely to influence the adoption of the studied technologies (as reported by authors of the reviewed publications) were also captured (i.e. challenges faced in the implementation of the study, the successes, and opportunities). We also identified factors that were likely to be confounders including those highlighted in the publications.

Quality assessment of individual studies

Because the papers were expected to be highly variable in content, reporting, and quality, we conducted a quality assessment based on the criteria developed by Alonso *et al.* (2016a), a broad and easy-to-apply criteria that has also been found to be suitable for use in data-poor contexts. It included the assessment of methodologies used in the study (i.e. were they consistent with good practice, was the writing coherent, and were data presented in a manner that enabled extraction) (Annex 1). This method of quality assessment is more subjective and cannot be relied on as a tool for risk-of-bias assessment (and was only used as a guide to identifying the publications we needed to consider in the study). We did not assess the quality of accepted abstracts.

Synthesis of results

There are several reasons why it was not possible to present summary estimates in the study. First, the diversity of the interventions, and second, the statistical tools reported were variable

Table 2. Summary of key data extracted from the literature on livestock-related interventions

Stage of review	What data were extracted	
Title and abstract	Publication details	Title, year of publication, author list, if original or review, if duplicate or not
Full paper review	Reviewers decision	Is the paper of acceptable or not acceptable quality
	Details related to the accepted paper	Year the study was done, region, country, area, if an urban, peri-urban or rural setting
	Description of the intervention	What the implementation consisted of, step in the value chain where applied, type of study, species involved, duration, etc.
	Description of the evaluation	If an evaluation was done, how long after the intervention, who did the evaluation, does the study take other interventions into account; does the study take other confounding variables into account, was there before and after data, was there an intervention and control group, were interventions randomly allocated to groups
	Assessment of effects	Which statistics were used, what was measured (improved knowledge, improved animal health, increased production, increased food consumption, etc.)
	What other data extracted	If an exit plan (plan for sustainability after the end of the project) was provided in the study, reported weaknesses, challenges, strengths and opportunities, summary of the paper conclusion, quality of the paper, etc.

(and ranged from simple descriptive analyses to tests of hypothesis – t^* tests, ANOVA, regression models, etc.). But even where the same tests were used, a great variability of outcomes across the studies was observed. Our analyses were thus descriptive; frequency tabulations and graphs were used to support the thematic description of the reviewed papers.

Results

Studies retrieved in the systematic process

The total number of articles included and excluded at each stage of the review is summarized in Fig. 1. Our scoping of the literature identified a total of 70 publications (56 full papers and 14 abstracts); 23 (32.8%) were from studies conducted in Asia, 45 (64.3%) in Africa, one (1.4%) in Latin America (Mexico), and one (1.4%) in multiple countries across the three regions. The publications ($n = 70$) described interventions in cattle (40%), other ruminants (17%), and poultry (17%) (Annex 2).

Data on specific areas where the studies had been implemented were missing in 26 of the publications reviewed. Most studies had been done in rural areas (94%; $n = 44$). Our search strategy aimed to capture papers published before May 2017, although we did not specify start periods, our search captured publications from as early as 1979 (Fig. 2). For 19 of the publications reviewed, it was not possible to determine when the intervention study had been conducted. One publication could not be dated.

Out of the 70 studies identified by our search protocol, many (32) did not provide data on the time interval between when they were implemented and when they were evaluated (implementation-to-evaluation period). An implementation-to-evaluation period of at least 12 months was observed in five of the studies reviewed; the longest time period reported was 48 months. For the remaining studies (33), the duration was variable, but less than 1 year.

Quality assessment and evaluation of interventions

The criteria developed by Alonso *et al.* (2016a) (Annex 1) was used to assess the quality of the reviewed publications ($n = 56$; 34 in Africa, 21 in Asia, and one from multiple continents).

Sixty-eight percent and 71% of the papers were judged as acceptable, for the Africa-based and Asia-based publications, respectively, which is in the range of acceptable publications (as reported by Alonso *et al.*, 2016a). A total of 17 publications were of unacceptable quality, on the basis of either study designs or approaches that were inadequately described, or writing that was too poor to follow and understand, or conclusions that were unsubstantiated, which made data extraction moot. We did not assess the quality of reviewed abstracts ($n = 14$) although there were indications that some of the abstracts were of poor quality. In one case, the abstract was found to be too unspecific to even determine if the study was original or a review. The eight additionally added publications (described later) were all judged to have good quality.

For the evaluations, we sought to understand if these had been planned while designing the intervention, and if the activity was implemented by the group that had implemented the intervention. Evaluation was considered in advance (or was judged to have been planned based on the methodology presented) in 51 of the reports representing 73% ($n = 70$) of publications found using the protocol. In seven studies, the evaluation seemed not to have been planned for at the start of the intervention. In some cases, a separate group of researchers was involved in the project evaluation exercise; in others, the implementers also evaluated the intervention. It was often difficult to determine how long interventions lasted (i.e. time from start to end); we were not able to determine this in 12 of the studies. Randomized controlled trial (RCT) was not an inclusion criteria, so studies with control groups and those presenting before-and-after comparisons were included. However, in most cases, it was difficult to determine if a comparison group for the intervention existed. Surprisingly, 30 studies provided no comparison group, raising questions on what the authors had relied on to determine the significance of any change. In yet another case, authors presented results of a comparison between two different interventions, but no control group was provided, again raising concerns on how these two performed when compared to a control.

Impact was assessed either using 'before-and-after' data (in 27 out of 70 publications) or using mid-term data (as observed in

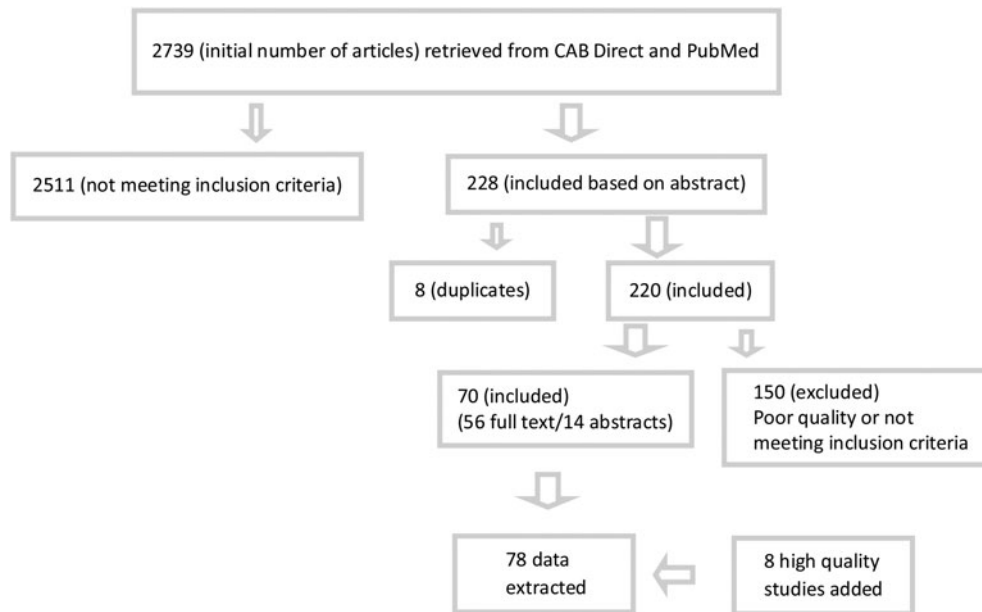


Fig. 1. Summary of the number of publications included or excluded at each stage of the scoping literature review process.

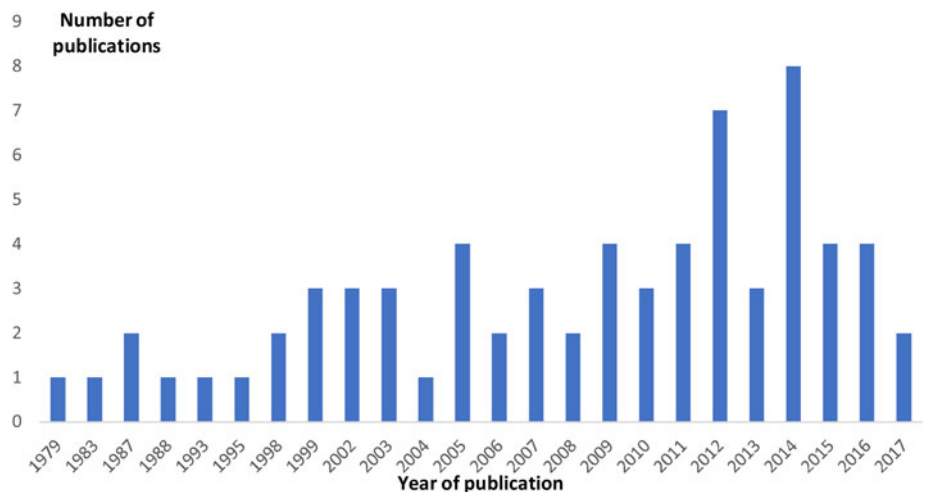


Fig. 2. Description of the number of publications included in the systematic review of literature on livestock-related interventions.

two publications). There was a mention of randomness in 18 of the publications reviewed; however, it was not always clear at what level in the study these had been implemented (if in the selection of participants for surveys, and if it applied to all samplings in the study, or if it meant the actual random assignment to the intervention or the control groups, and if individual, or clustered assignment). Only one publication considered the effects of other interventions in both the analyses and interpretation of the findings.

Addition of selected publications

After the screening process was over, we realized important publications had not been identified and this prompted us to deviate from the protocol to manually add these. Eight publications were added after the review process. These studies, according to the authors, represented a high quality of evaluation, and relied on RCT design to assess effects arising from the interventions.

These had not been detected in spite of the search syntax including the word 'trial'. Six of the publications were from studies conducted in Africa, one in Asia, and one in multiple countries across the three regions. The eight studies focused on cattle (2), pigs or pork (2), equids (1), poultry (1), and both cattle and small ruminants (2). Thus, in total, 15 studies, including the seven from the search component, used an RCT design which allowed the researchers to draw causal inferences, supporting the claim that the intervention actually caused the reported outcomes.

Syntheses of results

We observed a varied range of interventions, including (1) capacity building initiatives, varying from simple provision of information to long-term, hands-on training; (2) strategies to improve production, including feeding, disease control, breeding, etc.; (3) improving marketing; and (4) interventions that focused on

specific technologies (e.g. vaccines, vector repellants, etc.). These projects were implemented either singly or in combination, and were delivered through a variety of modalities including government extension services, non-governmental organizations, community development programs, vouchers redeemable against private sector services and inputs, and mass or social media. Training targeted different actors in the value chain: In a Newcastle disease (ND) control project, vaccinators were trained and supplied with vaccine (Danho *et al.*, 2006), while in a different project, farmers received animal health, nutrition, and marketing information that included vaccination, parasite management and bio-security, forage establishment, and husbandry. In the project by Heifer International, model farms were identified, given inputs to construct poultry houses, and provided with basic training on highly pathogenic avian influenza disease (i.e., symptoms, prevention, and control) (Bhandari *et al.*, 2011). The farmers would then participate in educating other farmers in the community, through demonstrations, informal discussions, exchange visits, and public education campaigns. Collaborators were issued training packages, which included syringes, needles, and educational promotion materials about early detection, reporting, and rapid response, and they were retrained every 6–12 months.

The goat improvement project not only provided farmers with starter goats but it also promoted the formation of farmer groups and established joint saving schemes (Ayele and Peacock, 2003). Women were given goats and encouraged to grow fodder. The project also promoted the establishment of goat groups and formation of credit schemes.

Overall, interventions were either research-based, development-based, or a combination of the two approaches. Assessment of different mastitis control interventions provides a good example of a research-based project, while those based on capacity building to support livelihood diversification (e.g. following drought periods) are development-based.

In-depth analysis of the selected high-quality studies

For the purpose of this analysis, we define ‘high-quality’ studies as RCTs, and included the seven studies from the initial review and the eight that were added during the review and analyses stage (for a total of 15 studies). Six of the studies were by authors, or co-authors from, the CGIAR (mostly from the International Livestock Research Institute (ILRI), Kenya). Nearly all these publications had lead or last authors from high-income countries, and almost all were co-authored by scientists from developing countries. Table 3 is a summary of the studies we considered as high-quality RCTs (an additional description is given in the follow-up text). The interventions have been categorized with reference to the role played by relevant value chains (vc), as those: (1) that were directly based on leveraging the value chain (i.e. direct vc); (2) not directly based on leveraging the value chain but with benefits that may have been mediated via value chains (i.e. indirect vc); and (3) with zero involvement of the value chain (i.e. no vc). The final outcome is also considered (on human health, animal health, and livelihoods).

The reporting format was also not uniform. In most papers, the results were presented in ways which made it difficult to analyze the uptake, magnitude of the effects, and the value of benefits and costs associated with the interventions. The papers lacked essential details as recommended in the CONSORT and REFLECT guidelines (Begg *et al.*, 1996; Schulz, 1997; Brand, 2009; O’Connor *et al.*, 2010).

Olney et al. study: improving childhood nutrition

The study by Olney *et al.* (2016) was implemented in Burkina Faso, involved women and aimed to improve their nutrition. The authors describe a large, well-designed, cluster-randomized trial of a 2-year intervention covering both agriculture and nutrition. Weaknesses were found in the manner in which data were analyzed and the way the results were presented, specifically: a failure to clearly report primary and secondary outcomes; making causal claims for secondary outcomes; multiple primary and secondary outcomes without adjusting significance levels for multiple comparisons; limited discussion of clinical significance; and making causal claims for sub-group analyses. Nonetheless, the study suggests, but does not demonstrate, that an integrated livestock-based intervention can have benefits on child and maternal health. Interestingly, another paper from the same study was not retrieved by our search syntax (Olney *et al.*, 2015), likely because poultry was only a minor part of the intervention and was not mentioned in the title, keywords, or abstract; however, both papers are discussed here.

Saenger et al. study: impact of milk quality information on behavior of dairy farmers

The authors evaluated the behavioral impact of sharing milk quality information with dairy farmers (Saenger *et al.*, 2014). As milk companies pay farmers based on the quality of milk, there is an incentive for them to under-report quality and therefore under-pay farmers. By introducing vouchers for third-party quality measurement, the company’s credibility with dairy farmers was improved. However, <10% of farmers opted for this information, although it was easy, cheap, and credible. Treatment farmers (who received information on milk quality) fed their animals more concentrates than control farmers; output also increased (when milk collection center was controlled for in the analyses). Primary and secondary outcomes were not specified. Overall costs were also not specified. By linking farmers to third-party quality checkers, under-reporting of milk quality was ruled out and farmers were able to allocate their resources more efficiently leading to higher outputs (and better incomes). The extent and uptake of the benefits were difficult to evaluate from the paper.

Four studies had a primary focus on livestock disease control (although one also evaluated income and livestock product consumption).

Henning et al. study: Newcastle disease (ND) vaccination and improved chick management

A large RCT on village chickens in Myanmar compared ND vaccination and improved chick management with a control (Henning *et al.*, 2009). The management changes included confined rearing of chicks with the hen under locally-designed coops combined with supplementary feeding with a creep feeder. A nested trial compared serological titers to ND in vaccinated and non-vaccinated birds. The first trial had seven outcomes, and primary and secondary outcomes were not specified. Vaccination against ND did not decrease crude mortality; however, overall infection pressure from ND was found to be low. Management changes lowered mortality, more birds were sold, and an income increase of \$US2.50 per month was observed. More households reported hatching chicks, and after a lag period of 7 months, they were also more likely to consume home-produced birds. External validity was addressed. Although multiple comparisons may be problematic, this study provides good evidence that

Table 3. Summary of 15 high-quality evaluations of livestock interventions using randomized controlled trial methods, classifying significant as effects with *P*-values <0.05

End outcome	Intervention	Step	Significant results	Not significant	Study
Human health	Integrated program providing livestock inputs and training and nutrition messages	Farm, indirect value chain	None	Primary outcomes: child anemia; anthropometrics	(Olney <i>et al.</i> , 2015)
			Secondary outcomes: prevalence of underweight mothers; empowerment	Secondary outcomes: intake meat/poultry; mean BMI; dietary diversity score	(Olney <i>et al.</i> , 2016)
Human health	Training on biosecurity, infrastructure change, awareness	Farm, no value chain	Primary outcomes: significant and substantial decreases in fish infection with zoonotic trematodes	None	(Madsen <i>et al.</i> , 2015)
Human well-being	Livestock, livestock training, vet services, savings, human health (farm, indirect value chain)	Farm, indirect value chain	2 consumption indicators, 3 financial indicators, mental health, more household time spent working, political awareness	Physical health; women's empowerment	(Banerjee <i>et al.</i> , 2015)
Human well-being	Livestock, livestock training, vet services, savings, human health	Farm, indirect value chain	2 consumption indicators, 3 financial indicators, mental health, more women's time spent working, political awareness, physical health	Household time spent working, women's empowerment	(Bandiera <i>et al.</i> , 2017)
Human well-being	Providing pigs, training and inputs	Farm, indirect value chain	Increase in livestock. Increase in loans. Improvement in 3 of 4 health measures	Domestic violence One of 4 health measures	(Glass <i>et al.</i> , 2017)
Human and animal health	Effect of (a) hygiene; (b) therapy; (c) hygiene and therapy on bovine mastitis	Farm, no value chain	Reduction of contagious pathogens from treatment c in one mode	Milk yield; somatic cell counts; clinical mastitis; infected quarter; contagious pathogens; environmental pathogens	(Omore <i>et al.</i> , 1999)
Animal and human health	Effect of a new Rift Valley fever vaccine on immunogenicity and safety	Farm, no value chain	High (>90%) immunogenicity in sheep and goats and moderate (>65%) immunogenicity in cattle	None	(Njenga <i>et al.</i> , 2015)
Animal health and human nutrition and income	Effect of Newcastle disease vaccination and improved management on poultry health, sero-conversion and livelihoods	Farm, indirect value chain	Higher antibody titers; farmer perception of less death from ND in vaccinated birds Lower mortality, more hatchings, higher sales, and higher consumption in chick management group	Crude mortality in ND vaccination group	(Henning <i>et al.</i> , 2009)
Animal health	Effect of anthelmintic treatment on equids	Farm, no value chain	Secondary outcomes: body condition score; fecal egg count; 5 out of 18 owner appreciation questions	Primary outcome: body weight Secondary outcome: 13 out of 18 owner appreciation questions	(Crane <i>et al.</i> , 2011)
Animal health	Effect of a repellent on reducing trypanosomosis	Farm, no value chain		Primary outcome: trypanosome infection incidence per animal per year	(Bett <i>et al.</i> , 2010)
Farmer capacity, animal health	Effect of a brochure on trypanosomosis treatment on farmer knowledge, practice and animal health	Farm, no value chain	Primary outcomes: increase in farmer knowledge and successful treatments	Primary outcome: increase in herd hematocrit	(Grace <i>et al.</i> , 2008b)
Farmer behavior	Effect of providing milk testing results	Farm, direct value chain	Increase in feed inputs to dairy cattle	None	(Saenger <i>et al.</i> , 2014)
Farmer knowledge	Effect of different communication methods on knowledge of mastitis	Farm, no value chain	Knowledge of mastitis increased	None	(Bell <i>et al.</i> , 2005)
Farmer knowledge	Effect of different communication methods on knowledge of equid health	Farm, no value chain	Knowledge of equid health increased	None	(Stringer <i>et al.</i> , 2011)

chicken management leads to better animal health and livelihood outcomes.

Crane et al. study: the effect of an anthelmintic program for working equids

A relatively large RCT was conducted on working equids in Morocco (Crane *et al.*, 2011). The study was characterized by a large loss to follow-up. The intervention was an anthelmintic treatment program involving the use of ivermectin and fenbendazole. Primary and secondary outcomes were not specified but it appears the major hypothesized effect was increased body weight in the dewormed horses. Although the observed differences in body weight were not significant, a significant effect on body condition score was reported. At the early stages of the study, the treatment group was perceived to have an improved health status, increased ability of the horses to work, and reduced pruritus. There was no adjustment for multiple comparisons. Fecal egg count was numerically lower also in the untreated controls compared to other studies, which could have affected the ability to find effects. This study provides no conclusive evidence that deworming can lead to better health outcomes in equids.

Madsen et al. study: control of fish-borne trematodes

A small, 2-year RCT was conducted on fish farms in Vietnam to control fish-borne zoonotic trematodes (Madsen *et al.*, 2015). The intervention included reducing egg contamination through the treatment of people and domestic animals, mud removal to reduce snail density, and reduction of infection in the juvenile striped catfish and giant gourami. Primary outcomes were stated and not excessive in number (i.e. infection status of fish and intensity of infection). Information was not provided on the costs or the amount of help given to the 14 intervention farmers. Although trematode infection in juvenile fish was reduced, levels in adult fish, human exposure, or prevalence were not assessed. This study provides good evidence that intensive disease control measures can reduce zoonotic disease in fish but gives no evidence on the feasibility or human health impacts of these measures.

Omoro et al. study: mastitis control strategies

A moderate-sized RCT conducted in smallholder dairy farms in Kenya compared three mastitis control strategies: (a) improved udder hygiene; (b) treatment of subclinical cases; and (c) a combination of the two (Omoro *et al.*, 1999). Primary and secondary outcomes were not stated. Six different indicators of mastitis were assessed and two different methods of comparison were used. In one model, there was an 18% decrease in environmental pathogens under treatment c, but no adjustment was made for multiple comparisons. Information was given on the cost of strategies. This study provides no evidence that mastitis control leads to better health in animals or yields economic benefits in smallholder systems.

Our search strategy failed to detect eight RCTs, four focused on disease control, two on livestock production, and two on knowledge transfer. One of the report authors was a member of the research team in three of these studies.

Bett et al. study: impact of tsetse repellent technology

A large, cluster-randomized trial among cattle owners in Kenya assessed the impact of a tsetse repellent technology on reduction in trypanosomosis (Bett *et al.*, 2010). One primary outcome was clearly stated. Unusually, the paper addressed the attractiveness

of intervention to end-users and not just effectiveness. A reduction of at least 50% was considered as the threshold that would make the technology a viable alternative to other options, but only 18% was achieved. This was attributed to the fragility of the technology and greater susceptibility of cattle under field conditions compared to earlier station experiments. This study is good evidence that a novel technology, such as this tsetse repellent, does not substantially reduce trypanosomosis.

Grace et al. study: impacts of rational drug use information

This large, cluster-randomized control trial among smallholder farmers in Mali assessed the impact of rational drug use information (Grace *et al.*, 2008b). Primary outcomes were specified, and included improvement in farmer knowledge of trypanosomosis treatment at 2 weeks and 5 months, successful treatment of sick animals by farmers and improvement in herd hematocrit levels at 5 months. Improvements were significant but modest to moderate for the first two outcomes (i.e. change in knowledge and animal treatment). Ten secondary outcomes were also assessed, of which two were significantly better in the intervention group. Adjustment for multiple outcomes was not done. External validity was addressed. Costs of the intervention were not given, but it was designed to be low cost, as information could be bundled with drug treatments. This study provides good evidence that providing information to farmers can improve their knowledge, practice, and health outcomes in their cattle.

Njenga et al. study: immunogenicity and safety of a new Rift Valley fever (RVF) vaccine

A moderate-sized RCT was conducted on commercial farms in Kenya, to assess the immunogenicity and safety of a new RVF vaccine (Njenga *et al.*, 2015). This was in response to concerns over the only existing commercial vaccine. The outcomes were clearly stated. The results show the new vaccine is safe to use and has high (>90%) immunogenicity in sheep and goats but moderate (>65%) immunogenicity in cattle. Although results were significant, immunogenicity was not satisfactory in cattle and further trials with different doses were recommended. The study provides good evidence on the safety and immunogenicity of the product but not on how protected the animals would be against infection.

Banerjee et al. study: improving nutrition and economic status of poor households

This graduation-based program was a large, six-country RCT which tested the effect of a package of livestock-based interventions on a range of outcomes (Banerjee *et al.*, 2015). The intervention was targeted to the ultra-poor population. There were 10,495 participants in Ethiopia, Ghana, Honduras, India, Pakistan, and Peru. The approach combined the transfer of a productive asset, usually livestock, with cash payments, training in livestock-keeping, follow-up visits, savings encouragement, and health education and/or services. The value of assets varied between sites, ranging from purchasing power parity (PPP) US\$451 to PPP US\$1228 per household. The primary goal was to substantially increase consumption among the very poor, which was achieved by the conclusion of the program and maintained 1 year later. Total costs varied from US\$1455 per household in India to US\$5962 in Pakistan, and estimated benefits were higher than the costs (except in Honduras where the chickens died), with a cost/benefit ratio as high as 433% in India. There were 10 primary outcomes, and this was not adjusted for in the analysis.

Bandiera et al. study: livestock asset transfer

A large RCT was conducted with ultra-poor households in Bangladesh (Bandiera *et al.*, 2017). This followed a similar approach to the above graduation-based program. There were six livestock asset bundles worth US\$560 and an associated support package of the same value. On average, benefits were 5.4 times more than costs.

Glass et al. study: livestock productive asset transfer

A moderate-sized RCT was conducted among rural households in the Congo (Glass *et al.*, 2017). The intervention included two piglets, training in pig keeping, biweekly home visits by trained staff, support for association meetings, pig health services, and 50 kg of feed. Primary outcomes are not described and there were nine outcomes clustered as: economic (2); physical and mental health (4); and domestic violence related (3). Economic (2) and health benefits (three out of four) were seen but information is not given on the cost of the intervention. This study provides moderate evidence that giving people livestock assets and training them on management increases possession of these assets, and the more interesting finding, that this improves their health but does not affect domestic violence.

Two studies evaluated the methods for improving farmer knowledge.

- (1) An RCT with smallholder farmers in Tanzania compared three different methods for imparting knowledge about mastitis (i.e., handout, meeting, or video) (Bell *et al.*, 2005). Five different combinations of methods were compared to the control. All showed an increase in knowledge of mastitis relative to the control. The village meeting was less effective and combining different methods had no advantage over the handout alone. This study provides good evidence that providing information to farmers improves their knowledge.
- (2) A second intervention aimed to improve farmer knowledge was also not detected by the search strategy. A cluster-RCT was used to evaluate and compare the effectiveness of three methods for imparting knowledge about equid health amongst rural Ethiopian working equid users (Stringer *et al.*, 2011). An audio program, a village meeting, and a handout all significantly improved knowledge relative to the control, with the audio program being the least effective.

Overall, six of the 15 individual RCTs did not find evidence that the intervention worked. These were: integrated programs and childhood nutrition; ND vaccination and reduced poultry mortality; RVF vaccination and immunity in cattle; equid deworming and improved body weight; mastitis interventions and reduced mastitis; tsetse control and reduced trypanosomiasis. Two of the six studies focused on a near-term, easy to obtain, but unimportant change, which was to improve knowledge through access to information; one study, focused on intermediate outcomes (i.e. reducing fish disease through improved biosecurity), failed to address feasibility and lacked external validity; one study reported significant, but intermediate or surrogate outcomes (i.e. immunogenicity of RVF vaccine in goats); another study showed significant and meaningful outcomes (i.e. providing pigs improves assets and health), but it was not likely to be sustainable because of the high cost of the intervention. Only two studies presented interventions with outcomes that were meaningful, with improvements to animal health and human livelihoods, and were likely to be sustainable and scalable because of

simplicity and low cost, involving improving poultry management and providing rational drug use information for farmer treatment of cattle, respectively.

Non-RCTs interventions and claims of impact

Our review included experimental studies with designs that were not RCTs, and although findings from these could suggest interventions led to outcome and impacts, we are less confident that positive findings, or failure to find effects, were real or due to study design weaknesses. We found several studies that randomly assigned a small number (two to six) of herds or villages to different intervention packages and also chose or randomly assigned a small number as controls. In these cases, the controls were counterfactual, but it is likely that because of the small group sizes involved, these could not yield results equivalent to those from RCTs that rely on prior power calculation to ensure an appropriate number of subjects. A study in Pakistan randomly assigned six settlements to either intervention or control groups (Rowland *et al.*, 2001). Another study attempted to compare four different interventions with only four subjects per group (Muhanguzi *et al.*, 2014). Other studies used difference-in-difference design where both the intervention group and the control group are compared at baseline and endline. A small number of studies used quasi-experimental designs, which rely on a control or counterfactual comparison group which has not been created by randomization. Overall, 55 studies reported positive findings following the implementation of the interventions, including improved knowledge (5), improved human health (2), improved animal health (15), and improved animal productivity (7). Eight publications reported improved food consumption while 18 reported improved economy or better livelihoods of the farmers. Negative findings were reported in two studies (i.e. a decrease in productivity or a completely failed intervention).

Uptake of interventions (barriers and bridges)

The success of any intervention is dependent on a number of factors (Mayne and Johnson, 2015). Based on the publications reviewed, factors such as the low adoption of technologies, inadequacies in the timing of interventions, and external factors such as drought or insecurity, can reduce the impact. Providing free or subsidized inputs is problematic, as continuation after the project usually requires farmers to undertake the full cost of inputs which is often beyond their means or willingness to pay, and this affects both sustainability and scalability. Culture and access can also limit adoption (Mathias and Mundy, 2010). The authors of the reviewed publications suggested factors to improve the success of interventions: targeting women to improve their participation (e.g. to provide incentives); establishing a good relationship with stakeholders (e.g., partners, farmers, etc.); having champions for each intervention; using participatory approaches and processes; considering farmer-led interventions; integrated approaches with multiple intervention strategies; and cost sharing and establishment of micro-financing mechanisms. The short duration of intervention or insufficient time between intervention and assessment, lack of power, risk of control groups having been influenced by intervention farmers, lack of a randomized controlled or experimental design, and the challenges in measuring outcomes that were influenced by factors other than the livestock intervention (e.g. childhood anemia) were the reasons for failure to observe an impact in some of the publications.

Discussion

In this review, we described how previous livestock interventions have been evaluated and highlighted the challenges in finding evidence to support their application. The papers with a rigorous design provided some insights into interventions involving value chains, and nearly half of the studies had outcomes relevant to human health or nutrition. However, in relation, the majority of the total screened literature had outcomes related to animal health, and a much lower proportion of studies considered human health outcomes. This may be because of the greater emphasis on RCTs and rigorous testing in health sciences, compared to agricultural sciences (Duflo *et al.*, 2011; Duflo and Banerjee, 2017).

Authors also used personal experience to identify additional documents during the analysis and development of the report. Given the design used, such papers were judged as from high-quality studies. It was surprising that these were not identified by the search syntax that included the word 'trial' in both databases and also contained six words that were related to evaluation; perhaps using a more complete or full search term would have helped. This suggests that evaluations of interventions may be especially difficult to retrieve because they are not clearly flagged in keywords, and their identification may require much more extensive searching or the use of additional methods such as expert information. Moreover, the use of RCTs for livestock interventions in LMICs is relatively recent and the lack of standardized approaches may hamper identification.

Our search strategy found 15 studies that met the quality criteria but only two presented interventions which had outcomes that were meaningful (i.e. improvements to animal health and human livelihoods) and were likely to be sustainable and scalable, based on the costs of interventions, possibility to continue after the end of the project, and the approach used. These two interventions, i.e. livestock transfer programs and giving farmers tools to manage diseases they consider as priorities, are good examples of what can be considered as promising livestock-based interventions, given their exceptionally strong evidence base. As reported by Sibbald and Roland (1998), RCTs are the most rigorous way of determining whether a cause-effect relation exists between treatment and outcome, and for assessing the cost-effectiveness of a treatment.

In evaluation science, it is commonplace that most interventions do not demonstrate outcomes or impact and that the better designed an intervention is, the less likely it is to find a significant difference (Rossi, 1987). Our review screened several thousand studies, analyzed 78 in-depth, and only 15 were found to be of sufficient quality, allowing for causal inferences to be drawn. Only two of the 15 interventions showed improvements that were not only significant but also meaningful and likely to be scalable. This is not unusual as major reviews in agriculture and nutrition have only found a handful of papers meeting quality criteria. For example, the recent SLR on the impacts of livestock product consumption on nutrition, undertaken jointly by ILRI and Chatham House, identified just eight studies of sufficient quality for causal inferences to be drawn (Alonso *et al.*, 2019).

While RCTs have for years been considered the gold standard in clinical studies, they have only recently been applied to agriculture (Duflo *et al.*, 2011; Duflo and Banerjee, 2017). Although they have been met with some criticism, most evaluators would agree that RCTs offer the best way of understanding the true impact of interventions. Typically, RCTs find fewer positive findings than

do less rigorous means of evaluation, which, in itself, supports the greater accuracy claimed for these studies. In this review, around half of the livestock RCTs analyzed in-depth did have some positive impact, which is higher than that found in some other fields. For example, of the 90 interventions evaluated in RCTs commissioned by the Institute of Education Sciences (IES) since 2002, approximately 90% were found to have weak or no positive effects, while in comparison, a majority of the non-RCT studies showed benefits (Coalition for Evidence-Based Policy, 2013).

We highlight a number of limitations (and bias) observed in our review:

- *Methodological flaws* – weaknesses identified in the sub-set of reviewed RCTs is consistent with what has been reported elsewhere in the literature. For example, a recent study compared effectiveness studies in animal and human health (Di Girolamo and Meursing Reynders, 2016). Based on the study findings, only 2% of the veterinary RCTs, versus 77% of the human RCTs, reported primary outcomes, random allocation, allocation concealment, and estimation methods.
- *Pooling of results and determination of summary statistics* – we did not perform any meta-analysis, due to differences between studies, missing information, and variation in outcome variables. The design weaknesses likely affected the validity of the individual study results and therefore also limited the extent to which the findings could be compared and generalized. In most cases, particularly for variables with missing data, it was difficult to ascertain if the criterion identified as important (e.g. 'does the study take other interventions into account?') had been done and not reported, or if it had not been considered at all.
- *Search strategy and missing publications* – our search strategy missed a relatively high proportion of the most rigorous studies (eight out of 15 RCTs), which was surprising. In one case, one publication from a large project was captured by our syntax, but a second project from the same research effort was not captured. Including keywords 'random', 'clinical trials', and 'control' would have yielded most of these additional intervention studies on PubMed, thus it is recommended for future studies.
- *Focus on research-oriented publications* – our review was based on peer-reviewed publications and, as described in the text, those that used RCT methodologies to evaluate impact were considered to be more credible. These criteria likely eliminated many of the reports, particularly those resulting from studies that used other approaches, including the development-based interventions. We do not ignore the importance of the wider large-scale development intervention; however, given the above explanation, agencies should consider supporting interventions that have been shown to work using RCT-based approaches. Studies with robust designs are needed to demonstrate that an effect is linked to an intervention (Alonso *et al.*, 2019). This would not only lead to scaling out of interventions with a higher likelihood of success, but would also provide a pathway for continuous reporting and improvement of interventions, particularly when these interventions are applied in countries with varied settings.
- *On-station and challenge trials* – these were excluded because of their lack of external validity. Unlike field trials, on-station experiments are implemented in restricted research environments. In challenge trials, subjects are exposed to the pathogen or hazard of interest with or without the intervention, and they

are often implemented as a step before a randomized controlled field trial is undertaken. Although some challenge trials are reported in the literature, especially for vaccines, they have limited use in animal health and are often poorly implemented and reported (Wisener *et al.*, 2014). However, better use of challenge trials could overcome the constraint we observed in three of the six disease-control RCTs where unexpectedly low disease pressure meant that the efficacy of the intervention was difficult to demonstrate (i.e. deworming horses, vaccines for ND, and vaccines for RVF).

- **Adoption studies** – a type of intervention study that does not always use controls, but rather seeks to understand the proportion of target populations that take up the intervention and factors influencing the uptake. There are many reasons why farmers may not adopt interventions (Guerin and Guerin, 1994) and uptake obviously impacts the success. Unfortunately, few studies in our review provided sufficient information to either estimate adoption rates or the extent (if any) of subsidy to the technology or the factors driving adoption or failure to adopt. Adoption was not included as a search term, which could explain the inability to identify more studies. The five examples below from the wider literature give a range of adoption for different technologies:
 - *East Coast fever (ECF)* is one of the most important cattle diseases in east Africa. A highly effective, but relatively expensive, vaccine is available in several countries. Lynen *et al.* (2012) implied uptake rates could be high (more than 50%) within a project context. Elsewhere, adoption is much lower: studies in Uganda and Tanzania found uptake of <10% (Kasibule, 2013; Yiryele, 2016). Comparing the known quantities of ECF vaccine produced (Perry, 2016) to the number of cattle at risk, it suggests that <1% of the cattle at risk had been vaccinated.
 - *Trypanosomiasis* is the most important disease of cattle in many parts of Africa. A high proportion of farmers (>80%) pay for treatments when animals are sick. These treatments are mostly obtained through the informal sector. However, community-based trypanosomiasis control, although technically effective and economically attractive, has been shown to be unsustainable because of farmers not willing to invest in preventative care and the high transaction costs of communal action (Catley and Leyland, 2001; Widjastuti *et al.*, 2015).
 - *Artificial insemination (AI)* is an effective way of improving cattle genetics. Kenya has one of the most advanced dairy sectors in Africa. Only 18% of the dairy herd is bred by AI and <0.05% of the beef cattle are under AI programs (Makoni *et al.*, 2015). Of these, 83% are carried out by private inseminators, 13% by dairy cooperatives and <4% by the public sector.
 - *Index-based livestock insurance* is a way of overcoming the challenges of high cost of verification in supplying insurance products to the poor. Payments are made on the basis of remote sensing. Yet even with subsidies, adoption remains disappointingly low, rarely above 30% of the intended population, across the several contexts in which it has been introduced and interest in the product tends to decline over time, so the initiative remains dependent on donor funding (Takahashi *et al.*, 2016).

Recommendations for investors

The quality and quantity of evidence was unfortunately limited, but based on our analysis, we make some recommendations on

best-bet livestock interventions, which can help improve future assessments. Investors should be aware that a great majority of interventions are not able to demonstrate effectiveness (the so-called ‘Iron Law of Evaluation’ (Rossi, 1987)). Although this is common across a broad range of disciplines and fields, it is especially true for social programs. Failure to understand this leads to unrealistic expectations and is an incentive for over-optimistic evaluation and reporting. Although high levels of positive outcomes warrant caution, there is a compelling evidence of positive impact for several livestock interventions: livestock transfer combined with a package of training and resources is one of the best-evaluated methods for improving the consumption of the very poor in communities; provision of livestock foods appears to have clear nutritional benefits especially in the first 1000 days (Alonso *et al.*, 2019); and giving farmers novel tools and capacities to improve the management of problems they see as a high priority such as high mortality or sickness from the visible disease. Other livestock evaluations seem promising but have not always been sufficiently evaluated, e.g. community animal health workers that appear to be able to deliver satisfactory services in a sustainable way. For some interventions, there is good evidence that they are not sustainable outside the project context or if continuously subsidized by donors, e.g., community-based trypanosomiasis control and livestock insurance interventions. One way to understand how interventions may work is theory-of-change, which also lists assumptions and requirements for the interventions to have an impact (Johnson *et al.*, 2015a, 2015b; Mayne and Johnson, 2015). We recommend that this is used to better understand adoption.

Livestock interventions appear to have an unusually broad range of benefits with the most substantial benefit being the contribution to food and nutrition security, which is especially important early in life (Alonso *et al.*, 2019). In addition, livestock are often owned and cared for by women who also often dominate food processing and preparation, and so livestock interventions may help empower women (Olney *et al.*, 2016). Livestock may however transmit zoonotic pathogens, which contribute to a high degree of the disease burden in developing countries (Engels and Savioli, 2006), and livestock products are probably the most important source of foodborne disease (McDermott and Grace, 2012). Zoonoses can promote poverty and are also disproportionately affecting the poorest populations (Grace *et al.*, 2017).

In livestock development, as for most other development sectors, there is a paucity of high-quality intervention studies: the same is observed for clinical veterinary interventions (Di Girolamo and Meursing Reynders, 2016). Evaluations that do not have an RCT design may be of limited utility as they can never make convincing causal links between the interventions and outcomes. Although methods such as cohort studies and propensity scoring can make causal inferences more plausible, they are not an adequate substitute for RCTs. Interventions which are costly or require high effort from beneficiaries should be subjected to high-quality evaluations. Where RCTs are used, researchers should use available quality guidelines to design and implement studies. RCTs are the ‘gold standard’, but they are not always appropriate (Scriven, 2008; Petticrew *et al.*, 2012; Vogt *et al.*, 2012). They are relatively complex and expensive and may not be suitable for answering questions such as ‘does providing farmers with information increase their knowledge of that information?’ or ‘does giving farmers free livestock and training on livestock-keeping increase their ownership of those

livestock?'. Some complex and situation-specific issues cannot easily be reduced to a question answerable by an RCT, and appropriate qualitative and quantitative methods should be used to investigate these issues. In addition, some well-proven interventions may not require RCTs (e.g. giving dogs rabies vaccination to reduce community risk of rabies); but we note that some seemingly obvious interventions were not able to demonstrate the impact in RCTs under some circumstances (e.g. ND vaccination in an ND endemic area did not reduce mortality).

Impact is complex, and a single RCT is not definitive. In our review, two interventions which are very likely to be beneficial (i.e. deworming of horses and vaccination of chickens), showed no impact when evaluated by a well-designed RCT. This may have resulted from the unusually low disease challenge, evaluation when the disease was less present, or from other reasons. A negative RCT can chill future studies, and thus it is best practice to base causality on multiple well-conducted RCTs.

Overall, RCTs considered in this review had many methodological problems. There was a tendency for more recent RCTs to be better designed. Recommendations to improve the design of future livestock RCTs are:

- Evaluations should include less immediate but more important outcomes; for example, change in practice as the result of capacity building, and not just focus on change in knowledge as a result of the provided information.
 - Evaluations should distinguish between primary and secondary outcomes, and causal inferences should not be based on differences in secondary outcomes.
 - Positive responses from farmers, or other beneficiaries, are more common than objective evidence of benefits and, because beneficiary reports are likely to be biased as a result of politeness norms, claims of impact based on farmer reports should be given less weight than those based on more objective indicators.
 - There is a high likelihood of spurious association when a large number of outcomes is involved and multiple comparisons are performed. This should be adjusted for or at least discussed in the analysis.
 - There is considerable divergence between RCTs implemented by nutritionists, economists, and epidemiologists, which creates barriers to interpreting studies. Initiatives aimed at bringing the various disciplines to a common understanding should thus be encouraged. Though not a focus in our study, we found nutritionists and economists to be less likely than epidemiologists to follow best practice methods and reporting as set out in guidelines.
 - Especially for evaluations of small sample size, external validity must be addressed. The participants in such studies may be very different from the targeted beneficiaries.
 - The importance (e.g. clinical or developmental significance) as well as the statistical significance of results should be discussed. Especially with large sample sizes, results may be statistically significant but of little real-world importance.
 - Registration of trials and protocols before they are conducted is strongly recommended, and information on accessing such item(s) should be provided in the evaluation reports and papers. It should be noted when outcomes change during the course of the study.
 - There are many best practice guidelines for conducting and reporting different types of evaluations and these should be followed and reported (see <https://www.equator-network.org>).
- Many evaluation studies are under-powered, which should be avoided.
 - Splitting one study into numerous publications (i.e. 'salami slicing' or segmented publication) can make interpretation more difficult and should be avoided or at least the linked papers should be abundantly cross-referenced.
 - Relatively few studies have information on the cost of interventions, yet this information is essential in understanding their potential uptake.

Conclusions

In developing countries, livestock provide a multitude of benefits to both farmers and other actors in the value chain. While many interventions have been implemented, with a focus on improving food and nutrition security, proper and rigorous scientific evaluations of these works have been sparse and often struggle to find significant effects. Methodological weaknesses, as observed in most of the studies reviewed, limits the extent to which findings can be generalized to the larger populations. RCTs may provide better evidence, but are difficult to perform, expensive, and often require better collaboration between disciplines to provide evidence on impacts at different levels (e.g., both nutrition and livelihoods). Reviews are useful for garnering evidence on livestock evaluations, but, because of the great diversity in reporting, may miss important studies.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1466252320000146>.

Acknowledgements. This work was funded through Agriculture for Food Security, AgriFoSe2030, Theme 4, by Sida, the Swedish International Development Agency, and with support from the donors to the CGIAR system.

References

- Alonso S, Dohoo I, Lindahl J, Verdugo C, Akuku I and Grace D (2016a) Prevalence of tuberculosis, brucellosis and trypanosomiasis in cattle in Tanzania: a systematic review and meta-analysis. *Animal Health Research Reviews* 17, 16–27.
- Alonso S, Lindahl J, Roesel K, Traore SG, Yobouet BA, Ndour AP, Carron M and Grace D (2016b) Where literature is scarce: observations and lessons learnt from four systematic reviews of zoonoses in African countries. *Animal Health Research Reviews* 17, 28–38.
- Alonso S, Dominguez-Salas P and Grace D (2019) The role of livestock products for nutrition in the first 1,000 days of life. *Animal Frontiers* 9, 24–31.
- Ayele Z and Peacock C (2003) Improving access to and consumption of animal source foods in rural households: the experiences of a women-focused goat development program in the highlands of Ethiopia. *The Journal of Nutrition* 133, 3981S–3986S. Available at <http://www.ncbi.nlm.nih.gov/pubmed/14672299> (Accessed 18 September 2017).
- Bandiera O, Burgess R, Das N, Gulesci S, Rasul I and Sulaiman M (2017) Labor markets and poverty in village economies. *The Quarterly Journal of Economics* 132, 811–870.
- Banerjee A, Duflo E, Goldberg N, Karlan D, Osei R, Parienté W, Shapiro J, Thuysbaert B and Udry C (2015) A multifaceted program causes lasting progress for the very poor: evidence from six countries. *Science* 348, 772–800. Available at <http://science.sciencemag.org/content/348/6236/1260799> (Accessed 29 September 2017).
- Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D and Stroup DF (1996) Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *Journal of the American Medical Association* 276, 637–639.

- Bell CE, French NP, Karimuribo E, Ogden NH, Bryant MJ, Swai EM, Kambarage DM and Fitzpatrick JL (2005) The effects of different knowledge-dissemination interventions on the mastitis knowledge of Tanzanian smallholder dairy farmers. *Preventive Veterinary Medicine* **72**, 237–251.
- Bett B, Randolph TFF, Irungu P, Nyamwaro SOO, Kitala P, Gathuma J, Grace D, Vale G, Hargrove J and McDermott J (2010) Field trial of a synthetic tsetse-repellent technology developed for the control of bovine trypanosomiasis in Kenya. *Preventive Veterinary Medicine* **97**, 220–227.
- Bhandari DP, Wollen TS and Lohani MN (2011) Preventing highly pathogenic avian influenza (HPAI) at the rural community level: a case study from Cambodia. *Tropical Animal Health and Production* **43**, 1071–1073.
- Brand RA (2009) Editorial: standards of reporting: the consort, quorum, and strobe guidelines. *Clinical Orthopaedics and Related Research* **467**, 1393–1394. doi: 10.1007/s11999-009-0786-x.
- Catley A and Leyland T (2001) Community participation and the delivery of veterinary services in Africa. *Preventive Veterinary Medicine* **49**, 95–113.
- Coalition for Evidence-Based Policy (2013) *Randomized Controlled Trials Commissioned by the Institute of Education Sciences Since 2002: How Many Found Positive versus Weak or No Effects*. Washington, DC: Coalition for Evidence-Based Policy.
- Crane MA, Khallaayoune K, Scantlebury C and Christley RM (2011) A randomized triple blind trial to assess the effect of an anthelmintic programme for working equids in Morocco. *BMC Veterinary Research* **7**, 1.
- Danho T, Kacou A, Bodjo S and Couacy-Hymann E (2006) 'Increase in family poultry production: the case of the Ivory Coast', in *Final research coordination meeting on improving farmyard poultry production in Africa: interventions and their economic assessment*; 24–28 May 2004. Vienna, Austria: Joint FAO/IAEA Division of Nuclear Techniques in Food and Agriculture, Animal Production and Health Section, pp. 107–115. Available at https://inis.iaea.org/search/search.aspx?orig_q=RN:37060714 (Accessed 5 November 2019).
- Di Girolamo N and Meursinghe Reynders R (2016) Deficiencies of effectiveness of intervention studies in veterinary medicine: a cross-sectional survey of ten leading veterinary and medical journals. *PeerJ* **4**, e1649.
- Duflo E and Banerjee AV (2017) *Handbook of Field Experiments*, vol. 1, 1st Edn. Edited by E. Duflo and A. V. Banerjee, 73–140. Amsterdam, Netherlands: Elsevier Academic Press Inc.
- Duflo EC, Kremer MR and Robinson JM (2011) Nudging farmers to use fertilizer: theory and experimental evidence from Kenya. *American Economic Review* **101**, 2350–2390.
- Engels D and Savioli L (2006) Reconsidering the underestimated burden caused by neglected tropical diseases. *Trends in Parasitology* **22**, 363–366.
- Glass N, Perrin NA, Kohli A, Campbell J and Remy MMM (2017) Randomised controlled trial of a livestock productive asset transfer programme to improve economic and health outcomes and reduce intimate partner violence in a postconflict setting. *BMJ Global Health* **2**, e000165.
- Grace D, Jost C, Macgregor-Skinner G and Mariner J (2008a) 'Participation of small farmers in animal health programmes', in *76th General Session Final Report*, 25–30 May 2008. Paris, France: OIE (World Organisation for Animal Health). Available at <https://cgspace.cgiar.org/bitstream/handle/10568/1139/OIE-FinalReport-76thGeneralSession.pdf?sequence=1> (Accessed: 6 April 2019).
- Grace D, Randolph TF, Diall O and Clausen P-H (2008b) Training farmers in rational drug-use improves their management of cattle trypanosomiasis: a cluster-randomised trial in south Mali. *Preventive Veterinary Medicine* **83**, 83–97.
- Grace D, Lindahl J, Wanyoike F, Bett B, Randolph T and Rich KM (2017) Poor livestock keepers: ecosystem-poverty-health interactions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **372**. Available at <http://rspb.royalsocietypublishing.org/content/372/1725/20160166> (Accessed 9 June 2017).
- Guerin L and Guerin T (1994) Constraints to the adoption of innovations in agricultural research and environmental management: a review. *Australian Journal of Experimental Agriculture* **34**, 549.
- Henning J, Morton J, Pym R, Hla T and Meers J (2009) Evaluation of strategies to improve village chicken production: controlled field trials to assess effects of Newcastle disease vaccination and altered chick rearing in Myanmar [corrected]. *Preventive Veterinary Medicine* **90**, 17–30.
- HLPE (2016) Sustainable agricultural development for food security and nutrition: what roles for livestock? A report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security, Rom. Rome, Italy.
- Johnson N, Atherstone C and Grace D (2015a) The potential of farm-level technologies and practices to contribute to reducing consumer exposure to aflatoxins: a theory of change analysis. IFPRI Discussion Paper 01452. Available at <http://papers.ssrn.com/abstract=2685215> (Accessed 1 April 2016).
- Johnson N, Mayne JR, Grace D and Wyatt AJ (2015b) How will training traders contribute to improved food safety in informal markets for meat and milk? A theory of change analysis. *IFPRI Discussion Paper* **1451**. Available at <http://papers.ssrn.com/abstract=2685229> (Accessed 7 March 2016).
- Kasibule D (2013) Impact of the immunisation of cattle against east coast fever in Bugabula County, Kamuli District, Uganda. Makerere University, Uganda. Available at <http://dspace.mak.ac.ug/handle/10570/2053> (Accessed 29 September 2017).
- Leroy JL and Frongillo EA (2007) Can interventions to promote animal production ameliorate undernutrition? *The Journal of Nutrition* **137**, 2311–2316. Available at <http://www.ncbi.nlm.nih.gov/pubmed/17885016> (Accessed 9 September 2017).
- Levitt SD and List JA (2009) Field experiments in economics: the past, the present, and the future. *European Economic Review* **53**, 1–18.
- Lindahl JF, Magnusson U and Grace D (2018) Urban livestock-keeping: contributions to food and nutrition security. *Reference Module in Food Science Encyclopedia of Food Security and Sustainability: Vol. 3, Sustainable food systems and agriculture*, Elsevier, 317–321. doi: 10.1016/B978-0-08-100596-5.21531-X.
- Lynen G, Yrjö-Koskinen AE, Bakunane C, Di Giulio G, Mlinga N, Khama I, Hanks J, Taylor NM, James AD, McKeever D, Peters AR and Rushton J (2012) East Coast fever immunisation field trial in crossbred dairy cattle in Hanang and Handeni districts in northern Tanzania. *Tropical Animal Health and Production* **44**, 567–572.
- Madsen H, Thien P, Nga H, Clausen J, Dalsgaard A and Murrell K (2015) Two-year intervention trial to control of fish-borne zoonotic trematodes in giant gourami (*Osphronemus goramy*) and striped catfish (*Pangasianodon hypophthalmus*) in nursery ponds in the Mekong Delta, Vietnam. *Acta Tropica* **152**, 201–207.
- Makoni N, Hamudikwanda H and Chatikobo P (2015) Artificial insemination and vaccine production value chains in Kenya. 1–44. Embassy of the Kingdom of The Netherlands in Nairobi, Kenya.
- 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, *BMJ (Clinical research ed.)*. London: British Medical Journal Publishing Group. doi: 10.1136/BMJ.D8222.
- Mathias E and Mundy P (2010) *Adding Value to Livestock Diversity*. Food and Agriculture organization of the UN. 107–134. Rome, Italy.
- Mayne J and Johnson N (2015) Using theories of change in the CGIAR research program on agriculture for nutrition and health. *Evaluation* **21**, 407–428.
- McDermott J and Grace D (2012) Agriculture-associated diseases: Adapting agriculture to improve human health. In Fan S and Pandya-Lorch R (eds), *Reshaping Agriculture for Nutrition and Health*. Washington, DC: IFPRI, pp. 12–103.
- Moher D, Liberati A, Tetzlaff J, Altman DG and Group TP (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* **6**, e1000097.
- Muhanguzi D, Picozzi K, Hatendorf J, Thrusfield M, Welburn S, Kabasa J and Waiswa C (2014) Collateral benefits of restricted insecticide application for control of African trypanosomiasis on *Theileria parva* in cattle: a randomized controlled trial. *Parasites & Vectors* **7**, 432.
- Njenga MK, Njagi L, Thumbi SM, Kahariri S, Githinji J, Omondi E, Baden A, Murithi M, Paweska J, Ithondeka PM, Ngeiywa KJ, Dungu B, Donadeu M and Munyua PM (2015) Randomized controlled field trial to assess the immunogenicity and safety of rift valley fever clone 13 vaccine in livestock. *PLoS Neglected Tropical Diseases* **9**, e0003550.

- O'Connor AM, Sargeant JM, Gardner IA, Dickson JS, Torrence ME, Dewey CE, Dohoo IR, Evans RB, Gray JT, Greiner M, Keefe G, Lefebvre SL, Morley PS, Ramirez A, Sicho W, Smith DR, Snedeker K, Sofos J, Ward MP and Wills R (2010) The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety by modifying the CONSORT statement. *Zoonoses and Public Health* 57, 95–104.
- OECD (2012) Glossary of key terms in evaluations and results based management. ISBN 92-64-08527-0. Organization for Economic Cooperation and Development/Development Assistance Committee. Paris, France.
- Olney DK, Pedehombga A, Ruel MT and Dillon A (2015) A 2-year integrated agriculture and nutrition and health behavior change communication program targeted to women in Burkina Faso reduces anemia, wasting, and diarrhea in children 3–12.9 months of age at baseline: a cluster-randomized controlled trial. *The Journal of Nutrition* 145, 1317–1324.
- Olney DK, Bliznashka L, Pedehombga A, Dillon A, Ruel MT and Heckert J (2016) A 2-year integrated agriculture and nutrition program targeted to mothers of young children in Burkina Faso reduces underweight among mothers and increases their empowerment: a cluster-randomized controlled trial. *The Journal of Nutrition* 146, 1109–1117.
- Omoro AO, McDermott JJ, Arimi SM and Kyule MN (1999) Impact of mastitis control measures on milk production and mastitis indicators in small-holder dairy farms in Kiambu District, Kenya. *Tropical Animal Health and Production* 31, 347–361. Available at <https://link.springer.com/article/10.1023%2FA%3A1005269007104?LI=true> (Accessed 18 September 2017).
- Perry B (2016) The control of East Coast fever of cattle by live parasite vaccination: a science-to-impact narrative. *One Health* 2, 103–114.
- Petticrew M, Chalabi Z and Jones DR (2012) To RCT or not to RCT: deciding when 'more evidence is needed' for public health policy and practice. *Journal of Epidemiology and Community Health* 66, 391–396.
- Randolph TF, Schelling E, Grace D, Nicholson CF, Leroy JL, Cole DC, Demment MW, Omoro A, Zinsstag J and Ruel M (2007) Role of livestock in human nutrition and health for poverty reduction in developing countries. *Journal of Animal Science* 85, 2788–2800.
- Rossi P (1987) The iron law of evaluation and other metallic rules. *Research in Social Problems and Public Policy* 4, 3–20.
- Rowland M, Durrani N, Kenward M, Mohammed N, Urahman H and Hewitt S (2001) Control of malaria in Pakistan by applying deltamethrin insecticide to cattle: a community-randomised trial. *The Lancet* 357, 1837–1841.
- Ruel MT, Quisumbing AR and Balagamwala M (2018) Nutrition-sensitive agriculture: What have we learned so far? *Global Food Security* 17, 128–153 (Elsevier).
- Saenger C, Torero M and Qaim M (2014) Impact of third-party contract enforcement in agricultural markets – a field experiment in Vietnam. *American Journal of Agricultural Economics* 96, 1220–1238.
- Schulz KF (1997) The quest for unbiased research: randomized clinical trials and the CONSORT reporting guidelines. *Annals of Neurology* 41, 569–573.
- Scriven M (2008) A summative evaluation of RCT methodology: an alternative approach to causal research. *Journal of Multidisciplinary Evaluation* 5. Available at http://journals.sfu.ca/jmde/index.php/jmde_1/article/download/160/186/0 (Accessed 6 April 2019).
- Sibbald B and Roland M (1998) Why are randomised controlled trials important? *BMJ* 316, 201.
- Stringer AP, Bell CE, Christley RRM, Tefera G, Gebreab F, Trawford K, Reed A, Pinchbeck GL, Tefera G, Reed K, Trawford A and Pinchbeck GL (2011) A cluster-randomised controlled trial to compare the effectiveness of different knowledge-transfer interventions for rural working equid users in Ethiopia. *Preventive Veterinary Medicine* 100, 90–99.
- Takahashi K, Ikegami M, Sheahan M and Barrett C (2016) Experimental evidence on the drivers of index-based livestock insurance demand in southern Ethiopia. *World Development* 78, 324–340.
- Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, Levac D, Ng C, Sharpe JP, Wilson K, Kenny M, Warren R, Wilson C, Stelfox HT and Straus SE (2016) A scoping review on the conduct and reporting of scoping reviews. *BMC Medical Research Methodology* 16, 15.
- Vogt WP, Gardner DC and Haeffele LM (2012) *When to Use What Research Design*. New York, NY: Guilford Press.
- Wallach MG, Ashash U, Michael A, Smith NC and Roberts B (2008) Field application of a subunit vaccine against an enteric protozoan disease. *PLoS ONE* 3, e3948.
- Widyastuti MDW, Bardosh KL, Sunandar, Basri C, Basuno E, Jatikusumah A, Arief RA, Putra AAG, Rukmantara A, Estoepangestie ATS, Willyanto I, Natakesuma IKG, Sumantra IP, Grace D, Unger F and Gilbert J (2015) On dogs, people, and a rabies epidemic: results from a sociocultural study in Bali, Indonesia. *Infectious Diseases of Poverty* 4, 30.
- Wisener LV, Sargeant JM, O'Connor AM, Faires MC and Glass-Kaaster SK (2014) The evidentiary value of challenge trials for three pre-harvest food safety topics: a systematic assessment. *Zoonoses and Public Health* 61, 449–476.
- World Bank (2019) *Historical Classification by Income*. World Bank Country and Lending Groups. Available at <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> (Accessed 4 November 2019).
- Yiryele M (2016) *Prevalence of Theileria parva and Trypanosome Infections in the Dry Season: A Case of Monduli District, Northern Tanzania*. Tanzania: Sokoine University Of Agriculture. Available at <http://suair.suanet.ac.tz:8080/xmlui/handle/123456789/1575> (Accessed 29 September 2017).