Systematic reviews of academic literature for evaluating the effectiveness of farm advisory services

Pierre P. Labarthe, Catherine E. Laurent, Thomas Andrieu, Aldo Mora Sanchez, Monica Caggiano

Prospects for Farmers' Support: Advisory Services in European AKIS
WP 2 – Advisory services within AKIS: International debates
Deliverable 2.2

Systematic reviews of academic literature for evaluating the effectiveness of farm advisory services

Preliminary findings based on a case study about farm advice and occupational health

Pierre LABARTHE
Catherine LAURENT
Thomas ANDRIEU
Aldo MORA
Monica CAGGIANO

National Research for Agricultural Research (INRA)
Research department "Sciences for Action and Development (SAD)

December, 2014

Content

List of Tables

List of Figures

Executive Summary

This report addresses the question of the evaluation of interventions based on farm advisory services. The aim is to provide guidelines for producing evidence about the effectiveness of such interventions. These guidelines have been tested on a specific case: the effectiveness of farm advice for reducing occupational exposure to pesticides of farm labour.

In the first section, we discuss the distinction between the different objectives which an evaluation of farm advisory programmes can endorse. We propose to distinguish three major goals for evaluation, by adopting the typology of Berriet et al. (2014): i) to measure the impact of farm advisory programmes; ii) to understand the causal path that generates changes through farm advisory programmes; and iii) to support learning processes for the stakeholders involved in a programme to in turn enhance its implementation.

A key point of this typology is that it allows for reflection on the methodology which is needed to gather the best possible evidence according to the goal pursued by the evaluation. The 'best' possible evidence should be:
a) socially relevant to those concerned (external validity);
b) based on types of evidence (evidence of presence, of difference-making, of mechanism) which is adequate for the purpose of the evaluation;
c) reliable (produced using rigorous methods, to ensure the highest degree of probative force, i.e. the highest level of evidence).

Numerous studies provide frameworks for describing the impact schemes (or causal paths) of farm advisory service interventions, however there is a clear knowledge gap regarding the actual measurement of their impact, especially in Europe.

The second section is focused on this specific question of impact assessment of interventions based on the support given to farm advisory services. It highlights the fact that many methods (experimental or quasi-experimental) are now available to implement impact assessments. It also gives some critical points for evaluating the internal quality of studies based on such methods and for identifying their limits of validity for practice. A key dimension is that the results of impact assessments of farm advice interventions might not be extrapolated beyond the very context where the study was implemented. This statement calls for a need for reviews that combine results from impact assessments implemented in various contexts.

The third section proposes methodological corner stones for the implementation of such systematic reviews about the effectiveness of farm advice. It describes the different steps for implementing this methodology, and proposes some examples of analytical tools such as: i) algorithms and diagrams to search and sort the papers; ii) tables to describe and evaluate the quality of papers; iii) criteria to combine evidence from different researches. We tested this methodology on a specific case: the effectiveness of farm advice regarding the reduction of occupational exposure to pesticides of farm labour. The case study is presented extensively in section 6.

The results highlight a specific knowledge gap in the European context. There are almost no studies on the impact of farm advice on health issues in European agriculture. Nevertheless, the systematic review reveals that some positive effects of interventions (training, advice…) on farmers and workers have been demonstrated in North America (Canada and United States of America).

1 Three main goals for evaluation initiatives[1]

The different reports of the WP3 of PRO AKIS, which consist of inventories of AKIS and farm advisory services for every country of the European Union (EU)[2], have revealed a global change in the conception of the role of the public sector in the provision of farm advice. There is clearly a trend towards more contracts between the Ministries in charge of agriculture and a diversity of associations and private actors for implementing specific advisory programs, mainly regarding environmental issues. This has led to a renewed interest on questions of the evaluation of advisory programs. But there is a clear lack of debate about the methods available to implement such evaluations, and about how to collect and combine relevant and robust evidence in that respect. Our aim is to contribute to this debate by providing guidelines that would help to choose between methods and to design protocols to implement such methods in practice.

A general objective of the evaluation process is to organise and analyse the information gathered about the programme concerned. Many methods exist to frame this objective, and this diversity may be confusing (Patton, 2008; Martin et al., 2011 for farm advice). Several classifications of these methods were proposed recently (Fitzpatrick et al., 2011; Rogers, 2008). A very common classification is the one proposed by OECD (table 1). But such classifications hardly enter into the detail of the method needed to gather information and empirical evidence so as to achieve each goal of evaluation.

Table 1. The principle of evaluation of the *Development Co-operation Directorate* (DAC) of the Organisation for Economic Co-operation and Development (OECD) (source: OECD 1991)

Criteria of evaluation	Description
Relevance	**The extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor.** • *To what extent are the objectives of the programme still valid?* • *Are the activities and outputs of the programme consistent with the overall goal and the attainment of its objectives? With the intended impacts and effects?*
Effectiveness	**A measure of the extent to which an aid activity attains its objectives**. • *To what extent were the objectives achieved / are likely to be achieved?* • *What were the major factors influencing the achievement or non-achievement of the objectives?*
Efficiency	**Efficiency measures the outputs -- qualitative and quantitative -- in relation to the inputs.** It is an economic term which signifies that the aid uses the least costly resources possible. • Were activities cost-efficient? Were objectives achieved on time? • Was the programme or project implemented in the most efficient way compared to alternatives?
Impact	**The positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended.** This involves the main impacts and effects that result from the activity on the local social, economic, environmental and other development indicators. • What has happened as a result of the programme or project? • What real difference has the activity made to the beneficiaries? For how many people?
Sustainability	**Sustainability is concerned with measuring whether the benefits of an activity are likely to continue after donor funding has been withdrawn**. • To what extent did the benefits of a programme or project continue after donor funding ceased? • What were the major factors which influenced the achievement or non-achievement of sustainability of the programme or project?

[1] This section is partly based on the paper written by M. Berriet-Sollec, P. Labarthe P. and C. Laurent (2014).
[2] http://www.proakis.eu/inventory

In this report, we adopted Stern (2004)'s idea that showed the heuristic value of an analysis that links evaluation methods with the purpose of the evaluation. We adopt here the typology proposed by Berriet et al. (2014) that classifies evaluation according to the three goals that might be pursued:

i) evaluation aimed at **measuring** the impact of farm advisory programmes;
ii) evaluation aimed at **understanding** the impact scheme (or causality framework) of farm advisory programmes;
iii) evaluation supporting **learning** between the stakeholders involved in the programme so as to enhance its implementation, and/or to favour the involvement of these stakeholders into a field of intervention.

A key point of this typology is that is allows for reflection on the methodology for gathering evidence that best suits the evaluation according to the goal pursued.The 'best' possible evidence should be:

i) socially relevant to those concerned (external validity);
ii) reliable (produced using rigorous methods, to ensure the highest degree of probative force, i.e. the highest level of evidence);
iii) based on types of evidence (evidence of presence, of difference-making, of mechanism) that are adequate for the purpose of the evaluation.

Broadly speaking, three types of empirical evidence may be necessary to evaluate interventions (Berriet et al., 2014):

i) ***Evidence of presence:***
- The description and verification of a thing which exists on the ground (e.g. species observed while building a botanical inventory to describe biodiversity). This type of evidence is used to build an agreement among different stakeholders on the state of the world (before and after the programme). This can be approached through a proxy (for instance, the number of footprints of individuals belonging to certain species).

ii) **Evidence of difference-making:**
- This can be evidence of effectiveness: evidence that a given action yields the desired result (e.g. reduced health expenses following the implementation of various prevention programmes).
- It can be also evidence of harm: obtained when adverse effects of an intervention have been looked for and found.
This type of evidence is used to measure an impact.

iii) ***Evidence of mechanism for a phenomenon:***
- This type of evidence is produced when the entities or the activities of a mechanism which produces a phenomenon, and the organisation of these entities and activities, are known (e.g. change of practices resulting from a financial incentive). This type of evidence provides information on the causal pathway to intervene upon for the goals of a public programme to be achieved. It is also used to design new programmes.

1.1 Goal 1: To learn – the evaluation is designed as a collective learning process

Many studies emphasize the importance of elements that support the use of evaluation. The aim here is mainly the appropriation of evaluation findings by different types of users (Patton, 2008). Evaluation is considered as an operational approach for improving further public action and decisions. This goal can be related to diverse participatory methods (Cousins and Whitmore, 1998; Mertens, 1999). This 'learning' objective can be paired with the goal of empowerment (Fetterman and Wandersman, 2005).

In the field of farm advisory services and AKIS, there are many research projects that advocate a switch towards a participative appraisal of the outcomes of participatory projects associating public and private actors in an innovation perspective (Moschitz, 2013).

One of the classical approaches of this learning perspective is embodied in the use of the Soft System Methodologies (SSM, Checkland, 1981) in the evaluation of farm advisory services and training programmes (Rohs and Navarro, 2008). The aim here is to provide actors with a method that makes it possible to draw a shared impact scheme of the programme so as to build a consensus about how to enhance the programme. SSM is designed to help stakeholders make the most effective decisions in uncertain and complex contexts (Checkland, 1981) where learning is the priority. Checkland and Scholes (1990) point out that SSM as a model is not intended to establish versions of reality. Instead the aim of the seven stages of SSM is to facilitate debate so that collective decisions and actions can be taken in problematic situations (see table 2).

These approaches raise several questions where the issue of evidence is concerned. Corroboration with facts and producing the best possible evidence do not appear to be at the heart of this conception/evaluation approach. Rather it aims to promote and structure debate between stakeholders to reach a consensual solution. In practice, however, significant problems may arise (Salner, 2000). In participatory workshops, for example, evidence is provided verbally by different stakeholders, and must be verified. Salner (2000) likens this method to journalism, in that it involves the verification of the opinions of different stakeholders so that '*analysis makes it possible to mount an argument for change which was not simply an intuitive reaction to a conversation held; it was an argument which could be explicitly retraced at any time with links to supporting evidence*' (Checkland and Scholes, 1990: 198–9).

Table 2. The seven stages of the SSM Methodology (Checkland, 1981)

Step of evaluation	Description
Step 1	Inquiring into the situation (identifying the problem using different communication techniques: brainstorming, interviews, participant observation, focus groups, etc.)
Step 2	Describing the situation (describing the context using a wide variety of sources)
Step 3	Defining HAS (identifying programme stakeholders, and interviewing them on the transformations they are expecting)
Step 4	Building conceptual models of the HAS (representing the relationships between stakeholders in the programme being designed or evaluated)
Step 5	Comparing the conceptual models with the real world (preparation of a presentation of the model for a debate with stakeholders)
Step 6	Defining desirable and feasible changes
Step 7	Implementation

It is assumed that the open and collective nature of the debate will secure the quality of the outcomes. Comparison with 'fact checking' in journalism, however, only holds true if the evidence presented is evidence of presence describing facts known through stakeholder practices or causal pathways that can be reached through experience. But arguments often go deeper and target the expected impact of programmes and may involve a complex causality diagram. Thus, these evaluation methods rely on evidence of presence but also on evidence of effectiveness and mechanisms. But they often do not formalize this integration and give little consideration to the control of the quality of this evidence.

1.2 Goal 2: To understand – the evaluation is designed to understand causal frameworks

Here, the aim of the evaluation is to identify and analyse the mechanisms through which the programme can produce the expected outcomes (or may create adverse effects). This second goal derives from studies that advocate for a better understanding of the theories underlying public programmes, and for a better understanding of the specific mechanisms by which they made an impact (Chen, 1990; Shadish et al., 1991). In practice, this raises the question of what knowledge can be used to provide a reliable empirical basis to implement these approaches (Pawson, 2006; Schwandt, 2003).

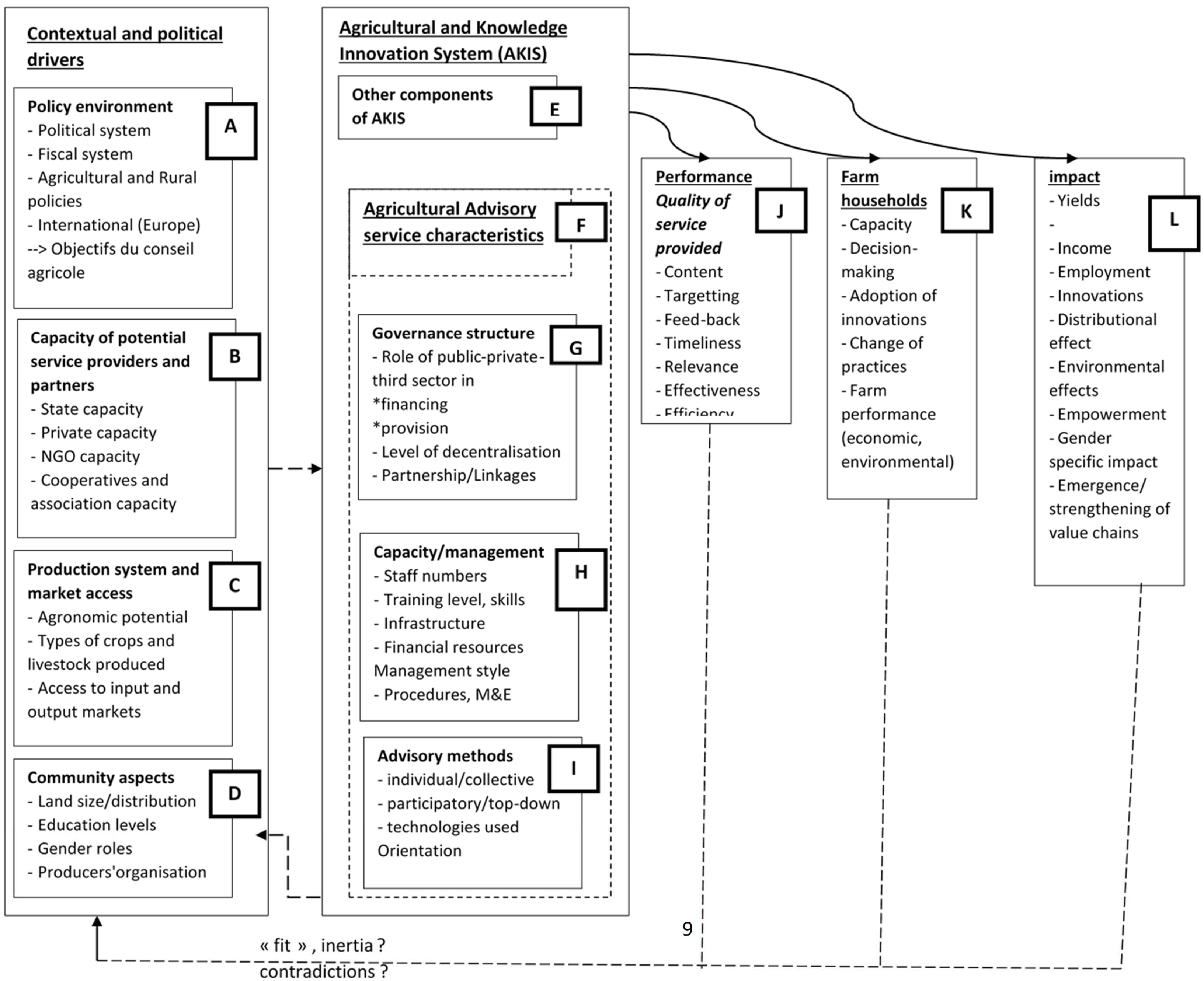
In our field of interest, there have been recently many studies that proposed impact pathways for agricultural advisory service programmes (Faure et al., 2013; Birner et al.; 2009; Labarthe and Laurent, 2011; Prager et al., 2014). Such frameworks are embedded in a debate "from best practice to best fit". The idea is that evaluation should not aim at identifying best solutions regardless of the contexts, but rather at understanding whether the theoretical causal frameworks of farm advisory programs fit with the conditions and contexts where they are used. Such methods often end in rather complex schemes that require the collection of extensive sets of data. Figure 1 shows, for instance, that the framework proposed by Birner et al. (2009) implies to describe:

- the contextual and political drivers of the farm advisory programme (including policy environment, production systems and market access, communities' characteristics…);
- the characteristics of the farm advisory systems (governance, capacity, methods…);
- the diversity of households that may benefit from the services…

Such a starting point implies that field work should be implemented at different scales (gathering data both about national policies and farmers' behaviour, needs, capabilities etc.). It also implies to combine different theoretical perspective (sociology, economics, agronomy, etc.). The causality diagrams formalized in these theoretical models are only ever partial representations of complex causal structures. Their predictive capacities vary according to the object under evaluation and the context; therefore one cannot replace the observation of the real effects (the production of evidence of effectiveness) with that of expected effects (derived from the means implemented in the programme).

Validating the theoretical construction of the impact scheme of farm advice requires an extensive and complex methodology to gather empirical evidence that might be hard to implement. In practice, such methods often combine different sources of evidence: experts' opinions, results from field surveys and interviews, use of descriptive statistics. In certain situations, it makes sense to rank these various sources of evidence, in order to assess the robustness of available evidence. However, the use of theoretical models to infer the effective impact of a programme, as sophisticated as they may be, is often limited. And the probative force of the evidence that is available may vary from the one theoretical standpoint to the next.

Figure 1. Impact Scheme of Farm Advisory Services (adapted from Birner et al. 2009)



1.3 Goal 3: To measure – the evaluation is designed to assess the effects of a programme

A last group of studies focuses on the quantification of programme impacts, often using micro-economic techniques (Rossi et al., 2004), in line with the work of Heckmann. An emblematic principle of this type of research is the identification of an experimental or quasi-experimental situation in which systematic reference to a counterfactual can be used to identify outcomes which are specific to the programme under evaluation (Banerjee and Duflo, 2009). This group of studies seeks to assess if a public intervention works (a measure of effect usually referred to as 'impact assessment')[3].

Such methods were applied to the case of farm advisory services, but essentially in the context of south countries where there is a pressure to demonstrate the impact of extension programmes funded by various private or public donors. These are all the more important as different conceptions of extension interventions compete: top-down vs. bottom-up, individual vs. group advice. In this context there have been, for instance, many studies implemented to measure the impact of specific conceptions or methods of delivery of farm advice, for instance the "Farmer Field School (FFS)" approach (van den Berg and Jiggins, 2007; Davis et al., 2012, Gotland et al.; 2004). The idea is to use econometric methods to measure whether farmers who benefited from FFS perform better than a control group who did not benefit from these services. Such methods are applied at a micro-level scale.

The following sections of this report will focus on the production of this type of evidence aimed at measuring the impact of farm advice programmes. These approaches are designed to obtain the highest possible level of evidence of difference-making (effectiveness or harmlessness) – and only this. In particular, an evidence of effectiveness may provide no information regarding the mechanism(s) that generate this effect.

To obtain high-level evidence of difference-making may seem simple or even simplistic. It is in fact quite challenging and involves costly practices which pose significant methodological and ethical problems. Nevertheless, it is the only way to obtain rigorous evidence of the actual impact of a public programme. The next section (section 2) is dedicated to this question of "internal quality" of studies aimed at producing evidence of difference making. It provides a checklist of the methodological elements that shall be integrated in the evaluation of studies that measure the impact of farm advice.

The significance of evidence of difference-making should not be overestimated. As they do not indicate which mechanisms rendered the programme action effective the results are limited to the context where they have been applied (Deaton, 2009; Banerjee and Duflo, 2009). Thus, there is a need to combine studies from different contexts as well as with studies on mechanisms in order to better understand under which conditions an intervention is effective. This can be achieved with methodologies of systematic reviews of academic literature. The section 3 will give some guidelines about how to implement such reviews. A case study is provided in section 6.

[3] Other studies, less frequent, address the question of the effectiveness of farm advice and agricultural R&D at a more macro scale. They try to evaluate the contribution of R&D national public investments on the productivity of the agricultural sector. Such studies have been applied to both contexts of North (Esposti 2000) and South countries (Evenson 1998). Others aim at measuring efficiency. This involves measuring the value of goods or services produced through public programs against the cost of their production. The goal is then to determine whether the program has produced as many benefits as possible given the resources it has at its disposal.

2 Issues to address for producing evidence of difference-making and evaluating the impact of farm advice programs[4]

This section is focused on the evaluation methods which aim to produce evidence of difference-making in order to evaluate the impact of farm advice programmes. In the PRO AKIS project we have defined farm advisory services as "the entire set of organizations that will enable the farmers to co-produce farm-level solutions by establishing service relationships with advisers so as to produce knowledge and enhance skills" (Labarthe et al., 2013, p.10). Measuring the impact of farm advice is thus a way to validate the fact that relevant and robust knowledge was indeed produced for the targeted population thanks to this "intervention". Farm advice is considered here as an intervention that intends to produce a change. In order to conduct this assessment, an indicator of the change must be agreed on (e.g. knowledge acquisition by farmers, changes in their practices, or level of income of their farms, etc.). Then, different methods are available to collect data so the impact of interventions can be assessed. They are often named as experimental or quasi-experimental methods. We will present the diversity of such approaches in the next sub-sections (section 2.1), as well as the associated limitations and specific problems. More specifically, we will address two questions: how to assess the quality of such approaches (the "internal validity" issue, section 2.2), and how two combine results from different studies implemented in various contexts (the "external validity" issue, section 2.3).

2.1 The diversity of experimental and quasi-experimental methods for producing evidence of difference-making and evaluating the impact of farm advice programmes

As a main goal experimental designs have to provide an unbiased estimate of the effect of an intervention, trying to isolate it from other effects. This is the definition of impact assessment. Impact is defined as the difference between the actual situation with the programme; and the situation that would have occurred without it. Producing such evidence of difference-making of a public programme thus requires the identification of an 'all else being equal' relationship between two variables: a proxy of the treatment or public programme under evaluation (here: farm advice programmes), and a proxy of the desired outcomes of that programme on a given population (here: farmers or farm workers). The main objective is thus to measure the difference between an observable situation and a counterfactual unobservable one. In practice, this is done by comparing, through proxies, the levels of the outcome in a population which received the treatment, with one in a population that did not receive the treatment. Different methods are available for this measurement: experimental or quasi experimental settings. The main difference lies in how they solve the attribution problem and create a "counterfactual situation" (with a control group) to compare against the situations of farmers or farm workers who actually benefited from the programme. This is the key question to assess the "internal validity" of scientific methods of impact assessment.

[4] This section is based on the Master Thesis of Aldo Mora (2013), funded by the PRO AKIS project.

Counterfactual analysis addresses the question of what would have happened if the intervention had not taken place. In this way it is possible to subtract the effect of factors which are alien to the programme itself, nonetheless affecting the selected outcomes. Such factors might be external (here: change in the natural or economic environment of the farm…), or internal, like changes in farmers' behaviour or beliefs that are not related directly to the advisory programme. The most common way to address this question is by means of a control group (farmers who did not benefited from the farm advice intervention). A control group should be similar in all relevant aspects to the treatment group, which is the group that received the farm advice intervention. Different econometric methods have been conceived to address this "counterfactual issue". Table 3 gives an overview of the different methods of impact assessment.

Table 3. Different methods for evaluating the impact of farm advice interventions (adapted from Mora, 2013)

Type of methods	Description
Experimental methods **<u>(Randomized Controlled Trials – RCT)</u>**	Randomized Controlled Trials Randomized Controlled Trials (RCTs) are often considered to provide the highest degree of internal validity in a study. Subjects are randomly selected and, before the intervention, they are randomly allocated to treatment or control group. After allocation, both groups should be treated exactly in the same way, with the exception of aspects related to the treatment itself. RCTs should, if possible, be blinded.
Quasi- Experimental methods. **<u>1. Simple and multiple linear regression</u>**	This is the simplest case of effect estimation. A vector of effects assigns a weight to each of the explanatory variables for predicting the outcome(s). This method models the relationship between one (simple) or several (multiple) dependent variables, or outcomes in this case, with a set of explanatory variables.
Quasi- Experimental methods. **<u>2. Instrumental variables</u>**	This method is used when controlled experiments are not feasible. Instrumental variables allow consistent estimation of a parameter (in this case the effect) when explanatory variables are correlated to the error term of a regression. However, an instrument is required. There are two main requirements for an instrument. It must be correlated with the explanatory variables and uncorrelated with the error term.
Quasi- Experimental methods. **<u>3. Propensity score</u>**	This technique estimates the effect by accounting for the explanatory variables or covariates that predict the treatment. The probability p of receiving the treatment T given the covariates X is calculated first : $p(x) = Pr(T = 1 \| X = x)$ It states that the conditional distribution of X given p(x) is independent of assignment to treatment. In other words, for a given value of p(x) the allocation looks as done randomly. Hence, this is a way to artificially randomize groups ex-post, by selecting balanced subgroups. Once p(x) is computed, groups can be subdivided in substrates of similar p(x) values, individuals can be matched or weighted according to their p(x) value, or regression can be performed using p(x) as a covariate.
Quasi- Experimental methods. **<u>4. Double difference</u>**	Same as Propensity score matching, but with a double difference: difference between the population that receive the treatment and the one that did not receive the treatment; and a difference across time (before and after the treatment).
Quasi- Experimental methods. **<u>5. Pipeline approach</u>**	The pipeline approach uses as comparison group beneficiaries already selected for an intervention, but that have not received it yet. It assumes they are comparable in terms of relevant variables. This can be further assessed by other matching methods.

Conceptually, the simplest way to achieve this similarity is by the random allocation of elements into any of the groups. Such experimental methods are called Randomized Control Trials (RCTs). Randomized experiments, when feasible, deliver the most reliable results. The latter, provided all conditions are set equal. For answering the same question, if a randomised experiment is feasible and no major threats are expected if implemented, it will yield more reliable results than a non-experimental design. Nevertheless, RCTs might not applicable in certain contexts, as allocating people randomly to a control group might be unfeasible or regarded as unethical, as it is the withholding of a treatment (here deciding which farmer or farm workers should benefit in priority from support to access knowledge and services). Then quasi-experiments can be carried out.

The fundamental difference between a randomized experiment and a quasi-experiment is that selection and allocation is random in the first case, whereas in a quasi-experiment the comparison is done using individuals that actually underwent the intervention (observational data) or groups in which randomization was not feasible. The idea is to build two similar groups, differing only by the fact that one received the intervention and the other did not. This allows a counterfactual reasoning. Differences in both groups are then related to the intervention only.

Quasi-experiments do not necessarily guarantee internal validity; however, they use information of actual conditions of real life environments. Quasi-experiments mimic the behaviour of randomized experiments; however, assumptions about causal relationships and/or observability are required. Selecting one or the other type of design depends on where more bias is expected. Randomized studies are considered to be the gold standard, however, if it is thought that experimental conditions will differ much from field conditions, or if randomization seems unethical or unfeasible, a quasi-experiment would be a good alternative option, provided the statistical assumptions behind it can be satisfied.

2.2 Methodological problems associated to the "internal validity" of impact assessments

The methods for impact assessment mentioned above raise many specific problems that should be looked at very carefully so as to guarantee the internal validity of the knowledge produced by the evaluation.

2.2.1 Selection bias

Selection bias arises when the sample is biased: control and experimental groups are not comparable (balanced in their distributions of both, observables and unobservables). It is important to have an unbiased sample and comparable groups for an adequate attribution, but also for being able to extrapolate results to the whole population. Selection bias may lead to spurious observed differences, due to the structure of the sample rather than to the intervention itself. A clear example is self-selection, when individuals voluntarily chose to receive an intervention. Self-selected individuals might be on average more motivated than the general population, and hence more willing to accomplish the goals. An observed effect would not thus correspond to the true effect if it is measured using a sample of self-selected individuals. This problem is all the more important for many farm advice programmes, which are based on the idea of a voluntary participation of farmers or farm workers to groups or networks facilitated by advisors.

2.2.2 Attrition

Attrition occurs when participants dropout of the study. There are two types of attrition, random attrition and non-random attrition. Random attrition takes place when people (here farmers or farm workers) abandon the study for reasons unrelated to the farm advice programme. Non-random attrition occurs when participants dropout for reasons related to the intervention. Random attrition is no source of bias, as the remaining units can be thought of the result of a random sampling, however, when non-random attrition takes place, the sample should be analysed as biased, given that attrition is a particular case of selection bias.

Random attrition occurs for instance with seasonal workers. When a group of workers leave an area due to their migration patterns rather than to reasons related to the intervention. Non-random attrition occurs for instance when participants abandon the programme due to lack of motivation, or because the intervention is not producing results. It is important to analyse non-random attrition, as ignoring it could bias the results, overestimating the effects of the intervention.

2.2.3 Crossover

Crossover refers to the possibility of members of the control or experimental group being affected by interventions other than those that they were allocated to. It has different names depending on the study design or source of external information. Spillover, or contagion in experimental designs, occurs when the control group benefits indirectly from the intervention. Contamination refers to the experimental group receiving other parallel interventions. Education is a recurrent tool in development interventions, and it is very likely (and hoped for!) that participants might share and spread knowledge with non participants. Large scale, long lasting studies are also more likely to be affected by crossover.

This problem is linked to an essential hypothesis of the methodologies of impact evaluation based on RCTs or semi-experimental evaluation: beneficiaries must not be influenced by the fact that non-beneficiaries do not benefit from the programme, and vice versa (Stable Unit Treatment Value Assumption – SUTVA). This hypothesis may be contrary to the causality frameworks underlying certain advisory service programmes, particularly those built on so-called diffusionist models (e.g. the World Bank's Train & Visit program): in theory, their effectiveness resides in the fact that farmers who directly receive advice will share acquired knowledge with those who have not.

2.2.4 Distributive effects

The aspect is related to the distributive effects of the evaluated policy. In most impact studies, the effect is calculated by looking at the difference between the average obtained by the group of individuals benefiting from the measure in a sample and that of the individuals who do not benefit. However, an average improvement for the target population can hide great inequalities or even aggravate these inequalities. Abadie et al. (2002) have shown for instance that a training programme for poor populations could result in an increase in the average income of the target populations, but have no effect on the poorest fraction of this population. As social cohesion and inclusion is an important dimension of agricultural and rural policies in the European Union, distributive effects should be looked at very carefully, especially as some researches evidenced some specific problems of access to farm advice and information for certain social groups of farmers, such as smallholders (Labarthe and Laurent, 2013).

2.2.5 Data sources and ability to find variables that describe the outcome of farm advice

It is always desirable to collect data from the ultimate effect that an intervention intends to achieve. However, this might be resource consuming or unfeasible. By directness it should be understood how close to the ultimate outcome the collected data is. This is particularly complicated in the case of interventions such as farm advice. Indeed, as illustrated in section 1.2, farm advice implies very complex impact pathways, between the intervention (providing advice to farmers) and its outcome at farm level or beyond.

Evaluating the impact of farm advisory services supposes the ability to identify a proxy of this outcome. Such a decisive step raises many questions, and very diverse options can be chosen (table 4). At which level should this outcome be selected (Van den Berg and Jiggins, 2007)? The level of farm performance (yield, income, etc.); the level of the adoption of innovations; or the level most directly affected by farm advisory services: farmers' knowledge and skills? The question then becomes how to express this knowledge and these skills in quantitative variables. In that respect, Godtland et al. (2004) have stressed the difficulties and limitations of their attempt to express farmers' knowledge through knowledge tests. Likewise, the effects of this proxy will have to be observable over relatively short durations (due to costs, RCTs are often used in one- to two-year population tests). However, in the case of farm advisory services, one can wonder whether this short-term measure makes sense due to certain mid- or long-term dimensions of learning processes.

Table 4. Examples of impact outcomes of farm advice intervention (source van den berg and Jiggins 2010)

Domain	Immediate impact	Developmental impact
Technical	Knowledge about ecology	More sustainable production
	Experimentation skills	Improved livelihoods
	Improved crop management	Ability to deal with risks, opportunities
	Pesticide reduction	Innovation
	Yield increase	More cost-effective production
	Profit increase	Reduced public health risks
	Risk reduction	Improved biodiversity
	Reduced water contamination	Improved marketability of produce
	Reduced pesticide poisoning	Poverty reduction
Social	Group building	Collaboration between farmers
	Communication skills	Farmer associations
	Problem solving skills	Community agenda setting
		Farmer study groups
		Formation of networks
		Farmer-to-farmer extension
		Area-wide action
Political	Farmer-extension linkage	Stronger access to service providers
	Negotiating skills	Improved leverage position
	Educational skills	Awareness campaigns
		Protests
		Policy change

2.3 Generalization and external validity of the results of impact assessments

Randomized experiments, often considered to be the gold standard, even when performed rigorously, take samples from a targeted (at least geographically) population. Decision makers might tend to try to generalize directly the results of randomized experiments, arguing their alleged superiority, without considering it refers to their internal validity only. But generalization also raises a question of external validity: the results of an impact assessment are valid only for populations that share the same characteristics regarding the programme as the population on which the impact assessment was applied. For different populations, if properties relevant for the outcome(s), either observable or unobservable differ significantly, results cannot be extrapolated directly (Cartwright, 2011; Cartwright and Hardie, 2012). If this is still warranted then an adjustment should be performed by studying rules, processes, causal paths and relevant variables (Deaton, 2009). It is perhaps important to remember that randomized experiments were mostly developed as a tool for medical research. Biological attributes, rules and processes seem to be rather stable and consistent across populations. A randomized medical study for a population is often safely generalized, unless there is a reason to believe there is a difference in attributes or processes between that and other populations.

This is not the case for intervention such as farm advisory programs, where many dimensions of the populations may vary from one context to another and influence the impact and impact scheme of farm advice. This includes cultural and social elements (farm structures, diversity of households, characteristics of the population in terms of education…) as well as economic ones (farm capital, production systems…). Thus, in that case, results from an impact assessment in a given context might not be generalized to a broader population, or to other contexts. This implies that it is necessary to find methods providing rigorous frameworks for combining different studies implemented in different contexts, and draw lessons about the impact of a given type of intervention. This is precisely the aim of systematic reviews that we present briefly in the next session, and that we implemented for a specific case: the impact of advice on farmers and farm workers exposure to pesticides (see section 6).

3 Building systematic reviews of the effectiveness of farm advisory programs based on academic literature

As stated in the previous section, it is very difficult to extrapolate the results of an impact assessment study beyond the specific context of the population on which it was applied. This means that, in order to produce broader knowledge about the effectiveness of a given type of farm advice intervention or about a given topic (health, environment…), it is necessary to combine evidence from different impact assessment studies and with qualitative knowledge about causal frameworks and mechanism. This is the goal of systematic reviews of literature. In this section, we present some general methodological principles and key steps of systematic reviews.

The aim of a systematic review is to create new knowledge by synthesizing existing evidence. In our case, we shall consider evidence of effectiveness of a given intervention, such as farm advice programmes. This implies to set rules for assessing the reliability of existing findings with regards to a specific question, and then to combine and integrate studies implemented in different contexts, which can present contradictory outcomes and findings. Such reviews also allow for the identification of unanswered questions and knowledge gaps, which need to be resolved to allow for further conclusions to be drawn.

Systematic reviews are methodologies that were first elaborated in the field of medical studies, with the aim to help practitioners to benefit from updated knowledge about the effectiveness of different treatments regarding specific diseases or health issues. Such reviews are not framed in an academic perspective. They intend to answer a very precise and specific question, with a clear practical dimension (typically: is the treatment A effective to cure the disease B?).

It should be noted that such methods have been formalised by different organisations, which have set some quality standards for implementing such systematic reviews and propose methodological guidelines and handbooks. The most important organisations in that respect include the Cochrane collaboration and the Campbell foundation[5]. An important dimension of such studies is that they should be explicit, transparent and replicable about the key steps of implementation of the review work. All of them follow the same key principles:

1) The **question** must be discussed with stakeholders and clearly identified before the review and in the report (it may concern the effectiveness of an intervention or another issue, for instance evidence of presence [e.g. biodiversity inventories]).

2) It is never possible to produce an exhaustive inventory of the literature BUT the limits of **the search strategy must be transparent and replicable** (list of database, exact queries, language) (all these criteria can be adapted to the question, the disciplines, the resources that are available).

3) The **inclusion criteria**[6] must be transparent and non-equivocal (they must be adapted to the question; it is possible to keep for the context analysis or other purpose some papers that do not meet these criteria, but they should not be mixed with the others).

4) The **quality criteria**[6] used to analyze and assess the papers must be transparent (these criteria may vary according to the target, the discipline... the review procedure may also differ according to the project (e.g. number of people reviewing the papers).

This method aims at providing synthetic knowledge for end users, hence it usually involves several participants at different stages. It is often organised following the different steps presented in the table 5.

Table 5. The steps for implementing systematic reviews

Step of the Systematic review	Description
	Step 1. Planning of the work
Phase 0	Identifying the practical knowledge needs that shall be addressed in the systematic review
Phase 1	Proposition of the proposal
Phase 2	Preparation of the search protocol (identifying relevant keywords, sets of academic databases) from which the review will be implemented
	Step 2. Implementation
Phase 3	Debates and agreement on the question to address (may involve interactions with practitioners)
Phase 4	Agreement on criteria of inclusion of studies to include in the study (internal quality, relevance…)
Phase 5	Evaluation of the quality of the study selected
Phase 6	Data Analysis
Phase 7	Synthesis
	Step 3. Writing of the report
Phase 8	Report and recommendation

[5] See for instance (http://www.cochrane.org/resources/handbook)
[6] See for instance http://www.plosone.org/static/guidelines#systematic

Various examples of systematic reviews protocols are available, for instance on the websites of the Campbell collaboration, which has for instance recently published a systematic review about the impact of the “Farmer Field School (FFS)” approach of extension services (Waddington et al. 2014). This publication is a good example of what could be achieved thanks to systematic reviews, and of the willingness to go one step beyond impact assessment. Waddington et al. (2014) try to combine findings from the impact assessments studies of FFS, with the construction of a theoretical causal frameworks of these farmers’ schools. This combination of qualitative and quantitative finding is really promising, and the methodological principles proposed by collective organisation are powerful tools to share knowledge and to produce knowledge bases that can be up-dated. However, they require a significant amount of work (about 10 months of work for one review).

A short example of this method is presented in the case study proposed in the section 6 (p. 22). This case study is based on a communication that was made to the 11th European IFSA Symposium (1-4 April 2014 in Berlin), by Catherine Laurent, Thomas Andrieu and Aldo Mora. It deals with the effectiveness of advisory services interventions aiming at protecting farm labour from pesticides exposure in the EU. It provides an example of implementation of the methodology of systematic review; and highlights learnings from a set of papers that reported interventions outsides the EU, which could inspire the design of evaluation schemes in the EU.

4 Conclusion

The WP3 of PRO AKIS project has provided some descriptions of the transformations of farm advisory services and AKIS across the different member states of the EU. These reports highlight some very important changes in the provision of services for farmers, with an increasing pluralism of the types of providers of services (public, private, farmer based organisations etc.), of methods and technologies supporting farm advice (individual vs group advice, use of Decision Support Tools…) and of themes for advice. There are also some shifts in the conception of public intervention, with a move towards more contracts between the ministries in charge of agriculture and a diversity of organisations implementing agricultural advisory programmes for farmers, integrating public good issues (about environment, health etc.).

This trend calls for a better use of the knowledge generated by the evaluation of advisory programs. This evaluation can support learning for and with the stakeholders involved, but also contribute to a better understanding of the causal scheme of the programmes as well as to measure their actual impact.

A key challenge towards better evaluation and a better use of evaluation is to specify the methodologies needed for gathering the best possible evidence according to the goal pursued by the evaluation. The ‘best’ possible evidence should be socially relevant to those concerned, adequate for the purpose of the evaluation and reliable.

If numerous studies carried out by researchers provide frameworks for describing the impact schemes (or causal paths) of farm advisory services interventions, there is a clear knowledge gap regarding the actual measurement of their impact, especially in Europe. This gap is partly an institutional one, as producing evidence of the effectiveness of a public programme does not appear as a priority neither for research (which focuses on understanding mechanisms) nor for administrations (which focus on accountability issues and on supporting learning with the stakeholders involved in public programs).

Nevertheless, there are some international initiatives that provide guidelines, tools and collective procedures to build systematic knowledge about the effectiveness of interventions. Such initiatives provide many examples of what can be achieved thanks to systematic reviews, and of the willingness to go one step beyond impact assessment. Some of them try to combine findings from impact assessments with the construction of theoretical causal frameworks of public investments in knowledge through education, training or advice. This combination of qualitative and quantitative finding is really promising. The methodological principles proposed by collective organisations are powerful tools to share knowledge and to produce knowledge bases that can be up-dated. However, they require a significant amount of work that could be supported by public and/or collective investments. These investments in evaluation could generate economic value and support a better use of public expenditure when their results are used appropriately.

5 References

Abadie A., Angrist J., Imbens G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings, *Econometrica* 70: 91–117.

Andrieu T. (2013). Quelles sont les formes d'intervention mises en oeuvre auprès des agriculteurs pour promouvoir l'adoption de pratiques d'utilisation des pesticides plus respectueuses de la santé des personnes travaillant dans l'agriculture? Rapport de recherche bibliographique, M1. Inra, AgroParistech.

Angrist J. D, Imbens G. W, Rubin D. B. (1996). Identification of causal effects using instrumental variables, *Journal of the American statistical Association* 91(434): 444–455.

Banerjee A.V., Duflo E. (2009). The experimental approach to development economics, *The Annual Review of Economics* 1: 151–78.

Berriet-Solliec M., Labarthe P., Laurent C. (2014). Goals of evaluation and types of evidence. *Evaluation* 20(2): 195-213.

Birner R., Davis K., Pender J., Nkonya E., Anandajayasekeram P., Ekboir J., Mbabu A., Spielman D., Horna D., Benin S., Cohen M. (2009). From best practice to best fit. A framework for analyzing pluralistic agricultural advisory services worldwide, *Journal of Agricultural Education and Extension* 15(4): 341-355.

Cartwright N. (2011). Evidence, external validity and explanatory relevance. In: Morgan GJ (ed.), Philosophy of Science Matters: The Philosophy of Peter Achinstein. New York: Oxford University Press, 15–28.

Cartwright N., Hardie J. (2012). Evidence-Based Policy: A Practical Guide to Doing It Better. Oxford: Oxford University Press.

Checkland P.B. (1981). System Thinking, System Practice. New York: John Wiley.

Checkland P.B., Scholes J. (1990). Soft Systems Methodology in Action. Chichester: John Wiley & Sons.

Chen H.T. (1990). Theory-Driven Evaluation. Newbury, CA: SAGE.

Cousins J.B., Whitmore E. (1998) Understanding and participatory evaluation, *New Directions for Evaluation* 80: 69–80.

Davis K.E. (2008). Extension in Sub-Saharan Africa: overview and assessment of past and current models, and future prospects, *Journal of International Agricultural and Extension Education* 15(3): 15–28.

Davis K.E., Nkonya E., Kato E., Mekonnen D.A., Odendo M., Miiro R., Nkuba J. (2012). Impact of farmer field schools on agricultural productivity and poverty in East Africa, *World Development* 40(2): 402–13.

Deaton A.S. (2009). Randomization in the tropics and the search for the elusive keys to economic development. National Bureau of Economic Search Working Paper: 14690. Cambridge, MA.

Duflo E., Krémer M. (2005). Use of randomization in the evaluation of development effectiveness. In: Pitman G, Feinstein O and Ingram G (eds), Evaluating Development Effectiveness. New Brunswick, NJ: Transaction Publishers, 205–32.

Espositi R. (2000). The impact of public R&D and extension expenditure on Italian agriculture: an application of a mixed parametric-nonparametric approach, *European review of agricultural economics* 27 (3): 365-384.

Evenson R. E. (1998). The economic contribution of agricultural extension to agricultural and rural development. In: Swanson E., Bentz R. and Sofranko A. (éd.), Improving agricultural extension. A reference manual. Rome: FAO, 1998.

Faure G., Desjeux Y., Gasselin P. (2012). New challenges in agricultural advisory services from a research perspective: a literature review, synthesis and research agenda, *Journal of Agricultural Education and Extension* 18(5): 461-492.

Fetterman D.M., Wandersman A. (2005) Empowerment Evaluation. Principles and Practice. New York: The Guilford Press.

Fitzpatrick J.L., Sanders J.R., Worthen B.R. (2011). Program Evaluation: Alternative Approaches and Practical Guidelines. Upper Saddle River, NJ: Pearson Education.

Godtland E.M., Sadoulet E., de Janvry A., Murgai R., Ortiz O. (2004) The impact of farmer field schools on knowledge and productivity: a study of potato farmers in the Peruvian Andes, *Economic Development and Cultural Change* 53(1): 63–92.

Grabisch M. (1995). Fuzzy integral in multicriteria decision making, *Fuzzy sets and systems* 69(3):279–298.

Grabisch M. (1996). The application of fuzzy integrals in multicriteria decision making, *European journal of operational research* 89(3):445–456.

Laurent C., Berriet-Solliec M., Labarthe P., Trouvé A. (2012). Evidence-based policy: de la médecine aux politiques agricoles?, *Notes et études socio-économiques* 36. http://www. agreste. agriculture. gouv. fr/publications/notes-et-etudes-socio-economiques/.

Laurent C., Baudry J., Berriet-Solliec M., Kirsch M., Perraud D., Tinel B., ... & Ricroch A. (2009). Pourquoi s' intéresser à la notion d'«evidence-based policy»?, *Revue Tiers Monde* 4: 853-873.

Labarthe P., Laurent C. (2013). Privatization of agricultural extension services in the EU: Towards a lack of adequate knowledge for small-scale farms? *Food Policy* 38: 240–252.

Labarthe P., Laurent C. (2011). Économie des services et politiques publiques de conseil agricole, *Cahiers Agriculture* 20(5) : 343-351.

Labarthe P., Caggiano M., Laurent C., Faure G., Cerf M. (with the collaboration of Cristovao A., Knierim A., Prager K., Sutherland L.A.) (2013).Concepts and theories available to describe the functioning and dynamics of agricultural advisory services.Learning for the inventory (PRO AKIS WP3), 2013.Deliverable 2.1 of the PRO AKIS FP7 EU Project. 21 p. + annexes.

Martin A., Sabine Gündel S., Apenteng E., Pound B. (2011). Review of Literature on Evaluation Methods Relevant to Extension, Lindau (Switzerland), GFRAS – Global Forum for Rural Advisory Services.

Mertens D. (1999). Inclusive evaluation: implications of transformative theory of evaluation, *American Journal of Evaluation* 20(1): 1–14.

Mora Sanchez A. (2013). Measuring the effectiveness of interventions: the farm advisory case. Thesis for the Erasmus mundus Master in complex systems. Inra, Ecole Polytechnique.

Moschitz H. (2013). From Project Management to Process Management - Effectively Organising Transdisciplinary Projects, *GAIA-Ecological Perspectives for Science and Society* 22 (3): 211-213.

OECD (1986). Glossary of Terms Used in Evaluation, in 'Methods and Procedures in Aid Evaluation.

OECD (1991). The DAC Principles for the Evaluation of Development Assistance.

OECD (2000). The Glossary of Evaluation and Results Based Management (RBM) Terms,

Patton M.Q. (2008). Utilization Focused Evaluation, 4th edn. Thousand Oaks, CA: SAGE.

Pawson R . (2006). Evidence-based Policy: A Realistic Perspective. London: SAGE.

Prager K., Creaney R., Lorenzo-Arribas A. (2014). Advisory services in the United Kingdom: Exploring 'fit for purpose' criteria. The James Hutton Institute, Craigiebuckler, Aberdeen, Scotland, UK, 19 pp, presentation at the 11th conference of the International Farming Systems Association (IFSA), Berlin, 1-4 April 2014.

Rochs F., Navarro M. (2008). Soft System Methodology: an intervention strategy, *Journal of International Agricultural and Extension Education* 15(3): 95–9.

Rogers P. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation* 14(1): 29–48.

Rossi P.H., Lipsey M.W., Freeman H.E. (2004). Evaluation: A Systematic Approach, 7th edn. Newbury Park, CA: SAGE.

Salner M. (2000). Beyond Checkland & Scholes: improving SSM, *Occasional Papers on Systemic Development* 11: 23–44.

Schwandt T. (2003) 'Back to the rough ground!' Beyond theory to practice in evaluation, *Evaluation* 9(3): 353–64S

Shadish W.R., Cook T.D., Leviton L.C. (1991). Foundations of Program Evaluation Theories of Practice. Newbury Park, CA: SAGE.

Vaessen J. (2010). Methodological and conceptual challenges of evaluating the impact of development interventions. PhD thesis, Maastricht.

Vaessen J., Todd D. (2008). Methodological challenges of evaluating the impact of the global environment facility's biodiversity program, *Evaluation and Program Planning*, 31(3):231–240.

Van den Berg H., Jiggins J. (2007). Investing in farmers: The impact of farmer field schools in relation to Integrated Pest Management, *World Development* 35(4): 663–87.

Waddington H., Snilstveit B., Hombrados J., Vojtkova M., Phillips D., Davies P., White H. (2014). Farmer Field Schools for Improving Farming Practices and Farmer Outcomes: A Systematic Review Campbell Systematic Reviews 2014:6 DOI: 10.4073/csr.2014.6

Wang-Sheng L. (2013). Propensity score matching and variations on the balancing test, *Empirical economics* 44(1):47–80, 2013.

6 Case study: an example of the different steps of systematic reviews

This section is based on the communication “Effectiveness of Advisory Services Interventions Aiming at Protecting Farm Labour from Pesticides Exposure in the EU”, presented by Catherine Laurent, Thomas Andrieu, and Aldo Mora Sanchez at the 11th European IFSA Symposium (1-4 April 2014, Berlin, Germany).

Abstract

- Context: The impact of agricultural pesticides on human health is a source of major concern. The EU directive on the sustainable use of pesticides (2009/128/EC) aims at the reduction of indirect exposure to pesticides (food residues, pollution of the environment...) but also the reduction of direct exposure (farm labour) because the use of pesticides may generate important health hazards for farm labour. Adequate training and advisory services are needed to improve the situation.

- Objectives: In order to inform policy making on this topic, the objective of this study was to systematically review the available evidence of the effectiveness of farm advisory interventions aimed at reducing the level of pesticide exposure of farm labour in Europe.

- Search strategy: We implemented a systematic approach using the following electronic databases: Web of Science, Medline, Cabi, Scielo, Cochrane database of systematic review, and Econlit.

- Inclusion criteria: We selected articles that involved farm advisory programs with occupational safety concerns dealing with pesticides and with an evaluation of the effectiveness of these interventions. All types of studies were considered for Europe (including monographs). A second set of papers was selected in non-European situations corresponding to interventions that were designed with control groups and explicit measures of effectiveness. Studies reporting only tests of machinery or individual protection equipment in controlled conditions were not included.

- Results: We could only identify 6 studies which met all of the inclusion criteria, but not a single one presented the evaluation of the effectiveness of an intervention with a robust methodology. Other studies provided contextual information. In the USA, Latin America and India several studies were implemented to assess the effectiveness of large scale prevention programs. Lessons from these experiences can be combined with contextual information in the EU in order to better assess the difficulties to transpose these programs to Europe.

- Conclusion: This review shows a dramatic lack of reliable evidence in the EU that could help design advisory programs to improve the protection of farm labour from pesticide exposure.

6.1 Context of the study

Agriculture is one of the most hazardous occupations in the EU (Eurostat 2007) and worldwide (ILO 2000). According to ILO's analysis "*the increasing use of machinery and of pesticides and other agrochemicals has aggravated the risks. In several countries, the fatal accident rate in agriculture is double the average for all other industries. Machinery such as tractors and harvesters cause the highest frequency and fatality rates of injury but exposure to pesticides and other agrochemicals constitute major occupational hazards which may result in poisoning and death and, in certain cases, work-related cancer and reproductive impairments*". This effect of pesticides on farm labour's health is documented in many reviews (e.g. Inserm 2013) and the possibility of adverse effects that were not identified until now (i.e. endocrinal disrupting properties of many pesticides) is now discussed.

This is why the EU directive on the sustainable use of pesticides (2009/128/EC) [7] aimed not only at reducing the use of pesticides to protect the environment but also at protecting the health of people, including those working in agriculture. When discussing the use of pesticides, the commission considers that it "*does involve risk, because most have inherent properties that can endanger health and the environment if not used properly. Human and animal health can be negatively affected through (i) direct exposure (e.g. industrial workers producing plant protection products and operators applying them) and (ii) indirect exposure (e.g. via their residues in agricultural produce and drinking water, or by exposure of bystanders or animals to spray drift when they are applied). This is why the EU seeks to ensure the correct use of pesticides or plant protection products and to maintain public awareness. In this respect, the Common Agricultural Policy includes measures that help promoting the sustainability in the use of plant protection products: decoupling, operational programmes of the fruit and vegetables regime, agri-environmental measures (e.g. support to integrated farming), training, use of farm advisory services.*"

(http://ec.europa.eu/agriculture/envir/pesticides/index_en.htm)

For the 2009/128/EC Directive, "*It is essential that Member States set up systems of both initial and additional training for distributors, advisors and professional users of pesticides and certification systems to record such training so that those who use or will use pesticides are fully aware of the potential risks to human health and the environment and of the appropriate measures to reduce those risks as much as possible. Training activities for professional users may be coordinated with those organised in the framework of Regulation (EC) No 1698/2005*". For this directive, ""*advisor" means any person who has acquired adequate knowledge and advises on pest management and the safe use of pesticides, in the context of a professional capacity or commercial service, including private self-employed and* public advisory services, commercial agents, food producers and retailers where applicable". In other words, all the components of the agricultural advisory information system (Akis) are expected to provide adequate information on occupational safety when advising on pesticide use because there is no barrier between productive activity and occupational health prevention. Moreover, the rationale of advisory services organisation should be consistent with the Council Regulation (EC) No 1698/2005 on support for rural development by the European Agricultural Fund for Rural Development, which makes it compulsory to have national strategic plans ensuring that above a certain threshold of subsidies, farmers will benefit from adequate farm advisory services.

Therefore, can we consider that everything is under control?

As a matter of fact, it is often believed that most industrialized countries provide adequate specialized training on occupational health services in agriculture through their extension services, farmers' associations, and other relevant institutions. However, in the EU such representation must be seriously questioned. First of all, many countries have the networks dealing with production and environmental issues are disconnected from those in charge of occupational health. Secondly, both advisory networks have a limited impact because many farmers have no access to them. Regarding technical advice, only a limited number of farms can benefit from face to face advice that would allow them to discuss the specific problem they meet to adapt their practices. Regarding occupational health, for most EU countries prevention is focused on farms with salaried workers, while hired non-family workers only represent 7.8% of those regularly working on the farms (EU actual economic briefs 2013). As a consequence, only a limited proportion of farms are eligible for this types of advice. Therefore, as mentioned in dozens of policy

[7] Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 *establishing a framework for Community action to achieve the sustainable use of pesticides,* in accordance with the Regulation (EC) 1698/2005 of 20 September 2005 *on support for rural development by the European Agricultural Fund for Rural Development*

documents and papers on pesticides hazards, there is an urgent need to develop effective advisory services to protect farm labour from pesticide hazards, in all types of farms.

What information could inform policy-making on this issue?

The aim of this communication is to contribute the answer for this question. Our objective was to systematically review the existing evidence of the effectiveness of farm advisory intervention that aim to reduce the level of exposure to pesticides of farm labour in Europe. We shall present the method (section 2) and the results (section 3) of this review. Then knowledge gained from foreign experiences (section 4) will be discussed in order to provide cornerstones for further investigations.

6.2 Search method

We looked for articles that involved farm advisory programmes with occupational safety objectives related to pesticides and that questioned the effectiveness of these interventions. Studies reporting only tests of machinery, clothing, personal protection equipment (PPE) etc. in controlled environments were not included. The results of such studies can be a useful input for tuning the content of advisory interventions, but they give no information on the actual effectiveness of an advisory intervention that would make use of this data in normal work conditions for several reasons. For instance, many studies show that (i) in working conditions some people do not wear PPE for various reasons, (ii) PPE properties may vary, due to heat, transpiration, etc.) (E.g. Lander, Hinke 1992, Garrigou et al. 2011). We searched the following electronic databases: Web of Science, Medline, Cabi, Scielo, Cochrane database of systematic reviews, and Econlit.

Search query : (prevention OR advi* OR stewardship OR conseil OR "farmers field schools") AND (pesticid*) AND (agric* OR farm*) AND (health) AND (Russia OR Ukraine OR France OR Spain OR Sweden OR Norway OR Germany OR Finland OR Poland OR Italy OR United Kingdom OR Romania OR Belarus OR Kazakhstan OR Greece OR Bulgaria OR Iceland OR Hungary OR Portugal OR Serbia OR Ireland OR Austria OR Czech Republic OR Ireland OR Georgia OR Lithuania OR Latvia OR Croatia OR Bosnia OR Herzegovina OR Slovakia OR Estonia OR Denmark OR Netherlands OR Switzerland OR Moldova OR Belgium OR Albania OR Macedonia OR Turkey OR Slovenia OR Montenegro OR Cyprus OR Azerbaijan OR Luxembourg OR Andorra OR Malta OR Liechtenstein OR San Marino OR Monaco OR Europe). Timespan=All years. Search language=English

Inclusion criteria	***Description***	***Relevant***	***Not relevant***
Population	All categories of farm labour, including casual labour, including people not directly employed by the farm but working on it.	Study dealing with at least one category of these people.	Characteristics of the population are not described or other types (e.g. gardeners).
Pesticides	All types of chemical products used against pests as long as they are used for farm activity (crops, livestock, seed protection, indoor insect control).	- Plant protection product as defined in Regulation (EC) No 1107/2009; - Biocide product as defined in Directive 98/8/EC; - Some veterinary drugs (for external parasites).	Products that cannot be source of human exposure (e.g. insecticides given to livestock through injection). Antibiotics.
Exposure situation	All situations generating an exposure to pesticides, i.e. a contact between a person and a pesticide (dermal, ocular, inhalation, ingestion), outside or inside the farm buildings, or even outside the farm (e.g. dipping of livestock in collective settings).	Borderline cases included: secondary exposure of farm workers' family (laundry, contamination of dust...).	Exposures that have no link with the activity of the farm.
Type of intervention	All advisory programs aiming at the adoption of practices likely to reduce the level of pesticide exposure of farm labour and assessing this impact.	Any method. Interventions for the development of integrated pest management (IPM) is relevant, provided that the impact on pesticides exposure for farm labour is assessed.	Interventions to reduce the impact of pesticides on the environment, on food residues etc. Field test of the efficacy of security devices, of machineries, of personal protection equipment (PPE).

Table 1. Inclusion criteria

Geographical descriptors were used to identify the papers concerning interventions in Europe more easily. All types of studies were considered for Europe (including monographs) and are analysed in this paper. A second set of papers were identified in non-European situations. They are not analysed systematically in this paper. We shall only consider some results of a sub selection of these papers, corresponding to interventions that were designed with control groups and explicit measures of effectiveness.

This search on electronic databases has limits as it does not reach unpublished studies, techno-administrative documents, publications with no English abstract that do exist in the EU. However, it provides us with a first estimate of the available literature that could inform public policies and contribute to Farm Advisory Systems adjustments.

6.3 Results of the electronic search

With no geographical restriction, the first step of the electronic search yielded 1515 papers. With restriction to European countries (including non EU countries, e.g. Russia, Turkey), 253 documents were found. After reading the abstracts, 42 documents were selected for in depth analysis. A large number of the papers that were rejected at this stage described exposures to pesticides and/or analysed their impact on farm workers' health and concluded to the necessity of developing advisory services to improve the prevention in this area, but do not present an existing intervention. Among the remaining 42 documents, 4 could not be analysed for linguistic reasons (articles in Russian and Roman), 4 were not scientific papers but administrative reports of professional bodies, 3 were rejected for other reasons (not focussed on Europe etc.). From the reference lists, 4 other documents were identified and added.

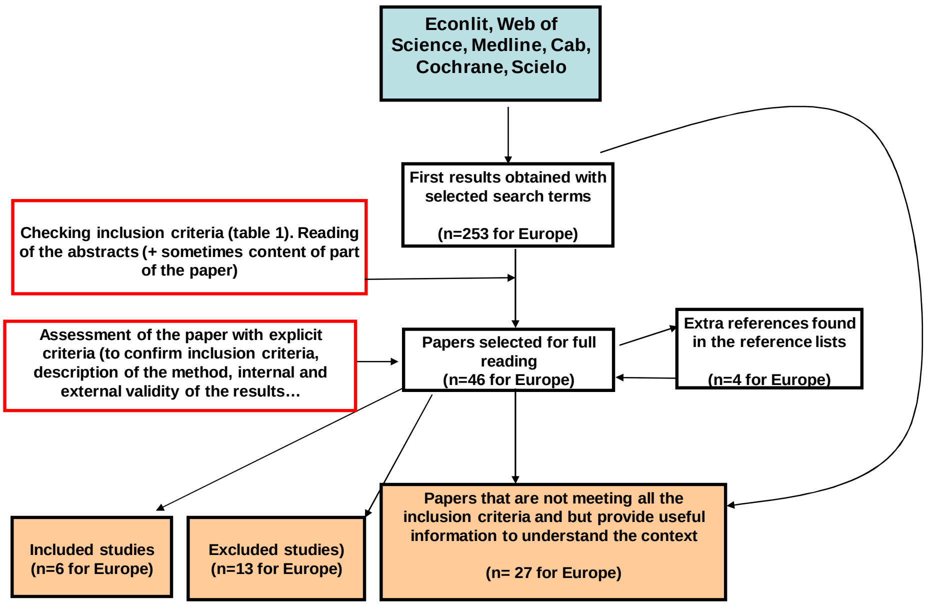


Figure 1. The different steps of the selection of the papers

6.4 Analysis

Finally, 6 papers meet the inclusion criteria in a broad perspective, but not a single one presents the assessment of the effectiveness of an advisory intervention with a robust methodology (precise indicator of effectiveness, methodology with a control group etc.).

- One is a review on the "Effectiveness of interventions in reducing pesticide overexposure and poisoning" that was published in 2000 by M. Keifer. Most of the studies that are analysed concern tests of the efficacy of personal protection equipment, clothing, etc. in very controlled conditions. One cited study (Lander; Hinke 1992) concerns Denmark but it is an epidemiological study that does not measure the impact of a specific advisory intervention. Therefore, Keifer concludes that "*we found no study that examined the effect of prevention programs on pesticide poisoning per se*".

- Three other articles correspond to successive reviews that were conducted by a network of researchers studying interventions for preventing injuries in agriculture (DeRoo et al. 2000, Lehtola et al. 2008, Rautinen et al. 2009). These reviews consider all types of injuries. Regarding pesticides, in 2009 Rautinen et al. concluded that "*Pesticides constitute a serious health hazard to farmers especially in developing countries. From our review, it becomes clear that measures to decrease these risks have seldom been evaluated. The included time series showed some evidence that banning of a pesticide does not lead to illegal or more dangerous use but that it can have a favourable effect on poisoning fatalities".*

The reference list does not identify interventions to reduce farm labour exposure to pesticides in Europe.

-Two other papers (van den Broucke, Colemont 2007, Colemont, van den Broucke 2008) are focussed on the identification of factors that contribute to unsafe and damaging behaviours among farmers in Belgium. They are based on the same set of observations. All occupational hazards are considered, including pesticide use. The level of education, the vocational training, and the safety training are considered in the variables that describe the farmers, and their link with risk behaviour is analysed. However, these articles do not include the assessment of the effectiveness of any specific intervention.

Authors	Type of study	Conclusion
Keifer M. 2000	**Review** "Effectiveness of interventions in reducing pesticide over exposure and poisoning"	Most of the identified studies concern tests of PPE and machinery in a controlled environment "We found no study that examined the effect of prevention programs on pesticide poisoning per se"
De Roo *et al.* 2000 **Lehtola *et al* 2008** **Rautinen et al. *2009***	**Systematic reviews**. Collect of papers with high level of proof. Interventions for presenting injuries in the agricultural industry	Rautinen 2009. “it becomes clear that measures to decrease these risks [pesticides health hazards] have been seldom evaluated"
van den Broucke, Colemont 2007, Colemont, van den Broucke 2008	**Cross section survey** (510 farmers) Determinants of occupational health related to behaviour in Belgium.	Correlations between behavourial risk factors and level of education (<0), and age (>0)

A first conclusion is that we could not identify any paper measuring the effectiveness of advisory methods that could result in less exposure to pesticides of farm labour in Europe. This issue is a blind spot for the research communities who work on Agricultural Knowledge and Information Systems (AKIS) in spite of societal and political demands, and in spite of the warnings from epidemiologists.

6.5 Learnings from foreign experiences: four corner stones for designing and assessing advisory interventions to reduce the pesticide exposure of the farm labour

The situation is slightly different in other countries. Several large scale studies were implemented to assess the effectiveness of advisory programs whose objective was to reduce the exposure of farm labour and their family to agricultural pesticides. It is not possible in this document to give the details of this part of the review. We have selected a few papers (one per program), 5 in the US (Perry, Layde 2003, Vela Acosta *et al*. 2005, Donham *et al*. 2007, Thomson *et al.* 2008, Arcury *et al*. 2009), 1 in Ecuador (Orozco *et al.* 2011), and 2 in India (Sam *et al.* 2008, Mancini *et al.* 2009) in order to draw some lessons from these experiences for the European case.

6.5.1 Beyond plant protection products

In the EU, chemical risk in farm activities is often synonymous with the risks associated with the use of plant protection products. However, a substance, for instance deltamethrin, which is a broad-spectrum insecticide, is not only used in various commercial specialities for plant protection. It can also be used for the treatment of livestock against external parasites, or as a biocide for indoor treatments. This is why we put together these three types of products under the denomination "pesticide" in our inclusion criteria, although they correspond to three different EU regulatory areas. In the available literature in the EU, several authors have a holistic approach of the farmers' activities and chemical risk (e.g. Colemont, van der Broucke 2008, Lopez Crespi et al. 1997, Subias Loren *et al*. 1995). However, most of the papers are focussed on crop protection products, even when they target people working on mixed farming systems. A literature review that was recently completed in France to inventory the available data on farm labour exposure to pesticides, could not identify a single paper related to the chemical risk associated with livestock activities, in spite of strong field evidence of such exposures (e.g. sheep dipping practices) (Anses 2014).

Yet, for decades now, the World Health Organization (WHO) has insisted on the risks associated with the use of biocides for insect control (both indoor and outdoor) (WHO 1979, 1990). It is also known that farm labour can cumulate different sources of exposure. For instance when Reynolds et al. (2007) tried to characterize the exposure to farm chemicals among a population of a county of Iowa (USA), they observed (n=395 people who ever mixed or applied pesticides) that crop insecticides were typically (median) used 10 days per year, livestock insecticides 10 days per year, and fumigants and crop storage insecticides 1 day per year. Although we have not yet identified articles measuring exposures to pesticides in mixed systems in the EU, there are enough analyses of farm practices to conclude that these situations of multiple sources of exposure exist in a significant part of EU farms with mixed production systems. Therefore, it would be logical to consider these three main sources of exposure when designing interventions.

6.5.2 Object of intervention and limits for generalisation of the results

Biological attributes, rules and processes (e.g. the biological impact of a certain dose of chemical) are likely to be more stable and consistent across populations and time than social attributes (e.g. the behaviours or the risk perception in a certain group of the population). Ultimately, all interventions aim at reducing or suppressing the exposure to pesticides but a very wide range of indicators is used to assess the effectiveness of the programmes (table 2). For most of the programs it is based on an evaluation of knowledge acquisition and /or a change of practices (reported or observed). In very few cases it involves a direct measure of biological characteristics such as the level of cholinesterase to trace the exposure to organophosphates (e.g. Thomson *et al*. 2008).

	What is measured (indicator of effectiveness)
1.Knowledge acquisition	**The improvement of the risk perception and the understanding of the rationale of the rules aiming at the reduction of exposures**, e.g. - to read and understand information on the toxicity of the products on the labels, - to integrate this information for making technical choices, - to be aware of the advantages and limits of hygiene rules and protection equipment, - ....
2. Behaviour / practices change	**The adoption of practices reducing exposure to pesticides**, e.g. - to wear adapted protection equipment in a proper way, - to follow adequate rules of hygiene, - to make an assessment of the necessity to use pesticides against alternative methods that take into account their toxicity, - ...
3. Level of exposure	**The actual exposure through direct measure** e.g. - measure of skin contamination at the end of the working day, - measures via biometrical indicators (e.g. level of acetyl cholinesterase when using organo-phosphate pesticides), - ...
4. Health impact	**The final impact of the programme on health and health related expenses** **-** number of injuries - health related expenses - ...

Table 2. Several ways of assessing the effectiveness on an intervention.

In all the cases, advisory interventions in real work conditions involve an exchange of knowledge and a change of behaviours. The causal structures of individual behaviours and the structure of the farm sector may deeply vary from one society to the next. For instance, if we consider the situation of migrant workers in EU farms, we can find some commonalities between the situations of Mexican migrant workers in the US and Polish migrant workers in the South of France (linguistic barriers...) but there are also huge differences (farm structures, level of workers' organisation etc.). This is why the measures of the effectiveness of advisory interventions in the US can provide some lessons to develop further interventions. However, they are not sufficient to generalise their results to the European situation even if they are obtained with robust methodologies (Berriet et al. 2014).

6.5.3 Target population

In the literature available concerning exposure of farm labour to pesticides in the EU, many papers mention the "farmers", i.e. the head of the farm and his (her) spouse, or "farm labour" without further description of this population. However, the choice of the target population is a very sensitive issue for these interventions and the knowledge of the characteristics of the different sub-groups is essential to secure the effectiveness of advisory programmes. Several factors need to be taken into account, for instance:
- Biological characteristics: pesticides may have stronger health impacts on certain groups of the population (e.g. children, teen-agers, pregnant women...).
- Occupational status: some types of workers may benefit from less protection (e.g. casual workers in certain areas...)
- Cultural background and level of education: specific characteristics must be considered by advisory services in particular: the linguistic skills of migrant workers [Vela at al. 2005], the differences in risk perception (according to gender [Arcury et al. 2009] and to cultural background)
- Specific issues related to particular practices or behaviours (for instance children behaviours on the farm [Thompson et al. 2008]).
This is why most of the papers reporting experiences outside the EU stress the need to identify the various groups who are concerned in order to adjust the interventions accordingly.

These recommendations are very relevant for the EU. For instance, thousands of migrant workers (from the EU or from other countries) work in places where they do not fully understand the language. They may have difficulties accessing safety messages (e.g. on the labels of the products or related to the PPE) and they are not conversant with the local safety regulation. Several studies show that this population is likely to suffer from unsafe conditions of pesticide use particularly in vineyards, horticulture, orchards, and Mediterranean agriculture (e.g. Laurent 2013 for the French Mediterranean Agriculture). Actually, all the papers that describe farm labour characteristics in relation with pesticide use in the EU show the large heterogeneity of farm labour population and the need to take it into account (e.g. Garcia *et al.* 2002, Lander, Hincke 1992). All categories of people working on the farm deserve specific attention, including farm advisors (Nicourt, Girault 2009), people who are not employed directly by the farm but work on it.

6.5.4 Institutional building of the occupational health dimension of the Farm Advisory Systems

The EU directive on the sustainable use of pesticides (2009/128/EC) states that in each Member State, the Farm Advisory System (Reg. 1698/2005) should provide adequate advice to meets its objectives, hence to guarantee a safe use of pesticides. The electronic search identified two studies on this topic at the national level, one in Sweden (Hoglund 1990) and a recent one in Ireland (McNamara, Phelan 2008). Most probably, some reports that could provide information exist in several State Members, in their national language. However, we could not identify a scientific paper or a report (for instance form the European Agency for Safety and Health at Work (EU-OSHA)) that would make an inventory of the national initiatives related to this dimension of the Farm Advisory Systems and analyse the various methodologies that are used.

Such studies would be useful to share information and experiences within the EU because foreign experiences show that various institutional frameworks and methodologies can be chosen to set up large scale prevention programs (e.g. "Safe Farm Certification" [Donham et al. 2007], programs based on Lay Health Advisors [Arcury et al.2009], specific programs for migrant workers [Vela Acosta et al. 2005], community-based method programs [Orozco et al. 2008], Farmers' Field Schools [Mancini et al. 2009], effects of a small-group educational intervention designed to increase PPE use and to reduce direct pesticide exposure [Perry, Layde 2003], etc.). The effectiveness of these methods can be measured and the advantage and limits of each method can be assessed (Andrieu 2013, Mora-Sanchez 2013). However, specific studies in the EU context are necessary because the evidence of effectiveness in a certain context is only valid for this population and its environment.

For the time being, this review shows a dramatic lack of reliable evidence in the EU that could help design advisory programs to improve the protection of farm labour from pesticide exposure.

6.6 References cited in the case study

Andrieu T. (2013). *Quelles sont les formes d'intervention mises en oeuvre auprès des agriculteurs pour promouvoir l'adoption de pratiques d'utilisation des pesticides plus respectueuses de la santé des personnes travaillant dans l'agriculture?* Rapport de recherche bibliographique, M1. Inra, AgroParistech.

Anses (2014). (Groupe de travail Travailleurs agricoles et pesticides). *Exposition des travailleurs agricoles aux pesticides. Revue systématique de la littérature scientifique disponible sur les expositions aux pesticides des travailleurs agricoles en France.* Version janvier 2014. https://www.anses.fr/fr/documents/ AIR2011sa0192Ra1.pdf

Arcury T, Marin A, Snively B, Hernández-Pelletier M, Quandt S (2009) Reducing farmworker residential pesticide exposure: evaluation of a lay health advisor intervention. *Health Promotion Practice* Vol. 10, No. 3, 447-455,. Doi : 10.1177/1524839907301409

Berriet-Solliec M., Labarthe P., Laurent C. (2014). Goals of evaluation and types of evidence*. Evaluation. The International Journal of Theory, Research and Practice.* Issue April. In press.

Colemont, A. and S. Van den Broucke (2008). Measuring determinants of occupational health related behavior in Flemish farmers: An application of the Theory of Planned Behavior. *Journal Of Safety Research* 39(1): 55-64.

DeRoo L., Risto MPH., Rautiainen M.S., (2000). A systematic review of farm safety interventions. *American Journal of Preventive Medicine*, 18 (4S), 51-62

Donham K.J., Rautiainen R.H. Lange J.L. (2007. Injury and Illness costs in the certified safe farm study. *The Journal of Rural Health.* vol 23, 4, 348-355.

Garcia A.M., Ramirez A., Lacasana M., 2002. Pesticide application practices in agricultural workers. *Gaceta sanitaria.* 16, 3. 236-40

Garrigou, A., I. Baldi, et al. (2011). Ergonomics contribution to chemical risks prevention: An ergotoxicological investigation of the effectiveness of coverall against plant pest risk in viticulture. *Applied Ergonomics* 42(2): 321-330.

Hoglund, S. (1990). Farmers Health And Safety Programs In Sweden. *American Journal Of Industrial Medicine* 18(4): 371-378.

ILO (International Labor Office) (2000). *Safety and health in agriculture*. Geneva. http://www.ilo.org/public/english/standards/relm/ilc/ilc88/rep-vi-1.htm#4.51.

Inserm (2013). *Pesticides: effets sur la santé.* Une expertise collective. http://www.inserm.fr /actualites/rubriques/actualites-societe/pesticides-effets-sur-la-sante-une-expertise-collective-de-l-inserm

Keifer M : Effectiveness of Interventions in reducing pesticide overexposure and poisonings. *Am J Prev Med* 18(4S):80–89, 2000

Lander F., Hinke K. (1992). Indoor Application of Cholinesterase Agents and the Influence of Personal Protection Uptake. Archive of Environmental Contamination and Toxicology. 22, 163-166

Laurent C. (2013). The Ambiguities of French Mediterranean Agriculture. Images of the multifunctional agriculture to mask social dumping? *Research in Rural Sociology and Development,* 19: 149–171.

Lehtola MM., Rautiainen RH., Day LM, Schonstein E., Suutarinen J., Salminen S., Verbeek JH. (2008). Effectiveness of interventions in preventing injuries in agriculture -a systematic review and meta-analysis. *Scandinavian Journal of Work, Environment &Health.* 34(5): 327-336.

Lopez Crespi, F., J. Brosa Luengo, et al. (1997). Descripción de los hábitos de los agricultores en relación con la aplicación de plaguicidas. *Atencion primaria / Sociedad Espanola de Medicina de Familia y Comunitaria* 20(10): 549-53.

MacNamara J., Phelan J., (2008). Developing the role of extension in farm safety and health in Ireland. *Proceedings of the 24th annual meeting of AIAEE,* 326-337.

Mancini F., Jiggins J., O'Malley M., (2009) Reducing the incidence of acute pesticide poisoning by educating farmers on integrated pest management in South India. *Int J Occup Environ Health* ; 15:143–151

Martin A., Sabine Gündel S., Apenteng E., Pound B. (2011). Review of Literature on Evaluation Methods Relevant to Extension, Lindau (Switzerland), GFRAS – Global Forum for Rural Advisory Services.

Mora Sanchez A. (2013). *Measuring the effectiveness of interventions: the farm advisory case*. Thesis for the Erasmus mundus Master in complex systems. Inra, Ecole Polytechnique.

Nicourt, C. and J. M. Girault (2009). The human cost of pesticides: how French vine growers and vine advisers deal with risk." *Vertigo - La Revue Electronique en Sciences de l'Environment* 9(3):.

Orozco F, Cole DC, Ibrahim S, Wanigartne S (2011). Health promotion outcomes associated with a community based project on pesticide use and handling among small farm households. *Health Promotion International,* Advance Access electronic publication February

Perry J, Layde P (2003) Farm pesticides: outcomes of a randomized controlled intervention to reduce risks. *Am J Prev Med* 24(4):310–315

Rautiainen R., Lehtola MM., Day LM., Schonstein E. Suutarinen J., Salminen S., Verbeek JH. (2009). Interventions for preventing injuries in the agricu*ltural industry (Review). The Cochrane Librairy, Issue 3.*

Reynolds S.J., Tadevosyan A., Fuortes L., Merchant J.A., Stromquist A.M., Burmeister L.F., Taylor C, Kelly K.M. *(*2007). Keokuk County Rural Health Study: Self-Reported Use of Agricultural Chemicals and Protective Equipment. *Journal of Agromedicine.* 45-54.

Sam K, Andrade H, Pradhan L, Pradhan A, Sones S, Rao P, Sudhakar C (2008) Effectiveness of an educational program to promote pesticide safety among pesticide handlers of South India. *Int Arch Occup Environ Health*

Subias Loren, P. J., M. A. Salvador Milian, et al. (1995). Symptoms related to the use of agricultural pesticides. The perspective from the primary care service. *Atencion primaria / Sociedad Espanola de Medicina de Familia y Comunitaria* 16(10): 615-7.

Thompson B., Coronado G.,Vigoren E.,Griffith W., Fenske R., Kissel J., Shirai J., Faustman E. (2008) Para Niños Saludables: a community intervention trial to reduce organophosphate pesticide exposure in children of farmworkers. *Environ Health Perspect* 116:687–694 van den Broucke, S. and A. Colemont (2007). Behavioral and nonbehavioral risk factors for occupational injuries and health problems among Belgian farmers. *Journal of Agromedicine* 16(4): 299-310.

Vela Acosta M., Chapman P., Bigelow P., Kennedy C., Buchan R. (2005) Measuring success in a pesticide risk reduction program among migrant farmworkers in Colorado. *Am J of Ind Med* 47:237–245, 2005.

Verberk, M. M., Brouwer D. H., Brouwer, E. J., Bruyzeel D.P., Bruyzeel, D. P., Emmen, H. H., Van Hemmen, J. J., Hooisma, J., Jonkman, E. J., Ruijten, M. W., Salle, H. J. (1990). Health effects of pesticides in the flower-bulb culture in Holland. *La Medicina del lavoro* 81(6): 530-41.

WHO, World Health Organisation (1979). *Safe use of pesticides.* Third Report of the WHO Expert Committee on Vector Biology and Control. World Health Organization Technical Report Series. WHO: 44 pp.-44 pp.

WHO, World Health Organisation (1990). *Public health impact of pesticides used in agriculture*. Public health impact of pesticides used in agriculture. WHO: 128 p.

6.7 Appendix: Quality assessment of each paper integrated in the review

Author, Date	
Title	
Journal/year	
reviewer / date of the review	
Topic of the study	
Relevant for other topic	

CONTEXT = (who did the study, financing, context, country, types of farm systems, period...)

......

OBJECTIVES =

.....

Are the objectives clearly presented		
YES	NO	NSP

TYPE of document (monograph, survey, RCT, review, analysis of the regulation, etc.

.....

Theoretical view point and analysis of the limits of this point of view

Disciplinary and theoretical choices are transparent	
The resulting simplifications are explained	
The factors influencing the researchers' position are mentioned (e.g. personal links to some actors, financing of the program, interactions in participatory approaches, etc.).	
Ethical issues …	
Others	

Types of exposures (par ex. application, reentry, livestock handling, in-door treatments, outside exposition in contaminated areas…)

DESCRIPTION of the exposure

Place	
Period	
Number of people	
Average Age	

Type of person	Présence O/N	n=	tasks
permanent salaried			
Casual labour			
Family labour			
Other (neighbour, "community" etc.)			
women			
maen			
Childrens (âges ?)			
Migrants (status ?)			

Type of farming systems	
productions	
Size	
Technical indicators	

The caracteristics of the population are clearly described			
Yes	NO	NSP	Not discussed

Inclusion criteria are clearly described and adequate to the tobjective			
Yes	No	NSP	Not discussed

COMMENTS

The types of pesticides is clearly specified			
Yes	No	NSP	Not discussed

Effectiveness of the intervention

Is the objective of the intervention specified (e.g. PPE, hygiene, etc.) ?	
Is the type of intervention clearly described? (par ex. type de conseil, modalité de délivrance du conseil, périodicité…)	
Is there an indicator to measure the impact?	
Quality of the measure (sampling, directness, accuracy of the measure, etc.)	

COMMENTS

Main results

...…

Is there a discussion of the limits of validity of the results?

………

MAIN CONCLUSIONS

Conclusion consistent with the method? (Y / N)	
Limits discussed in the conclusions?	
If extrapolations are proposed, are they justified	
Conclusions aiming at influencing practices? Y/ N Is it consistent with the characteristics of the study?	
Conclusions aiming at influencing policy making? Y/ N Is it consistent with the characteristics of the study?	
Comments	