An FAO e-mail conference on approaches and methodologies in ex post impact assessment of agricultural research: The moderator's summary

John Ruane
Research and Extension Unit,
UN Food and Agriculture Organization (FAO),
Viale delle Terme di Caracalla,
00153 Rome, Italy

Executive Summary

From 5 May to 1 June 2014, the UN Food and Agriculture Organization (FAO) hosted a moderated e-mail conference on "Approaches and methodologies in ex post impact assessment of agricultural research: Experiences, lessons learned and perspectives". It was organized by FAO as part of its contributions to an EU-FP7 project called IMPRESA that is looking at the impacts of agricultural research in Europe.

A total of 618 people subscribed to the conference and, of these, 59 (i.e. 10%) submitted at least one message. Of the 109 messages that were posted, about 30% each came from people living in Africa and Europe; 15% from Asia; 13% from Latin America and the Caribbean; 7% from North America and 6% from Oceania. Messages came from people living in 38 different countries, and 58% were posted by people living in developing countries. A total of 25, 23 and 22% of messages were from people working in universities, national research centres and as independent consultants respectively while 14, 7, 5 and 5% came from people working in the international agricultural research system, inter-governmental organizations, non-governmental organizations and government ministries/bodies respectively.

This document summarizes the main issues that were discussed by the participants during this 4-week conference on ex post impact assessment of agricultural research (epIA-AR). They paid most attention to issues related to micro-level epIA-AR, which typically looks at the impacts of a specific research project in one part of a country, rather than to macro-level epIA-AR, which typically looks at the impacts of investments in agricultural research (or in one of its sectors) at the national level.

Participation was one of the main topics of discussion in the conference. Contributors strongly supported participation of key stakeholders and beneficiaries in epIA-AR processes. They noted that appropriate communication tools were needed to ensure full participation of some stakeholders, such as small farmers. Participation in the epIA-AR processes was perceived as a positive feature in itself and one which also had practical advantages for carrying out epIA-AR, such as enabling improved access to useful data and making it easier to communicate the eventual findings of the impact assessment with key stakeholders and beneficiaries.

Two main kinds of methodologies are used in epIA-AR: quantitative, typically involving data recording and statistical testing, and qualitative, typically involving interviews of key individuals or group discussions. During the conference, while some differences in opinion were expressed regarding the relative merits of using quantitative versus qualitative methods for epIA-AR, there was general agreement about the mutual benefits to be gained from using both. Difficulties in getting data needed for quantitative methods, particularly in developing countries, were discussed.

The challenges of attributing changes that occur within complex and dynamic agricultural systems to specific research-derived interventions were underlined by several participants. They noted, for example, that impacts attributed to a specific research project may be partially the consequence of other interventions in the same target area that were not considered in the analysis. Given the weaknesses they perceived with attribution analysis, some participants emphasized the merits of contribution analysis as an alternative.

The importance of getting relevant and reliable data for epIA-AR was underlined on many occasions throughout the conference. The need for planning to ensure that appropriate data are collected and available for evaluation studies in the future was underlined. Practical challenges were described in getting data for indicators of the social and environmental impacts of agricultural research as well as getting reliable data on adoption rates, because farmers often modify the research-derived innovations to their own situations. To be able to assess the impacts of research institutions, there was a call to collect relevant baseline data.

While the main focus of the e-mail conference was on 'how' to do epIA-AR, there was also ample discussion regarding why and when it should be done and what kind of impacts should be assessed.

Regarding why, the two main objectives mentioned by participants were accountability and learning. It was argued that the objectives might influence what epIA-AR methodology to use (quantitative versus qualitative) or what kind of research initiatives to assess.

Regarding when, while participants seemed to fully support the idea that impacts refer to long-term effects of the research-derived intervention, discussions indicated a lack of consensus regarding what exactly 'long-term' means and consequently when epIA-AR should be carried out. Discussions also indicated that the length of time between completion of the research project and undertaking the epIA-AR could influence the impacts that are found.

Regarding what, several participants underlined that, in addition to economic impacts, it was also necessary to assess the impacts of agricultural research on the non-economic dimensions and they described some of the initiatives that had taken place and approaches used to assess these non-economic impacts. The strengths and weaknesses of using the sustainable livelihoods approach for epIA-AR, encompassing both the economic and non-economic impacts within a single holistic framework, were described.

Three specific methodological issues were also debated in some detail during the conference. The first concerned the merits of social network analysis as a tool in epIA-AR, where participants expressed divided opinions. Second was the validity of propensity score matching to deal with potential 'selection bias' when farmers themselves decide whether to adopt a given innovation, where participants proposed a number of more suitable quantitative methods. Third was about the appropriate criteria to use when choosing case studies for epIA-AR analysis.

Finally, several messages in the conference bemoaned the lack of resources and capacities for epIA-AR and that the subject area was not given sufficient importance by researchers and research institutions.

1. Background to the e-mail conference

The UN Food and Agriculture Organization (FAO) is one of nine partners participating in a 38-month collaborative research project called IMPRESA (http://www.impresa-project.eu/) which began in November 2013 and is funded by the European Union (EU) Seventh Framework Programme. The project aims to assess the economic, environmental and social impacts of research on European agricultural and food systems (Chartier and Midmore, 2014).

IMPRESA explores the connection between research and innovation in agriculture and addresses questions regarding how scientific research affects agriculture and why certain insights lead to innovation and impact, while others do not; how research-derived innovations reach farmers; what the causal chains are for agro-ecological and socio-economic impacts; and what factors enable, foster or limit the effectiveness of agricultural research.

In particular, the project will, inter alia,
- Survey the trends, sources and objectives of agricultural research across Europe

- Carry out macro-level impact assessments, studying the impacts that agricultural research had on productivity and other relevant economic, environmental and social dimensions in selected countries
- Carry out micro-level impact assessments, through in-depth analysis of the impacts that agricultural research had in six specific case studies

The project is therefore looking at the ex post impacts of agricultural research, where the term 'ex post' refers to the fact that the assessment is done after the intervention (which is derived from agricultural research) has been completed (this is in contrast to ex ante impact assessment, which aims to predict the impacts that a research-led intervention may have in the future). In addition, IMPRESA considers ex post impact assessment (epIA) of agricultural research from various perspectives, from the national level down to that of specific research projects while looking at the impacts of agricultural research on economic as well as various non-economic indicators. These different perspectives can be captured using a wide range of potential methodologies and approaches which may have different statistical and conceptual properties, data requirements and practical implications. Large numbers of both refereed and non-refereed publications are dedicated each year to application of the different methodologies, in epIA of agricultural research (epIA-AR) and of other areas, particularly for micro-level epIA [see, for example, Gertler et al. (2011), Khandker et al. (2010) and Walker et al. (2008) and references therein]. In this fast-evolving area, it was therefore appropriate that at the early stages of the IMPRESA project this e-mail conference was held to get an updated overview of the different experiences, lessons learned and perspectives that people have about the approaches and methodologies used in epIA-AR.

The conference took place as one of the activities within the first IMPRESA work package (on Concept Development and Learning), which is led by FAO and has as one of its main aims the establishment of a common framework to update concepts and methodologies for impact assessment of agricultural research. In the week before the conference, the moderator sent the conference Background Document (Ruane, 2014) to the subscribers. The document's aim was to provide an easily-understandable introduction and brief overview of the approaches and methodologies used in epIA-AR. It began with an introduction to the general area of evaluation, with definitions of the main relevant terms, and then narrowed down to epIA, one component of the whole evaluation package. It proceeded to describe the main approaches and methodologies used for carrying out epIA-AR. These were classified into two broad groups: macro-level assessments, typically looking at the impacts of agricultural research (or one of its sectors) at the national level, and micro-level assessments, typically looking at the impacts of a specific research-derived innovation in a given region of a country. The document also briefly considered how the epIA-AR results were communicated to donors and policy-makers and how they subsequently used them. It concluded with a series of questions which participants were asked to address during the conference.

The conference ran from on 5 May to 1 June 2014. In his opening message to the conference, the moderator welcomed the participants to the conference and briefly reminded them of some of its main guidelines, i.e. that participants should introduce themselves briefly in their first posting to the conference; that they are assumed to be speaking on their own behalf and not on behalf of their employers; and that messages should not exceed 600 words. The conference was open for anyone to join and there were just under 500 subscribers when the conference began which rose to 618 on 1 June when the conference finished. Of these, 59 (i.e. 10%) submitted at least one message.

In their introductions, people posting messages typically provided their full work address and a description of their professional background and current occupation. Based on this, an analysis was carried out by country, geographical area and work. Note, the analysis is based on where people were living when they posted the message and does not indicate where they come from originally.

Of the 109 messages that were posted, 30% came from people living in Africa; 29% from Europe; 15% from Asia; 13% from Latin America and the Caribbean; 7% from North America and 6% from Oceania. A total of 63 messages (i.e. 58%) were posted by people living in developing countries.

The messages came from people living in 38 different countries. The greatest number came from people living in Uruguay, Switzerland, the United States of America, Australia, Ghana, India, Hungary, Indonesia, Italy, the Netherlands, Kenya and Nigeria (all with four messages or more).

Participants in the conference also came from a wide range of work environments. Of the 109 messages, 25% were from people working in universities; 23% from national research centres; 22% from independent consultants; 14% from people working in the international agricultural research system, mostly in Consultative Group on International Agricultural Research (CGIAR) centres; 7% from people working in inter-governmental organizations (mostly FAO); 5% from non-governmental organizations (NGOs) and 5% from people working in government ministries or bodies.

2. Summary of the main issues discussed in the conference

For people wishing to consult the original messages, it is recommended to visit http://www.fao.org/fileadmin/user_upload/research_extension/docs/impactmsgs.pdf (630 KB), where the messages have been formatted for easier reading (e.g. the moderator's comments are in italics). Alternatively, they are available at https://listserv.fao.org/cgi-bin/wa?A0=Impact-L where the message archives are searchable, with a 'free text' search button on the right hand side of the webpage. The original messages can also be viewed in chronological order (latest on top) for May and June 2014 at https://listserv.fao.org/cgi-bin/wa?A1=ind1405&L=Impact-L&O=D&H=0&D=1&T=1 and https://listserv.fao.org/cgi-bin/wa?A1=ind1406&L=Impact-L&O=D&H=0&D=1&T=1 respectively.

Messages were numbered from 1 to 109 in order of posting to allow easy cross referencing during the conference. Some of the individual messages are referred to in this document, where the number of the message is provided within brackets with the author's surname. For messages written by more than one person (i.e. nrs. 31, 72 and 88), only the first author's name is provided here. All of the messages posted during the conference can be read in their entirety using the web links provided above.

In this Section, the goal is to summarize discussions on the main issues that received most attention in the conference. These issues are:
- Participation of stakeholders in the epIA-AR processes (Section 2.1);
- Merits of quantitative versus qualitative methods (2.2);
- Difficulties of attributing changes to research-derived interventions (2.3);
- Importance of getting good data for epIA-AR analysis (2.4);
- Motives (2.5) and timing (2.6) of epIA-AR;
- Analysis of non-economic versus economic impacts in epIA-AR (2.7);
- Three specific methodological issues, i.e. social network analysis, selection bias and criteria for choosing case studies (2.8);
- Lack of resources and capacities to carry out epIA-AR (2.9).

Finally, Section 2.10 summarizes a number of other specific issues that were discussed briefly.

Throughout the conference, most discussion focused on issues related to micro-level rather than macro-level epIA-AR.

2.1 Participation of stakeholders and beneficiaries in the impact assessment processes

Participation was one of the main topics of discussion in the conference. Contributors strongly supported participation of key stakeholders and beneficiaries in epIA-AR processes. They noted that appropriate communication tools were needed to ensure full participation of some stakeholders, such as small farmers. Participation in the epIA-AR processes was perceived as a positive feature in itself and one which also had practical advantages for carrying out epIA-AR, such as enabling improved access to useful data and making it easier to communicate the eventual findings of the epIA-AR with key stakeholders and beneficiaries.

As mentioned, contributors strongly supported participation of key stakeholders and beneficiaries in the epIA-AR processes (e.g. Nguyen, 33; Augustyn, 47; Yapi, 51; Vergara, 53; Pareja, 55; Barjolle, 61). For example, Kavoi (36) stated there was a need to "embrace participation/involvement of all concerned stakeholders, especially the farming community on one side because they are better placed to state whether or not the project under review has had any impact on their livelihoods or not and if they could provide evidence to support their information". Yapi (51), supported by Pareja (55), characterized epIA as a multi-disciplinary and participatory undertaking because it needed contributions from all the stakeholders who participated in development of the technology (including plant breeders, agronomists, socio-economists, research managers/administrators and farmers), dissemination of the technology (extension technicians, seed centres and distribution agents, NGOs and farmers) and funding agencies.

Contributors noted that key stakeholders and beneficiaries were often able to participate to different degrees in the epIA-AR processes. Nguyen (5 and 18), supported by Chikezie (105), pointed out that, even when participatory tools have been used in impact assessment, local communities have not been able to participate fully in some cases because local languages were not used. Nguyen (18), supported by Augustyn (47), like Murimi (27), pointed out that appropriate communication tools are essential when dealing with smallholders and Nguyen (54) described a positive example of their use for epIA-AR in the Northwest Highlands of Vietnam. Fieldsend (80) suggested that facilitated group learning was a useful tool, providing an inclusive environment for communication with farmers.

Nguyen (18, 33) differentiated between participation in epIA-AR as "a means", where key stakeholders and beneficiaries are informed and asked to respond to impact assessment questions, or as "an end", where they are empowered to make their assessment and own decisions on how impact could be sustained respectively. Whereas the first might involve people participating in passive ways in impact assessments they considered being carried out for outsiders, he argued that participatory impact assessment should not only help to generate information and statistical data but should also empower local communities towards sustainability (Nguyen, 33).

Participation of stakeholders in epIA-AR processes make it possible to access new data (Olaniyan, 12; Barjolle, 61; Nguyen, 33), although some contributors cautioned about potential bias in the data collected. Mfegue (32), supported by Nguyen (54), suggested that farmers may give biased answers to impact assessment questionnaires if they believe there may be new projects in the pipeline. Tinsley (34) also argued that participatory procedures could easily be manipulated if, for example, participants were asked leading questions. Garrett (37) emphasized, however, that good data can be collected from participatory approaches in epIA but that data validation was very important – "the main idea though is to ensure that no single question is without at least one other validating data point". He explained how bias might arise when using participatory approaches and underlined the importance of training and practice in using them. Kavoi (36), like Tinsley (34), argued that data from participatory approaches should also be complemented by additional independent inputs, concluding that "a combination of the two is more likely to provide evidence-based findings and conclusions of the impact assessment being carried out".

Contributors also argued that one of the advantages of key stakeholders and beneficiaries participating in the epIA-AR processes was that it would make it easier to communicate the findings of the epIA-AR with them. Nguyen (33, 54) and Sheikh (59) noted that these findings are often not shared with key stakeholders such as farmers and local extension actors, and Nguyen (54) argued that the best way to do this was to involve them in the epIA process. Similarly, to get the various stakeholders to read and understand the reports emerging from impact assessment, Augustyn (47) suggested that participatory processes be used to involve them in the epIA processes. To communicate the epIA findings with local stakeholders, Nguyen (54) suggested that it could be done in the field or the most suitable location for them, soon after the epIA processes were completed, and that simple and understandable language should be used. For donors and policy-makers, he suggested that video material or written publications could be produced.

During the running of research projects, it was also argued that stakeholders should participate in monitoring and evaluation (Kavoi, 45; Perez, 57; Barret, 97). Vergara (53) described an example from the Philippines where the farmer-beneficiaries were in charge of monitoring the results of new technologies adopted in the field, concluding that "farmers who have been involved since the planning stage up to implementation and monitoring and evaluation are more likely to have a feeling of ownership for the project and more likely to continue implementing the project once funding is already terminated".

2.2 Quantitative versus qualitative methods

As described in the Background Document (Ruane, 2014), methodologies used in epIA-AR can be grouped into those that are quantitative or qualitative, where quantitative methods can be roughly defined as those which produce data in the form of numbers, typically involving data recording and statistical testing, while qualitative methods tend to produce data expressed in prose or textual forms, typically involving interviews of key individuals or group discussions. The quantitative and qualitative methods have different strengths and weaknesses and their deployment involves use of different skills by the people carrying out the epIA-AR.

During the conference, while some differences in opinion were expressed regarding the relative merits of using qualitative versus quantitative methods in epIA-AR, there was general agreement about the mutual benefits from using both. Difficulties in getting data needed for quantitative methods, particularly in developing countries, were discussed.

Several participants advised using both quantitative and qualitative methods (Yapi, 23; Weeraratna, 26; Nguyen, 33; Mukoyi, 87). For example, for the smallholders' context, Olaniyan (12) argued that quantitative assessment is important but it does not necessarily capture all the essential data needed for impact assessment of an agricultural research or development project and that it might need to be complemented with qualitative approaches such as story-telling and participatory group discussions. In a similar vein, Stigler (15) argued that qualitative methods can enrich quantitative designs for epIA-AR by e.g. using semi-structured interviews to determine whether, and to what degree, the quantitative data might be statistically biased and thus help to identify solutions to overcome such problems.

Whereas Stradiot (25) stated that he would measure the impact of research by asking questions and listening, Suryadarma (70, 83) underlined instead the need for quantitative methods, concluding that as the main purpose of an impact evaluation is to measure effect size (how big the impact is) then "quantitative methods must be used. Without quantitative estimates, there is no impact evaluation" (Suryadarma, 70).

Pareja (76) thought this conclusion sounded good in theory but was impossible to apply rigorously in practice because, apart from the economic dimension where quantitative data might be available, data were often lacking for the social and environmental dimensions, especially in developing countries. Suryadarma (83) accepted the point about data scarcity in developing countries, but argued that many social and environmental outcomes are still quantifiable. Pareja (89) agreed about the need for quantitative data for the social dimension but was concerned with the issue of attribution, as so many other factors could influence social impacts, and argued that qualitative methods could be used to identify some social impacts. Barret (97) noted that the potential solution of using quantitative methods for the economic dimension and qualitative methods for other dimensions of impact might sound reasonable in theory but, "in practice the question of data prevails. Adaptation to the evaluation topic and to the local conditions seems a basic rule".

Stigler (95) also reacted to the conclusion of Suryadarma (70), arguing that several authors recently questioned the possibility of using quantitative methods to produce "single number estimates" unless very strong and somewhat unrealistic assumptions were made. Furthermore, he argued that the whole quantitative impact evaluation methodology was looking at the "effects of causes" or, in most cases, the single causal effect of a single cause, rather than the "causes of effects", to seek the causes of

observed effects. Suryadarma (100), after providing an example he described as a rigorous quantitative impact evaluation enriched by a similarly rigorous qualitative evaluation, summarized that "the issue is not about qualitative vs quantitative methods. The best impact assessment studies that I know of combine many methods, building the body of evidence, explaining the intermediate outcomes, consider external validity, and internalizing (as much as possible) the complexities and contexts - including taking into account other causal factors".

Several participants mentioned the possibility of using impact pathways and theories of change in epIA-AR (e.g. Mayne, 8; Nguyen, 5, 18, 33; Stigler, 30) and some of them discussed their application with respect to quantitative or qualitative methods. Stigler (95) said he was aware of only one quantitative study which had explicitly investigated an impact pathway with more than one arrow linking output to impact and asked if there were others. In response, Suryadarma (100) described an example to assess interventions to improve educational quality in Indonesia, in which he had participated, noting that the project had used a theory of change developed at the start of the project.

2.3 Attribution of changes to research-derived interventions

As described in the Background Document, the issue of attribution, i.e. ascribing a causal link between observed changes and a specific intervention (OECD-DAC, 2010), is a central one in epIA. During the conference, the difficulties of attributing the changes that occur within complex and dynamic agricultural systems to specific research-derived interventions were underlined by several participants. Given the weaknesses they perceived with attribution analysis, some participants emphasized the merits of contribution analysis as an alternative.

Khatiwada (2), Perez (7) and Taye (63, 84, 92) described the complex and dynamic nature of agricultural systems, where there are many interlinked stakeholders and, particularly for agricultural research and extension, there are many actors and sources of information and technology with complex interactions and inputs (Taye, 63). Khatiwada (2), supported by Issaka (3), wondered how all the complexities in such a multi-actor system could be included in the epIA strategies while Taye (63, 84, 92) questioned the role of attribution studies in such complex and dynamic situations, arguing that "the innovation development to dissemination to adoption to impact continuum is characterized by complex and dynamic processes and multi-institutional and plural actors with interweaved roles and contributions. This makes attributing impact to research and extension difficult" (Taye, 84). In a similar vein, Pareja (48) argued that it is usually difficult to attribute impact to a particular body because although an innovation might be the result of an initial institutional effort that led the research work, other stakeholders such as universities, farmer organizations and private companies might also have contributed to its development.

It was argued that such difficulties could lead to misleading results. Taye (84) stated that attribution analysis in research and extension interventions could over-estimate the impacts. Nguyen (5), supported by Mfegue (32), similarly suggested from his review that the impacts from single research projects were sometimes over-estimated due to ignoring overlapping impacts from other initiatives in the same area. Yapi (21) agreed, giving an example where the economic impacts of a new sorghum variety in Chad were boosted by the presence of a FAO-supported seed station in the same area and by the governmental road system which facilitated extension efforts and contributed to the variety's wide adoption by farmers. Yapi (21), however, argued that the problem was related more to transparency than to the approach itself, and that it could be addressed in the epIA-AR by acknowledging all the other projects that could contribute to the impacts and indicating how such contributions were treated in the process. Issaka (28) agreed about the difficulties of attributing impacts to a single project in a target area, but suggested that problems arise because impact assessment specialists are often asked by donors or others to demonstrate what can be attributed to a specific intervention.

Given the weaknesses they perceived with attribution analysis, some participants underlined the advantages of contribution analysis as an alternative (Taye, 84 and 92; Stigler, 95; Barret, 97). Whereas quantitative methods tend to be used for the former, qualitative methods tend to be used for

the latter. For Taye (84), contribution analysis "offers a step-by-step approach designed to arrive at conclusions about the contribution of an intervention to particular outcomes. The essential value of contribution analysis is that it offers an approach designed to reduce uncertainty about the contribution the intervention is making to the observed results through an increased understanding of why the observed results have occurred (or not!) and the roles played by the intervention and other internal and external factors". Suryadarma (86) argued, however, that proper attribution analysis could also be used to provide in-depth understanding into why observed outcomes occurred as a result of an intervention.

Taye (92) underlined that he was not interested in attribution versus contribution as an academic debate as each had their merits, but argued that when assessing complex and dynamic interventions, as in the case of research and extension, most agree that attribution is not always possible and that contribution analysis is appropriate. Stigler (95) felt that contribution analysis could give the same insights as serious quantitative methods but with the "supplementary advantage that it claims to also uncover the causes of the effects". Barret (97) argued that from the French Agricultural Research Centre for International Development's (CIRAD) experience of dealing with complex and long-term case studies, "attribution seems a fallacy" and that the contribution approach could fit with the agricultural innovation system (AIS) framework they had adopted for their projects. Khatiwada (101) informed the conference about recent guidelines to the Donor Committee for Enterprise Development (DCED) Standard for Results Measurement, which also covered the topic of attribution.

2.4 Importance of getting good data for impact assessment

The importance of having relevant and reliable data for epIA-AR was underlined on many occasions throughout the conference. The need for planning to ensure that appropriate data are collected and available for evaluation studies in the future was underlined. Practical challenges were described in getting data for indicators of the social and environmental impacts of agricultural research as well as getting reliable data on adoption rates, because farmers often modify the research-derived innovations to their own situations. To be able to assess the impacts of research institutions, there was a call to collect relevant baseline data.

As Pareja (48) wrote, one of the lessons learned from their work on epIA-AR in Uruguay (Pareja et al., 2011) was that "the main limitation for impact evaluation is the absence of reliable information", from solid baseline data through to data on adoption rates and on evidence-based economic, social and environmental impacts. Several people emphasized that even at the research project design stage, sufficient attention should be given to planning its evaluation in the future (Nguyen, 18; Lwezaura, 38; Sonnino, 43; Marquardt, 49). For example, De la Fuente (42) described how she had collected baseline data for a Chilean project on sustainable production of native forests in order to assess in the future the impacts of a specific intervention (a government law). Kavoi (45) said the planning should encompass the kind of data to be collected (when, by whom and for whom); how the collected data will be analysed and the results shared. Perez (57), supported by Magingxa (79), added that there was also a need to go beyond a static assessment with predefined plans and goals because there might be opportunities to change and innovate during the lifetime of a research project and that these changes would need to be tracked and documented. Data recorded during the project, together with the baseline and endline data, would therefore be used for the final project impact evaluation (Perez, 57).

Three particular challenges regarding data for epIA-AR were highlighted. First, Pareja (48, 66, 76 and 89) noted that data for indicators of social and environmental impacts are lacking or difficult to gather, particularly in developing countries. Suryadarma (83) accepted Pareja's point about data scarcity in developing countries, but argued that macro-level data on many social and environmental indicators were nevertheless available, such as the World Development Indicators from the World Bank.

Second, the difficulties in getting reliable adoption data for use in epIA-AR studies were also mentioned, because research-derived innovations may be modified in different ways by different farmers to suit their particular situation (Tinsley, 35; Lwezaura, 38; Yapi, 41; Pareja, 50). This may make it difficult for evaluators to recognize that a given innovation has been adopted by farmers

(Pareja, 66). To deal with this reality in epIA, Yapi (41) suggested that different classifications of the innovation could be used, with "full size" meaning there was full adoption without any modifications and "medium size" or "small sizes" denoting the extent of the modifications made by farmers before adoption. Pareja (50) suggested, however, that in practice it might not be possible to differentiate between full and partial adoption of an innovation.

Third, Pareja (66) noted that baseline studies are conducted for most development projects but not for research institutions, such as national agricultural research systems or CGIAR centres, which limits the possibilities of being able to assess their impacts. He therefore argued that these institutions should have their own socio-economic units to gather and standardize data, establish baselines and monitor adoption rates over time (Pareja, 48, 66 and 89).

At the national level, Lele (69), supported by Reader (74), underlined the importance of routine data collection, monitoring and evaluation which feeds into policy. She argued that countries need sustained long term capacity to monitor changes in the external and internal environment in a package of areas ranging from research, extension and agricultural education to agricultural finance and market access as together they determine productivity growth.

Finally, in order to make data collection for epIA-AR easier, Wolf (94) suggested that Current Research Information Systems (CRIS), informational tools that are increasingly used to provide, manage and disseminate research information, might be expanded to include information relevant to the impacts of research on society.

2.5 Motives for doing impact assessment

Many of the messages considered the reasons why assessment of the impacts of agricultural research should be carried out (Nguyen, 5, 33 and 54; Lopez, 16; Yapi, 22, 23 and 41; Weeraratna, 26; Issaka, 28; Stevenson, 46; Pareja, 55; Suryadarma, 70 and 100; Stigler, 95; Barret, 97; and Ntakirutimana, 102). The two main objectives mentioned were accountability and learning. It was argued that the objectives might influence what epIA-AR methodology to use (quantitative versus qualitative) or what kind of research initiatives to assess.

The first main reason mentioned was accountability, to investigate whether the resources spent on agricultural research by donors and other research funding bodies had been used properly and wisely and resulted in satisfactory impacts for these investments (e.g. Pareja, 55). As Weeraratna (26) highlighted, this is important because the financial and human resources invested in agricultural research worldwide are very substantial, involving an estimated US$50 billion in 2009. For agencies which have funded specific research projects, epIA-AR allows them to consider whether the investments they made were worthwhile (Yapi, 23). For example, Henríquez (56) described recent evaluations showing how projects supported by the Regional Fund for Agricultural Technology (FONTAGRO) had strengthened research and institutions in Latin American and Caribbean countries and said that these positive results were being used to encourage decision makers to increase public investment in agricultural R&D.

The other main reason mentioned was learning - about which research-derived interventions work the best (Suryadarma, 100) or might be suitable for replication or scaling up (Stigler, 95) and so that donors might be able to identify beneficial areas for future investment (Yapi, 22).

Stigler (95) argued that the goal of epIA could influence the methodologies used. If epIA was carried out for accountability purposes, he suggested that quantitative methods giving more precise estimates but with less analysis of the mechanisms might be most suitable for donors while if carried out to answer policy questions regarding issues such as replication or scaling up, qualitative methods with less precise estimates but with better insights into the mechanisms by which impacts were achieved might be preferred. In response, Suryadarma (100) argued that the goal of epIA-AR is not

accountability but that it enables us to learn which interventions are most successful and that to compare different interventions, quantitative analysis is needed.

Barret (97) argued instead that the main objective of epIA could influence the subject of the evaluation. She said the main objective of impact evaluation at CIRAD, where she worked, was learning and answering beneficiaries and stakeholders' questions and so CIRAD had chosen to evaluate the impacts of sets of projects/programmes. If alternatively the main objective had been accountability to donors, she said they might have chosen to evaluate the impacts of specific research institutions (Barret, 97).

2.6 When to carry out impact assessment

As mentioned in Section 1.3 of the Background Document, the commonly-used OECD-DAC (2010) definition of impacts refers to them as the "positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended". While participants seemed to fully support the idea that impact refers to long-term effects (Garrett, 4; Steane, 11; Lwezaura, 38; Pareja, 58, 85 and 98; Yapi, 91), discussions indicated a lack of consensus regarding what exactly 'long-term' means and consequently when epIA-AR should be carried out. They also indicated that the length of time between completion of the research project and undertaking the epIA-AR could influence the kind of impacts that are found.

Nguyen (5) reviewed existing impact assessment practices in the Northwest Highlands of Vietnam and found that most only involved short-term impact assessment although he noted that research for development often takes a long time to achieve impacts. Stradiot (25), using an example from greenhouse production in Europe, cautioned that it can take a long time before research-derived innovations are adopted by farmers. One of the lessons that Pareja (48) shared from their impact assessment study of the National Institute for Agricultural Research (INIA) in Uruguay (Pareja et al., 2011) was that, although impacts of innovations are measured some years after farmers have adopted them, there are also recently released technologies showing results (not yet impacts) that need to be considered in the epIA-AR as well as others in the pipeline that need to be seen as "potential future impacts".

Yapi (23) argued that the timing of epIA-AR is crucial and, noting that farmers often learn from observing their neighbours, said that epIA-AR would estimate greater impacts if done some years after a successful project was completed rather than directly afterwards (to allow enough time for spread of the technology to farmers who learned from others). Olaniyan (12), on the other hand, argued that impacts estimated in later years might be lower than the short-term impacts because farmers might drop a new innovation if eventual financial or other incentives related to its adoption disappear once a project finished. Lwezaura (38) pointed out that one of the challenges in this area is that farmers might not adopt the innovation consistently over different years. He also reflected on the number of years that should pass before doing epIA of a completed research project/programme and suggested that it would depend on the nature of the research-derived intervention. For example, if the intervention represented improved maize varieties which mature, go to market and change the farmers' income quickly, he argued the epIA-AR could be done sooner (Lwezaura, 38).

On a separate but related issue, Coli (31) also noted that agricultural R&D encompasses different kinds of activities (basic research, applied research and experimental development) and that the time lag between the R&D activity and assessment of its impacts should depend on the activity (e.g. would be longest for basic research).

2.7 Assessing non-economic impacts versus economic impacts

As written in Section 2.2.5 of the Background Document, most of the work in epIA-AR continues to focus on economic impacts but there is increasing interest in assessing the impacts of agricultural research on environmental (including natural resources management and sustainability), social

(including poverty and hunger alleviation), government policy and other non-economic dimensions, although these assessments can represent considerable data and methodological challenges.

Discussions in the conference backed this up. Several participants underlined that, in addition to economic impacts, there is also the need to assess the impacts on the non-economic dimensions and they described some of the initiatives that had taken place and approaches used to assess the non-economic impacts. The strengths and weaknesses of using the sustainable livelihoods approach for epIA-AR, encompassing both the economic and non-economic impacts within a single holistic framework, were also described.

The importance of assessing the impacts of agricultural research on economic indicators was underlined on several occasions (e.g. Yapi, 22; Pareja, 55; Khatiwada, 60). Weeraratna (26) and Nguyen (5, 65) noted that the current focus of epIA-AR studies was on economic impact but called for more emphasis on the other dimensions, because current practices tended to "ignore human, social, physical and natural impacts" (Nguyen, 5). According to Pareja (48, 66) it is, however, much simpler to assess economic than social and environmental impacts because, firstly, indicators of economic impacts are easy to identify and have been standardized whereas limited work has been done on developing reliable and measurable indicators of social and environmental impacts and, secondly, reliable data on these latter dimensions are not available or difficult to collect. He urged therefore that a menu of indicators should be developed and made available to agricultural research institutions to measure the economic, social and environmental impacts resulting from their innovations.

He later described the approach and indicators used to assess the social (Pareja, 85, 89) and institutional (Pareja, 99) impacts of agricultural research by INIA in Uruguay (Pareja et al., 2011). To assess the social impacts, three aspects were prioritized: the connection between INIA's decision makers and staff and the final users of its products; the perception of technology adopters about social changes related to the technologies; and the perception of INIA's researchers about the social impacts of their research (Pareja, 85). Regarding institutional impacts, Pareja (99) noted that research interventions might result in new or improved farmer organizations, as was documented in the INIA study, and that the sustainability of such new farmer organizations was often underestimated.

In order to encompass both the economic and non-economic impacts within a single holistic framework, Nguyen (5, 65) noted that there had been discussion about using the sustainable livelihoods framework to assess the impacts of agricultural research for development. He described the advantages and strengths, as well as some limitations and challenges, to adopting the framework for impact assessment (Nguyen, 65). Pareja (66) agreed that the framework could be used for this purpose but emphasized that consideration had to be given to the limitations mentioned by Nguyen (65) as well as additional ones. Zaal (73) recognized all of these limitations and described how partners in Burkina Faso, Ghana and the Netherlands had tried to develop a systematic way to address these limitations through a method called Participatory Assessment of Development (PADev).

2.8 Three specific methodological issues: Social network analysis, selection bias and criteria for choosing case studies

Three specific methodological issues were discussed in some detail during the conference.

The first was the merit of social network analysis (SNA) as a tool in epIA-AR, where participants expressed divided opinions. As described by Stigler (72), SNA analyses the links between individuals in a network, usually through network maps and various indices of clustering and concentration, centrality and power within the network. Perez (96) provided a clear illustration of how SNA might be applied in a practical farming situation. While recognizing its benefits for documenting the quality and quantity of interactions between stakeholders, Stigler (72, 88) was not convinced about its merits for epIA because of its unclear relevance to achieving impacts; the availability of cheaper and easier alternatives; and its static approach. In response, Affognon (78) argued it could be very helpful for epIA-AR, as knowledge and information dissemination are important for agricultural productivity and

rural incomes. Perez (90) similarly thought it was very useful for epIA, as it can be used to document the diversity and types of interactions between stakeholders. Slavova (104), on the other hand, was not convinced about its usefulness as she had concerns about the reliability of the data used and the relevance of SNA results to the creation and diffusion of innovation. She also argued that less sophisticated methods could be used to understand relationships in a network. Marquardt (49) was potentially interested in using SNA and Garrett (77) noted that people at a university in the United States were using it as a tool for epIA. The possibilities of using electronic fora to overcome problems of costs and of being static were discussed by Garrett (77) and Stigler (88).

The second was the appropriate quantitative method to apply when there is 'selection bias'. As described in Section 2.2.1 of the conference Background Document, when farmers themselves decide whether to adopt a research-derived innovation there may be selection bias because the farmers who decide to adopt might be different from those who do not. Several quasi-experimental methods exist to deal with selection bias, including propensity score matching (PSM). Ba (13) described a project in Senegal for which she was planning to carry out an epIA-AR, where new farming techniques had been introduced and the farmers themselves had decided whether to adopt them, and she asked whether PSM was the most suitable method to use to deal with selection bias. In response, participants described the theoretical background and weaknesses of PSM and proposed a number of alternative quantitative methods they considered to be more suitable (Stevenson, 14; Suryadarma, 64 and 70; Bonou, 81). Stigler (15) also suggested how qualitative methods might be used to support the quantitative methods e.g. to investigate the kind of selection bias introduced and the variables that might be recorded to adjust for it.

The third related to the appropriate criteria used to choose case studies for epIA-AR. Midmore (20) was concerned that, when deciding which case studies to analysis for epIA-AR, researchers might be tempted to choose ones they found interesting and which supported hypotheses they believed to be true. He said there was certain evidence of this bias from the international relations field and proposed that in evaluating the impacts of agricultural science research, there should be an active search for negative case studies to balance the selection of positive ones. In response, Garrett (29) recommended that, to control for bias, all cases should be recorded and that selection of case studies should then be at random. Stigler (44) reported the views of Walker et al. (2008) who argued instead that, because most agricultural research projects had little or no impacts while a minority had major success, random selection of case studies was not appropriate and that it was better to document key success stories.

In choosing a sample of case studies for analysis, Stigler (44) pointed out that a key consideration was whether the aim was to generalize from the results to other contexts or simply to identify the lessons learned etc. from those particular case studies. Barret (97) said that CIRAD was planning to select 12 case studies from 60 but she did not indicate the selection criteria that would be used. She noted that an important issue to consider was whether to use a standardized epIA-AR methodology for all of them or a set of methods to adapt to the diversity of the different case studies. Fawki (106) proposed that a case study databank be established, describing the methodologies, key factors and weak points for each one, which could be used to get an overview of approaches that might be applicable for a specific case study. Olapede (52), supported by Fawki (106), suggested that systematic reviews and meta-analysis were two important tools that could be used to critically examine the variable results that might emerge from analysis of different case studies.

2.9 Lack of resources and capacities to carry out impact assessment

Several messages in the conference bemoaned the lack of resources and capacities for epIA-AR and that the subject area was not given sufficient importance by researchers and research institutions.

Nasar (17) argued that there are a substantial number of multi-institutional and multinational agricultural research programmes which are competently monitored in most cases but that very little attention or funding is granted to their epIA-AR. He also suggested that expertise to carry out well-defined and statistically valid epIA was largely absent and that FAO and governments should establish

independently functioning institutions with expertise in epIA to assist. Similarly, Babu (93) maintained that the capacity for conducting impact assessment was grossly missing, and not fully recognized and appreciated. The need for training in use of specific epIA methodologies was underlined by Ba (13) and Garrett (37).

In developing countries, participants noted that lack of research funding (Sheikh, 59; Yakubu, 75) and institutional/infrastructural weaknesses [Alkhazraji (71), supported by Yakubu (75) and Mukoyi (87)], made it difficult to evaluate agricultural research. From her experiences in research projects for small farmers, Mfegue (32) stated that epIA-AR is often omitted or, if included in the project package, its allocated time tends to be cut for different reasons, such as scientists wishing to implement new projects. Nguyen (18) noted that resources allocated to impact assessment were often limited, which could restrict the full participation of key stakeholders and beneficiaries in the epIA process because participatory sessions were often time-consuming and costly.

Garrett (37) argued that, for many people, evaluation is considered unimportant and something that has to be tolerated. In Nigeria, Yakubu (75) said the government does not prioritize evaluation of research output. Pareja (66) underlined the need to streamline the "evaluation culture" within research institutions as many still refuse to adopt a strict, evidence-based evaluation system of individual researchers, projects, programs and institutions.

From the large numbers of individuals who subscribed to the e-mail conference and from correspondence with many of them before and during the conference, the moderator's impression is that there is very large demand for clear and easily-accessible knowledge in the complex and fast evolving area of epIA-AR. This indicates that appreciation of the importance of impact assessment, as well as the demand for capacity development in this area, are both strong.

2.10 Other issues

Sections 2.1 to 2.9 summarize discussions of the main issues that received most attention during the conference. A number of other specific issues relevant to epIA-AR were also briefly discussed, including:
- that in macro-level epIA-AR studies where research expenditure is the independent variable, it may be necessary to consider the relationship between public or private research funding when trying to attribute effects because the amount of public research funding might affect the amount of private research funding and vice versa (Coli, 31; Midmore, 82).
- whether epIA methods developed for one agricultural sector (crop, livestock, forestry, fishery or agro-industry) could be transferred to others, where Murimi (27) thought this was very possible while Muralidharan (103) expected for the forestry sector that, given the longer time horizon and specific issues involved (relevance to indigenous people, biodiversity and climate change), the methods used would differ compared to the other sectors. Arunachalam (42) wanted specific protocols for impact evaluation of seeds and planting materials of improved/local varieties.
- the use of total factor productivity or partial productivity measures in epIA-AR (Knickel, 19; Pareja, 55; Lele, 69) and the need to consider that all agricultural research initiatives result in capacity development, which can increase productivity independently of the specific results of the research initiative being evaluated (De Meyer, 67).

Several messages also considered how agricultural research could be improved to achieve greater impacts, for example through better research planning to ensure that research projects focus on the real needs and operational conditions of farmers or by changing the incentives for researchers so they prioritize achievement of impacts rather than development of new technologies or publication of scientific articles (Ntsefong, 6; Augustyn, 9; Tinsley, 10, 35 and 108; Lopez, 16; Stradiot, 25; Garrett, 29; Yapi, 51; Mukoyi, 87; Wolf, 94; Ragnekar, 109).

3. References, abbreviations and acknowledgements

Chartier, O. and P. Midmore. 2014. European scientific research on agriculture: Trends, sources and objectives. Poster presented at the 14th Congress of the European Association of Agricultural Economists, Ljubljana, 26-29 August 2014. http://www.impresa-project.eu/fileadmin/user_upload/IMPRESA/Filesharing/EAAE_Poster_4.pdf (430 KB).

Gertler P.J., Martinez S., Premand P., Rawlings, L.B. and C.M.J. Vermeersch. 2011. Impact evaluation in practice. World Bank. http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf (3 MB)

Khandker, S.R, Koolwal, G.B. and H.A. Samad. 2010. Handbook on impact evaluation: Quantitative methods and practices. World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/2693/520990PUB0EPI1101Official0Use0Only1.pdf (3 MB).

OECD-DAC. 2010. Glossary of key terms in evaluation and results based management. Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC). http://www.oecd.org/dac/evaluation/2754804.pdf (600 KB)

Pareja, M., Bervejillo, J., Bianco, M., Ruíz, A. and A. Torres. 2011. Evaluación de los impactos económicos, sociales, ambientales e institucionales de 20 años de inversión en investigación e innovación agropecuaria por parte del Instituto Nacional de Investigación Agropecuaria (INIA) - Uruguay. Instituto Interamericano de Cooperación para la Agricultura (IICA). http://www.iica.int/Esp/regiones/sur/uruguay/Paginas/Innovacion/EvaluacionImpactoMultidimensional.aspx

Ruane, J. 2014. Background document to the FAO e-mail conference on "Approaches and methodologies in ex post impact assessment of agricultural research: Experiences, lessons learned and perspectives" (5 May to 1 June 2014). http://www.fao.org/docrep/019/as549e/as549e.pdf (120 KB)

Walker, T., Maredia, M., Kelley, T., La Rovere, R., Templeton, D., Thiele, G. and B. Douthwaite. 2008. Strategic guidance for ex post impact assessment of agricultural research. CGIAR Independent Science and Partnership Council (ISPC). http://www.fao.org/docrep/011/i0276e/i0276e00.htm

ABBREVIATIONS: CGIAR = Consultative Group on International Agricultural Research; CIRAD = French Agricultural Research Centre for International Development; EpIA = Ex post impact assessment; EpIA-AR = Ex post impact assessment of agricultural research; FAO = UN Food and Agriculture Organization; IMPRESA = Impact of Research on EU Agriculture (an EU-FP7 project); INIA = National Institute for Agricultural Research (in Uruguay); PSM = Propensity score matching; SNA = Social network analysis.

- Document published 15 September 2014.
- Recommended citation for this publication:
Ruane, J. 2014. An FAO e-mail conference on approaches and methodologies in ex post impact assessment of agricultural research: The moderator's summary. FAO. http://www.fao.org/nr/research-extension-systems/res-home/news/detail/en/c/217706/