

Document: EB 2012/107/INF.7  
Date: 4 December 2012  
Distribution: Public  
Original: English

**E**



Enabling poor rural people  
to overcome poverty

## **Methodologies for Impact Assessments for IFAD9**

### **Note to Executive Board representatives**

#### Focal points:

#### Technical questions:

**Thomas Elhaut**

Director, Statistics and Studies for Development  
Tel.: +39 06 5459 2491  
e-mail: t.elhaut@ifad.org

**Marzia Perilli**

Associate Researcher, Statistics and Studies for  
Development  
Tel.: +39 06 5459 2781  
e-mail: m.perilli@ifad.org

#### Dispatch of documentation:

**Deirdre McGrenra**

Head, Governing Bodies Office  
Tel.: +39 06 5459 2374  
e-mail: gb\_office@ifad.org

Executive Board — 107<sup>th</sup> Session  
Rome, 12-13 December, 2012

---

**For: Information**

## Contents

<b>Abbreviations and acronyms</b>	<b>ii</b>
<b>Executive summary</b>	<b>iii</b>
<b>Methodologies for Impact Assessments for IFAD9</b>	<b>1</b>
A. Purpose of this information paper	1
B. IFAD9 commitments related to impact evaluation	1
C. State of monitoring and impact evaluation in IFAD supported country programmes	2
D. Learning from experience with impact evaluations for agricultural development projects	4
E. Methodologies for IFAD9 impact evaluations	6
F. Partnerships and anchoring M&E and IE in national systems	15
G. Financing M&E and Impact Evaluations	16
H. Conclusion	18

## Abbreviations and acronyms

<b>3iE</b>	International Initiative for Impact Evaluation
<b>ASAP</b>	Adaptation for Smallholder Agriculture Programme
<b>ATAI Agricultural</b>	Technology Adoption Initiative
<b>BMGF</b>	Bill and Melinda Gates Foundation
<b>CAMARI</b>	Programme for Capacity Building in Managing for Results and Impact
<b>CEVAL</b>	Center for Evaluation in the University of the Saarlands U.K Department for International Development
<b>IaDB</b>	Food and Agriculture Organization Inter-American Development Bank
<b>IE</b>	Impact Evaluation
<b>IFPRI</b>	International Food Policy Research Institute
<b>KM</b>	Knowledge Management
<b>LSMS</b>	Living Standards Measurement Study
<b>MCC</b>	Millennium Challenge Corporation
<b>M&amp;E</b>	Monitoring and Evaluation
<b>NONIE</b>	Network of Networks on Impact Evaluation
<b>OECD</b>	Organization for Economic Co-operation and Development
<b>PREVAL</b>	Regional Platform for Evaluation Capacity Building in Latin America and the Caribbean
<b>PROSPERER</b>	Support Programme for Rural Microenterprise Poles and Regional Economies
<b>PSM</b>	Propensity Score Matching
<b>PSR</b>	Project Status Report
<b>RCT</b>	Randomized Controlled Trial
<b>RIMS</b>	Results and Impact Management System
<b>RMF</b>	Results Measurements Framework
<b>USAID</b>	U.S Agency for International Development
<b>USG</b>	U.S Government
<b>WB</b>	World Bank
<b>WFP</b>	World Food Programme

## Executive summary

1. This paper responds to the IFAD9 deliverable requiring an information paper to the December 2012 Executive Board on the methodologies IFAD will employ in carrying out impact assessments and in measuring 4 impact-level indicators: household asset ownership index, length of the hungry season, child malnutrition and number of people moved out of poverty. The targets for the first three indicators are to be "tracked" and in relation to the fourth indicator of poverty the targets are: 90 million people received services from IFAD-supported projects, cumulatively from 2010 onwards to 2015; and 80 million of these people moved out of poverty.
2. By end-2015 IFAD will conduct, synthesise and report on approximately 30 project impact surveys, of which three to six will use Randomised Controlled Trials (RCT) or other similarly rigorous methodology.
3. A range of coordinated initiatives is underway to improve IFAD's performance in monitoring & evaluation (M&E) and impact evaluation (IE). With the current proactive thinking, planning and action, IFAD is well placed to improve its M&E and IE performance and is therefore in a reasonable position to meet the IFAD9 impact evaluation related commitments.
4. A number of institutions have summarised IE experiences in authoritative publications. These findings, especially those related to agricultural and rural development for poverty reduction and food security, guide IFAD's impact evaluation work: namely, its approach and its methodologies.
5. IFAD will use a mix of methodologies (mainly quantitative) for measuring and reporting on results and impacts. Standing project monitoring systems will be used to report on the number of people that received services from IFAD supported programmes. Expert institutions will be invited to carry-out 30 rigorous impact evaluations in order to identify the number of people moved out of income-poverty, define the respective impact pathways for learning purposes, and extrapolate poverty reduction impacts from a statistically sound project sample framework.
6. The report, by end-2015, will synthesise the 30 impact evaluations: aggregate the measured results and impacts, summarise the lessons learned on impact pathways and advise on rigour and cost-effectiveness of different impact evaluation methodologies.
7. IFAD will conduct the 30 impact evaluations in close collaboration with expert institutions and other development partners, and will seek to work within national M&E systems and strengthen them in the process.
8. To finance these impact evaluations IFAD will combine country programme resources, grant resources and supplementary financing from development partners.

## Methodologies for Impact Assessments for IFAD9

### A. Purpose of this information paper

1. This information paper responds to Article II GC Resolution 166/XXXV<sup>1</sup> and to the corresponding IFAD9 commitment matrix endorsed by the Governing Council<sup>2</sup> to “present an information paper to the Executive Board on the methodologies IFAD will employ in carrying out impact assessments and in measuring the new impact-level indicators introduced in the Results Measurement Framework (RMF) 2013-2015” (para 45, bullet 2).

### B. IFAD9 commitments related to impact evaluation

2. The IFAD9 Consultation recommended an “enhanced thrust on impact evaluation”, which was endorsed by the Governing Council (page iii, para 4, bullet 4). “The enhanced thrust on impact evaluation and measurement represents the most significant improvement” to IFAD’s results management approach. For this purpose 4 indicators have been specified (page 12, para 42):
  - (a) **Household asset** ownership index, as a proxy for the income of target group households;
  - (b) Length of the **hungry season**, as a measure of food security for target group households;
  - (c) **Child malnutrition**, as a measure of nutrition security of target group individuals; and
  - (d) Number of people moved **out of poverty**, relative to a defined poverty line.
3. The RMF 2013-2015 (annex II, pages 22 and 23) leaves the targets for the first three indicators to be “tracked”, and in relation to the fourth indicator it specifies the targets for:
  - (a) Outreach (or efficiency): 90 million people received services from IFAD-supported projects, cumulatively from 2010 to 2015 (6 years);
  - (b) Impact (or effectiveness): 80 million people moved out of poverty, cumulatively from 2010 to 2015 (6 years).
4. The IFAD9R (para 18, bullet 2) also calls for initiative to enhance indicators to measure impact and results in gender equality and women’s empowerment.
5. In addition to the commitment highlighted in paragraph 1 above, the IFAD9R (para 23 and 45) specifically calls for:
  - (a) Conducting, synthesising and reporting on approximately 30 project impact surveys over the IFAD9 period. Three to six of these will use randomised control trials (RCT) or other similarly robust or **rigorous methodology**, depending on cost-sharing opportunities, interest and availability of institutions specialised in impact evaluation to support this work (para 45, bullet 4);
  - (b) Raising the level of compliance with the requirement for projects to have a baseline survey by the end of their first year of implementation (para 45, bullet 1);

<sup>1</sup> Annex V of GC 35/L.4. II. Measuring results, effectiveness and efficiency . (a) During the replenishment period, the Results Measurement Framework set forth in annex II to the Ninth Replenishment report shall constitute a systematic approach to management, monitoring and measurement to ensure that the intended results have the greatest likelihood of being achieved.

(b) In order to raise the capacity of the Fund to effectively and efficiently manage on-going operations and deliver the programme of work, the Executive Board and the President shall adopt the measures and undertake the actions set forth in annex I to the Ninth Replenishment report.

<sup>2</sup> GC 35/L.4, Annex I to the “Report of the Consultation on the Ninth Replenishment of IFAD’s Resources” (covering the 2013 – 2015 period) hereafter referred to as the IFAD9R – for: IFAD9 Report.

- (c) Strengthening national M&E systems by enhancing the capacity of project management staff and implementing partners, particularly at start-up and early project implementation through the systematic engagement of M&E experts during design and supervision missions (para 23);
  - (d) Active pursuit of partnerships with institutions specialised in impact evaluation, and resource mobilisation to develop adequate internal capacity to conduct and manage impact evaluation work (para 45, bullet 3).
6. These IFAD9 commitments call for two sets of management initiatives: **immediate actions** for 2013 to 2015 and **longer term capacity building actions** for 2013 and beyond:
- (a) Immediate action focuses on organising up to 30 rigorous impact evaluations to be synthesised and presented to the Executive Board in December 2015. This encompasses enhanced compliance with the current RIMS requirements, building partnerships with centres of excellence in impact evaluation and working with national systems and capacities. **This is the core focus of this information paper.**
  - (b) The longer term actions will also be initiated in 2013 and will run simultaneously with the immediate actions. These will entail improvements in IFAD's M&E systems, defining the most appropriate impact evaluation methodologies and approaches (relevance, rigour and cost-effectiveness), country level capacity building and expanding partnerships with centres of excellence in the field of impact evaluation.

### **C. State of monitoring and impact evaluation in IFAD supported country programmes**

7. As a matter of policy, IFAD funded programmes are underpinned by a logical framework - a hierarchy of objectives, with assorted measurable indicators - since 2003 linked to the RIMS (IFAD's mandatory Results and Impact Management System). To administer this logframe as a results-based management instrument, IFAD's detailed programme designs provide financing for project M&E functions, typically in the order of 3% of programme costs, mostly as an integral part of Project Management functions.
8. The 2012 portfolio reviews of the regional divisions show, however, that in aggregate, implementation of M&E functions falls short of design, even though RIMS compliance continues to improve. There are delays in the establishment of M&E units and in the appointment of M&E staff. Staff numbers, terms of reference, competencies and experience do not always meet requirements. M&E activities are not carried out systematically, which results in under-disbursement of resources allocated. The role of M&E as an instrument for Managing for Development Results (MfDR) is not always recognised by national Project Directors. M&E implementation and RIMS compliance are partial, in part due to shortcomings in design, and in part due to perceptions of RIMS as an IFAD-owned instrument, not always integrated in national M&E systems. Compliance with baseline, mid-term and completion RIMS survey requirements is partial. Relevance, adequacy and quality of data is variable, with a focus on input and outputs, less on outcomes and impact. The quality of surveys and the pertinence of analyses are uneven; and the timing of reports is inconsistent. This limits the use of M&E results for the purpose of: identifying impact pathways; impact attribution; learning and performance enhancement; defining scaling-up pathways; evidence-based policy dialogue.
9. Supervision and implementation support missions, as well as follow-up, especially by IFAD Country Office staff, contribute to improving M&E performance. Project Status Reports (PSRs), completed by the Country Programme Managers at the end of supervision missions, show that systematic follow-up improves M&E performance scores over time. In addition, PMD and the Regional Divisions are pursuing a number

of structural initiatives to improve M&E performance and RIMS compliance. Below is a sample of what is being done by the regional divisions:

- i. The **Asia and the Pacific (APR)** division is disseminating and supporting a M&E/KM Toolkit; is replicating the Madagascar model for country programme M&E (Nepal); experiments rolling surveys, carrying out random selection in a phased approach, for divisible or modular projects; has piloted Annual Outcome Surveys (India) as well as Annual Thematic Surveys (Bangladesh); has piloted in Vietnam, with IFPRI, a RIMS+ methodology based on an expanded and more flexible questionnaire using a control group, right-sizing both treatment and control groups and using GPS.
  - ii. The **East and Southern Africa (ESA)** division has is introducing participatory M&E processes, particularly in decentralised government systems where M&E performance may fall short of requirement (Malawi, Rwanda); has established M&E Thematic Groups In-Country (Madagascar); is studying how to capitalise on the experience with measuring monetary poverty reduction related to specific micro-enterprise models (PROSPERER<sup>3</sup>, Madagascar).
  - iii. The **Latin America and the Caribbean (LAC)** division is exploring the potential for using data generated within country statistical systems in M&E processes in order to apply it at a large scale in LAC (Nicaragua); integrates the results of country's Living Standards Measurement Studies (LSMS) in recent COSOP documents for analytical and M&E purposes; and will be piloting an RCT of financial inclusion (Peru).
  - iv. The **Near East and North Africa (NEN)** division has introduced peer-to-peer support (the Sudan) among projects; mobilises its regional social network of national M&E experts (trained under the CAMARI<sup>4</sup> grant); has developed hands-on approaches and M&E tools for measuring development effectiveness (funded by the CEVAL<sup>5</sup> grant); and is piloting "the Planner", an Excel-based application to tabulate results data from projects (Turkey).
  - v. The **West and Central Africa (WCA)** division is anchoring M&E functions in national systems (Sierra Leone). Most importantly, the division has been a pioneer in partnering with world-class research institutions in implementing RCTs, less for ex-ante experimental design than for concurrent assessment of effectiveness/impact of specific project interventions for IFAD's target groups (Sierra Leone and Ghana).
10. In view of the difficulties of monitoring and evaluating adaptation to climate change, the **Environment and Climate Division (ECD)** coordinated Adaptation for Smallholder Adaptation Programme (ASAP) is expected to develop indicators for measuring resilience; will make use of the Geographic Information System (GIS) to better monitor landscape use. The ASAP provides for the use of experimental methods (RCTs) to test adaptation approaches, for this reason a 3IE (International Initiative for Impact Evaluation) conducted workshop, was organised jointly by ECD and SKM, on methods to evaluate adaptation to climate change.
11. **PMD** provides technical support to the regional divisions for compliance with M&E and RIMS requirements, and part of this is coordinated through a standing departmental RIMS Task Force. PMD is addressing the M&E/IE related issues, in addition to SMART<sup>6</sup>(er) design of indicators, the size of the treatment and control groups, the merit of panel data and a more flexible design of questionnaires for the RIMS surveys to cater for differentiated learning agendas by project type. Most

<sup>3</sup> Support Programme for Rural Microenterprise Poles and Regional Economies

<sup>4</sup> Programme for Capacity Building in Managing for Results and Impact

<sup>5</sup> Center for Evaluation in the University of the Saarlands

<sup>6</sup> Specific, Measurable, Attainable, Relevant, Timely

importantly, PMD's 2013-2015 Medium Term Plan (MTP) is to set out regional targets for poverty reduction (80 million people) and numbers of people reached (90 million), applying approaches used for defining the aggregate IFAD9 targets.

12. The **Strategy and Knowledge Management Department (SKM)** is set to support PMD in the above activities. SKM is completing the recruitment of staff with expertise in M&E design, impact evaluation and econometric analysis, to provide methodological, technical and analytical support, as well as support in training activities to PMD for the IFAD9 IE agenda.
13. The **Independent Office of Evaluation of IFAD (IOE)** will: i) systematically participate in in-house discussions on impact evaluations; ii) contribute to reviews of approach papers/TORs as well as draft final reports of the impact evaluations conducted; iii) undertake one (up to 3) impact evaluations between 2013-2015; and iv) assist in providing ratings (on an agreed scale) of the impact evaluations conducted, in accordance with the harmonization agreement between IOE and IFAD Management.
14. The challenges related to M&E/IE are not specific to IFAD alone. According to Brookings' scaling-up work with IFAD, the Fund may be doing better than other development partners. However, this does not leave room for complacency. With the current pro-active thinking, planning and action, IFAD is well placed to improve its M&E and IE performance, probably even in the short-run, and is therefore in a reasonable position to meet the essence of the IFAD9 IE related commitments to: test different IE methods, with approximately 30 impact surveys and report on them with a synthesis report by end-2015.

#### **D. Learning from experience with impact evaluations for agricultural development projects**

15. A number of institutions<sup>7</sup> have summarised IE experiences in authoritative publications. These findings, especially those related to agricultural and rural development for poverty reduction and food security, guide the IFAD9 IE work. In turn, IFAD's 2015 synthesis report will seek to meet the standards to contribute to this recognised body of knowledge. As highlighted generally<sup>8</sup>, two categories of factors make evaluation in the agricultural sector challenging: 1) agricultural development projects pose sector-specific challenges for rigorous impact evaluation; and 2) evaluation approaches may cause challenges for flexible implementation of projects.
16. Agricultural development projects pose challenges for rigorous impact evaluation. The nature of projects, working through governments and with participatory approaches, makes conducting IEs challenging. Results and impacts in agricultural and rural development have long gestation periods and are highly responsive to **crop cycles and seasonality**, which determines the windows for IE surveys. Agriculture is sensitive to temperature, rainfall and **weather shocks**, conditions that change from year to year and that can be location specific; this poses challenges for control group selection and for IE. Agricultural development and project implementation approaches evolve significantly over the course of a project cycle in response to changing **market conditions** (e.g. the 2008 food crisis and the ensuing price volatility). While adjusting implementation approaches makes interventions more effective, it challenges the validity of evaluations and reduces the potential for learning what really works. For projects that finance a range of integrated interventions (e.g. farmer training, access to credit, and infrastructure investments and value chain support) **sequencing** is important for achieving the desired outcomes, and this complicates timing and approach for IE. **Spill-over or demonstration effects** are common in agricultural development projects, can be

<sup>7</sup> Most notably MCC, WB, laDB, AsDB, OECD, IFPRI, 3IE, CGAP, DFID, NONIE

<sup>8</sup> and quite specifically by MCC in its recent agricultural IE review



quite large and are often an explicit component of program logic, particularly in the case of agricultural projects where technology (such as planting in rows and weeding instead of broadcast seeding) is easily transferred; this poses IE methodology challenges. Project components that target agriculture-related **national-level policy reform or institutional change**, such as new rural institutions, water use or land tenure regulations, credit regulations, improved agricultural services or improved linkages to export channels for targeted value chains, are not generally conducive to identifying a within-country counterfactual. An important constraint for impact evaluations is the degree of **self-selection** involved in the participation of beneficiaries in the agricultural development projects. This implies that only certain types of farmers may choose to participate in a given project, requiring the evaluation to isolate the impacts of a given intervention from the influence of unobservable characteristics of the individuals that chose to participate.

17. Evaluation approaches, in turn, can cause challenges for implementing agricultural development projects. Implementers face incentives to meet various levels of development targets, regardless of how this affects the evaluation methodology. On the other hand evaluators have incentives to adhere to evaluation approaches, regardless of how this may limit flexibility of implementers to adapt to changing conditions and new information. This is especially the case for experimental design or Randomised Controlled Trials (RCTs). Strict adherence to a rigorous impact evaluation methodology, defined at project start-up, may limit implementers' **ability to adapt implementation approaches** in response to changing conditions or new information. While impact evaluations can be designed to be remarkably robust to explicitly planned ranges of adaptation by implementers, an intervention without such planned adaptation structure poses a serious evaluation and learning risk. It may also pose a significant investment risk for the financier. The correct selection and proper management of treatment and control groups challenges project implementers (especially for panel data recording). **Randomized selection of communities or beneficiaries** may be socio-politically challenging or mean that the project may not work simultaneously with all eligible beneficiaries, but instead pursue a pipeline or portfolio approach. Finally, there is the question of **incentives for project implementers** to support a rigorous evaluation of project results and impacts versus reporting on monitored input delivery and outputs. Often project implementers are unclear about **costs and value of IE** and particularly the level of efforts required from implementers to design and implement a rigorous impact evaluation can be underestimated.
18. In **designing its approach and methodology** for each of the 30 impact evaluations, IFAD will factor-in these challenges and related lessons learned. As recommended, the design of each impact evaluation methodology will define at the outset what is most important to be achieved on the ground, what exactly needs to be learnt, and how to integrate the two, with room for adaptation and trade-offs. The evaluation methodology design will then:
  - (a) focus on the most important learning goals;
  - (b) define the stable or essential components of the project that are indispensable to the project's theory of change, and anticipate possible change;
  - (c) ensure the causal pathway to be tested is clear and reflected in the evaluation approach;
  - (d) determine what components of the program logic are evaluable and have a credible counterfactual; and
  - (e) if necessary, look for opportunities to do smaller-scale evaluations within the overall project to specifically target learning about the effectiveness of one intervention relative to another;

- (f) explicitly assess, through modelling, how non-evaluable components can affect outcomes, to capture the interactions between program components;
  - (g) keep in mind the sector-specific challenges of agricultural development projects - synchronise IE with crop cycles and seasonality; ensure that evaluation sampling errors properly account for effects of location-specific micro-climates; minimize bias by ensuring adequate sample sizes, and use of appropriate evaluation methods to account for observable and unobservable differences between participants and non-participants; capture spill-over effects that blur differences between treatment and control groups, through explicit statistical modelling;
  - (h) factor-in scale and complexity; and
  - (i) where the programme logic supports a multi-faceted package of interventions, consider limiting, if not the size of the intervention, the scope of the evaluation;
  - (j) capture also long-term impacts, in addition to project completion impacts, through modelling.
19. As called for in the IFAD9R and as recommended by the above referenced experience, IFAD will apply and learn from a **range of methodologies** for the 30 impact evaluations. This will permit a closer match with specificities of the projects to be evaluated, the state of their RIMS compliance, and the specific learning agenda, while guaranteeing the integrity and coherence of the overall impact measurement and learning initiative.
20. In order to **ensure commitment and useful impact evaluation outcomes**, in the planning of the 30 evaluations IFAD will align incentives of all stakeholders by: engaging early and communicating often; planning together with all partners; supporting implementers in finding value in impact evaluations; being explicit about costs and required commitment and efforts; creating clarity and incentives through contracts.

## **E. Methodologies for IFAD9 impact evaluations**

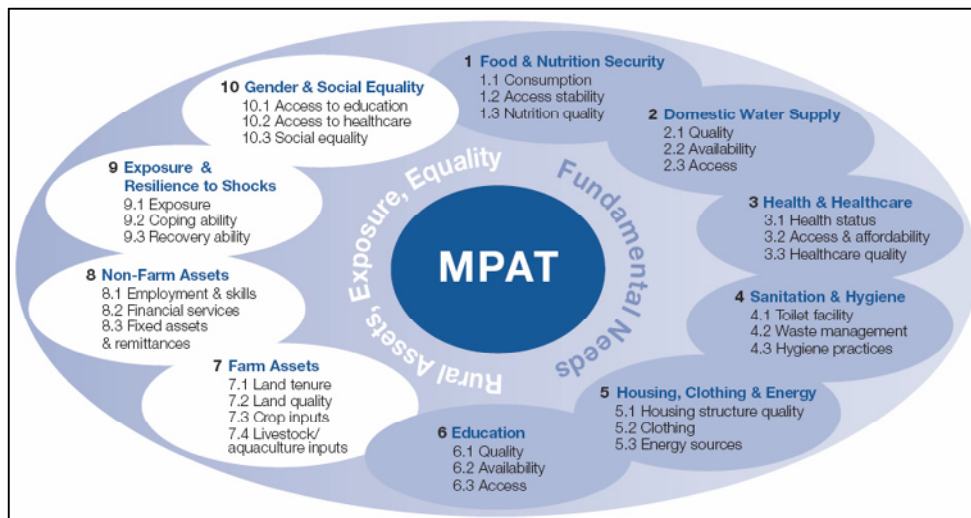
21. With these lessons for successful approaches to impact evaluation in mind, and before presenting key methodology features for IFAD's 30 impact evaluations, it is useful to reiterate the purpose of this IE initiative. IFAD contributes significant resources to agricultural and rural development for rural poverty reduction and food security. It needs to demonstrate its results, quantify impacts that are attributable to its country programme activities, contribute to learning about which approaches work best to increase incomes and reduce poverty, and be accountable for a cost-effective use of scarce development resources. A better understanding of impact pathways will underpin replication and mission critical scaling-up pathways, as well as policy dialogue. The learning from the application of different methodologies (including their rigour and cost-effectiveness trade-offs) and from identifying different causality chains will enhance IFAD's and its partners' development effectiveness and impact evaluation capacities. The cornerstone for this effort is rigorous impact evaluations that measure the change in a development outcome that is attributable to a defined intervention, through the use of a counterfactual, in order to distinguish between change that happens with a project and that which happens without the project.
22. While this is normally done on a project by project basis, with individual projects as the unit of account, IFAD also wants to aggregate this information for the entire 2010-2015 portfolio in order to compare and in order to demonstrate aggregate development effectiveness, while differentiating for country and project specificities.

### **E.1 Which results and impacts will be measured, by whom, and how?**

23. The IFAD9R (and para 2 to 4 above) spells out the results and impacts to be reported on. Between 2010 and 2015 IFAD will ensure that, cumulatively:
- (a) 90 million people received services from IFAD-supported projects; and
  - (b) 80 million people were moved out of poverty.
24. **90 million people reached.** There is consensus that current work on improving M&E systems design and implementation, compliance with annual RIMS reporting requirements and quality assurance for RIMS reporting and data entry, will allow the Project M&E Units to monitor and report accurately on numbers of people reached, as per RIMS+ policy. These figures will be aggregated for the entire IFAD portfolio.
25. The counting of people receiving services from IFAD financed programmes will include, in a differentiated and gender disaggregated manner, **direct beneficiaries as well as indirect beneficiaries** (only if credibly identified through reliable surveys). It may be overstressing the capacity of project management units that do not register individual beneficiaries to filter out rigorously multiple project-service benefits, which may lead to an overestimation of people reached. Documented correction estimates will seek to reduce such risk for double-counting. The benefits would relate to the **full programme**, not just the IFAD financing share. The count of benefits would relate to the **entire portfolio ongoing** between 2010 and 2015, on a cumulative basis, and not just the projects closed between 2010 and 2015.
26. **80 million moved people out of poverty.** Poverty is complex in its statistics and its dynamics, as illustrated in the graphic below drawn from the IFAD sponsored work for the Multidimensional Poverty Assessment Tool<sup>9</sup>. IFAD9R looks at poverty as a level of income below "a defined poverty line", and identifies complementary indicators of household asset ownership, length of the hungry season and child malnutrition – which may also serve as proxies for poverty and for the purpose of statistical inference and modelling. It should be highlighted that measuring and comparing income-poverty is notoriously complex, but in view of the need for comparison and aggregation, an income-based definition of poverty is useful.

---

<sup>9</sup> The Multidimensional Poverty Assessment (MPA) project was a collaborative international initiative led by IFAD to develop, test and pilot a new tool for local level rural poverty assessment. MPAT is a multi-purpose tool that can be used to assess and support rural poverty alleviation efforts in developing countries. MPAT takes a step back from assessment methods that are focused on economic- and consumption-oriented indicators and strives to provide an overview of fundamental and relatively universal dimensions of rural livelihoods and rural life, and thus of rural poverty. MPAT is a survey-based (household and village level) thematic indicator primarily designed to support monitoring and evaluation, targeting, and prioritization efforts at the local level. However, MPAT also has many other uses, such as: making in-country and cross-country comparisons; supporting project design; facilitating policy dialogue and national programme support; raising awareness among a variety of stakeholders; empowering beneficiaries; and providing for innumerable secondary data analysis with the survey datasets. MPAT allows project managers, government officials and others to determine which dimensions of rural livelihoods likely require support and, more generally, whether an enabling environment is in place to allow rural residents to pursue their livelihood goals.



**Figure 1** Organizational diagram of MPAT's components and subcomponents.  
Source: MPAT book.

27. Contrary to the number of people reached, Project Management Units (PMU) would not generally be in a position to rigorously measure or estimate how many people have moved out of poverty. First, because not all (past) project designs specify the indicator and level of poverty to be targeted in an equally rigorous and measurable manner. Second, because PMU are not systematically in a position (timing and capacity-wise) to measure the impact of projects on poverty levels. This requires rigorous and survey-based assessments; which are better outsourced to institutions that have recognised impact evaluation expertise.
28. In view of the above and taking into account the fact that IFAD supports poverty reduction in different types of countries (LIC, MIC, fragile states, ...) at different stages in their agricultural and rural transformation, a range of poverty indicators will be combined for (income-poverty) impact evaluation purposes. In consistency and compliance with RIMS+ policy, poverty reduction will be measured as follows:
- For projects that have well defined income-poverty lines or other measurable poverty lines, impact will be measured in terms of the **approved income-poverty line** or the income-level equivalent of the other types of approved indicators, estimated through statistical inference.
  - For projects without such poverty lines, impact evaluations will use either the national poverty line or the **USD 1.25 a day, as well as the USD 2.5 a day international poverty lines**, as appropriate in view of the country or project specificities. If not directly observable, RIMS-based proxy indicators will be used to infer such income-poverty impacts. Reporting will differentiate these different levels of poverty reduction within the aggregation, ensuring also purchasing power comparability.
  - As required in IFAD9R, the indicators of household asset ownership, length of the hungry season and child malnutrition will also be measured in their own right and may complement the income-poverty data. However, there will be instances where these indicators will be measured and used to complement or even statistically infer income-poverty and income-poverty reduction, through modelling. Selective attempts will be made, where possible, to use also **consumption data** to infer poverty and poverty reduction.
  - If credibly identified through reliable surveys, indirect benefits and spill-over or demonstration effects will also be integrated in the aggregation of impacts, in a discernible manner.

- (e) The need to measure impacts and results in gender equality and women's empowerment will be addressed through specific evaluation questions, as well as, basically, through gender differentiated survey data collection and analysis. This may involve the application of the **Women Empowerment in Agriculture Index (WEAI)**<sup>10</sup>, already tested in IFAD.
29. IFAD9R is not explicit regarding the meaning of "moving out of poverty". For the purpose of this initiative, "moving out of poverty" will mean **any level** above the set poverty-line<sup>11</sup>, **at the time of the survey**, irrespective of durability or risk resilience<sup>12</sup> of the impact, for which evaluations would need to be ex-post, rather than at completion. The number of people surveyed as moved out of poverty will be aggregated and reported on for IFAD's (entire 2010 - 2015) portfolio, ensuring purchasing-power comparability and differentiating the different levels of poverty reduction according to poverty lines adopted. The different IE reports, as well as the synthesis report, will be explicit about margins of error which are bound to be significant given the nature of projects and/or the state of data availability. As per expert advice<sup>13</sup>, aggregation across project-level evaluations will also use the **Poverty Reduction Rate of Return (PRRR)** analysis, measuring increases in incomes of the specific number of poor people benefiting from the project(s).
30. Ideally, impact surveys would be carried out systematically for all of IFAD's 2010-2015 portfolio. This would not be useful, as some of the projects would be at very early stages of project implementation; these would be able to show numbers of people receiving services, but poverty reduction impacts would not be observable yet. It would also be extremely costly to carry out a full portfolio impact assessment. Therefore the 2015 reporting on 80 million people moved out of poverty will need to be based on the findings of the sample of about 30 impact evaluations, actually planned for the specific purpose of learning about impact pathways. The findings of these 30 rigorous impact evaluations would then be extrapolated to the entire portfolio, and this requires a number of rigour conditions to be met in these 30 impact evaluations, and especially in terms of the statistical representativeness of the sample of projects chosen.

## **E.2. 30 rigorous impact evaluations for learning and extrapolation purposes**

31. In order to ensure the dual purpose of the 30 impact evaluations planned (systematic learning about impact pathways and rigorous extrapolation of impacts to the entire portfolio) it is important to define: what we seek to learn; which set of projects and components (through factor analysis) combine optimally both learning opportunities and statistical representativeness; which methods to use for the analysis of survey data; how to assure quality of the impact evaluations; and how to report.

<sup>10</sup> The WEAI is constructed using the Alkire Foster Method developed by Sabina Alkire, director of the Oxford Poverty and Human Development Initiative (OPHI) at the University of Oxford, and James Foster of George Washington University and OPHI. A method for measuring multidimensional poverty, well-being, and inequality, it measures outcomes at the individual level (person or household) against multiple criteria (domains and/or dimensions and indicators). The WEAI monitors 5 domains: production, resources, income, leadership and time. The method is flexible and can be applied to measure poverty or well-being, to target services or conditional cash transfers, and to design and sequence interventions. Different domains (for example, education) and indicators (for example, how many years of education a person has) can be chosen depending on the context and purpose of the exercise.

[http://www.ifpri.org/sites/default/files/publications/weai\\_brochure.pdf](http://www.ifpri.org/sites/default/files/publications/weai_brochure.pdf)

<sup>11</sup> A recent CGAP and Ford Foundation sponsored summary of RCTs for asset-transfer based programmes for graduation out of extreme-poverty indicates impacts in the order of 10 - 15% increases in monthly incomes (to be confirmed in finalising the preliminary results), and this may become a benchmark as IFAD reports on poverty impacts, in 2015, together with other benchmarks

<sup>12</sup> However, in the different IE reports as well as the synthesis paper, use will be made of the work done by Tim Frankenberger (TANGO) on resilience analysis and resilience programming.

<sup>13</sup> Maximo Torero, IFPRI

***What do we seek to learn? Which impact pathways do we want to understand?***

32. The most important first step – both for successful implementation and evaluation – is to have a clear picture of what a programme aims to achieve, and how planned interventions are expected to lead to that outcome. This programme logic, or theory of change<sup>14</sup>, is the starting point for designing an implementation approach as well as an evaluation. It is critical to ensure that the development hypotheses evaluations are testing are well articulated and the potential causal pathway is well-defined.
33. To support this requirement IFAD has worked with a group of institutions<sup>15</sup> on a shared **Food Security Learning Agenda (FSLA)**, to be completed by end 2012<sup>16</sup>.
34. The dimensions of this living learning agenda, which are aligned with IFAD's core thematic focus areas, consist of:
  - (a) Improved Livelihoods, with a focus on Rural Productivity;
  - (b) Enhanced Nutrition and Dietary Quality;
  - (c) Increased Resilience of Vulnerable Populations;
  - (d) Improved Research, Innovation, and Commercialization for Agriculture and Nutrition;
  - (e) Enhanced Management of Natural Resources and Adaptation to Climate Change;
  - (f) Expanded Markets and Value Chains;
  - (g) Improved Policies and Institutions for Food Security and an Enabling Environment;
  - (h) Improved Gender Equality and Women's Empowerment.
35. For each dimension the FSLA states the development issues, poses the questions that need answers and suggests a number of indicators. As such this learning agenda will guide the design of individual project evaluations, without straightjacketing them and will ensure that all specific RIMS+ survey based evaluation findings can eventually be aggregated and synthesised in a coherent body of knowledge on agricultural and rural development pathways out of poverty. The FSLA supports and integrates the RIMS based processes and their outcomes; it does not substitute to it, nor to the pre-existing project specific theories of change.

***Which projects to evaluate?***

36. The IFAD9R specifies that impact surveys will be carried-out **for completed projects, with baseline surveys** undertaken at the time of project start-up.
37. The universe of projects to select from counts 355 ongoing projects (closing between 2010-2015), plus about 100 projects that may be approved between now and end-2015; i.e. a total of about 455 projects. Of these projects 206, or 45%, have been closed since 2010 or are currently set to close before end-2015. A sample of 30 projects to be evaluated, would represent 7% of the total portfolio, and a significant 14.5% of the closed/closing projects.
38. While the overlap between closed/closing projects and projects with baseline surveys and/or completion surveys needs to be ascertained<sup>17</sup>, 103 projects have baseline surveys (56 of which are considered usable; and another 38 may be usable with some review and data complements); 17 projects have completion surveys; and a

<sup>14</sup> A coherent articulation of how and why a given intervention will lead to specified change

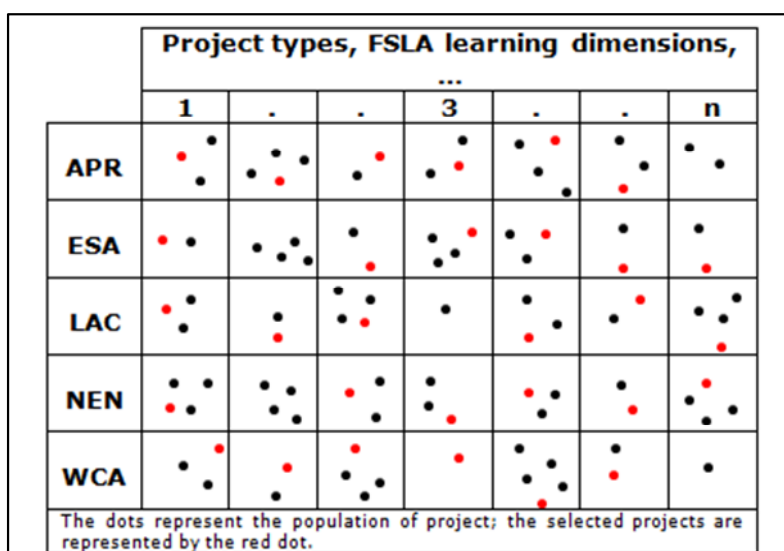
<sup>15</sup> Initially USAID (with other USG institutions and programmes such as MCC and US Bureau of Statistics), BMGF, WB, DFID, FAO, WFP and expanded with IFPRI, JPAL, OECD, DANIDA, and others.

<sup>16</sup> available in draft on request

<sup>17</sup> Since RIMS became effective only in 2004, and with a typical project duration period of 6 to 7 years, the overlap between (2010-2015) closed/closing projects and projects with RIMS surveys is estimated to be high

limited number of recently closed projects actually completed impact evaluations. A sample of 30 projects to be evaluated, would represent 30% of the portfolio that have usable baseline surveys.

39. A formal selection process will establish the list of the 30 projects to be evaluated, from the above 2 lists and especially from their overlap. There are of course a number of "low-hanging fruits", in terms of impact evaluations already planned or being considered, also in some cases in terms of meeting the funding requirements. These include, as a limited list of examples:
- projects with recently completed project completion impact evaluations, subject to the robustness of the methodology applied;
  - 3 experimental designs planned by ECD under the ASAP programme;
  - 6 impact evaluations, considered for execution by 3IE under its agricultural window, with supplementary funding<sup>18</sup>; and potentially also some under 3IE's climate change window;
  - 2 impact evaluations, with IMI funding, possibly in Vietnam and Madagascar;
  - 3 project evaluations to be conducted by IOE between 2013 and 2015, starting with the Sri Lanka Dry Zone Livelihood Support and Partnership Programme (DZLSPP);
  - an update of the econometric study of impacts of IFAD supported programmes in China;
  - the RUMEP programme in the Philippines, where the authorities look for a rigorous project evaluation, in view of scaling-up.
40. From the perspective of learning about impact pathways within the FSLA framework, a purposive selection process of this type may well ensure the evaluation of the strategically most pertinent projects. However, with a view to extrapolation of impacts from these 30 impact evaluations to IFAD's entire portfolio, there is a need for an approach that assures statistical representativeness of the selected sample, and this requires a degree of randomisation of selection.



**Figure 2** The grid. Source: IFAD.

<sup>18</sup> Potentially from DFID and BMGF and currently under early exploration

41. Expert advice<sup>19</sup> recommends IFAD to develop a grid (see Figure 2 above), with (e.g.) regions (5) on the horizontal axis and (e.g.) project types or FSLA learning dimensions (6) on the vertical axis. The eligible group of closed/closing projects with baseline surveys, as well as completion surveys or completion evaluations<sup>20</sup> would be entered in one of the cells taking into account that:
- (a) categories should be both exhaustive and mutually exclusive, so that every project fits into one box in the grid and no projects are left without boxes;
  - (b) there should be a sense that the expected impacts within a cell are homogenous in some way, so that a subset of projects within that box could be taken as representative of the overall set of projects in the box; and
  - (c) one needs to be able to count the total number of beneficiaries for every project in each box, although not every project will be evaluated.
42. With this in place, the further process runs as follows:
- (a) select the subset of projects within each box intended to be evaluated using ex-post methods. The most rigorous way would be to select them randomly, but this may not prove practicable once data requirements, existence of baselines, presence of survey firms on the ground and so on are taken into account<sup>21</sup>. The most practicable way to move forward is therefore to state ex-ante a set of data requirements that selected projects must satisfy, and then from within the subset of projects that fulfil these data requirements the group that is most representative of the overall box would be picked;
  - (b) every box must have at least one evaluable project in it, which argues against an extended grid that would have too many boxes with few observations in each – and 30 cells may be too many;
  - (c) the evaluation will then be conducted and will return an Intention to Treat Effect (ITT) that gives the change in each indicator for every intended beneficiary;
  - (d) with an ITT for each box and each indicator, and knowing how many intended beneficiaries there are in each box, the grid can now be used to calculate institution-wide impacts by multiplying the ITT times the number of intended beneficiaries per box, and summing across boxes;
  - (e) consideration may also be given to add an element of probability proportional to size (PPS) in the sampling<sup>22</sup>.
43. This approach is critical for the statistical validity of the extrapolation and will be fine-tuned in the near future, with continued expert advice, in order to ensure the quality of the sampling framework as well as the rigour of the evaluation and the extrapolation.

***Which methods will be applied?***

44. The wide range of methods available for (rigorous) impact evaluations<sup>23</sup> can be categorised as follows:

<sup>19</sup> Howard White, 3IE; Craig McIntosh, ATAI; Raghav Gaiha, MIT; Maximo Torero, IFPRI

<sup>20</sup> Non IFAD initiated programmes need not to comply with RIMS requirements, as per RIMS policy. For those programmes equivalent baselines or other surveys are also considered acceptable for this initiative. This would also apply to IFAD initiated projects that have equivalent surveys.

<sup>21</sup> There is a hypothesis that the IFAD9R decision of selecting projects with baseline surveys biases the sample towards better managed projects. This hypothesis will be tested and addressed in the reports.

<sup>22</sup> In this context, and without affecting the integrity of the grid based sampling, efforts will be made to include IFAD's participation in Sector Wide Approach (SWAP) programmes or national programmes initiated by other IFI's and cofinanced by IFAD.

<sup>23</sup> Summarised most notably in: NONIE Guidance on Impact Evaluation (NONIE, 2009), Broadening the Range of Designs and Methods for Impact Evaluations (DFID, 2012), Handbook on Impact Evaluations (WB, 2010).



- (a) **Quantitative impact evaluations:** an analysis based on a representative survey of treatment group and a comparison group, before and after the intervention. There is a range of techniques under this heading - *experimental design techniques* or Randomised Controlled Trials (RCT); *quasi-experimental techniques*, including double difference, single difference, instrumental variables, judgmental matching, pipeline approach, propensity score matching, reflexive comparisons and simulated counterfactual methods; and *regression based techniques* – regression discontinuity;
- (b) **Qualitative and participatory impact evaluations:** an analysis based on participatory methods amongst beneficiaries; and
- (c) **Theory-based approaches:** an analysis tracing the underlying theory of change, the logframe from inputs to outcomes, using a mix of methods to establish causal linkages.
45. Although these are different approaches, good impact evaluations will typically use a **mixed-method approach**. Qualitative analysis, even if not a full-blown participatory analysis, helps provide valuable context. A theory-based approach helps build the story around the intervention and understand why it worked or not. However quantitative methods generally give a more authoritative indication of the counterfactual and impact on outcomes. IFAD's 30 impact evaluations will seek to test the different approaches, in various combinations, depending on project specificities.
46. However, as there is a need for aggregation of impacts for the entire 2010-2015 portfolio, IFAD's 30 impact evaluations will privilege quantitative approaches, while qualitative and theory-based methodology elements will be crucial for the (FSLA guided) learning agenda, on impact and scaling-up pathways.
47. The cornerstone of quantitative impact evaluation is data collection from a statistically representative sample, using a structured questionnaire such as the RIMS+ survey questionnaire. Data are usually collected before the intervention (baseline) and after (completion survey); however, conducting a midterm survey offers an advantage. Data should be collected from both the affected population (the treatment group) and a comparison group<sup>24</sup>. Project impact is then calculated as either a single difference (difference in outcome between project and control after the intervention), or double difference (the difference in the change in outcomes for the project and control before and after the intervention).
48. The main quantitative approaches are:
- (a) **Experimental (randomized) approaches:** The most rigorous method for measuring attributable project impacts, "the gold standard"<sup>25</sup>, is through randomized control trials (RCTs). Experimental (or randomized) evaluation design requires that the eligible population be identified and then a random sample of those treated. For example, only 200 farms are chosen at random to be included in the project out of 1,200 eligible farms in 10 project districts. Because RCTs identify groups of individuals (control groups) that will not receive project interventions, evaluators can compare what happened to beneficiaries with the project versus what happened absent the project, potentially over a long period of time. The untreated (or a random sample of

<sup>24</sup> Selection of an appropriate comparison group is one of the main challenges in impact evaluation. They should be identical to the treatment group except that the latter receive the intervention and the former do not. In practice this is difficult to achieve for two reasons. First, beneficiaries of the intervention may be selected (or self-select) on the basis of certain characteristics. If these characteristics are observed then a comparison group with the same characteristics can be selected. But if they are unobserved then in principle only a randomized approach can eliminate selection bias. Second, the comparison group may be contaminated either by spill-over effects from the intervention or a similar intervention being undertaken in the comparison area by another agency.

<sup>25</sup> It should be considered that any of the impact evaluation methods, when properly executed and used as a management tool or to improve programme design and enhance its effectiveness, become in a sense experimental design methods, and RCTs (only) constitute the gold standard.

the untreated) are a valid comparison group since there should be no systematic difference between their characteristics and those of the treatment group. Experimental methods have the strong appeal of avoiding an otherwise unknown bias from selectivity, but are in practice only applicable to a narrow range of the interventions supported by development agencies. This use of a counterfactual creates the greatest opportunity for learning what works, and for measuring project impacts that accrue over time. However, it can be challenging to apply RCT methodology when project targets are so large that it is hard to find a comparable control group, or when demand for project interventions does not exceed the planned scope of the project. In addition, RCTs require that implementers and evaluators have significant amounts of information about all potential beneficiaries (both treatment and control groups) from the outset, which can be challenging when project design is fast. When strict RCTs are not feasible for socio-political or other reasons, a portfolio or randomised roll-out approach may be applied.

- (b) **Pipeline:** The pipeline approach takes as the comparison group individuals, households or communities which have been selected to participate in the project, but not yet done so. Clearly the approach can only be used for activities which continue beyond the end of the project being evaluated.
  - (c) **Propensity score matching:** Selection may be based on a set of characteristics rather than just one. Hence the comparison group needs to be matched on all these characteristics. PSM uses statistical modelling to identify a group of individuals, households or firms with the same observable characteristics as those participating in the project. The potential problem with PSM is that facing all quasi-experimental approaches: selection on unobservables. Unobservables, which simply affect project outcomes and are constant over time, can be swept out by taking double difference estimates. But if they are time variant, or correlated with both selection and outcomes, then biased estimates will result.
  - (d) **Regression-based approaches**<sup>26</sup>: The regression based approach models the determinants of outcomes and possibly also models the determinants themselves. The approach has the advantage of flexibility – it does not lump different activities under the single heading of ‘the intervention’ – and automatically incorporates differing intensities of participation. It is only when the treatment is a simple, homogenous activity that dummy and mean comparison approaches are appropriate. However, the adoption of the regression-based approach does not mean that problems of selection bias are removed. Hence the need to address them. Where selection is based on observables then this is readily done.
49. The fact that the 30 project evaluations for IFAD9 are meant to report on impacts of closed projects, limits the scope for RCTs in this IE initiative to fast gestation projects, at very early stages of implementation. Therefore a broader range of methods will be applied. In addition, the demanding conditions for proper RCTs (a control group, appropriate sample size for the treatment and control groups, and data panelling) are not systematically in place. Therefore, when applying other methodologies, the IE work will include improvement or construction/reconstruction of a control frame (using other data sets such as a census, LSMS, or other survey based assessments<sup>27</sup> and/or through modelling and simulation). Shoestring evaluation techniques will also be applied, where necessary. Factor analysis will permit to isolate the impact of different types of components/activities.

<sup>26</sup> Regression discontinuity design evaluations, regression modelling

<sup>27</sup> Such as, for example, the data base for the Latin American Center for Rural Development (RIMISP) sponsored “Poverty and Inequality – Latin American Report 2011”

50. In synthesising the 30 impact evaluations, the findings from the rigorous impact assessments will be complemented with and framed in the context of **meta-analyses** or possibly even the more demanding **systematic reviews** that would provide the story called for in the FSLA.

***How do we assure quality?***

51. IFAD needs to ensure highest quality of its impact evaluations, a level of quality that would meet standards such as for the Cochrane Collaboration or, for agriculture, the Campbell Collaboration. To this effect IFAD will continue to work with centres of excellence in rigorous impact evaluation and will engage a continuous flow of services from 3IE (or, in the case of a conflict of interest, another expert institution<sup>28</sup>) to provide technical support, review and quality assurance of evaluation designs, surveys and analyses. This quality assurance service would be funded through a small grant, not exceeding USD 500,000, covering the IE quality assurance needs for 3 years (2013, 2014 and 2015). In addition, IFAD will seek to mobilise peer review processes for QA purposes.

**E.3. Reporting**

52. The reporting on this initiative will be at three levels: a) in the context of the annual RIDE; b) possibly a short progress report on the entire process (not yet the findings) in the context of an IFAD9 mid-term review in 2014, if any; and finally, c) the overall synthesis report, by end-2015.
53. The 2015 substantive synthesis report will:
- (a) summarise and aggregate the results and impacts identified in the 30 impact evaluations;
  - (b) synthesise the lessons learnt on impact pathways, within the framework of the FSLA;
  - (c) synthesise recommendations for policy engagement that may emerge from the 30 IEs and their synthesis<sup>29</sup>;
  - (d) make recommendations for further development of RIMS; and
  - (e) advise on impact evaluation methodologies, their feasibility, rigour and cost-effectiveness.

**F. Partnerships and anchoring M&E and IE in national systems**

54. The IFAD9R calls for: an active pursuit of partnerships with institutions specialised in impact evaluation; as well as the strengthening of national M&E systems by enhancing the capacity of project management staff and implementing partners, particularly at start-up and early project implementation through the systematic engagement of M&E experts during design and supervision missions.
55. **Partnerships.** Progress has been made in partnership development, especially with centres of excellence which provide state-of-the-art knowledge on M&E and Impact Evaluation (IE). As suggested in IFAD9R, IFAD will continue to work in partnership with such institutions and initiatives. IFAD will also seek to broaden the rigorous IE expertise base, beyond the limited number of premier centres.
56. In addition, on 27 March 2012 a group of institutions (initially USAID, USG/MCC, BMGF, WB, DFID, FAO, WFP and expanded with IFPRI, JPAL, OECD, DANIDA, and others) initiated a partnership to learn from each other and improve their performance in programme monitoring and impact evaluation, in evidence-based knowledge management and in scaling-up tested innovations to improving food security and poverty reduction on the ground. This partnership aims for joint learning, sharing scarce resources and competencies, and promoting best

<sup>28</sup> Including the FSLA peer review.

<sup>29</sup> While this IFAD9 IE initiative of 30 impact evaluations will not extend directly to policy evaluations *per se* (evaluations of policies), the synthesis will report on the implications of the findings for policy (evaluations for policy).

approaches to food security . The partnership is informal and open, for like-minded institutions to join<sup>30</sup>; and it functions as an informal operations-focused network. The partners are now perusing a shared Food Security Learning Agenda (FSLA), common indicators, a set of principles and a peer review process.

57. **Country capacity building.** A number of recent COSOPs (e.g. Nicaragua, Niger) or project designs (e.g. Peru, Brazil) explicitly plan for building national capacities in M&E and IE, and for anchoring M&E and IE work in national systems or institutions. This will constitute the main approach for meeting this IFAD9R requirement. The QA will ensure the mainstreaming of this best practice. In turn, support to national statistical and M&E systems will also contribute to enhance M&E and IE capacity building at the project/programme level.
58. A complementary approach to contribute to this objective, is strengthening national statistics and M&E systems through supporting international initiatives. IFAD is engaged in 3 partnership initiatives that will also contribute to this objective: the Global Strategy to Improve Agriculture and Rural Statistics (GSIARS); the World Agriculture Watch (WAW); and the Agricultural Transformation Index (ATI):
- (a) **The GSIARS.** As an important part of the M&E and IE improvement programme consists of strengthening the national capacities. IFAD will look for opportunities to link its M&E and impact related work to national institutional capacities (for instance the Bureau of Statistics, the Planning Authority, etc.). This requires a prior availability of the required capabilities within these institutions. For this purpose, IFAD is actively engaged in the FAO's GSIARS with a view to dovetail IFAD's country programmes with the national capacity building programmes.
- (b) **The WAW.** The aim of WAW is to conduct comparative analyses on various scales, by organizing a network of national observation centres covering the different types of agriculture, their dynamics and their impact in terms of sustainable development. This will involve characterizing production structures and analysing their impacts using environmental, social and economic indicators. Three levels will be taken into account: agricultural production units, territories and markets. This systemic approach and the involvement of organized players in its governance are what make the initiative original. The initiative was taken by the Government of France and CIRAD, and is managed and housed in FAO. IFAD is actively engaged in the WAW initiative with a view to dovetail IFAD's country programmes with the work of the national observatories. The observatories will provide data required to construct counterfactuals.
- (c) **The ATI.** In light of the rapid transition of the global agricultural and food system it is felt necessary to develop a set of indicators that combine into a composite index that reflects the transformation in agriculture. The Government of Denmark with the USA Government (USG), in the context of the G8, have taken the initiative to bring together a wide range of stakeholders to develop and implement the ATI within the next 2-3 years. IFAD is actively engaged in the ATI initiative with a view to dovetail IFAD's country programmes with the country level work on ATI.

## G. Financing M&E and Impact Evaluations

59. An important explanation for the inadequacy in the coverage and numbers of IE consists of their high costs and the efforts of institutions to enhance cost efficiency gains in their budgets. Costs for rigorous impact evaluations are reported to typically range from USD 500,000 to 1 million. Inadequate attention is paid to the benefits of

---

<sup>30</sup> While the core of the partnership is among the participating institutions, the partners will also reach out to other organizations, centres of excellence, think tanks, observatories and programmes concerned with impact, economic and financial analysis, statistics, household and other surveys.

IE work versus its costs. It is therefore necessary to develop a strategy to finance IE activities and their improvement. An important consideration for this strategy consists of the fact that many institutions, especially the bilateral institutions and the NGOs/foundations, integrate the cost of IEs and related staff costs into the actual programme funding versus their administrative budgets. The IFAD9R calls for an effective resource mobilisation strategy to develop adequate internal human capacity and financial resources to conduct and manage this impact evaluation work.

60. An approach reportedly adopted by the Inter-American Development Bank (IaDB) sheds an interesting light on this matter. Because IE and related knowledge is considered to be a global public good, IaDB introduced a systematic holdback from the loan and grant proceeds (5%) which are pooled in an IE "account" managed by the institution. This mechanism allows for systematic IE across the portfolio and insures that each IE cumulatively contributes to the general understanding of impact pathways and scaling up pathways, as a public good.
61. The approach to finance the 30 impact evaluations for IFAD9 consists of several elements, including:
- (a) **Building IEs in project costs.** The first option for financing IE activities consists of explicitly building these costs into the project costs, like any other M&E costs. For grant funded programmes this is easily applicable, whereas in loan funded programmes governments would need to be convinced of the merit of financing IE from the loan proceeds, as a global public good. The mechanism to secure access to these resources could consist of blank withdrawal applications.
  - (b) **Providing IE as Reimbursable Technical Assistance (RTA).** IFAD is proceeding on a new approach for the financing of services provided by IFAD to governments during country programme design as well as implementation. In addition to technical advisory services (e.g. private sector involvement with smallholders); policy advice in the agricultural and rural development sector (e.g. development of 'climate smart' approaches for smallholders); donor/private/public sector coordination in agricultural sector; research services and delivery of training; IFAD is also recommending the financing of impact evaluation and results management in the smallholder sector through this RTA mechanism: a new and, as such, still untested instrument.
  - (c) **Supplementary funding from specifically interested donors.** Individual donors have shown interest in financing a selective range of IEs (for instance, in a specific country or on a specific theme or using a specific methodology). This would include the possible selective participation in Requests for Proposals (RFP) by interested donors<sup>31</sup>. While these individual initiatives help financing the required activities there is a risk of fragmentation and duplication. This approach also limits the scope for synergies, cost effectiveness gains and knowledge integration. The approach that would address those issues consists of the development of a medium-term IE programme, the establishment of a multi-donor trust fund and the channelling of these resources through the trust fund in the process of being established by PMD.
  - (d) **Partnership – cost sharing.** One of the advantages of long term strategic partnerships with other stakeholders in the field of M&E and IE, consists of potentially sharing resources and carrying out joint IE activities. It is for instance envisaged that the M&E partnership (Para 56) will lead to enhancing the resource envelope for IEs within the partnership as well as a better sharing of those resources.

---

<sup>31</sup> E.g. RFP from BMGF, 3iE, ATAI, etc.

## H. Conclusion

62. In conclusion:

- (a) By end-2015 IFAD will conduct, synthesise and report on approximately 30 project impact surveys, of which three to six will use randomised controlled trials (RCT) or other similarly rigorous methodology.
- (b) Four impact-level indicators have been identified: household asset ownership index, length of the hungry season, child malnutrition and number of people moved out of poverty. The targets for the first three indicators are to be "tracked" and in relation to the fourth indicator of poverty the targets are: 90 million people received services from IFAD-supported projects, cumulatively from 2010 onwards to 2015; and 80 million of these people moved out of poverty.
- (c) A range of coordinated initiatives is underway to improve IFAD's performance in M&E and IE. With the current proactive thinking, planning and action, IFAD is well placed to improve its M&E and IE performance and is therefore in a reasonable position to meet the IFAD9 impact evaluation related commitments.
- (d) A number of institutions have summarised IE experiences in authoritative publications. These findings, especially those related to agricultural and rural development for poverty reduction and food security, guide the impact evaluation work: its approach and its methodologies.
- (e) IFAD will use a mix of, mainly quantitative, methodologies for measuring and reporting on results and impacts. Standing project monitoring arrangements will be used to report on numbers of people that received services from IFAD supported programmes. Expert institutions will be invited to carry-out 30 rigorous impact evaluations in order to identify the number of people moved out of income-poverty, define the respective impact pathways for learning purposes, and extrapolate poverty reduction impacts from a statistically sound project sample framework.
- (f) The report, by end-2015, will synthesise the 30 impact evaluations: aggregate the measured results and impacts, summarise the lessons learnt on impact pathways; and advise on rigour and cost-effectiveness of different impact evaluation methodologies.
- (g) IFAD will conduct the 30 impact evaluations in close collaboration with expert institutions and other development partners, and will seek to work within national M&E systems and strengthen them in the process.
- (h) To finance these impact evaluations IFAD will combine country programme resources, grant resources and supplementary financing from development partners.