GUIDELINES

ASSESSMENT OF RDP RESULTS: HOW TO PREPARE FOR REPORTING ON EVALUATION IN 2017

FINAL DRAFT – November 2015

The Evaluation Helpdesk is responsible for the evaluation function within the European Network for Rural Development (ENRD) by providing guidance on the evaluation of RDPs and policies falling under the remit and guidance of DG AGRI's Unit E.4 'Evaluation and studies' of the European Commission (EC). In order to improve the evaluation of EU rural development policy the Evaluation Helpdesk supports all evaluation stakeholders, in particular DG AGRI, national authorities, RDP managing authorities and evaluators, through the development and dissemination of appropriate methodologies and tools; the collection and exchange of good practices; capacity building, and communicating with network members on evaluation related topics.

Additional information about the activities of European Evaluation Helpdesk for Rural Development is available on the Internet through the Europa server (http://enrd.ec.europa.eu).

GUIDELINES

ASSESSMENT OF RDP RESULTS: HOW TO PREPARE FOR REPORTING ON EVALUATION IN 2017

FINAL DRAFT – November 2015

CONTENT

TABLES, FIGURES AND TOOLS

FOREWORD AND ACKNOWLEDGMENTS

(to be developed).

INTRODUCTION

Starting in June 2016, and each year until 2024, Member States shall submit to the Commission an annual implementation report (AIR). The AIR provides the information about the implementation of rural development programme, as well as the evaluation plan.

The AIR submitted in 2017 shall also include the quantification of programme achievements, in particular through the assessment of the result indicators (including complementary result indicators), and answers to relevant evaluation questions.

The Technical Handbook of the CMEF of the CAP 2014 – 2020 and Annexes provides general guidance as well as detailed fiches for each result indicator to be reported in the AIR submitted in 2017 and 2019. Each fiche contains the link to the RD priority and the focus area, the definition of indicator, unit of measurement, the methodology for calculation, data requirements and sources, including those collected via operations database, point and frequency of data collection, and the means of data transmission to the Commission. The definition and detailed description of common indicators are provided in the legal framework and the above-mentioned fiches.

However, in order to support the managing authorities and evaluators in using them, and to ensure consistency across all Member States and RDPs, it was agreed that additional guidance would be needed. This should address questions which are frequently asked by evaluation stakeholders in the Member States, such as:

- How to use complementary result indicators in the assessment of RD interventions in the AIR submitted in 2017? How to attribute results to the RDP interventions?

- How to address the assessment of additional contributions of RD measures (secondary effects) to focus areas under which they have not been programmed but still might influence them?

- How to address challenges caused by the flexibility in programming of EAFRD interventions in individual RDPs?

- How to deal with gaps in the monitoring and evaluation system linked to each individual RDP caused by a flexible programming approach?

- How to develop programme specific result indicators and related evaluation questions?

- How to report on the assessment of RDP achievements in 2017?

- How to communicate and disseminate evaluation results and the implementation of the evaluation plan?

With a view to comply with legal requirements and enable Member States to establish a robust monitoring and evaluation system, capable of generating the required information to be reported in the AIR submitted in 2017, the above-mentioned questions should be answered at an early stage of programme implementation.

Against this background the Annual Work Programme 2015 of the Evaluation Helpdesk foresaw to build up on existing legal framework and guidance, and develop support for assist Managing Authorities, paying agencies and evaluators in reporting on the RDP achievements in the enhanced AIR to be submitted in 2017, and beyond.

Under the guidance of DG AGRI Unit E.4 a Thematic Working Group consisting of qualified evaluation experts from the Member States and the EU was established in order to:

- Examine the challenges related to the reporting on evaluation in the AIR submitted in 2017, in consultation with stakeholders;

- Discuss possible solutions to overcome these challenges considering methodological and practical issues;

- Identify effective approaches to assess the progress in achieving objectives of the RDP, in particular through the assessment of the result indicators, including the complementary result indicators, and observable secondary effects on other result indicators in 2017 and answering related evaluation questions;

- Develop guidelines for preparing and drafting the evaluation and results sections of the enhanced AIR submitted in 2017.

Work has been carried out in a series of thematic workshops, desk-research, drafting, editing and lay outing. The draft guidance document was discussed with the Expert Group on Monitoring and Evaluating the CAP in several stages: during its 7th meeting on 2^5th June 2015 the Expert Group discussed the outline of the guidelines. A Sounding Board composed of Expert Group members discussed the 2nd draft of the guidelines in September 2015. During the 8th meeting of Expert Group the final draft of the document was discussed.

The present guidelines have been structured in several parts:

- **PART I** (mainly for Managing Authority) provides information and recommendations on what needs to be report in the AIR 2017. Specific attention was given to the governance and management of evaluation including the involvement of various actors and communication of evaluation results.

- **PART II** (For MA, PA and evaluators) provides detailed guidance through the preparation, structuring and conduction phase of evaluation, including methodological guidance and recommendations. This part leads the reader through the process of answering all common and programme specific evaluation questions for the entire programming period (with a specific focus on the information to be reported in the AIR in 2017).

- **Part III – Annexes** consisting of several useful tools such as the proposed SFC template for point 7 of the AIR submitted in 2017, an overview of reporting requirements, check-lists for self-assessment of the quality of the evaluation report, etc.

Table 1. Evaluation process and support provided in guidance

Phase	Working Step	Who?	Chapter	Expectded outcome
Planning	(see EP guidance)	MA	6.	Evaluation Plan
Preparing	Re-visit the RDP intervention logic	MA, Ev,	6.1	Revisited intervention logic
	Link intervention logic to evaluation elements	MA, Ev	6.2	Defined evaluation elements (indicator fiches, programme-specific indicators and EQs) Consistent evaluation framework RDP-specific evaluation elements
Structuring	Set up a consistent evaluation approach	MA, Ev.	7.1	Evaluation approach selected Evaluation method selected
	Establish the evidence for evaluation	MA, Ev.	7.2	inventory of existing data sources, data gap analysis; arrangements for data provision and access to data, databases and data management for evaluation
Conduct-ing	Observing	Ev.	8.1.1	Data and information for evaluation
	Analysing	Ev.	8.1.2	calculated values of indicators (gross, net)
	Judging	Ev.	8.1.3	Answers to EQs, conclusions, recommendations
	Reporting	MA	1	AIR submitted in 2017
	Dissemination	MA	4.1	Various dissemination products
	Follow-up	MA	4.2	

The following figure illustrates, the logic for the separation of the content between Part I, II and III of the guidelines.
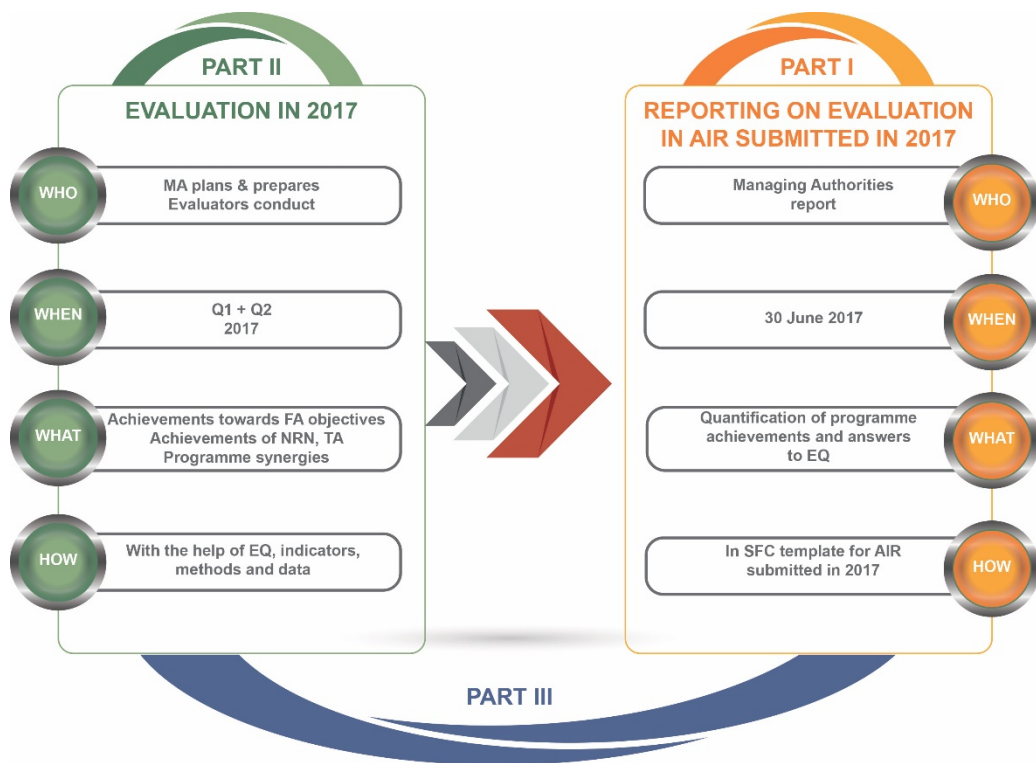
PART I – FOCUSING, MANAGING AND REPORTING ON EVALUATION

The programming period of 2014-2020, will include two AIRs, which will be submitted in 2017 and 2019 and combine both monitoring and evaluation elements. The monitoring elements of the 2017 and 2019 AIRs are identical to previous AIRs, however more and increasingly complex evaluation elements will be included as the programming period advances.

Figure 1. Content of Annual Implementation Reports[1]

ALL AIRs

Programme implementation
- Financial data
- Common and programme-specific indicators and quantified target values
- Achievements towards the milestones set in the performance framework (from 2017)

Evaluation Plan implementation
Issues affecting the RDP performance
Technical assistance
Programme publicity
Implementation of FI

AIR SUBMITTED IN 2017

Fulfilment of ex ante conditionalities
(or in 2016 where relevant)
Implementation of sub-programmes
Quantification of programme achievements in particular through assessment of CRI and relevant EQ
Actions to related to promotion of equality and non-discrimination, sustainable development & role of the partners

AIR SUBMITTED IN 2019

Progress towards EU2020
Assessment of the programme's net contribution to changes in CAP impact indicator values, and relevant EQs

Progress made in ensuring integrated approach

Source: European Evaluation Helpdesk for rural development, 2015

---

[1] Commission implementing regulation No 808/2014, Annex VII,

1   REPORTING REQUIREMENTS IN 2017

Monitoring and evaluation requirements in the AIR submitted in 2017:

The overall reporting requirements for 2017 combine monitoring and evaluation elements. The MA will gather monitoring information on various issues (e.g. progress on implementing the Evaluation Plan[2], financial commitments and expenditures[3], data on indicators[4] and progress towards targets). In addition the MA will:

- Organise the evaluation (e.g. to quantify complementary result indicators, answer evaluation questions[5,] assess the programme achievements, and provide conclusions and recommendations).
- Ensure monitoring information and other available data will be used as input to inform on the quantification and assessment of programme achievements.
- Report on the evaluation in the AIR submitted in 2017, on the basis of the report on evaluation[6.] This is the focus of these guidelines.

1.1   What is required in relation to reporting on evaluation in 2017?

The AIR submitted in 2017 must report on RDP evaluation[7]. The evaluation to be reported on must be carried out by internal or external experts that are functionally independent from the MA.[8] The reporting itself, i.e. summarising the main findings and recommendations can be done by the MA.

The main focus of the evaluation in 2017 is reporting on results. In this respect the AIR submitted in 2017 should include the following information stemming from evaluation activities:

> Reporting and quantification of programme achievements, in particular through the assessment of the complementary result indicators, and relevant evaluation questions.

The reporting requirements on evaluation in 2017 can be further analysed as follows:

- Summary of evaluation through the quantification and assessment of programme achievements denoted by result indicators (complementary result indicators[9], programme specific indicators and additional indicators/information). The quantification and assessment of indicators will be based on primary and secondary effects of RDP operations.[10] Values of result indicators should relate to completed operations where possible. If there are no completed operations, it may

---

[2] Commission implementing regulation No 808/2014, Annex VII, point 2
[3] Commission implementing regulation No 808/2014, Annex VII, point 1a
[4] Commission implementing regulation No 808/2014, Annex VII, point 1b
[5] Commission implementing regulation No 808/2014, Annex VII, point 7
[6] Report provided by evaluators to the MA on the basis of which the information on point 7 of the AIR is provided
[7] Commission implementing regulation No 808/2014, Annex VII, point 7
[8] Regulation (EU) No 1303/2013, Art. 54.3
[9] Commission implementing regulation No 808/2014, Art. 14.1 b) and Annex VII, point 7
[10] additional contributions of measures to focus areas, other than those under which they are programmed

be necessary, depending on the stage of implementation, to assess selected operations.

- Answer relevant (focus area related) evaluation questions[11] based on the assessment of the above indicators.
- Reporting period: The information should cover the years 2014-2016. The compilation of the report and the evaluation tasks (hereinafter: evaluation in 2017) has to take place in the first half of 2017.

1.2 How should Member States report on evaluation in 2017?

The reporting on evaluation in the AIR submitted in 2017 will be done through the AIR SFC template for each of the common evaluation questions separately. The SFC template will contain space for each of the programme specific evaluation questions, which Member States will have to complete in case the programme specific evaluation questions have been employed.

Point 7 of the SFC template for AIR submitted in 2017 will contain the following items:

- **Description of the system for answering the common evaluation question**: e.g. common and additional (if needed) indicators used for this purpose, as well as relevant methods and data used, etc.
- **Values of common and additional indicators:** Comments to calculations of common and additional indicators, and findings of triangulation and interpretation of these values, in the context of the RDP. In cases where the values of the indicators cannot be calculated (e.g. low up take), justify why and explain what was the alternative means used to answer the evaluation question e.g. studies, expert opinion, theory of change, etc.
- Calculate values (gross and net, if available) of all common, additional and programme specific indicators (the latest values should be used in case the programme specific evaluation questions have been employed and the answers should be provided with programme specific indicators).
- Answers to the CEQ

A proposed SFC template can be found in PART III of the guidelines as Annex 1.

How should the information to be provided in the respective part of the SFC template for the AIR submitted in 2017 be developed?

Guidance for the preparation, structure and treatment of the evaluation to be reported in the AIR submitted in 2017 are described in detail in PART II of the guidelines. In brief the development of required information should include:

- Assessment of RDP results, calculated on the basis of data from the monitoring system (common result indicators) and additional evaluation activities (e.g. complementary result indicators and programme specific indicators). The additional data gathering should also involve comparisons between beneficiaries

---

[11] Regulation (EU) No 1303/2013, Art. 50 and Commission implementing regulation No 808/2014, Annex VII, point 7

and non-beneficiaries in order to attribute the RDP achievements to the programme interventions.

- Validation of findings with qualitative research, if needed.
- Use of any further findings of various evaluations conducted before 2017, which provide evidence with respect to RDP results. , should be taken in consideration.
- Answers to common and programme specific evaluation questions related to FA and other aspects of RDP (synergies, Technical Assistance, NRNs) based on the evidence collected in the assessment of result indicators and additional information.
- Formulation of conclusions and recommendations for the improvement of RDP design and implementation based on the findings and answers to evaluation questions.

What kind of evidence should be used in the evaluation in 2017?

- Monitoring data on beneficiaries. It is considered to be a good practice to include in the operations database also the collection of data for result indicators.
- Data on non-beneficiaries from national/regional statistics, FADN, annual accounts, etc.
- Qualitative information can complement quantitative data in case of scarcity of the latter, e.g. to answer evaluation questions with respect to Technical Assistance, NRNs, delivery mechanisms, etc.
- In case of low uptake, the AIR submitted in 2017 may need to rely on more qualitative evidence and shift the focus from indicators to, for instance, the establishment of monitoring and evaluation systems, and implementation arrangements, etc.

Recommended considerations:

- Different forms of support (financial instruments) applied in the implementation of RDP measures can also affect the programme´s achievements and the net values of result indicators. Therefore the comparison of the effects of various financial instruments (in cases which they are applied) may be part of the assessment.
- Using sample data from monitoring systems (application forms before the project start and payment request after the project is finalised) will facilitate the assessment of secondary effects of measures on other Focus Areas.
- Although not explicitly required by legal acts, the assessment of programme delivery mechanisms is considered to be a good practice. Delivery mechanisms, i.e. the set of processes and procedures that ensure that policy objectives become concrete actions[12], is seen as one of the factors which affect the performance of the programme and the measures taken[13]. Among issues, which should be analysed with respect to the RDP delivery mechanism and its effectiveness and efficiency, are:
  - o Targeting the EAFRD support to selected groups of beneficiaries,

---

[12] http://enrd.ec.europa.eu/enrd-static/policy-in-action/improving-implementation/delivery-mechanisms/en/delivery-mechanisms_en.html
[13] Commission implementing regulation No 808/2014, Annex VII, point 3

- o Selected types of support (financial instruments, simplified cost options) with respect to measures,
- o Information dissemination of EAFRD support to beneficiaries,
- o Application, selection, contracting and payment procedures,
- o Monitoring mechanisms and feedback from beneficiaries on the application of selected measures etc.

How to handle evaluation in 2017 in case there is low uptake?

If a programme approval or start was delayed severely, there will be little or no monitoring data on completed operations by beneficiaries to assess result indicators. In such a case, it is necessary to assess reasons of low uptake and take into consideration any information available on potential beneficiaries, applications, existing/ongoing contracts and explain the situation why result indicators could not be calculated as required. Relevant common and programme specific result indicators should be calculated for those measures and focus areas, which have shown sufficient programme uptake.

The legal acts also require reporting in the AIR on evaluations that have been done in the programme area during the previous years. This covers all studies relevant for a given RDP area. For example there could be studies conducted by research institutes or universities in the area of climate change, biodiversity, local development, or business development, which provide useful information on RDP beneficiaries and territories. Findings of such studies have to be summarised in the AIR. They can be further used for evaluation purposes in case of low uptake.

How to handle evaluation in 2017 in the case of small programmes?

Small programmes with a low number of beneficiaries (typically in multi-regional Member States) may find it difficult to quantify result indicators and answer the relevant evaluation questions using quantitative methods, due to a lack of data. Qualitative research should nonetheless be conducted to answer the EQs. Furthermore, the proposal should focus the assessment on the level of achievement of RDP targets and milestones in these AIRs.

Further reading

Guidelines: Establishing and implementing the Evaluation plan of 2014-2020 RDPs, Evaluation Helpdesk 2014-2020, PART I, Chapter 3.4 and PART II, Chapter 6.4, Brussels, 2015, http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications

Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDPs, PART II, Chapter 1, 4 and 5 and Part III, Annexes 5 and 6, Evaluation Helpdesk 2007 – 2013, Brussels, 2015,
http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

1.3   Reporting on the evaluation of RDP specific elements

Flexibility in programming and its implications on evaluation

Flexibility in programming is one of the new characteristics of the 2014-2020 programming period. It aims to strengthen the strategic approach and to increase the effectiveness, efficiency and performance of rural development policy. In practical terms, Member States have flexibility to develop programme specific objectives, and to adapt the combination of measures. Measures are no longer attributed to specific "axes" as in the past, but can be flexibly programmed and combined under the Union priorities/focus areas and programme specific objectives. The underlying rationale is that Member States should be enabled to mix and combine measures under focus areas in a way that better reflects their specific rural development needs.

Flexibility in programming is helpful in the RDP monitoring and evaluation system. Namely, it is useful to develop programme-specific evaluation questions and indicators to complement the monitoring and evaluation system, tailored to concrete RDPs.

The figure below illustrates the linkages between objectives, and common and programme specific evaluation elements connected with the programme intervention logic.

Figure 2. The common and programme specific monitoring and evaluation system
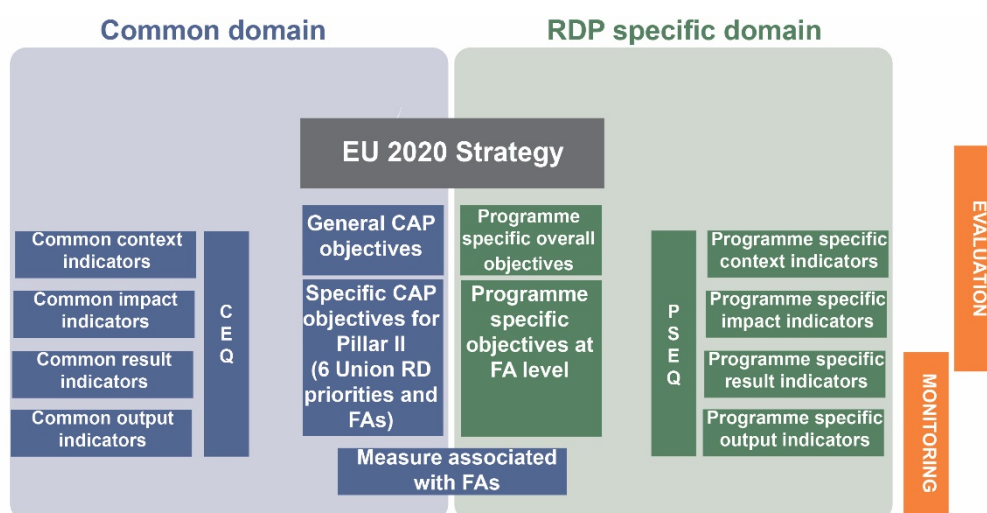
Source: European Evaluation Helpdesk for Rural Development, 2015

The MA may need to develop additional evaluation elements to assess aspects that are of interest for the MA. These could include the assessment of the NRN, RDP delivery mechanisms, issues supported via technical assistance, such as administration and management, communication, etc. The specific evaluation topics also require the development of specific EQs, judgment criteria and means to collect evidence to answer the EQ.

It is best to define the programme specific evaluation elements already in the programme design phase and include them in the evaluation plan. As an alternative, the programme-specific elements can be designed later, e.g. before or at an early stage of programme implementation to ensure suitable data will be available early on. Further programme specific elements can also be developed when conducting the actual evaluations, in case gaps in the M&E framework are still detected by the evaluators.

The principle of proportionality should be respected when developing and reporting on programme-specific monitoring and evaluation elements. This implies that the utility of additional information reported should be balanced against the resources required to provide this additional information.

PARTII, Chapter 6 provides detailed guidance on how to develop programme specific evaluation questions and indicators.

Reporting of programme specific M&E elements

Reporting on evaluation requires the use of both common and programme-specific evaluation questions and indicators[14]. The standard AIR are the main channel for Member States to report on the achievements of the RDP by referring to financial data, indicators, and quantified target values. In the AIR submitted in 2017, the RDP achievements should be quantified through the assessment of common and programme specific results indicators[15]. The evidence collected via result indicators helps to answer focus area related common and programme specific questions.

---

[14] Regulation (EU) No 1303/2013, Art. 50.2
[15] Commission Implementing Regulation No 808/2014, Annex VII, point 7

2 LINKS TO OTHER EU REPORTING REQUIREMENTS AND REPORTING REQUIREMENTS BEYOND 2017

2.1 Links and synergies between the AIR 2017 and other EU reporting requirements

The AIR submitted in 2017 has various links with other reports at Member State and EU levels submitted in 2017, or 2018 respectively, which inform on the implementation and progress of programmes financed by the ESI Funds and the CAP. (See Figure 3).

Figure 3. Links between AIR submitted in 2017 and other EU reporting requirements in 2017
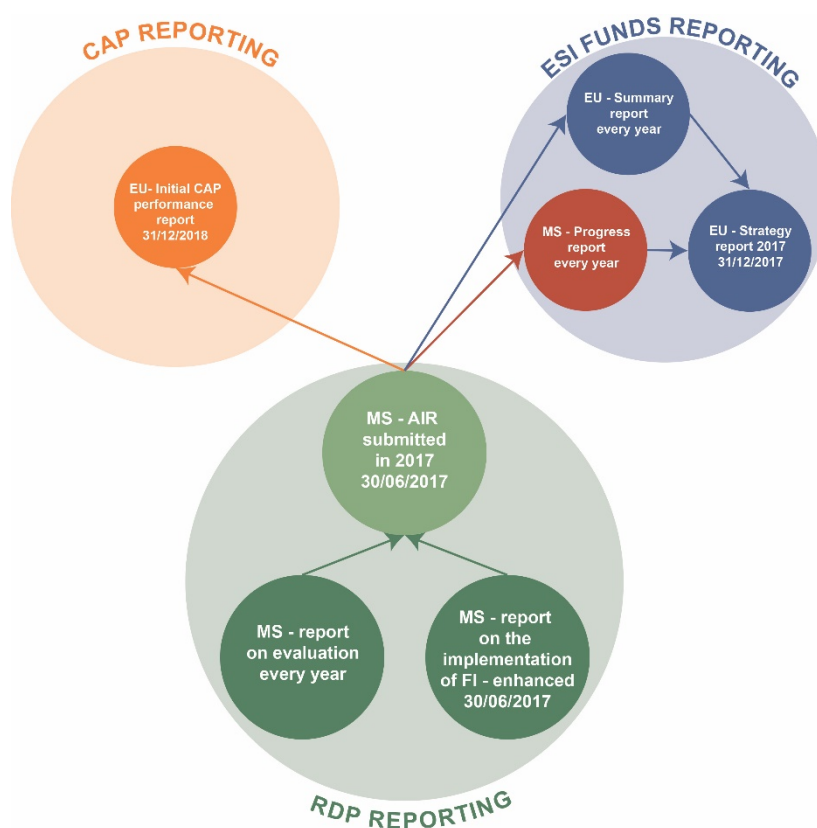
Source: European Evaluation Helpdesk for Rural Development, 2015

The links between AIRs and other reporting requirements in years 2016 – 2019 are illustrated in the figure above and summarised in the following table.

Figure 4. Overview of reporting requirements and links 2016 – 2019
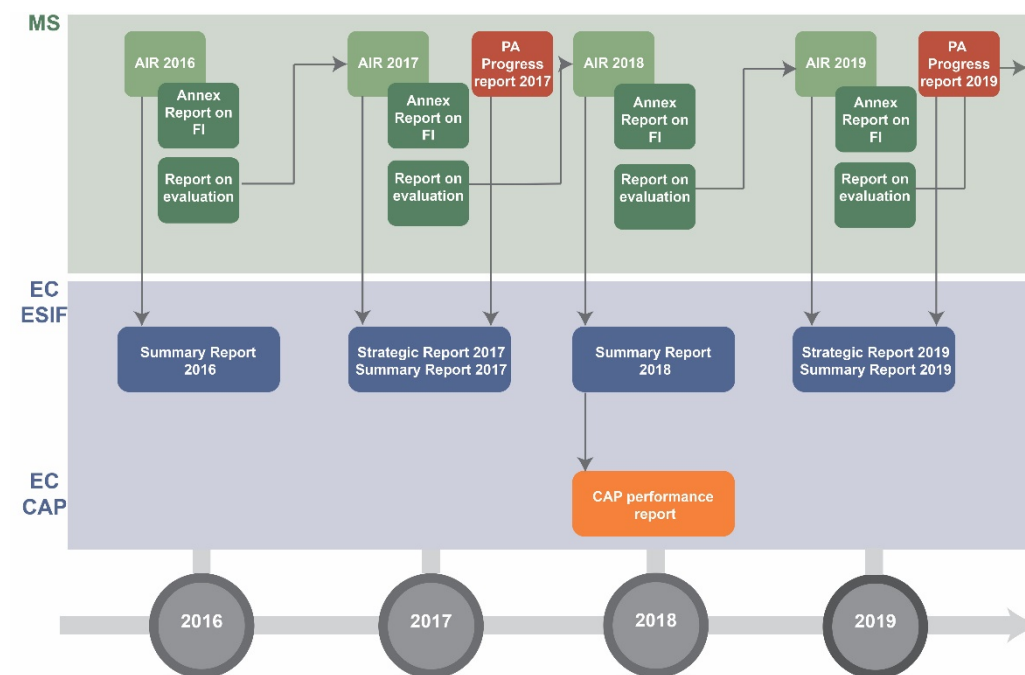
Table 2. Links between AIR 2017 and other EU reporting requirements

Reporting requirements	Legal source	Links and synergies with AIR submitted in 2017
RDP related reporting		
Reporting on evaluation (RDP level) The MAs shall ensure that evaluations are carried out according to the Evaluation Plan (in particular those to assess effectiveness, efficiency, and results. At least once during the programming period, an evaluation shall assess how support from the ESI Funds have contributed to the objectives for each priority. All evaluations shall be examined by the Monitoring Committee and sent to the Commission.	1303/2013, Art. 50.2 and 56.1 Commission Implementing Regulation No 808/2014, Annex VII, point 2 d) and e)	AIR must contain the synthesis of the findings of all RDP evaluations that may have become available during the previous financial year. AIR 2017 has to include evaluation elements, such as answers to EQs and assessment of progress in terms of result indicators related to focus areas.
Report on the implementation of financial instruments (RDP level) The MA shall send to the Commission each year a specific report covering the operations comprising financial instruments. This report includes information inter alia on implementation arrangements of the financial instrument, identification of bodies implementing financial instruments, financial information on the financial instrument, the performance of the financial instrument and contribution to the achievement of indicators of the priority or measure concerned.	1303/2013, Article 46.2 Commission Implementing Regulation No 808/2014, Annex VII, point 10	This report on implementation of financial instruments is annexed to the AIR. The AIR submitted in 2017 should additionally contain information on the progress of achieving the expected leverage effect of investments made by the financial instrument and the contribution to the achievement of indicators of the priority or measure concerned.
CAP related reporting		
CAP performance report (EU level) The Commission presents the initial report on the implementation of Article 110 of Regulation 1306/2013. This report concerns the monitoring and evaluation of the CAP and the first results of its performance. This report shall be presented to the European Parliament and the Council by 31 December 2018.	1306/2013, Article 110.5	The AIR 2017 will include the synthesis of the findings of all evaluations of the RDP as well as changes in the values of result indicators. This body of information will therefore form one of the information sources on the performance of rural development measures for the preparation of the CAP report.
Partnership Agreement related reporting		
Progress report on PA (MS level) By 31 August 2017, the MS shall submit to the Commission the progress report on the implementation of the Partnership Agreement as at 31 December 2016.	1303/2013, Article 52.1	The progress report relates to the implementation of ESI Funds via operational programmes under the Partnership Agreement. The AIR 2017 provides the input in relation to the implementation of EAFRD RDP results and the RDP contribution towards policy objectives (where appropriate).

Reporting requirements	Legal source	Links and synergies with AIR submitted in 2017
Strategic report (EU level) In 2017 the Commission shall prepare a strategic report summarising the progress reports of the Member States. The strategic report shall be submitted to the European Parliament, the Council and the European Economic and Social Committee and the Committee of the Regions by 31 December 2017.	1303/2013, Art. 53.2	The information from the AIR submitted in 2017, including reporting on evaluation, is fed into the progress report on the PA submitted in 2017, which is synthetized in the Strategic report at the EU level.
ESI Fund related reporting		
Summary report (EU level) The Commission shall submit a summary report in relation to ESI Fund programmes each year from 2016 to the European Parliament, the Council and the European Economic and Social Committee and the Committee of the Regions.	1303/2013, Art. 53.1	The AIR submitted in 2017 also provides the input to the summary reports on the implementation of ESI Funds, submitted each year. In 2017 the summary report will form part of the strategic report.

2.2 Links between reporting in 2017 and the overall reporting requirements on rural development in 2014-2020

Reporting requirements on the evaluation in 2017 relate closely with the reporting on evaluation across the entire programming period. Since evaluation tasks become more complex and sophisticated towards the end of the programming period, it is necessary to properly set up and develop the monitoring and evaluation system. In this way, it will be possible to provide information for the evaluation across the programming period and the ex post evaluation. The summary of reporting requirements across the programming period 2014-2020 can be found in the PART III of the guidelines as Annex 2.

Further reading

Guidelines: Establishing and implementing the Evaluation plan of 2014-2020 RDPs, Evaluation Helpdesk 2014-2020, PART I, Chapter 5.5 and PART III, Annex 14, Brussels, 2015, http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications

3   ENSURING GOOD QUALITY EVALUATION IN 2017

3.1   Careful planning and preparation of evaluation

There is limited time to conduct the evaluation tasks to be reported in 2017. Hence, prior planning and preparation is vital for the MA in order to be able to provide the necessary information in the enhanced AIR and to submit it to the Commission by the end of June 2017.

In the current programming period, the MAs are **required to plan evaluations** from the beginning, with the help of an evaluation plan, which is part of the programme [16]. As the evaluation plan was developed during the programming phase, it typically covers only minimum requirements at a general level. In this case, it is good to consider complementing the evaluation plan with an internal, more detailed planning document, containing more comprehensive information on the planned evaluation activities, topics and their timing.

In order to facilitate reporting and conduct evaluation in 2017, it is vital to identify the evaluation needs and activities related to the RDP early on. Therefore, relevant stakeholders need to invest resources and make **careful preparation** of evaluation methods either before the programme starts or at early stages of implementation.

The figure below provides an overview of the evaluation process.

Figure 5. Overview of the evaluation process
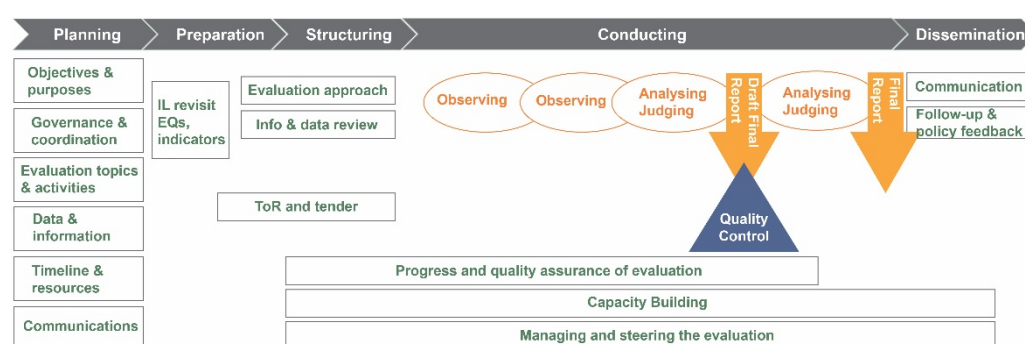
Source: European Evaluation Helpdesk for Rural Development, 2015

A detailed description of preparing, structuring and conducting the evaluation can be found in PART II of the guidelines.

The evaluation process is already documented in other guidelines [17] and therefore will not be further elaborated on in this document. Some important aspects do however

---

[16] Regulation (EU) No 1305/2013, Art. 8.1 g)
[17] Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDPs, Part I, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

require further explanation (such as quality assurance, quality control, communication, follow-up and policy feedback) and are highlighted below.

3.2 Quality assurance and quality control across the evaluation process

It is a **good practice** to build quality assurance (QA) and quality control (QC) components into the evaluation process.

Quality assurance focuses on the process and is a proactive way of preventing defects in the outcome. Quality assurance includes a notion of observation, development of quality standards, and continuous improvement.

Quality control, on the other hand, is product oriented and ensures that the outcome is what was expected. It is typically performed at the end of the process.

Quality control of evaluations is the responsibility of the Managing Authority of the programme. MAs have many different means to safeguard quality, e.g. drafting of precise ToR, selecting qualified evaluators, setting up evaluation steering groups, keeping regular contact with data providers, communicating regularly with evaluators, and requiring a high standard of evaluation reporting, etc. The capacity of the **evaluator** to use advanced evaluation methods also influences the quality of evaluation. The ability to bridge existing data gaps provide sound interpretation of evaluation findings and evidence-based answers to evaluation questions, conclusions and applicable policy recommendations. Other parties who may contribute to the improvement of the quality of evaluation are:

- Steering Committees (if established), can ensure both the professional capacity of evaluators and the quality of evaluation reports;
- Paying Agencies can assist in providing higher quality data for evaluations;
- Programme beneficiaries and their associations, who show interest in hearing evidence-based recommendations can help in the implementation of the programme;
- The general public in their position as tax payers, can demand solid information on the results of public spending.

3.3 Quality assurance in the evaluation process

The most typical **quality assurance tools** used in the evaluation process are checklists, content lists, general and detailed process maps. A set of tools has been developed to assure the quality of each phase of the evaluation process. These tools can be developed separately, but it may be useful for the MA and other stakeholders involved in the evaluation (e.g. evaluation steering group members) to draft a complete evaluation quality assurance handbook, with detailed process maps (sequencing tasks and responsibilities of main stakeholders to be followed at each phase) and checklists for each phase. A complete handbook for assuring the quality of evaluation can also

serve as a capacity-building manual. Maintaining institutional memory with regard to evaluations is vital.

Quality assurance in planning evaluation

Once the overall process and timetable for the evaluation are outlined, evaluation needs must be examined. Typically, a concept note is written to improve the focus of the evaluation planning and to prepare for the ToR. A typical content list for a concept note includes: the topic, timing, scope, and key areas of focus, as well as stakeholder roles.

Communication and capacity building plans for the evaluation should be drafted at the planning phase to maximise the quality of the evaluation and the use of evaluation results. A content list or a checklist can be used to ensure the inclusion of major stakeholders, usage of correct communication channels and tools, identification of communication responsibilities, and optimal timing of communication.

Evaluation capacity building:

- combines the improvement of evaluation knowledge and skills of individuals with the strengthening of the organisational evaluation-related mechanisms, namely established systems and processes;

- relates to human resources, which begins with the identification of the target audiences (e.g. MA, PA, policy makers, NSU, LAGs and other NRN members) and the assessment of their evaluation skills and information needs, as well as knowledge levels. Once it is clear what the main capacity building needs related to evaluation are, it is possible to design the capacity building actions, their timing, and who is responsible for them;

- in relation to organisations, means improving organisational learning and support mechanisms related to evaluation. Organisational learning means that the organisation observes its actions, documents its processes and tries to improve on all fronts. When the processes are documented, it is possible to set quality standards for them. The processes are also documented so that the processes and responsibilities are clear and knowledge remains in the organisation even in the event of staff changes. Quality assurance and utilisation of evaluation findings are integral parts of organisational learning. Content lists and checklists can further be used in assuring the quality of the capacity-building plan.

Quality assurance in preparing evaluation

PART II, Chapter 6 of the guidelines elaborates on what needs to be done to ensure the high quality of evaluation. The most vital issues are to attain a high quality of the intervention logic, evaluation questions, indicators, evaluation approach, information and data review, evaluation focus, data gaps, and the extent to which the topic can be evaluated. These issues influence the drafting of the ToR and the contract drawn up with the evaluator. A systematic, critical scrutiny of the proposed evaluation questions done through using a checklist with guiding questions such as: "Do the EQs cover all the objectives and evaluation needs? Are questions clearly formulated? Can these EQs be answered using the data that is available or which can be gathered?"

When **drafting the ToR**, it is important to also include a quality assessment grid against which the quality of the final report will be judged. Once the ToR has been drafted, it should be subject to a quality check in terms of required content, quality of the content, and the process. The required content can be checked against a checklist (see Annex 4 Check-list for self-assessment of the quality of the evaluation report, Part III). The checklist can also include the quality elements related to the content (accuracy, adequacy, relevance, and clarity) as well as the assessment of the proposed evaluation process and timeline.

Quality assurance in structuring and conducting evaluation

The quality assurance process starts with the implementation of evaluation. The ToR shall be used as the reference point at this stage.

It is also good to develop quality standards and a checklist for the inception report (including criteria on content such as methodological approach and data collection methods, as well as a criteria for the operational plan and the process of evaluation). The inception report can also include a section on quality assurance by the evaluator. Furthermore, it is equally important for the client to ensure that internal and external processes are in place for the timely delivery of materials to the evaluator.

The client and/or steering group can also develop and use checklists to scrutinise the content and quality of the progress and draft final reports. These tools can help to improve the quality and focus of the final report, as well as to give systematic and constructive feedback to the evaluator.

Quality assurance in dissemination of evaluation findings

The final report should be subject to a quality assessment, preferably using the assessment criteria included in the tender. A good quality criteria for the final evaluation report should include items relating to the evaluation process (relevance, timeliness, inclusiveness), normative issues (focus on independence and impartiality of the evaluator), as well as technical criteria (relevance of the evaluation, appropriate design, reliable data, sound analysis, credible findings, evidence-based answers to evaluation questions, valid conclusions, helpful recommendations, and report clarity)[18]. A draft checklist for the assessment of the quality of the evaluation report is included as an example in the Part III of this document.

The MA should draw up a plan, timetable and a checklist for following up evaluation recommendations. The fulfilment of the plan should be assessed at regular intervals. It is good practise for the MA and the evaluators to give mutual feedback at the end of the evaluation. The MA may also wish to assess the quality of the evaluation communication at the end of the process.

---

[18] Annex 6 'Quality Assessment Form', DG Markt Guide to Evaluating Legislation, pages 87-97
http://ec.europa.eu/dgs/internal_market/evaluation/evaluation/index_en.htm

Further reading

Guidelines: Establishing and implementing the Evaluation plan of 2014-2020 RDPs, Chapter 5.3, Evaluation Helpdesk 2014-2020, Brussels, 2015, http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications

Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDPs, PARTI Chapter 1.2 and Part III, Annex 4, Evaluation Helpdesk 2007 – 2013, Brussels, 2015, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

4   DISSEMINATION, COMMUNICATION AND FOLLOW UP OF EVALUATION RESULTS

4.1   Dissemination and communication of evaluation in 2017

The value of evaluation depends on its use, which means to disseminate, communicate, and use the evaluation findings to improve the policy.

In terms of dissemination, it is fundamental that the reporting on evaluation in the AIR submitted in 2017 is made public, e.g. on the website of the MA. Furthermore, to increase transparency, the evaluation report should be disseminated to the participants of the evaluation process and all the relevant stakeholders. As a good practice, it is recommended to write a citizens' summary of the main findings of the evaluation. It might also be useful to translate the citizens' summary to English.

Communication occurs throughout the evaluation process, but the main communication effort comes at the end, after the results and recommendations have been finalised. The communication actions should follow the communication plan developed at the beginning of the evaluation process. The main focus of the communication should be on results and achievements of the RDP. In case of low uptake of measures, results of other studies linked to RDP (e.g. water efficiency) could be included in the communication.

Specific information needs with respect to various target audiences for dissemination and communication of evaluation findings can be found in the table in the PART III of the guidelines as Annex 3.

The evaluation communication plan should be monitored and assessed to check its efficiency and effectiveness in delivering the key messages to the target audiences. The assessment of the communication plan on evaluation findings can be covered by the overall RDP evaluation. Further information on dissemination and communication of evaluation findings can be found in other Evaluation Helpdesk guidelines[19].

4.2   Follow-up of evaluation findings

Evaluation as part of the governance of the EU programmes constitutes a strategic management tool. When used effectively, the follow-up on evaluation findings results in:

- strengthening the use of evaluations;
- stimulating an evaluation culture based on the organizational learning and enhancing the liability for results;
- facilitating the discussion about programme evaluation;

---

[19] Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDPs, Evaluation Helpdesk 2007 – 2013, Brussels, 2015, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

- motivating stakeholders and programme managers to actively support and participate in the RDP performance improvement; and
- enhancement of public policies.

If the evaluation fulfils this role, all involved stakeholders, and particularly Managing Authorities and policy makers, need to pay considerable attention to evaluation findings and recommendations. Additionally, to receive sufficient attention from stakeholders, evaluations must be of high quality and effectively and efficiently present the evaluation findings.

Good quality evaluations produce better quality findings and operational recommendations that are more easily implemented, thus more likely to contribute to better policy design and delivery.

The minimum requirements for the Evaluation Plan require that the MA describe the mechanisms established to follow-up on the use of evaluation results.[20] The following table provides a tool for following up on the recommendations of evaluations, describing the management response, the actions planned and the implementation of actions. These tables have to distinguish between recommendations that depend on programme internal factors and recommendations that depend on external factors.

It is recommended to present these tables in the AIR, in order to safeguard certain levels of formalisation, the transparency in using evaluation findings and rigorous implementation of the recommendations.

The tables also establish a tangible commitment between the MA and the evaluators to develop an action plan for the implementation of recommendations.

Table 3. Evaluation follow-up matrix

WHAT?	WHEN?	WHO?	HOW?	WHAT HAS BEEN DONE?
What are the major evaluation findings? What was concluded/ recommended? What is the expected change? How does the MA see it? What should be done?	Time frame for follow-up, final date of follow-up	Responsibility for follow-up Others involved	Format of follow-up	Status of follow-up, date
Finding- conclusion - Recommendation 1: firmly based on evidence and analysis, clearly formulated, pragmatic Expected change:	quarterly, annual, biannual, etc	actor, unit	organisation's annual work planning / reporting, AIR, MC, SG, etc.	Status codes: not yet started ongoing finalised

---

[20] Commission implementing Regulation (EU) 808/2014, Annex I, part I, point 9

WHAT?	WHEN?	WHO?	HOW?	WHAT HAS BEEN DONE?
according to the evaluation report MA response: accepted, partially accepted, rejected – provide reasons Key actions:				
Finding- conclusion - Recommendation 2: Expected change: MA response: Key actions:				

Further reading

Guidelines: Establishing and implementing the Evaluation plan of 2014-2020 RDPs, Chapter 3.6, Evaluation Helpdesk 2014-2020, Brussels, 2015, enrd.ec.europa.eu/en/evaluation-helpdesks-publications

Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDPs, Chapter 1.2.3, Evaluation Helpdesk 2007 – 2013, Brussels, 2015, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

PART II ASSESSING RDP ACHIEVEMENTS

While PART I of the guidelines defined what needs to be reported in the AIR 2017 and was elaborated mostly for Managing Authorities, PART II is dedicated to a broader spectrum of evaluation stakeholders (Managing Authorities, evaluators, but also PA, NRN, LAGs etc.). This part focuses on how to prepare, structure and conduct the evaluations necessary to answer all common and programme specific evaluation questions for the entire programming period (with a specific focus on the information to be reported in the AIR in 2017).

5   SETTING UP THE SYSTEM TO ANSWER EVALUATION QUESTIONS

Evaluation questions (EQs) are an important component in the evaluation of rural development policy. They define the focus of evaluation in relation to EU and programme specific policy objectives and help to demonstrate the progress, impact, achievements, effectiveness, efficiency and relevance of rural development programmes[21].

Common evaluation questions for rural development (CEQ)[22] were designed by the EC in collaboration with Members States. As part of the common monitoring and evaluation system for rural development, CEQ enhance the comparability of evaluation results across Europe. They provide support to Member States in the evaluation of achievements of their rural development programmes (RDP) towards the EU policy objectives and programme results and impacts. There are three types of CEQ: 1) focus area related CEQ, 2) CEQ related to EU level objectives and 3) CEQ related to other aspects of the RDP.

Whenever the set of CEQ is not considered to be sufficient for capturing programme specific needs, priorities and objectives, Member States may additionally develop programme specific evaluation questions (PSEQ).

It is a legal requirement to answer the common and programme specific evaluation questions[23] at different points in time during the programming period and also after programme implementation. Member States shall provide evidence-based answers:

- In the annual implementation report (AIR) **submitted in 2017**, evaluation questions related to the RDP focus areas and evaluation questions related to other RDP aspects (synergies, TA and NRN) shall be answered by means of common and programme specific result indicators,
- In the AIR **submitted in 2019, and in the ex post evaluation report**, all common and programme specific evaluation questions shall be answered by means of common and programme specific result and impact indicators.

In order to provide answers to evaluation questions, which show the real RDP results, impacts and achievements towards the policy objectives at a given time, it is vital to assess net values of result and impact indicators. This calls for the application of advanced evaluation methods and requires high quality data and information.

Although in 2017 "only" CEQ related to focus areas and related to other RDP aspects should be answered, stakeholders in the Member States are well advised to start collecting already right from the beginning the evidence for answering *all* evaluation questions. This requires to collect data and information for all indicators (results and

---

[21]Regulation (EU) No 1303/2013, Art 54 (1); Regulation (EU) No 1305/2013 Art 68 (a)
[22] WD: Common Evaluation Questions for Rural Development programmes 2014-2020,
      https://enrd.ec.europa.eu/en/evaluation-helpdesks-publications
[23] Commission Implementing Regulation (EU) 808/2014, Art. 14.1c), Annex VII, Point 7,

impacts) necessary to answer the evaluation questions and includes also baseline values of indicators.

PART II elaborates on major steps in setting up the system to answer evaluation questions for the entire programming period (and at any evaluation milestones – in 2017, 2019 and ex post), highlighting the importance to carefully prepare and structure the evaluation prior to conducting the evaluation, to finally answer all evaluation questions and to formulate conclusions and recommendations for improved policy design and implementation.

The following major steps to set up the evaluation system for collecting the evidence to answer evaluation questions are elaborated in detail in PART II:

With respect to preparing the evaluation:

- Re-visit the RDP intervention logic and validate or review the findings of the ex ante evaluation as first step prior to conducting any evaluation.
- Link intervention logic to evaluation elements:
  - o Define evaluation elements, making sure that all terms concerning objectives, evaluation questions (both common and programme specific) and indicators (both common and programme specific) are consistent and formulated in such a way that everyone can understand; if this is not the case, ensure proper definitions.
  - o Check the consistency between intervention logic and evaluation elements, ensuring that all RDP achievements can be evaluated, proper evaluation questions are asked and that proposed indicators allow collecting right evidence to answer evaluation questions in a robust way.
  - o Develop additional indicators to answer common evaluation questions, ensuring that sufficient evidence is collected for those CEQ, linked focus areas, where achievements are measured with common target indicators.
  - o Develop programme specific evaluation questions and indicators, filling the gaps in case the existing evaluation questions (both common and programme specific) and indicators (both common and programme specific) are not sufficient to capture all programme achievements.

With respect to structuring the evaluation:

- Set up evaluation approach, deciding on most suitable evaluation approach and the set of evaluation methods which would allow providing robust answers to evaluation questions. Preference is given to a mixed evaluation approach, based on counterfactuals and the mixture of advanced quantitative and qualitative evaluation methods.
- Ensure the availability of data and information for evaluation, starting with screening of all available data and information sources and making sure that they fit for the RDP evaluation purpose, establishing the data management and collection system, linking databases etc.

With respect to conducting the evaluation:

- Observing, collecting available data and filling the data and information gaps with collection of additional data (e.g. via surveys etc.)
- Analysing, calculating and assessing the indicators, ideally in net values, triangulating and interpreting the evaluation findings (e.g. values of indicators), asking why the evaluation findings are as they are.
- Judging, answering the evaluation questions and providing the conclusions and recommendations for the improvement of RDP design and implementation.

Figure 6. Major steps in setting up the system to answer evaluation questions
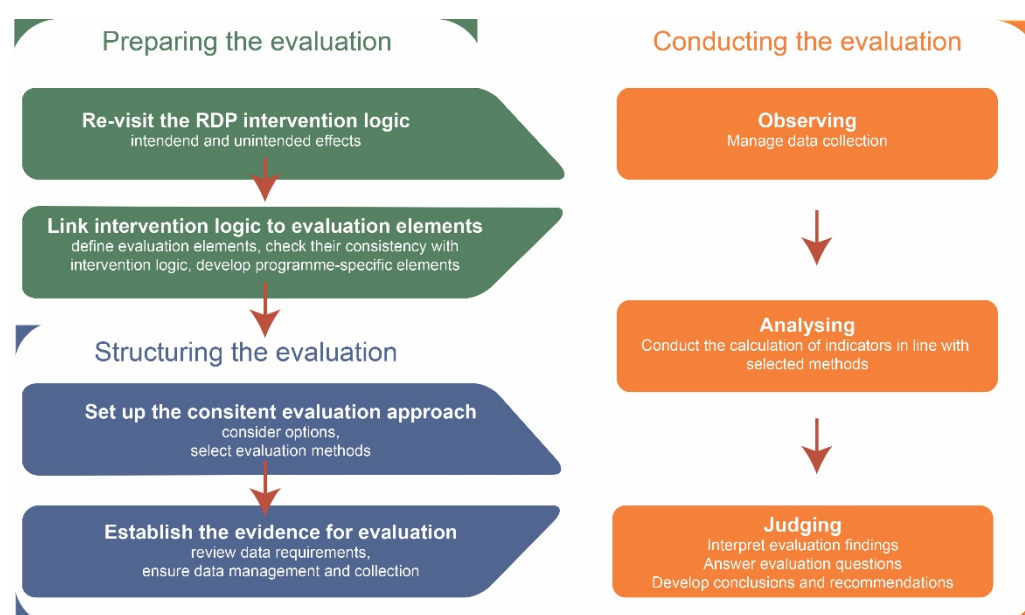
Source: European Evaluation Helpdesk for Rural Development. 2015

Responsibilities of various evaluation stakeholders are highlighted for each of the above mentioned steps and links with the ToR are explained as well. Special attention is given to specificities with respect to Leader, Technical assistance, NRN, the situation of low programme up take and in case of small RDPs in terms of budget and planned operations (proportionality).

Detailed guidance on how to specifically answer the CEQ to be reported in 2017[24] is furthermore provided in CEQ templates in PART III.

The table below illustrates at what point in time the three different types of CEQs need to be prepared, structured, observed, analysed and judged.

---

[24] In the AIR submitted in 2017 the CEQ related to FA (No 1 – 18) and CEQ related to other RDP aspects (No 19 – 21) shoud be answered. In the AIR submitted in 2019 all CEQ shoudl be answered.

Evaluation process	years									
	2014-2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Preparing evaluation	CEQ-FA	CEQ-FA		CEQ-FA					CEQ-FA	
	CEQ–OA	CEQ–OA		CEQ–OA					CEQ–OA	
	CEQ–EU	CEQ–EU		CEQ–EU					CEQ–EU	
Structuring evaluation	CEQ-FA	CEQ-FA		CEQ-FA					CEQ-FA	
	CEQ–OA	CEQ–OA		CEQ–OA					CEQ–OA	
	CEQ–EU	CEQ–EU		CEQ–EU					CEQ–EU	
Conducting evaluation — Observing		CEQ-FA	CEQ-FA	CEQ-FA	CEQ-FA	CEQ-FA	CEQ-FA	CEQ-FA	CEQ-FA	CEQ-FA
		CEQ–OA	CEQ–OA	CEQ–OA	CEQ–OA	CEQ–OA	CEQ–OA	CEQ–OA	CEQ–OA	CEQ–OA
		CEQ–EU	CEQ–EU	CEQ–EU	CEQ–EU	CEQ–EU	CEQ–EU	CEQ–EU	CEQ–EU	CEQ–EU
Conducting evaluation — Analysing		CEQ-FA	CEQ-FA	CEQ-FA	CEQ-FA				CEQ-FA	CEQ-FA
		CEQ–OA	CEQ–OA	CEQ–OA	CEQ–OA				CEQ–OA	CEQ–OA
				CEQ–EU	CEQ–EU				CEQ–EU	CEQ–EU
Conducting evaluation — Judging			CEQ-FA		CEQ-FA					CEQ-FA
			CEQ–OA		CEQ–OA					CEQ–OA
					CEQ–EU					CEQ–EU
Reporting evaluation			CEQ-FA		CEQ-FA					CEQ-FA
			CEQ–OA		CEQ–OA					CEQ–OA
					CEQ–EU					CEQ–EU

CEQ related to FA: CEQ-FA
CEQ related to other aspects: CEQ-OA
CEQ related to EU objectives: CEQ–EU

6   PREPARING THE EVALUATION

In the current programming period Member States are required to plan evaluations right from the beginning with the help of an evaluation plan, which is part of the programme[25]. Moreover, some Member States have complemented the Evaluation Plan in the RDP also with an internal, more detailed planning document, which contains more comprehensive information on the planned evaluation activities, topics and their timing.

Once the RDP has been approved, the MA in collaboration with other relevant evaluation stakeholders should therefore start to prepare the evaluation to be reported in 2017 and later on. The main emphasis should be put on setting up the system to answer the evaluation questions.

The first report on evaluation in the AIR submitted in 2017 requires to report only on common evaluation questions, which relate to focus areas and to other RDP aspects. However, it is highly recommended to invest sufficient time and resources in preparing the system to answer all common and programme specific evaluation questions and to start at an early stage with this process. This is especially important with a view to ensure the availability of high quality data and information (including baseline data).

Figure 7. Preparing the evaluation as part of the evaluation process.
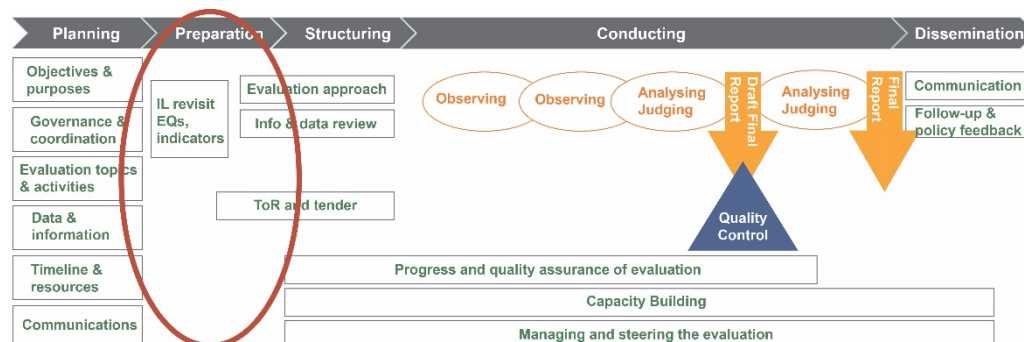
Source: European Evaluation Helpdesk for Rural Development, 2015

What needs to be covered when preparing the evaluation in 2017 and later?

In order to prepare for reporting on evaluation in 2017 (and for the entire programming period), the programme authorities should ensure that:

- The management of evaluation is established, staffed, equipped with necessary knowledge and skills and that all relevant stakeholders[26] are contacted and present in relevant working groups or steering groups for single evaluations (see PART I, chapter 3)

---

[25] Regulation (EU) No 1305/2013, Art. 8.1 g)
[26] Managing Authority, paying agency, evaluation experts, evaluators, data providers, monitoring committee members

- The RDP intervention logic, its EU common and programme specific objectives and expected effects are clearly understood and validated by relevant stakeholders (see PART II, chapter 6.1)

- EU common and programme specific evaluation elements (evaluation questions and indicators) and their relation to RDP intervention logic (overall and FA specific objectives, measure) are identified and known by relevant stakeholders and that they are able to judge on the sufficiency of evaluation elements to capture all RDP effects (see PART II, chapter 6.2),

- All the terms used in the RDP objectives, focus areas, measures, evaluation questions (and their judgment criteria) and indicators are defined (see PART II, chapter 5.2),

- The relevant evaluation stakeholders are familiar with evaluation approaches[27] and their suitability to serve the purpose of evaluation, respecting the data availability, quality and frequency (see PART II, chapter 7.1),

- The monitoring system is adjusted to the evaluation needs, e.g. data on result indicators and additional information are collected from beneficiaries (see PART II, chapter 7.2),

- Existing data sources, providers, arrangements to adjust data to RDP needs and data gaps are recognised and procedures how to fill the gaps defined (see PART II, chapter 7.3),

- Capacities of relevant stakeholders to manage and conduct the evaluation is ensured and continuously built.

Whenever these preparatory point are completed, the actual evaluation exercise can start or, in case the MA wishes to contract an external evaluator, the Terms of Reference can be drafted.

Further reading

Guidelines: Establishing and implementing the Evaluation plan of 2014-2020 RDPs, Evaluation Helpdesk 2014-2020, PART I, Chapter 3.4 and PART II, Chapter 6.4, Brussels, 2015, http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications
Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDPs, PART II, Chapter 1, 4 and 5 and Part III, Annexes 5 and 6, Evaluation Helpdesk 2007 – 2013, Brussels, 2015,
 http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

6.1   Revisit the RDP intervention logic

To revisit the intervention logic may become necessary, in case ex ante findings have not have been sufficiently robust, or not fully taken into consideration when building the RDP. Moreover, changes in the RDP context may have affected the relevance of the intervention logic with regard to the needs that had been identified at the time of

---

[27] Theory based, quantitative, qualitative, mixed approaches

programme design. Other needs may also have become more relevant at the time of RDP evaluation.

Recommended working steps:

Only if the above-mentioned changes have been identified, it is recommended to

- revisit the ex ante findings on the RDP´s intervention logic´s coherence and relevance, and
- appraise if intended direct RDP effects on beneficiaries are still valid and expected synergies between priorities at the level of the entire programme can be achieved[28] (➔ see tool to appraise various programme effects in the PART III)

Expected outcome: Revisited intervention logic

Revisit ex ante findings and indirect programme effects

Already during programme design the ex ante evaluator´s task has been amongst others to:

- examine the intervention logic´s [29] external coherence with EU policy objectives (the RDP contribution to EU 2020 strategy[30] and CAP[31]).
- the relevance in addressing the most important needs of the programme territory, deriving from the SWOT analysis and needs assessment, and
- the internal coherence between programme objectives, planned inputs (budget and forms of support), combination of measures, and expected RDP´s outputs, results and impacts.

The intervention logic should not only be appraised for direct and intended programme effects, but also for several types of other effects which may play an important role in the programme´s performance. These effects may influence the intended RDP achievements and expected results and impacts in positive or negative way. Although it will later be the task of the evaluator to net out programme effects, programme authorities need to develop an understanding of what their programme may "produce" apart from what has been planned. The RDP intervention logic or its priorities, focus areas and measures can produce indeed indirect effects[32], which may be:

- intended (identified in the context analysis) or

---

[28] This appraisal is also very important for the assessment of programme synergies and answering the common evaluation question 19: "To what extent have the synergies among priorities and focus areas enhanced the effectiveness of the RDP?"

[29] Also read Part II, Chapters 1 and 2 of the „Getting the most of your RDP: guidelines for the ex ante evaluation of 2014-2020 RDPs http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

[30] Union Strategy for smart, sustainable and inclusive growth http://ec.europa.eu/europe2020/index_en.htm

[31] http://ec.europa.eu/agriculture/cap-post-2013/

[32] Detailed description of programme effects can be found in Guidelines for ex post evaluation of 2007-2013 RDPs, Chapter 4.2.2 Challenge: identification of programme effects, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

- unintended (unforeseen in the context analysis and not flagged with the context indicator).

Indirect programmes effects to be looked at during revisiting the RDP intervention logic are:

- **Leverage effects** are the propensity for public interventions to induce private spending among direct beneficiaries (e.g. farmer receiving subsidy can better invest in property, which is not part of business). Leverage effects are usually unintended.

- **Deadweight loss effects** are changes observed in the economic, environmental or social situation of programme beneficiaries which would even have occurred without the intervention (e.g. farmers would invest anyway without subsidy later or with their own money, or they would use loans). Deadweight losses are usually unintended effects.

- **General equilibrium effects** occur when programme interventions positively or negatively affect RDP non-participants. They usually play a more important role in the evaluation of large programmes than in the evaluation of small programmes and include:

  o *Multiplier effects* resulting from increased income and consumption generated by the RDP. Multiplier effects are cumulative and take into account the fact that a part of the income generated is spent again and generates other income, and so on in several successive cycles. In environmental terms, the intended support of selected species may lead to support of other species as well. Multiplier effects are positive, often intended and expected.

  o *Displacement effects* occur in a programme area at the expense of other areas. For example, the programme support is affecting positively the employment at the expense of increasing unemployment in neighbouring areas. In environmental terms, the intended support of HNV farmland may lead to deterioration of farmland in neighbouring areas. Displacement effects might be unintended (if they cause further regional disparities) or intended (if they contribute to balancing disparities among regions).

  o *Substitution effects* are obtained in favour of direct programme beneficiaries but at the expense of units that do not qualify or participate in a given programme. For example, the irrigation support of small farmers may lead to their increasing competitiveness towards non-supported large farms. Substitutions effects can be unintended (if they support those already reach), or intended (if they contribute to balancing the socio-economic situation in the programme area)

Moreover, there are effects between measures, focus areas and priorities, which should have be taken in consideration already during the programme design and ex ante valuation and therefore be appraised, if necessary, when preparing the evaluation. These effects are:

- **Additional contributions** of measures to focus areas other than those under which they have been programmed – **secondary effects**. The legal framework asks for flagging the intended secondary effects during the programme design/ex ante

evaluation. The validity of flagging might be revisited during the preparation of evaluation again, and corrected if necessary.

- **Transverse effects,** which are horizontal effects between measures, or between focus areas or between priorities. Positive transverse effects are often called **synergies.** Transverse effects occur, if measures, focus areas or priorities are weakening or fostering each other in their effects. Transverse effects might be intended or unintended. Within the RDP there is often the scope for a policy instrument in one area to impact on another. For example, policy means in Areas of Natural Constraint can in fact enhance the performance of local firms and (possibly) lead to rural economic growth and job creation. As for synergies, the example might be advisory services which help farmers to invest in the right water management systems, which efficiency increases more with advisors´ assistance.

The table below summarises the programme effects.

	RD priority/ Impacts	RD focus area/ Results of measures and measure combinations	RD measure
RD priority/ Impacts	Transverse effects (positive = synergies and negative) Intended/unintended		
RD focus area/ Results of measures and measure combinations	Primary/secondary Direct/indirect Intended/unintended	Transverse effects (positive = synergies and negative) Intended/unintended	
RD measure	Primary/secondary Direct/indirect Intended/unintended	Primary/secondary Direct/indirect Intended/unintended	Transverse effects (positive = synergies and negative) Intended/unintended

A tool to appraise various programme effects can be found in Annex 7 (PART III of the guidelines).

Specificities of Leader

Leader is programmed under focus area 6B, where all primary effects and contributions to the hierarchy of objectives should be expected. However, Leader is implemented via local development strategies (LDS), which are in fact similar to small programmes that support a broad range of operations. This scope varies among Member States/regions: In some cases the LDS may deliver only RDP measures, or all measures listed in the Regulation (EU) No 1305/2013, or LAGs can invent their own measures. This implies that projects financed out of LDS measures will contribute to a range of focus areas, not just FA 6B (secondary effects of Leader). The *Working Document Data item list for Pillar II operations database* recognises two types of contributions of Leader projects to FA (Data items relevant for LEADER to be collected for each project):

- Predominant FA to which the project contributes
- Priority(ies)/FA(s) to which operations have additional contributions

Therefore, when appraising Leader in the RDP intervention logic, it is important to look at Leader secondary effects within the spectrum of all those focus areas where contributions via LDS are expected, confirm their validity, and appraise possible synergies or transverse effects.

Another specificity of Leader is linked to the **Leader method** expressed in 7 principles (partnership, bottom-up, multi-sector and area-based strategies, innovation, networking and cooperation) and the **Leader added value**, which goes beyond the Leader method, e.g. in the form of increased social capital or improved local governance. In many cases the Leader principles and the Leader added value are neither articulated as Leader-related programme specific objectives, nor are the mentioned as evaluation topic in the Evaluation Plan. Although in these cases they are not directly intended, they still happen and should therefore be considered when revisiting Leader in the RDP intervention logic. This helps to look at these effects (positive and negative) and to articulate them. This may lead to additional evaluation topics with respect to Leader.

Specificities of technical assistance (TA)

Technical assistance, financed at up to 4 % of the total amount of the rural development programme, is a horizontal measure which applies at the initiative of the Member State[33] and supports actions for:

- preparation, management, monitoring, evaluation, information and communication, networking, complaint resolution, and control and audit,
- reduction of the administrative burden on beneficiaries, including electronic data exchange systems, and
- reinforcement of the capacity of
  - Member State authorities and beneficiaries to administer and use the EAFRD

---

[33] Regulation (EU) No 1305/2013, Art. 51

o Relevant partners in line with the indicative areas, themes and good practices concerning how the competent authorities of the Member States may use the EAFRD to strengthen the institutional capacity of relevant partners in accordance with the legal framework and to support exchange of good practices between such partners [34].

Technical assistance is not necessarily a "visible" part of the RDP intervention logic but still supports the implementation of the RDP and contributes to the achievement of RDP objectives. However, in case more specific topics in the evaluation of actions for TA are envisaged at a given time (2017, 2019, ex post), it is considered a good practice to set up an intervention logic these topics, formulate objectives, expected outputs, results and impacts.

While re-visiting the intervention logic, it is recommended to re-check if the objective linked to the TA evaluation topic and expected effects of planned activities and budgets are still valid.

Specificities of national rural network (NRN)

The NRN is one important TA action. It groups the organisations and administrations involved in rural development with the aim: a) to increase the involvement of stakeholders in the implementation of rural development; b) to improve the quality of implementation of rural development programmes; c) to inform the broader public and potential beneficiaries on rural development policy and funding opportunities; and d) to foster innovation in agriculture, food production, forestry and rural areas [35]. The NRN, as other TA actions, constitutes a horizontal support contributing to all RDP objectives. As the NRN has got an action plan and a structure (NSU and network) it considered to be a good practice to specifically articulate the NRN intervention logic by taking into account also RDP objectives. It is moreover essential to establish a specific NRN intervention logic in case the MA included the evaluation of NRN into the Evaluation Plan or the MA/NRN intends contract a separate evaluation of NRN during or after the programming period. NRN Programmes, applied in multi-regional Member States, must be designed around a proper NRN intervention logic [36].

The practice shows that some Member States have formulated an own intervention logic for NRN already during the programme design. In these cases the intervention logic should be re-visited together with the RDP intervention logic. The same applies for the NRNPs.

In the absence of an own NRN intervention logic, it is recommended that the MA or the NSU, in co-operation with the MA, formulate the NRN intervention logic during the preparation phase of the evaluation. This requires the formulation of NRN specific

---

[34] Regulation (EU) No 1303/2013, Art. 5, 59
[35] Regulation (EU) No 1305/2013, Art. 54
[36] Getting the most of your RDP: ex ante evaluation guidelines of 2014-2020 RDPs, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

objectives, expected results and impacts (NRN added value going beyond the envisioned NRN common and programme specific objectives).

Responsibilities

- Evaluation experts within the Ministry of Agriculture (e.g. Evaluation Unit) and/or evaluators.

Relevance for ToR

- This step is not part of the ToR if the evaluation experts within the Ministry of Agriculture have the capacity to review the relevance and coherence of the intervention logic, e.g. with the help of a research institute, and in case of NRN, the experts and/or the NSU have the capacity to formulate the NRN intervention logic.

- This step is part of the ToR if the "in house" evaluation experts/NSUs do not have internal capacity to conduct this step.

Further reading

Getting the most from your RDP: Guidelines for the ex ante evaluation of 2014-2020 RDPs, PART II, Chapter 1 and 2, http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications.

Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDP, PART II, Chapter 1, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html.

Metis/WIFO/AEIDL (2014). Investment Support under the Rural Development Policy. Final Report. European Commission, Brussels.

Psaltopoulos, D., Phimister, E., Ratinger, T., Roberts, D., Skuras, D., Balamou, E., Bednarikova, Z., Espinosa, M., Gomez y Paloma, S., Mary, S., Nohel, F. and Santini, F. (2012). Ex ante Spatial Policy Impact Analysis of the Rural Development Policy in European Rural Areas.

(RURAL ECMOD), JRC Scientific and Technical Reports, Office for Official Publications of the European Communities, Luxembourg

6.2   Link intervention logic to evaluation elements

6.2.1 Define evaluation elements: evaluation questions, judgement criteria, indicators

In order to provide robust answers to evaluation questions it is important to ensure that all common and programme specific evaluation elements (evaluation questions, judgement criteria, indicators) are complete, clear and well defined. Furthermore, all terms used in these evaluation elements should be clear and understandable for the involved stakeholders.

Recommended working steps:

- **Examine if all terms used in the formulation of CEQ, their judgment criteria and common indicators are clear and understandable:** Appraise the clarity of all terms proposed in the EC Working Document: "*Common evaluation questions for rural development programmes 2014-2020*" and check if additional judgment criteria and definitions are still needed. Check if all terms used in common indicators are clear and sufficiently defined.

- **Examine if all terms used in the formulation in PSEQ, their judgment criteria, and programme specific indicators are clear and understandable**: This step is conducted if the RDP already contains PSEQ and PSI. In case the RDP contains PSEQ and judgment criteria, it is important to check if they define the PSEQ well in line with the expected RDP success. The clarity of all terms used in PSEQ and judgment criteria shall be examined and their definitions provided. Judgment criteria for PSEQ should be developed in case they are not already existing. Appraise if all terms used to formulate the programme specific indicators are clear and if they enable to provide sound definition needed to collect data. Define all terms in case of lack of clarity.

- **Develop fiches for all programme specific indicators:** In case fiches for programme specific indicators are already existing, check the clarity of all terms, e.g. with respect to measurement unit/formula, suitability of calculation methods, and accessibility of data in the required format. If this is not the case, develop fiches for all programme specific indicators.

Expected output: Revised intervention logic, indicator fiches, programme-specific indicators and evaluation questions, additional indicators (if needed)

The relation between the various evaluation elements is illustrated in the following figure.

Figure 8. Relation between evaluation elements

Source: European Evaluation Network for rural development 2007- 2013

Developing judgement criteria for evaluation questions

Judgment criteria further define evaluation questions and help to specify the expected success of the rural development policy interventions at the level of focus areas, EU objectives and other specific RDP aspects. Judgment criteria should be formulated in a clear manner and all terms used should be well defined. For CEQ, the judgment criteria are proposed in the *Working paper: Common Evaluation Questions for rural development programmes 2014-2020*[37]. Member States may decide to specify further the success of RDP interventions with additional judgment criteria for CEQs. Terms used in the common evaluation questions are usually defined in the DG AGRI Glossary[38]. If not, other official glossaries – at EU, national or regional level - can be used for this purpose.

Programme specific evaluation questions (PSEQ) are formulated either by Managing authorities (in the Evaluation Plan during the programme design or later) or by evaluators (during the RDP evaluation). Similarly to CEQ, all terms used in programme specific evaluation questions should be clear and well defined. As for CEQ, it is recommended to define judgment criteria also for PSEQ.

Table 4. Example of judgment criteria for focus area related common and programme specific evaluation questions

Common evaluation question	The success of intervention will be judged with (judgment criteria):
To what extent have RDP interventions supported innovation, cooperation and the development of the knowledge base in rural areas?[39]	Commonly proposed judgment criteria[40]: • RDP projects have been innovative and based on developed knowledge • Operational groups have been created • Variety of partners involved in EIP operational groups • Innovative actions have been implemented and disseminated by the EIP operational groups

---

[37] http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications
[38] http://ec.europa.eu/agriculture/glossary/index_en.htm
[39] Working paper: Common evaluation questions for Rural Development Programmes 2014-2020, Annex 1, CEQ No 1
[40] Working paper: Common evaluation questions for Rural Development Programmes 2014-2020, Annex 1,

Common evaluation question	The success of intervention will be judged with (judgment criteria):
	Example of additional judgment criteria (added by Member States, if needed): • The share of partners in operational groups from research and academia represents at least 50%.
Programme specific evaluation question	**The success of the intervention will be judged with (judgment criteria):**
To what extent have RDP interventions prevented youth exodus from rural areas?	• RDP projects encouraged the establishment of businesses by young people • RDP projects facilitated the employment of young people • RDP projects supported the participation of youth in rural development activities

Developing indicator fiches

Evaluation questions are answered by means of indicators. The legal acts and the CMES distinguish between:

- **Common indicators**[41] for context (describe the programme environment, facilitate the SWOT and needs assessment), output (measure the RDP outputs at measure level), results (measure RDP results at focus area level) and impact (measure RDP impacts at programme level) are elaborated in detail in various EC working documents in the form of indicator fiches[42]. Common indicators are used to answer CEQ.
- **Programme specific indicators** are developed by stakeholders in the Member States (Managing authorities, evaluators) as context, output, result and impact indicators in order to answer PSEQ. Ideally these indicators respect that RACER or SMART criteria.

Additional indicators are developed whenever common indicators cannot answer the CEQ in a satisfactory manner or do not capture all effects specified by the proposed judgment criteria.

Indicator fiches for programme specific indicators and additional indicators should be defined following the structure of those designed for the common indicators. The fiche shall provide at least:

- The name and definition of the indicator
- The link to the RDP objective/focus area, and evaluation question
- The measurement unit or formula of the indicator,
- The method for calculation
- The types, sources and accessibility of data to calculate the indicator

---

[41] Commission implementing regulation (EU) No 808/2014, Art. 14.1 b) and Annex IV.
[42] Working documents: Impact indicators, Common context indicators fiches, Target indicators for Pillar II, Complementary result indicators for Pillar II.

For common indicators the evaluation stakeholders (Managing authorities, evaluators) shall examine the indicator fiches[43], clarity of definition, measurement unit and formula, accessibility of the indicator with respect to data sources and collection, as well as the method proposed for the calculation of the indicator[44]. In case of programme specific and additional indicators, programme authorities or other evaluation stakeholders (evaluators) are responsible for the clarity of all terms used in the respective indicator fiches.

Specificities of Leader

The CMES 2014-2020 contains the following common indicators for Leader (output and target)[45]:

Indicator	Output	Target
Population covered by LAG	O18	T21
Number of LAGs selected	O19	
Number of Leader project supported	O21	
No of cooperation projects supported	O21	
No and types of project promoters	O22	
Unique identification number of LAG involved in cooperation project	O23	
No of jobs created		T23

Considering the flexibility in applying Leader in the Member States, the above mentioned common indicators might not capture all specificities of Leader and its expected/intended effects (including the added value of Leader). Therefore Member States may have developed Leader related programme specific evaluation questions and indicators in the RDP´s Indicator Plan or the Evaluation Plan. In such a case all terms used in programme specific questions and indicators should be well defined and indicators should be elaborated in the form of a fiche. If neither programme specific questions nor specific indicators for Leader have been developed but specific evaluation topics in relation to Leader are still envisioned, they may be defined at later stages (see the chapter 6.2.3 Develop RDP specific evaluation elements).

Specificities of TA

In case the Member State included in the evaluation plan evaluation topics, which relate to actions envisioned under technical assistance, these topics may be accompanied by clearly defined programme specific evaluation questions and indicators. However,

---

[43] Fiches published by the EC in WD: Impact indicators, Target indicator fiches for PIllar II, Complementary result indicator fiches for Pillar II

[44] PART III of the guidelines – Tool box provides the updated fiches for all result indicators linked to the RD focus areas and specified in the Annex IV of the Commission implementing regulation (EU) No 808/2014 (complementary result indicators and target indicators marked as result indicators)

[45] Commission implementing regulation (EU) No 808/2014, annex IV

if this is not the case, the above elements can be formulated later. Chapter 6.2.3 'Develop RDP specific evaluation elements' provides more guidance in this respect.

Specificities of NRN

A similar situation applies to common evaluation elements for NRN evaluation. There is only one common evaluation question defined: „To what extent has the national rural network contributed to achieving the objectives laid down in Regulation (EU) No 1305/2013, Art 54(2)?". This evaluation question is supported by three common output indicators[46]:

Indicator	Output
Number of thematic and analytical exchanges set up with the support of NRN	O24
Number of NRN communication tools	O25
Number of ENRD activities in which the NRN has participated	O26

Because of this limited set of common evaluation elements , Member States are well advised to define additional indicators to answer the above-mentioned common evaluation question, and possible programme specific evaluation questions.

It is evident, that programme specific evaluation questions and indicators (including result and impact indicators) should be developed for NRNPs in order to measure the expected added value. As in the case of Leader, all programme specific indicators should be developed in form of indicator fiches (see chapter 6.2.3 'Develop RDP specific evaluation elements').

Responsibilities

- Managing authorities are responsible for defining all programme specific elements mentioned in the RDP, its Evaluation plan or in other internal evaluation planning documents. In case of common evaluation questions and indicators, the EU legal acts, EC working documents, and guidelines must be taken into consideration when defining the common evaluation elements.
  a) Evaluation experts/evaluators are responsible for defining the programme specific evaluation questions and indicators developed additionally during the RDP evaluation. Additional programme specific evaluation elements are designed when those presented in the RDP are not sufficient to capture all programme effects or assess all evaluation topics.

Relevance for ToR

- This step is not part of the ToR if programme authorities or "in house evaluation experts" defined properly all common and programme specific evaluation elements either in the evaluation plan, internal evaluation planning documents or other relevant documents.

---

[46] Commission Implementing Act (EU) No 808/2014, Annex IV and WD: Data items list for Pillar II operations data base'

- This step is part of the ToR if the evaluation elements have to be completed and defined after the intervention logic appraisal (additional programme specific elements)

Further reading

Getting the most from your RDP: Guidelines for the ex ante evaluation of 2014-2020 RDPs, PART II, Chapter 3, link (to be added after revised version will be put on the web).
Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDP, PART II, Chapter 1, 2 and 3 link: http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html.
Working paper: Common Evaluation Questions for rural development programmes 2014-2020, link:http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications
Working document: Common context indicators fiches
Working document: Impact indicators
Working document: Target indicators fiches for Pillar II
Working document: Complementary result indicators for Pillar II
Working paper: Data items list for Pillar II operations database (Outputs and targets)
Working document: Rural development programing and target setting (2014-2020)

6.2.2 Check consistency of evaluation questions and indicators with RDP intervention logic

In order to answer common and programme specific evaluation questions in a correct way, their consistency with the intervention logic and indicators must be ensured. The assessment of the horizontal and vertical consistency between objectives, evaluation questions and indicators helps to judge on the use of the defined indicators to measure the achievements against objectives and to answer the evaluation questions. This step furthermore helps to identify potential gaps and to identify ways to overcome them.

Recommended working steps:

- **Re-visit the ex ante findings with respect to the consistency between RDP objectives, evaluation questions and indicators:** If gaps are identified, proceed with the following working steps.
- **Check the horizontal consistency between RDP objectives, evaluation questions and indicators:** The triangular consistency among objectives, evaluation questions, and indicators should be well established at each level of the intervention logic. Identify gaps in the ability of evaluation questions to capture RDP effects towards objective achievements and in the ability of indicators to answer evaluation questions. Propose solutions to bridge gaps. (→ Tool 1 Matrix for checking horizontal consistency at focus area level (example).
- **Examine the vertical consistency in the hierarchy of objectives, evaluation questions and indicators:** Check if the answers to lower level (related to focus area) evaluation questions can provide useful and sufficient information on

programme results in order to allow to answer higher level evaluation questions. Appraise if indicators at lower level provide sufficient evidence to conduct the analysis at higher level, e.g. if proposed common and programme specific result indicators are able to provide sufficient information to assess programme´s impacts. In case of identified gaps make suggestions how to bridge them. (→ Tool 2 Matrix for vertical consistency check (example))
Expected outcome: consistent evaluation framework

Using ex ante evaluation as starting point

As a first consistency check had already been part of the ex ante evaluation it should be used as a starting point. However, the ex ante evaluation´s findings may not have been fully considered by programme authorities and the programme context may also have changed. It is therefore considered good practice to repeat this assessment.

Figure 9. Horizontal and vertical consistency between objectives, evaluation questions and indicators
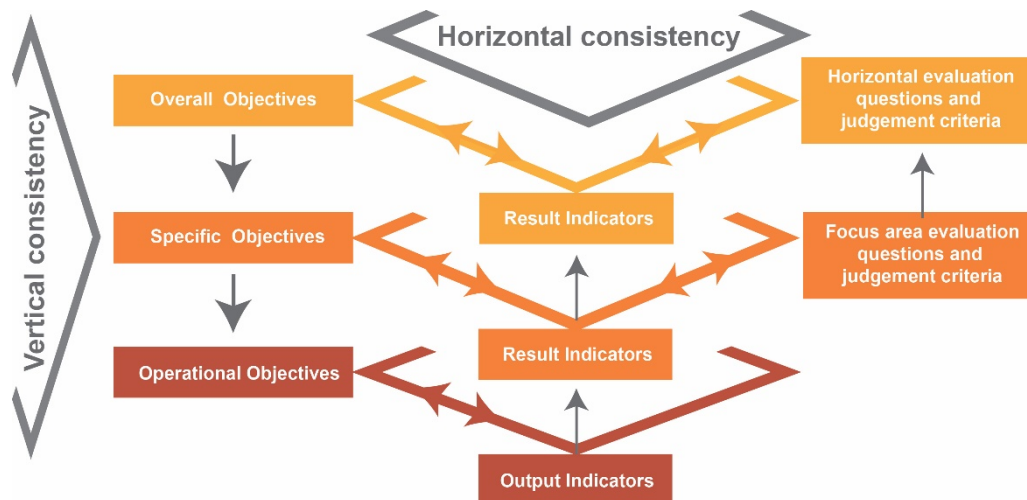
Source: European Evaluation Helpdesk for Rural Development, 2015

Checking the horizontal consistency

The horizontal consistency reflects the relation between EU and RDP policy objectives, common and programme specific evaluation questions and indicators at result and impact level and covers:

- Consistency of overall programme objectives with EU objectives related evaluation questions and impact indicators, which are used in the assessment of RDP achievements towards EU objectives and programme impacts. In case the RDP contains RDP specific overall objectives their consistency with programme specific evaluation questions and programme specific impact indicators should be examined as well.

- Consistency between RD focus areas´ objectives, focus area related evaluation questions and common result indicators. These indicators help to assess to what extent the specific objectives have been achieved within the group of programme´s beneficiaries – programme results. Any gaps in the consistency between CEQ and common indicators should be identified and bridged by proposing additional indicators (see templates for common evaluation questions in separate document). In case the RDP contains programme specific FA related objectives, programme specific evaluation questions and indicators (e.g. mentioned in the evaluation plan/indicator plan), their consistency needs to be examined as well.

When testing the horizontal consistency between objectives, evaluation questions and indicators the following guiding questions should be verified:

- To what extent do the evaluation questions allow a sound assessment of achievements of RDP objectives?

- To what extent does the evidence collected by means of indicators enable to answer the evaluation questions?

- Which programme specific evaluation questions and indicators still need to be developed to fill gaps in the consistency?

- Which additional information needs to be collected to make the additional programme specific indicators operational?

For an effective identification of potential gaps in the horizontal consistency, the following tool can be used:

Tool 1.   Matrix for checking horizontal consistency at focus area level (example)

Objective (FA related)	Common evaluation question	Judgment criteria	Indicator(s)	Identified gaps [47]	Additional indicators for filling identified gaps
6B: Fostering local development in rural areas	CEQ 17: To what extent have RDP interventions supported local development in rural areas?	Services and local infrastructure in rural areas have improved • Access to services and local infrastructure has increased in rural areas • Rural people have participated in local actions • Rural people have benefited from local actions • Employment opportunities have been created via local development strategies • Rural territory and population covered by LAGs has increased	% of rural population covered by local development strategies (FA 6B – Result indicator) • Jobs created in supported projects (Leader) (FA 6B – Result indicator) • % of rural population benefiting from improved services/ infrastructures (FA 6B – Result indicator)	Evidence collected via common indicators does not show the Leader contribution to improved services/infrastructure	• Number of projects/ initiatives supported by the Local Development Strategy • % of RDP expenditure in Leader measures with respect to total RDP expenditure

Checking the vertical consistency

The vertical consistency check follows the hierarchy of objectives and assesses in particular

- if the achievement of operational objectives leads to the achievement of specific and consequently of overall objectives.
- if answers to Evaluation Questions at lower level (related to focus areas) provide useful and sufficient information on programme results and can be used to answer evaluation questions at higher level (related to EU objectives).
- if the information collected via result indicators can be used in the assessment of impact indicators.

For the appraisal of the vertical consistency, the following guiding questions can be used:

---

[47] This is the case when the common elements are neither able to answer the relevant evaluation questions nor to be used later in the assessment of impacts

- To what extent can the evidence collected by means of common and programme specific indicators at lower/micro level (output for results and result for impacts) provide sufficient evidence to carry out an assessment of impacts at higher/macro level?
- Which gaps can be identified in the vertical set of common and programme specific indicators?
- Which indicators/additional data collection shall be proposed to fill these gaps?

To check the vertical consistency between various levels of indicators and to fill the identified gaps, the following table can be proposed:

Tool 2.   Matrix for vertical consistency check (example)

Impact indicator	Result indicators related to the impact indicator	Identified gaps	Filling the identified gaps (programme specific indicators and additional information)
I14: Rural employment rate	6A: Jobs created in supported projects (also under 6B (for LEADER))	No means to collect information on the employment rate for beneficiaries in supported sectors	Employment rate in supported sectors

Specificities of Leader

In the case of Leader, the consistency check between the RDP intervention logic and evaluation elements does not only cover focus area 6B but also those focus areas, where LDS operations are likely to contribute. If the programme contains programme specific objectives for Leader (e.g. concerning specific effects of Leader, its added value, implementation of Leader method etc.) their consistency with programme specific questions and indicators should also be checked. If programme specific evaluation questions and indicators do not exist, they can still be developed during the programming period.

Specificities of TA

If Member States included in the evaluation plan also evaluation topics which relate to actions supported by TA, such as management and administration (including the delivery mechanism), RDP communication, capacity building, etc. and these are equipped with programme specific evaluation questions and indicators, the consistency check has to cover them as well. They can be outlined already in the evaluation plan, or various evaluation stakeholders may propose them during the programming period.

In case evaluation topics linked to actions supported by technical assistance are not equipped with programme specific evaluation questions and indicators, these can be developed also later during the programming period.

As an example, the intervention logic for communication and information exchange with evaluation stakeholders and related evaluation elements is show in the below:

Figure 10. The intervention logic and evaluation elements for communication and information exchange
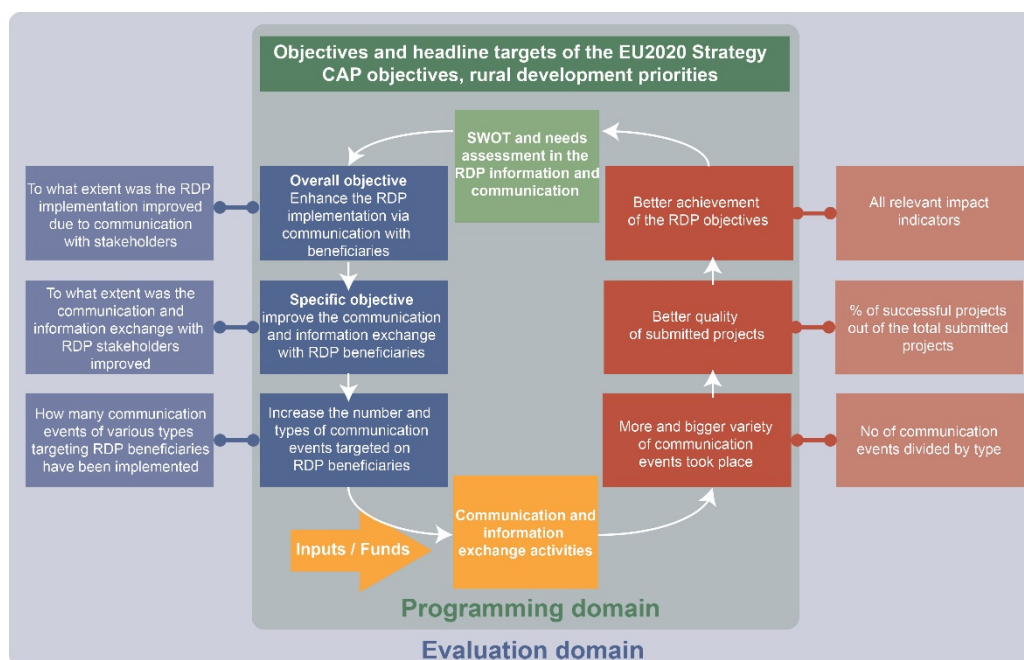
Source: European Evaluation Helpdesk for Rural Development, 2015

Specificities of NRN

In case the programme contains programme specific evaluation questions and indicators in relation to NRN, their consistency with NRN objectives should be checked as in the case of the RDP. This will be done in each case for NRNP.

If the RDP does not contain NRN specific evaluation questions and indicators, chapter 6.2.3 'Develop RDP specific evaluation elements' provides general guidance. Concrete advice with respect to NRN evaluation elements will be provided in separate NRN evaluation guidelines[48].

Responsibilities:

- Evaluation experts within the Ministry of Agriculture (e.g. Evaluation Unit) and/or
- Evaluators

Relevance for ToR:

- This step is not part of the ToR if the evaluation experts within the Ministry of Agriculture have the capacity to assess the consistency of the intervention logic with common and programmes specific evaluation questions and indicators and

---

[48] Information can be also found on: enrd.ec.europa.eu/enrd-static/evaluation/good-practices-workshops/national-rural-networks/en/national-rural-networks_en.html

if all programme specific evaluation questions and indicators are already developed and defined,

- This step is part of the ToR in case the programme specific evaluation elements are not properly developed or defined, or if the "in house" evaluation experts do not have capacity to conduct this step and define programme specific elements if needed.

Further reading

Getting the most from your RDP: Guidelines for the ex ante evaluation of 2014-2020 RDPs, PART II, Chapter 2 and 3,
Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDP, PART II, Chapter 1, 2 and 3, link: http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html
Working paper: Common Evaluation Questions for rural development programmes 2014-2020, link:http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications

6.2.3 Develop RDP specific evaluation elements

Whenever the scrutiny of the consistency between the intervention logic and evaluation elements identifies any major gaps, it is necessary to develop programme specific evaluation questions and indicators, as well as additional indicators. This will enable a correct and comprehensive assessment of programme results and impacts.

Recommended working steps

- **Development of (additional) PSEQs**: If the previous consistency check (chapter 6.2.2) between the SWOT analysis, intervention logic and evaluation elements has identified very important programme specific issues to be evaluated and not covered by already existing PSEQ, new PSEQ can be developed. (➔Tool 3 Working procedure for filling the gaps in evaluation elements using the SWOT analysis)

- **Development of (additional) programme specific indicators:** Newly developed PSEQs can determine the formulation of new programme specific indicators, to capture the primary, secondary, intended and unintended effects within the area of the newly identified evaluation topic. However, new indicators should be developed only if common and already existing programme specific indicators are not sufficient to answer the newly developed PSEQ.

- **Re-checking consistency between the complete set of evaluation elements and the intervention logic**: The consistency of newly developed programme specific evaluation elements with the RDP intervention logic shall be re- checked and verified.

Expected outcome:

- Comprehensive list of common and programme specific result and impact indicators able to capturing all direct and indirect RDP effects. This list is the

> base for identification of data needs, as well for the establishment the data management.

Additional indicators versus programme specific indicators

The current regulation requires the use of the common set of output, context, result, and impact indicators in order to answer the pre-defined common evaluation questions[49.] In case these are not fully consistent with the proposed judgment criteria, the *Working Document Common evaluation questions for RDPs 2013-2020* is suggesting to use additional information, which can be collected by means of additional indicators. Additional indicators are defined in the Member States and should not be mixed with programme specific indicators, which are used to measure programme specific effects.

Focus area related CEQ and additional indicators

Several common evaluation questions related to focus areas are linked with common result indicators which measure the percentage of all supported units under a given RD support scheme (in fact output indicators, also called target indicators). In most cases the above mentioned common evaluation questions cannot be answered with this type of indicator in a satisfactory manner. In such a case additional indicators will have to be employed already in the RDP evaluation of 2017.

Concrete advice how to develop additional indicators in answering FA related common evaluation questions can be found in (templates for CEQ, provided as separate file).

Programme specific evaluation questions and indicators

Programme specific evaluation questions and indicators are developed in order to capture programme specific effects, particularly in case the RDP contains specific objectives and priorities. PSEQ can also be developed in order to break down the rather general CEQs in order to provide more specific evidence. Moreover, programme specific evaluation questions can be formulated for Leader and specific evaluation topics presented in the evaluation plan (e.g. TA, NRN, delivery mechanisms, etc.).

Practice has shown that Member States have so far developed programme specific indicators rather than PSEQ. If there are gaps in the consistency between the RDP intervention logic and evaluation elements (see chapter 6.2.2), programme specific evaluation questions and indicators should be formulated.

Programme specific elements in RDPs

The review of approved RDPs shows that Member States often use programme specific context indicators in the description of the initial sectoral, environmental and socio-economic conditions characterizing a given programming area (see: SWOT analysis

---

[49] see: Annex IV and V of common implementing regulation No 808/2014

included in RDP document[50]), whenever the common context indicators are not sufficient to provide a realistic picture. Also, many programme specific result indicators are already formulated by the MA (lists of these indicators are available in annexes of the approved RDP). Yet, from the perspective of evaluation it may appear that these indicators are incomplete and/or internally inconsistent and have to be re-examined during the consistency check between IL and evaluation elements, adjusted and complemented, e.g. in order to assess all important programme specific effects, including potential programme specific indirect, secondary, unexpected or negative effects. In all these cases relevant PSEQs and a set of consistent PSIs have to be formulated.

Development of (additional) programme specific evaluation questions

PSEQs are linked to programme specific objectives in precise terms. The more precise the objective is, the easier is the formulation of the evaluation question and the more straightforward it becomes to link individual programme interventions to specific outcomes. In RDPs itself, programme specific objectives are not necessarily accompanied by PSEQ. When developing PSEQs the task is to identify to what extent:

- CEQ reflect the programme-specific objectives and expected effects of the intervention logic;
- CEQ, related judgement criteria and indicators will enable the evaluator to capture the full range of achievements of programme-specific objectives of the particular RDP and the programme-specific effects;

The formulation of PSEQs should:

- Reflect changes over time resulting from implementation of specific programme elements (e.g. specific programme measures) or specific characteristics of beneficiaries or programming area (e.g. age, gender, type of farms, environmental conditions, etc.).
- Specify an activity and an area of interest that can be clearly measured or observed.
- Be formulated in a manner that points to only one distinct aspect or change and establishes a clear causal relationship between the programme and a desired change in outcome (effect) (i.e. "To what extent did the change happen due to the intervention?'').
- Be clear, specific, and straightforward and should be phrased in terms of capturing the contribution of a given programme to the programme-specific objectives sought in terms of the programme results and impacts.
- Take into consideration individual programme objectives and specificities of a given programming area.

---

[50] Article ... of ... provides that RD programmes contain an analysis of the situation in terms of strengths and weaknesses (SWOT analysis) for the programming area or policy domain concerned.

It can be expected that in each individual RDP the consistency-check of its intervention logic with evaluation elements (including its linkages to the SWOT analysis) will lead to:

- A break-down of the CEQ in more specific PSEQ,
- The formulation of new PSEQs, which focus the evaluation on newly identified evaluation topics, and
- a fine-tuning of already existing programme specific evaluation questions as they work through the data design and data collection phase.

As regards the process of developing programme specific evaluation questions is recommended that the evaluator involves more evaluation stakeholders, e.g. in the form of systematic structured interviews or a workshop with programme stakeholders (including practitioners, beneficiaries, partners and policy makers). This will help to formulate relevant PSEQs by capturing basic trends in "possible" and "experienced" programme effects. By suggesting areas where the programme includes a range of unintended or uncertain effects, the evaluator and the stakeholders can draw on an earlier developed theory of change[51]. Clearly, at the first stage of evaluation, a preliminary qualitative assessment is essential because it can provide valuable insiders' perspectives and lead to the formulation of important PSEQs focused on programme performance, especially regarding its positive, negative, intended, unintended effects.

Development of (additional) programme specific indicators

PSEQs should be answered with programme specific indicators, by providing measurable information of individual aspects of the programme performance at an appropriate level of accuracy. Programme specific indicators should be developed in line with RACER criteria[52] (relevant, accepted, credible, easy, and robust) and/or SMART criteria (specific, measurable, achievable, relevant, time-bound). Programme specific indicators can be formulated as context, output, and results indicators and should be coherent with PSEQ, the IL and the SWOT analysis (see the section below). PSIs should show unintended and secondary programme effects. Their development should follow the procedures described in Chapter 6.1.

To sum up, the newly developed programme specific indicators may therefore comprise:

- result and impact indicators used to answer newly developed PSEQ, which derive from:
  - o Programme specific objectives
  - o SWOT analysis,
  - o Identified secondary and unintended effects,
  - o Breakdown of "old" PSEQ and CEQ.

---

[51] At this stage, an indicative analysis can also include a preliminary assessment of the scope of "observed" effects which at a later stage should be a subject to more rigorous quantitative verification using causal inference approaches (see Chapter … below).

[52] Technical Handbook on the monitoring and evaluation framework of the CAP 2014 - 2020

- context and result indicators enabling analysis of programme impacts.

Tool 3.   Working procedure for filling the gaps in evaluation elements using the SWOT analysis

The gaps in evaluation elements can be identified via consistency of existing PSEQs and programme specific indicators with individual aspects of the SWOT analysis included in the RDP. The SWOT analysis is based on judgments, and it is therefore subjective and qualitative by nature. The SWOT analysis can also be used for evaluation, e.g. as a tool/instrument to identify those areas where key improvements due to the programme can be expected.

The analytical procedure can be as follows:

First it is important to understand that weaknesses and strengths identified by the SWOT analysis can be controlled by the programme, i.e. the RDP is always expected to affect them. By contrast, opportunities and threats are mostly external aspects, which are usually out of the control of the RDP and are determined by its general socio/economic/environmental endowment (i.e. they usually remain unaffected by the programme).

Furthermore, it is important to keep in mind that one of the main reasons why a given RDP was installed was to:

Reduce identified sectorial, environmental and socio-economic weaknesses, and

Maintain and enforce strengths.

Second, the PSEQ can be formulated on the basis of the SWOT analysis in the form of cause- (given RDP) and effects- (results and impacts) relations, by asking: "To what extent has the RDP contributed to reduction of weaknesses stated in the SWOT analysis, and NOT affected the strengths mentioned in the SWOT analysis?"

While the first part of the PSEQ focuses on intended programme effects, the second part of the question stresses the avoidance of negative unintended effects. Given the above, answering these type of questions will most likely require the construction and use of a set of indicators, which will allow to measure the above mentioned intended and unintended effects.

In specific cases, an extended evaluation may seek to provide an answer to the question: to what extent have opportunities and threats mentioned in the SWOT analysis been affected by a given programme/measure. Yet, this will only be possible if the magnitude and the scale of the programme had been assessed as substantial from a macro-economic perspective.

Third, translate weaknesses and strengths mentioned in the SWOT into specific result and impact indicators enabling answering the PSEQs.

The SWOT analysis is usually based on the analysis of context indicators reflecting the state of the economic, social and environmental situation in a given territory prior to a RDP intervention. However, as the RDP is implemented, the performance of context indicators representing sectorial weaknesses and strengths will be affected by the programme itself as well as by other exogenous factors. While performance of specific context (and result) indicators measured at a micro-level (e.g. labour productivity in agriculture) will reflect inter alia the micro-effect of a given programme and measure, the performance of similar context (and impact) indicators, measured for a given territory (e.g. labour productivity in agriculture), can be expected to change, inter alia, due to programme impacts.

The context indicators should therefore serve two purposes:

Contribute to the identification and measurement of strengths and weaknesses within the region, as basis for the SWOT analysis, and

Help to analyse impacts achieved within the programme in light of the general economic, social, structural or environmental trends.

Context indicators showing weaknesses and strengths of rural areas targeted by the specific RDP are expected to improve, or at least not to deteriorate. The task is therefore to identify all the main weaknesses and strengths and express them in the form of relevant programme specific result and impact indicators.

Fourth, compare the list of common and already existing programme specific result and impact indicators with newly developed programme specific result and impact indicators (with indicators derived from the SWOT analysis).

Existing programme-specific indicators, already included in the RDP, should be the subject to scrutiny of their suitability to assess programme specific effects during the preparatory and structuring stage of the evaluation (Chapter 6.2).

Re-checking consistency between the complete set of evaluation elements and the intervention logic

The newly developed PSEQ and programme specific indicators extend the scope of the evaluation by asking inter alia how effective and efficient the RDP programme has been in addressing needs and priorities of a given programme area, expressed in newly identified programme specific objectives/topics. They should therefore be consistent with the IL and the already existing evaluation elements and show in detail the interactions between the priorities, focus areas and measures, on one side, and the linkages between expected programme specific outputs, results and impacts, on the other. Newly developed PSEQs and indicators allow collecting data and information to capture the primary, secondary, expected/unexpected, intended/unintended effects of the RDP interventions (see Annex 7, tool for qualitative appraisal of RDP intervention's effects, (PART III of guidelines).

The final set of common, additional and programme specific indicators for the RDP should allow to answer all common and programmes specific evaluation questions and reflect the RDP national/regional and EU priorities and objectives[53].

Specificities of Leader

Often the Leader related programme specific evaluation topics are formulated already in the Evaluation plan or at later stages during the programme implementation. Less often programme authorities formulate the same time also Leader related programme specific evaluation questions and indicators. Usually the questions and indicators, which relate to Leader evaluation topics are developed additionally either by Managing authorities or by evaluators. Similar rules, described above, are applied also in the case of formulation of Leader specific evaluation questions and indicators.

Leader specific evaluation topics may have been formulated also at LAG level in relation to LDS, but these evaluation topics are not the subject of this guidelines, which deals only with RDP elements.

---

[53] Capturing the success of your RDP: Guideliens for te ex post evalatin of 2007-2013 RDPs,

Specificities of TA

Managing authorities may decide to assess specific topics in relation to actions supported by TA. As mentioned in the previous steps, each time such a topic is identified, the programme specific objective, evaluation question and indicators should be formulated either by the managing authority itself or by evaluators selected to assess this topic.

Specificities of NRN

Often the NRN specific evaluation elements are formulated only after a decision is taken to evaluate the NRN. Once the NRN related evaluation, needs, topics and NRN objectives have been defined, the NRN specific evaluation questions are formulated as consistent with NRN related objectives and shall be further specified with judgment criteria. NRN specific indicators are formulated in consistency with the above judgment criteria. Concrete advice with respect to NRN evaluation elements will be provided in the NRN evaluation guidelines, to be published early 2016.

Responsibilities:

- Managing authorities with the support of "in house" evaluation experts may define programme specific evaluation questions and indicators in the RDP design stage and include them in the Evaluation Plan and Indicator Plan. They may also define programme specific evaluation elements at later stages during the programme implementation, at the occasion of a newly identified evaluation need or topic, which cannot be covered by common or existing programme specific evaluation questions.

- Evaluators may define additional programme specific evaluation questions and indicators, in case of:
  - Gaps in consistency between the RDP IL and common and programme specific elements (development of additional evaluation questions and indicators), or
  - Gaps in coherence of the RDP IL with the territorial and SWOT analysis, and needs assessment, or
  - Common and programme specific evaluation elements cannot fully capture RDP specific effects, or
  - Identified significant unintended IL effects, which are not covered by evaluation questions and indicators.

Relevance for ToR:

- This step is part of the ToR within the evaluators´ competence.

Further reading

Getting the most from your RDP: Guidelines for the ex ante evaluation of 2014-2020 RDPs, PART II, Chapter 2 and 3.
Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDP, PART II, Chapter 1, 2 and 3

http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html
Guidelines: Establishing and implementing the evaluation plan of 2014 - 2020 RDPs, Part II, Chapter 6, http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications
WD: Common Evaluation Questions for Rural Development Programmes 2014-2020, http://enrd.ec.europa.eu/en/evaluation-helpdesks-publications

7 STRUCTURING THE EVALUATION

In the structuring phase, the evaluation stakeholders set up the evaluation approach and establish the basics for collecting the evidence necessary to answer the evaluation questions. For this purpose, the evaluation methods are selected and combined; information needs for the assessment of common and programme specific indicators are identified; data sources are screened and provision are made in order to get data and information in the required format for RDP evaluation.

There is however no sharp borderline between preparing and structuring the evaluation. In some Member States Managing Authorities define already in the ToR a certain evaluation approach or even an evaluation method, whereas in other Member States it is up to the evaluators to propose the approach.

The decision on the selection of the evaluation approach ideally stays with the Managing Authority or evaluation experts within the ministry. In case the evaluators propose the evaluation methods, it is recommended that the Managing Authority closely follows the selection of the evaluation approach and methods. This is important also with a view to later judge on the quality of the outcomes of the evaluation.

As for conducting and reporting the evaluation in the AIR submitted in 2017 the structuring phase should be preferably focus on the legal requirement[54] linked to this particular report: "reporting and quantification of programme achievements, in particular through assessment of the complementary result indicators, and relevant evaluation questions", meaning those related to RD focus areas and other RDP aspects (operational performance, NRN and TA).

In spite the focus in these guidelines is on reporting the evaluation in 2017, it is recommended not to limit the structuring phase only to common evaluation questions No 1 – 21, but to prepare for all RDP related evaluation questions. The procedure is similar. Such an approach pays back later, when data will be needed to calculate impact indicators and answer horizontal evaluation questions – common EU objectives and overall RDP related evaluation questions at later stages (in 2019 and during the ex post).

7.1 Set up a consistent evaluation approach

7.1.1 Selection of suitable evaluation approach to the evaluation of RDP results in 2017

The decision on the evaluation approach (e.g. theory of change, quantitative, qualitative and mixed approaches) is one of the key steps in preparing the evaluation. The selected approach has to capture and quantify programme achievements, in particular through the assessment of result indicators and

---

[54] Commission implementing regulation (EU) No 808/2014, Annex VII, point 7, 1st and 2nd para

answering focus-area related questions. The approach has implications on data requirements, but also on the quality and robustness of evaluation findings.

Recommended working steps:

- **Review what the evaluation approach needs to be able to capture for the specific evaluation in 2017**. List those aspects that the evaluation approach needs to fulfil in 2017: quantification of programme achievements, assessment of programme results (direct programme effects, indirect programme effects, secondary effects, synergies and transverse effects, proportionality of analysis, assessment of counterfactual situation).

- **Review various evaluation approaches based on quality criteria:** Different evaluation approaches (theory of change, quantitative, qualitative and mixed approaches) should be critically reviewed by taking into consideration the criteria for selection of evaluation approaches. This can be done, e.g. by assigning scores (1-5) to each separate criterion (e.g. causality, selection bias, etc.) and by summing up the obtained results. Furthermore, scores might be adjusted by assigning to each of the above criteria individual weights (the latter show the "theoretical importance" of an individual criterion vs. other criterion). (Tool 4 - Criteria for selection of evaluation approach)

- **Assess adequateness of approaches for assessment of programme effectiveness, efficiency**: Under this step the MA and evaluators may consider various aspects of practicability of each evaluation approach by taking into consideration their suitability for analysis of programme effectiveness, efficiency at various levels, i.e. micro-level, regional-level, macro-level; as well as ability to provide results at various descriptive scales, i.e. nominal scale, ordinal scale and cardinal scale. Here, evaluators may assign scores to individual approaches by combining individual criteria, e.g. ability to assess programme effectiveness at micro-level using cardinal scale (score=5), etc.

- **Assess adequateness of approaches for analysing achievements of RDP objectives**: Under this step evaluators may consider other aspects of the evaluation approach's practicability by answering the question: which evaluation approach is most suitable for the analysis of common, horizontal and specific objectives of individual RDPs. Or, which approach appears as the most advantageous concerning the analysis of a particular type of objective, e.g. horizontal objective? This can also be done, e.g. by assigning scores (1-5) to each criterion mentioned above.

Expected outcome:

Decision on evaluation approach.

What needs to be captured by the evaluation approach?

The main focus of the RDP evaluation in 2017 will be the quantification of programme achievements, in particular through the assessment of result indicators and answering focus-area related questions. This quantification will reflect the real uptake of the RDP and will be based on all operations collected through the operations database by the end of 2016.

The assessment of programme results[55] (to be reported on in the AIR 2017) requires to take into consideration:

- Direct programme effects which occur at the level of programme beneficiaries (to see them involves comparison of effects between programme beneficiaries and non-beneficiaries), and
- Indirect programme effects (e.g. deadweight loss, leverage effects, etc.), which also occur at the level of programme beneficiaries; they have to be assessed on a case by case basis, depending on indicators and likelihood of indirect effects to appear.
- Secondary effects of measures to focus areas under which they are not originally programmed,
- Synergy and transverse effects between RDP measures, focus area and priorities.

More information on programme direct and indirect effects can be found in Chapter 6.1 Re-visit the RDP intervention logic. The tool to flag specific RDP effects can be found in PART III, Annex 7).

The level of RDP implementation (uptake) may affect the assessment of programme results. In case of few or no programme beneficiaries selected/treated it would be disproportionate to conduct a complex analysis.

Proportionality in the assessment of programme results (and impacts) is claimed by small and regionalised Member States. However, as regards the evaluation to be reported in 2017, proportionality is less at stake than for the later assessment of impacts. We are looking at programme effects within its context, whatever the size of the RDP may be. Yet, the approaches and methods can be simpler compared to "large" RDPs.

In 2019, proportionality can be important for the smallest programmes. In their case the assessment could be limited to gross impacts or programme results as the primary source of evidence to answer the evaluation questions related to EU level objectives.

Challenges in the assessment of programme results

Before selecting the evaluation approach a number of challenges must be addressed[56]:

- The scope of socio-economic and environmental objectives is very broad because of the heterogeneity of rural areas and their specific strengths and weaknesses;
- The programme effects have to be measured at various levels starting from single farms/firms up to the programme area and so cover relatively small areas but also the whole region or country and data may not be available;

---

[55] Assessment of impacts of a given RDP at territorial or macro level (which should be reported in AIR submitted in 2019 and ex-post evaluation) will require in addition to the above effects also consideration of other indirect programme effects (e.g. displacement effects, multiplier effects and other general equilibrium effects) occurring at the level of regions affected by a given RDP

[56] See EUROPEAN COMMISSION - Directorate-General for Agriculture and Rural Development – Unit E.4 (2014): Investment Support under Rural Development Policy. Final Report. Brussels. ISBN 978-92-79-35314-7

- The assessment should establish the causal relation between the programme and observable value of indicators (not easy since programme effects are not directly observable),

- Programmes operate in the real world, and therefore evaluation in some cases is confronted by limited data availability which may call for the use of less data-demanding methods. Although easier to implement, such methods alone are not ideal from an analytical point of view.

The main methodological challenge in the assessment of programme results (at a micro-level) in 2017 will be to answer the question: "What would have happened to the respective programme beneficiaries /area without the programme?" through providing the evidence of a true cause-and-effect link between the values of observed indicators and the RDP.

This question cannot be answered straightforwardly: It is not easy to establish whether causal relationships between the programme and the observable values of the indicators exist as also other factors (independent from the RDP) may simultaneously affect it. For example, the change in agriculture output/AWU observed for the group of programme beneficiaries can be affected by the change in input and output prices, managerial skills of beneficiary farmers, factor endowment, etc.

This causal inference problem can be solved empirically by finding an appropriate counterfactual. While finding suitable counterfactual is more easily achieved for RDP measures focused on sectorial and socio-economic effects, it proves to be more complicated for most of the environmental measures affecting biodiversity, water quality, HNV, climate change, etc. Due to the complexity and site specificity of programmes oriented preferably at the environment, the identification of control groups/areas and the establishment of a situation with and without the programme may become especially difficult.

Selection of the most advantageous evaluation approach

There are four basic evaluation approaches: a) theory of change, b) quantitative, c) qualitative and d) mixed approaches. The description of these approaches, and their applicability to assess RDP results in 2017 is summarized in the Ex-post evaluation guidelines of RDPs 2007-2013.

The evaluation practice shows that a mixed-method approach involving the integration of rigorous (e.g. quasi-experimental) quantitative and qualitative methodologies is the most adequate empirical approach to evaluate RDPs provided that sufficient data on beneficiaries can be collected. Mixed methods evaluations seek to integrate social science disciplines with quantitative (counterfactuals) and qualitative approaches to theory, data collection, data analysis and interpretation.

Mixed-methods approaches can help to develop more comprehensive evidence of programme results (and impacts) and for example can be used to distinguish between implementation failure and theory failure in the event of a lack of intended

results/impacts. The key strength of this approach is the ability to provide a triangulation of specific methods and data.

One of the main areas in the RDP evaluation where mixed-methods are especially applicable is the verification and an in-depth analysis of the main reasons for high, medium or low effectiveness of programme support. If the results from different methods converge, then inferences about the character and magnitude of these impacts will be stronger. If they diverge, mixed-methods can provide a more objective explanation of factors behind it.

The selection of a robust evaluation approach must take into account the limitations and requirements of different methods. A number of criteria should be considered:

Tool 4.   Criteria for selection of evaluation approach

Criteria	Description
Validity of results	Users of evaluations want results that are reliable, scientifically sound, robust and valid. The validity of results can only be checked with a transparent research process. Attention must be paid to restrictions to data access which may impair a high level of transparency. In order to improve the validity of results, authorities commissioning evaluation studies should put emphasis on as much transparency as possible.
Scale of measurement	Different methods produce results at different scales: quantitative methods (e.g. econometric, input-output, computational models) provide results in numbers (cardinal scales), while qualitative methods produce results on ordinal or nominal scales. When results need to be expressed in cardinal scales, the choice of methods is limited because qualitative and participatory methods and theory based and descriptive approaches allow ordinal statements at best. Quantitative methods would therefore be strongly preferred in this case.
Ability to analyse the counterfactual	A good evaluation should always develop a counterfactual. The counterfactual situation is the conceived and observed scenario that is used to compare firms/farms/territory/etc. with and without programme support. Although most methods can be used to analyse counterfactual scenarios, the challenge is to make judgements about an effect of the programme which by definition cannot be directly observed in most cases. Quantitative methods are more systematic in building counterfactuals; however, some qualitative methods can also include a counterfactual in their design.
Ability to identify and test causal relations	One of the biggest challenges in evaluation is to identify and test causal relations between the policy interventions and the outcomes. Only a small set of methods (typically econometric models) are suited to provide such results. If such results are not available, assumptions about causal effects need to be made, possibly through the application of qualitative methods.
Ability to link output and result with impact indicators	A good evaluation, even if carried out at an early stage of programme implementation, where impacts are not yet discernible, should be designed with a long-term perspective. This implies establishing mechanisms for linking outputs and results to impacts which at a later stage have to be assessed in relation to the intervention logic of the programme.
Ability to consider unintended effects	Several unintended effects are important in order to evaluate programme support. These include leverage, substitution, displacement effect and deadweight (windfall profit). Although most methods can be used to take account of leverage and deadweight effect, only econometric methods can be used to quantify their size.
Time horizon of the intervention	Many interventions, especially investments, have a period of operation that spans over many years (e.g. the planting of an orchard). The methods used to analyse the effects of such interventions are different from approaches that analyse interventions with short-term payoffs (e.g. an automatic milking system). It therefore has to be recognised that the effects of some interventions will be measurable only with a considerable delay that may go well beyond the programme period.

Responsibilities:

- The Managing Authority (e.g. in collaboration with the evaluation experts within the Ministry) will propose the evaluation approach for the assessment of the RDP. If programme authorities have relevant capacity, they can also propose evaluation methods.
- Evaluators elaborate further on the evaluation approach and evaluation methods used to accomplish the evaluation tasks.

Relevance for ToR

- This step is part of the ToR: evaluators must elaborate on the proposed evaluation approach, select and apply the most suitable evaluation methods.

Further reading

Capturing the success of your RDP: guidelines for ex poste evaluation of 2007-2013 RDPs, Chapter 4, Part II, enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

7.1.2 Select evaluation methods and their combination

Although the main responsibility for the choice and use of evaluation methods lies on evaluators, the Managing Authorities may however express a preference for certain methods in the ToR. In each case, the choice of evaluation methods should respect the criteria for conducting a high quality evaluation. The application of different methods, even using the same data set, may lead to different results, including the estimated magnitude of programme effects and even its sign (+/-). In consequence, the application of inadequate evaluation methods and techniques which are unable to control the selection bias and eliminate other systematic errors, may considerably obstruct an evidence-based policy-making.

Recommended working steps:

- **Verify the applicability of a given evaluation method in the context of a pre-selected evaluation approach**: The recommended mixed evaluation approach combines various quantitative and qualitative approaches with a theory based approach. It is crucial that at least some of the selected evaluation methods enable to assess the true programme effects and attribute observed changes to the intervention.

- **Review the ability of a method to meet basic evaluation standards**: Conducting rigorous evaluations of impacts (a mix of quantitative methods based on counterfactual analysis and credible qualitative methods) is crucial for policy learning Pre-selected methods should meet basic evaluation standards (rigour, credibility, reliability, robustness, validity, and transparency).

- **Consider budget, time and data constraints**: Early thinking on evaluation design can facilitate better budget planning, save resources and improve data quality and quantity. If this is not the case, the evaluator may face severe budget, time and data constraints, which may act as disincentives to conduct rigorous evaluations. In general, there are several options for carrying out sound evaluation under budget, time and data constraints.

- **Selection of adequate methods**: An appropriate bundle of methods can minimise potential bias and systematic errors in evaluations. It is therefore recommended to carry out additional surveys in order to collect additional data on characteristics and performance of programme beneficiaries and non-beneficiaries instead of taking the risk of overstretching the interpretation of programme results based on simple yet biased evaluation techniques.

Expected outcome: Selected evaluation methods

What should be considered in the selection of suitable evaluation methods?

In the process of selecting a suitable method, the following issues are crucial[57]:

- Users of evaluation have an interest that the results are reliable, scientifically sound, robust, and valid. When using these criteria, it is possible to classify

---

[57] See: Investment Support under RD policy, final report 2014.

methods and evaluation techniques with respect to their appropriateness to evaluate individual RD measures. Various methods meet the criteria of soundness, robustness, and validity at different levels. Each method has specific data requirements, strengths, and limitations. There is no best method with respect to the quality of all criteria used in evaluation[58]. Restriction on access of individual data may decrease the level of transparency but model specification, estimation techniques, and test statistics should be transparent even in such cases.

- Methods that can provide quantitative results are generally preferred; however, other methods may still be useful when there are restrictions in the use of quantitative methods (e.g. data gaps etc.)

- The majority of quantitative evaluation techniques are also not directly substitutable with each other. For example, a counterfactual approach that can be applied to the evaluation of programme results at micro-level cannot be replaced by another quantitative technique, e.g. Input-Output method which is applicable for the assessment of specific programme effects at regional- or macro-level.

- The results of different methods are expressed in different scales: e.g. counterfactual econometric and quantitative methods (e.g. Propensity-Score Matching[59], Input-Output or programming methods) provide results on cardinal scales, the other (e.g. qualitative) on ordinal scales or on nominal scales. When results have to be expressed in numbers (cardinal scale), the scope of methods is limited because qualitative, theory-based and descriptive approaches allow ordinal statements at best.

- One of the biggest challenges in evaluation is to identify and test causal relations between the policy intervention and the outcomes. Only a small set of methods (typically econometric counterfactual models) are suited to provide exact results in this respect. Key in testing causal relation is the development of a counterfactual in order see the effect of the programme, which cannot be directly observed. In general, there are two alternatives: results can be based on statistical evidence of randomised controlled trials or based on adequate counterfactual econometric assessments (e.g. PSM). In case causal effects cannot be identified by observations, assumptions on causality need to be made possibly through the application of qualitative methods. The latter approach captures causal relations conceptually, but does not provide robust information on the causal relationship.

- Results based on counterfactual approaches (PSM) should be tested by other methods in order to improve the validity of the results. The combination of methods contributes to the validity of results. For example, empirical evaluation studies reveal that the qualitative method can set the context and contribute to the development of hypotheses, which can subsequently be tested with quantitative methods. The results of quantitative methods can be validated and causal relationships explained in more in-depth by applying qualitative methodologies.

---

[58] In scientific literature not all types of approaches are equally well accepted and the robustness of results may be questioned if results are based on a small sample size or a model specification that has not undergone peer reviews or if results were obtained in a non-transparent manner. In scientific literature not all types of approaches are equally well accepted and the robustness of results may be questioned if results are based on a small sample size or a model specification that has not undergone peer reviews or if results were obtained in a non-transparent manner. In scientific literature not all types of approaches are equally well accepted and the robustness of results may be questioned if results are based on a small sample size or a model specification that has not undergone peer reviews or if results were obtained in a non-transparent manner.

[59] http://betterevaluation.org/evaluation-options/propensity_scores

- Indirect effects such as leverage, substitution, displacement effect and deadweight (windfall profit) are very important to consider when specific measures are evaluated. Any method can be used to take account of leverage and deadweight effect, but only counterfactual econometric methods (e.g. PSM) can be used to quantify their size.

- Qualitative methods can be very helpful in developing hypotheses that can be tested or further explored with quantitative methods. For example, qualitative methods should be utilized to explore the range of possible RDP unintended and indirect effects. The magnitude, scale and importance of the above effects should however be verified by using quantitative approaches.

- Qualitative and quantitative methods are complementary and should not be regarded as substitutes. Quantitative methods are also complementary in many respects because some important results (e.g. deadweight, leverage or multiplier) can only be estimated using specific method/techniques.

- Given the above, it is advisable that the evaluation team chooses an evaluation design not based on a single method but on a bundle of evaluation methods and techniques which should be internally consistent and complementary to each other. Moreover, all selected methods should meet the basic evaluation quality criteria, i.e. rigour, reliability, robustness, transparency, validity, and practicability. The evaluator should try to select a strong evaluation design (consisting of a bundle of evaluation techniques and methods), bearing in mind time, money and practicability constraints (including data availability)[60].

Specificities of Leader

Leader has to be seen as a multidimensional evaluation topic.

Leader is evaluated at the RDP level as the measure programmed under the FA 6B, but contributing to a range of RDP focus areas and priorities through local development strategies. However Leader is not only the subsidy, which is delivered via the local development strategy, but it is also the method which helps to create and develop partnerships (by integrating various sectors and areas of rural life), to foster participation of local people in development (bottom-up), innovation, cooperation and networking. Furthermore, Leader can bring additional value, such as increased social capital, enhanced utilisation of local resources, improved local governance etc. Leader is also part of CLLD where several funds work together towards achieving EU level objectives. All these aspects and others may become evaluation topics at national level.

Leader should be evaluated at local/LAG level. Here again the evaluation is multidimensional, covering the evaluation of LDS, assessment of the application of 7 principles and Leader added value for rural areas.

---

[60] In view of the above constraints considered as disincentives to conduct rigorous evaluations, evaluators tend to simplify evaluation design by using so called "naïve" approaches. However, by doing this, one has to be aware that there are at least two common evaluation designs, which do not qualify as sound impact evaluation designs. These are: a) pre- and post-intervention comparison of programme beneficiaries, and b) post intervention analysis of programme beneficiaries without a control group. see: World Bank, 2006, 2010.

Figure 11. Leader as an evaluation topic.
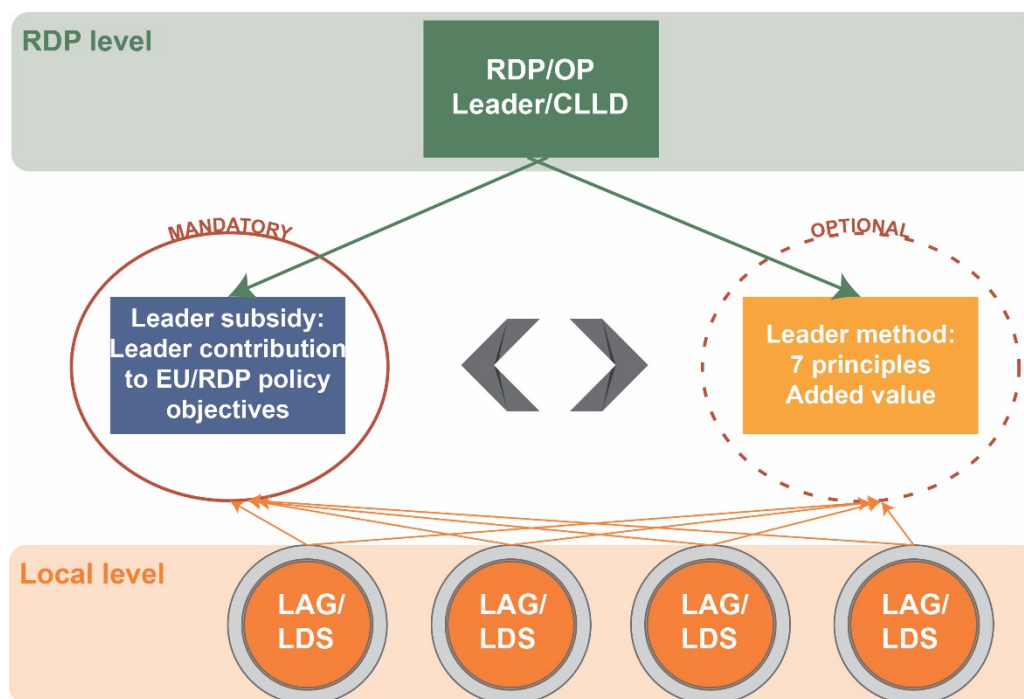
Source: European Evaluation Helpdesk for Rural Development, 2015

The complexity of Leader as an evaluation topic requires a variety of quantitative and qualitative evaluation methods and combinations thereof. To capture the effects of the Leader measure (subsidy) quantitative approaches will be used. In this respect Leader operations and their contributions to focus areas are assessed together with other RDP operations by means of result indicators. Qualitative methods are used to triangulate quantitative findings. In case of other evaluation topics such as the delivery of the Leader method or the Leader added value, most qualitative approach will be more likely applied. It is important to recognise that the counterfactual is also a suitable tool to see programme effects caused by Leader and can be used in quantitative, qualitative and mixed approaches. In countries, where only parts of territory are covered by LAGs, control groups could be selected within the same region with similar characteristics as LAGs. In "Leader mainstream" countries (the whole territory is covered by LAGs), LAGs with similar characteristics and different types of operations could be compared when assessing RDP Leader effects.

Specificities of TA and NRN

Specific interventions such as TA and NRN often require specific considerations in the selection of both the evaluation approach and methods. Although mixed approaches are proposed, in the majority of cases qualitative methods will be applied (focus groups, interviews, case studies etc.), accompanied by the assessment of monitoring data. More information is provided in this respect in the NRN evaluation guidelines.

Responsibilities:

- Managing authorities with the help of evaluation experts within the Ministry of Agriculture may express their preference for certain evaluation methods, which in the ToR.

Relevance for ToR:

- This step is fully part of the ToR.

Further reading

Capturing the success of your RDP: guidelines for ex post evaluation of 2007-2013 RDPs, Chapter 4, Part II, enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

7.2   Establish the evidence for evaluation

7.2.1 Review data requirements for the chosen methods

High quality and timely available data is essential in order to apply suitable methods to accomplish all required evaluation tasks and to obtain robust evaluation findings. While different methods have different data needs, the evaluator also has to develop strategies to overcome data gaps.

Recommended working steps

- **Check the availability of data for all common and programme specific indicators and of financial data:** Identify data gaps (quality and availability). In case the data for certain common indicators are missing and cannot be obtained in a cost effective way, proxy indicators should be used.

- **Check the availability of data to calculate target values of indicators:** Examine if the data to calculate target values is in place. Some expected values, e.g., expected leverage of investments, values should be set after specific research work has been carried out.

- **Decide which data is needed to calculate the net values of result (in 2017) and impact (2019) indicators and check their availability**: In this step, the data to conduct the counterfactual, calculate net values of indicators in line with selected methods, and additional data needed to answer the evaluation questions, should be identified. This contains also the data for control groups, relevant sector data or any type of data which is needed for applying selected evaluation methods. Existing sources/databases of required data should be identified and assessed for their suitability in RDP evaluation.

- **Make necessary arrangement for ensuring data availability:** Existing data might not be in the format required for the RDP assessment. Therefore it is necessary to make arrangements to obtain data for evaluation from existing databases, e.g. contract data providers, agree on data sets in specific format, apply legal procedure if necessary etc.

- **Decide on additional data to be primarily collected:** In case existing databases do not satisfy the evaluation needs, data shall be collected primarily. Data for beneficiaries can be collected via the existing monitoring system. For control groups surveys can be carried out selecting the samples in line with the chosen evaluation method.

- **Propose adjustments of the monitoring system in order to make it better fit for evaluation:** A monitoring system, which collects data on beneficiaries also for the purpose of evaluation is a cost effective tool to obtain data even for common and programme specific result and impact indicators. As such, it can be use as source for counterfactual analysis providing data in suitable format and time.

Expected outcome: inventory of existing data sources, data gap analysis; arrangements for data provision and access,

Evaluation tasks and data needs

Data requirements for evaluation are linked to evaluation tasks and evaluation needs, which are expressed in the form of evaluation questions answered by means of indicators. As such, data requirements are rooted in the intervention logic of the programme and its objectives, which address the "why" and "how" of the policy. The "why" describes the policy need and, to some extent, sets the quantified targets of the policy/measure. The "how" addresses the way by which instruments and resources will be used to achieve targets.

The following table presents the evaluation tasks to be reported in the AIR to be submitted in 2017[61] and respective data requirements.

Table 5. Evaluation task and data needs in 2017

Evaluation tasks	Types of data needed Relevant Data Needs	Time to collect	Legal Base
Assessment of the information and progress towards achieving the objectives of the programme: reporting and quantification of programme achievements, in particular through the assessment of the complementary result indicators and relevant evaluation questions	Data to calculate net values of result indicators Additional data needed to answer FA evaluation questions	2017	Commission Implementing Regulation (EU) No 808/2014, Annex VII, point 7

---

[61] Comission Implementing Regulation (EU) No 808/2014, Annex VII, point 7

Types of data in RDP evaluation

The following table provides an overview of the different types of data used in the evaluation of RDPs.

Table 6. Types of data used in rural development evaluation

Data	Sources of available data sets	Provider	Availability	Necessary arrangements
Primary data on beneficiaries	Operations database/monitoring	PA	Yes without additional costs	None
	Surveys	Evaluators	Yes with additional costs	Evaluator prepares and conducts the survey
	Income statements	Tax office	Yes, with additional costs	Agreement with the tax office
Primary data on non-beneficiaries	FADN (ensuring anonymity)	MS responsible body	Yes	None
	National/regional statistics (ensuring anonymity)	Statistical office	Yes, no additional costs	Request for data sets
	Surveys	Evaluators	Yes, with additional costs	Evaluator prepares and conducts the survey
Secondary data on beneficiaries and non-beneficiaries	FADN	EU	Yes	None
	National/regional statistics	Statistical office	Yes, no additional costs	Request for data sets
	Surveys	Evaluators	Yes, with additional costs	Evaluator prepares and conducts the survey
Other data	Research reports	Research institutes	Yes, with additional costs	Contract with the research institute
	Other reports	NGO	Yes, with additional costs	Contract with the NGO

Evaluation methods and data needs

The evaluation approach and methods chosen have consequences on data needs, as the data availability has implications on the selection of methods. The following table provides an overview of data requirements in relation to selected methods.

Table 7. Methods and related data needs (example)

Method	Data needs	Existing data sources/providers (links is possible)	Arrangements to ensure access (if needed	Data gaps /collection by evaluator
Quantitative				
Micro Methods				
Counterfactual surveys	The micro-level data may consist of bookkeeping data, survey data or both. Collected data should clearly identify programme beneficiaries and the level of support they received from individual RD measures (e.g. measure X, etc.). Ideally, micro-level data panel should comprise no less than 150 beneficiaries enterprises, farms or holdings for each of analysed measure (or a group of measures) and 2-3 times more non-beneficiaries.	Data at micro level (from beneficiaries and non-beneficiaries) should be collected on the basis of secondary data and/or own surveys. The FADN database combined with an anonymous data from the PA can be used for this purpose. http://ec.europa.eu/agriculture/ricaprod/	Arrangements carried out with Liaison Agencies: http://ec.europa.eu/agriculture/ricaprod/liaisonagency_en.cfm	Data on support from other programmes from the relevant authorities or the single Paying Agency.
Tracer surveys	Tracer studies address businesses, workers and trainees and attempt to trace through time the changing status of the survey's subject and record the time span of staying in a specific status.	Primary data collection	No access rules applied	All data to be collected by evaluator.
Macro Methods				
Input-Output (I-O) and related techniques	I-O data needs include a regional I-O model and RDP expenditure distinguished into demand for the product of model economic sectors and data on the change of productive capacity attributed to RDP projects. In the case of a SAM, data needs include a regional (rural) I-O model, data specific to disaggregation of economic activities, production and	National I-O tables from Eurostat. http://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/data/database	No access rules applied	National account, Consumer Expenditure Surveys and FADN data to complement the I-O and produce the SAM

Method	Data needs	Existing data sources/providers (links is possible)	Arrangements to ensure access (if needed	Data gaps /collection by evaluator
	households, and inter-institutional and factor-institution flows.			
CGE models	For Recursive Dynamic CGE model, data needs are more specific to the construction of the model rather than the RDP model input.	Does not apply, unless ready to use CGEs are used like the GTAP of Purdue University at: https://www.gtap.agecon.purdue.edu/	No access rules applied	I-O Table, National accounts, Consumer Expenditure Survey, Labour Force Survey and FADN data
Mostly Qualitative				
Focus groups and elite interviews	Focus groups and elite interviews are used when the evaluator aims at recording informed opinions by key stakeholders either through common participation and interaction (focus group) or individually (elite interview).	Does not apply	No access rules applied	All data to be collected by evaluator.
Mixed				
Quality of Life and of Local Governance, and Satisfaction Surveys	Surveys targeting quality of life in rural areas mainly refer to the provision, access and use of services and utilities. Local governance studies and life satisfaction surveys may also support the quality of life evaluation.	There are existing roll on surveys at a global as well as European scale. The problem is that they rarely distinguish between rural and urban areas and, of course, it is difficult to isolate net effects.	Eurostat SILK rules apply for anonymized data but the country-level samples are relatively small. The OECD does not provide access to micro data.	All data to be collected by evaluator.

Addressing data gaps

Data gaps may be caused by various reasons: no systematic data collection, data only available for limited levels (e.g. national, not regional), missing data-sets on certain context indicators (e.g. environmental indicators).

There are different ways in order to address data gaps. Typical approaches include additional research and surveys, use of additional databases (e.g. national, regional statistics) or a more intensive use of qualitative methods as an alternative to hard data. However, the most effective and proactive approach to deal with data gaps is to construct baseline databases and establish the necessary data infrastructure from early stages of programme implementation /development. If not defined otherwise data are collected primarily by evaluators.

The requirements for evaluation in the current programming period underline the need for Member States to put in place the necessary procedures for data collection and management for evaluations[62]. Managing Authorities are well advised to incorporate any data needs related to programme-specific elements in their monitoring and evaluation systems as early as possible.

Costs

To collect additional data is costly, therefore it is important that stakeholders (mainly MA) know well all the existing data sources at EU, national and regional level and are able to provide arrangements to bring them to the required format. The need for of additionally collected data by evaluators is affecting the evaluation budget.

Role of stakeholders in ensuring data availability

Several stakeholders should be involved in ensuring the availability of data for RDP evaluation. The communication among them should be ensured mainly by the MA or the by the evaluation steering group, if established. The following table offers an overview of how these stakeholders can support data availability and quality.

Table 8. Evaluation stakeholders and their role in ensuring data availability and quality[63]

Stakeholder	Description of role in the data management and collection
MA	Draw up a detailed evaluation plan with the monitoring and evaluation system
	Ensure a secure electronic system and a monitoring system for all common and programme specific indicators
	Establish appropriate data infrastructure, including the necessary resources for collecting, storing and updating data
	Establish clear and transparent data collection and management procedures

---

[62] Regulation (EU) No 1303/2013, Article 54 (2) and Regulation (EU) No 1305/2013, Article 76 (2)
[63] Capturing the success of your RDP: Guidelines for the ex post evaluation of 2007-2013 RDPs,http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

Stakeholder	Description of role in the data management and collection
	Screen all available data sources, communicate with providers, and make arrangements that data are available in format and quality needed for the RDP evaluation
	Ensure also communication among data providers
	Establish the processes for ensuring timely implementation of evaluations and reporting on them
	Design a good concept note and the ToR taking into account the available budget for evaluation, the evaluation scope and the preferred evaluation approach
PA	Establish clear and transparent procedures for the recording of monitoring information on applications, supported projects, payments and control
	Set up simple, transparent and fast procedures for the MA's and the evaluator's access to data
	Establish a common data system to facilitate access to data or an interface to facilitate the transfer and handling of data between the PA and MA
	Ensure link to other databases, if required
Evaluation Steering Group (if established, otherwise MA)	Facilitate and coordinate stakeholders consultation in evaluation
	Establish a process for checking and ensuring the relevance of M&E activities to programme needs
	Ensure a Steering Group composition of experts in evaluation and rural development and experts who can provide advice on data availability, information and relevant contacts to evaluators
	Potentially involve beneficiaries in the Steering Group to facilitate access to micro data and contacts of evaluators with beneficiaries
Working Groups (if established)	Ensure that working groups are composed of thematic and specialist experts so they can advise evaluators on specific sectoral data availability (e.g. environmental issues such as water protection or nature conservation) and other data issues (e.g. on Leader delivery)
	Establish TWGs according to the needs of each evaluation in order to enhance its quality
LAGs	Establish processes for bringing the LAG local knowledge and contacts to the service of the evaluation
	Establish processes for feeding in self-evaluations and the outcomes of local development strategies into the evaluation
	Encourage the participation of LAG members in evaluation steering groups
NRNs	Establish an effective and transparent system for disseminating evaluation results
	Facilitate contacts and provide advice to evaluators on alternative data sources and other data requirements
	Establish mechanisms to transfer knowledge on M&E from one country or region to another, including the provision of capacity building to evaluators
	Develop a system for the establishment of regional proxies when only national data is available for indicators
Regional governments and agencies	Establish a process of communication with the MA and PA to provide help in addressing any data requirements
	Adapt their data collection systems to the monitoring and data collection requirements of RDPs and feed into or build an interface with the MA systems
Data providers	Ensure early involvement of data providers (national statistical office, relevant ministries, research institutes, etc.) in M&E from the planning stage

Stakeholder	Description of role in the data management and collection
	Provide data of relevance to RDPs and the results of research on relevant topics to evaluators
	Provide expert knowledge and even collect specific monitoring data for the MA (possibly on a contractual basis)
	Encourage the participation of data providers in evaluation steering groups and/or the MC
Evaluators	Ensure capacities and expertise appropriate to the evaluation type and topic
	Set up evaluation teams with capacities to implement the necessary evaluation methods (quantitative, qualitative)
	Establish a continuous interface with the MA for the timely collection of data, the identification of data gaps and solutions to bridge them
	Evaluators often have a lot of past experience with data collection and access to different data sources. Such experience can be brought into the evaluation teams to address any difficulties to access data.

Specificities of Leader

Data on operations implemented through LDS are collected in the operations database as any other RDP output data (also for target indicators of the focus area where Leader operations are contributing to with secondary effects). The challenge will be to collect data which relate to complementary result indicators or additional indicators used to answer evaluation questions related to FA, where LDS is contributing. Here a sampling will need to be applied and the data will be collected through surveys on LDS beneficiaries at LAG level.

In case of a LDS self-standing situation the LAG should develop feasible indicators and ensure the data collection itself.

Specificities of TA

It is expected that mainly qualitative methods will be used in the evaluation of TA actions. This will require to collect mainly information from RDP stakeholders using qualitative tools, such as interviews, surveys, focus groups etc. Data on accomplished actions divided by type and information needed in the assessment may be collected by programme authorities and used by the evaluator in addition to qualitative information.

Specificities of NRN

In case of NRN, the data on activities implemented via the action plan, will be used from the monitoring system. The evaluator will have to use a sampling in order to obtain data for NRN specific result indicators. In addition, it is expected that qualitative information will be collected and analysed using qualitative methods. More information will be provided in this respect in the NRN evaluation guidelines.

Responsibilities

- Responsibilities are described in detail in the table above.

Relevance for ToR

- This step is part of the ToR and reflects the responsibilities of evaluators

7.2.2 Managing and collecting data for evaluation

The fundamental condition of effective data management is the establishment of functioning and solid databases, which allow to collect data of high quality and timely availability. For each evaluation exercise it is important to create a database, which starts with the baseline data concerning all common and programme specific indicators (context, output, results, including the complementary results indicators, targets, and financial data) on both – RDP beneficiaries and their control groups. , The baseline database needs to be linked and harmonised with the monitoring system (operations database), which collects both - monitoring and evaluation data on beneficiaries.

Recommended working steps:

- **Allocate clear data collection responsibilities in order to clarify who is responsible to collect data for various types of indicators** (Context, Output, Result-Target, Complementary Result, and Impact):
  - collecting primary disaggregated data at application level, or through the payment request after project finalisation
  - transmitting or uploading disaggregated data at pre-defined time and frequency
  - inspecting data for abnormal and missing values (quality control)
  - aggregating data according to spatial units and the Member State level
  - storing data

- **Allocate clear data retrieval responsibilities:** Many indicators (e.g. impact indicators) demand data from Eurostat supported by national statistics. Other evaluation exercises may require anonymized data from European wide Eurostat surveys, e.g., the FADN in order to establish counterfactuals or a simple sampling frame. At this stage, the MA should consult sources from Eurostat and national statistics:
  - retrieving data aggregated or disaggregated at pre-defined time and frequency;
  - inspecting data for abnormal and missing values and communicating with the relevant statistical authorities;
  - storing data and produce a metafile with the information concerning data.

- **Decide on the necessary primary surveys to be conducted by the evaluator:** At this stage, the evaluator should have a clear idea of which of the data that are not collected via databases but are necessary to be collected. There will be secondary or unintended effects, which may be judged as non-important and

for which a decision not to collect data is made. Field surveys (of any kind) are expensive and time consuming so, make sure that the data under consideration cannot be retrieved from statistical sources and you cannot utilize a secondary analysis of recently collected data. At this stage you should decide on:

- o The scope and objectives of the survey
- o The type of the survey (qualitative-quantitative) taking into account the decided evaluation methodology
- o Statistical data which may serve as a sampling frame (e.g., FADN)
- o Opportunities to merge surveys and minimize the time and cost

- **Plan field surveys**: Field surveys should be planned as early as possible because many of the required baseline data may be retrieved at the time of application.

- **Plan qualitative exercises:** Most qualitative survey exercises are based on key actors and informed agents. During the time of the programme's public consultation or the implementation of its various measures, the relevant authorities may have spotted key persons to be included in focus groups, experts to be conducted via in-depth interviews or simply members of key stakeholders that have followed the evolution of the programme and thus can express an informed opinion. All these contact information should be kept in database of "potential" participants to qualitative studies.

- **Monitor the database:** The database will required the cooperation among local and regional authorities, various agencies and the central Managing Authority/ Payment Agency. The flow of data that are regularly collected or retrieved, and the data from surveys should be monitored with milestones starting from the time an AIR should be delivered and going backwards. Close monitoring will flag delays, inconsistencies and confusion or responsibilities very early in the data collection process. For each indicator and for each planned survey there must be an electronic system that will inform and alert the evaluators of any delays or abnormalities. Thus there will be enough time to take a corrective action either as concerns the flow of data or their quality.

- **Create and maintain a dynamic database:** Database management and maintenance including security of personal data is of the outmost importance. The evaluators should have an access to this database limited with the competence of the given RDP and liaise with database managers.

Expected outcome: databases and data management for evaluation

Databases as the ground for proper data management

The evaluator should examine if a specific evaluation task can be accomplished by making use of the existing data sources, or if additional data needs to be collected. Additional data might be retrieved from other existing databases, may be collected primarily for the purposes of the evaluation, or may be obtained in the combination of both.

For example, the calculation of certain impact indicators calls for a dynamic link with Eurostat as well as national databases. Finally, the response to certain evaluation

questions and definitely a response to the question of RDP' effects may require targeted evaluation research based on counterfactuals (net effects) or macroeconomic modelling (economy wide effects). This calls for a well-planned and tightly monitored system of data collection, data retrieval and survey implementation sets a framework of a dynamic and evolving data management system.

The baseline database should be solid, but at the same time allow for adjustments which may become necessary due to changing data needs across the programming period. For example, programme specific indicators could be modified or newly added at later stages of programme implementation. This also means to identify and collect more types of data through existing databases and to have a well- functioning and flexible data management system.

Data primarily collected for the evaluation or accessed from existing databases together with the baseline database form the dynamic database, which will feed the evaluation process with appropriate data. In the dynamic database, data will be collected across the time respecting a panel of indicators. This database will allow also adding data in order to facilitate evaluation of RDP indirect and unintended effects and provide links with other databases, such as GISs, national/regional statistics etc. to respond to more complex data queries. The following figure provides a decision-making flow for data requirements respecting the purpose of the evaluation.

Figure 12. Decision making process for data requirements and data management.
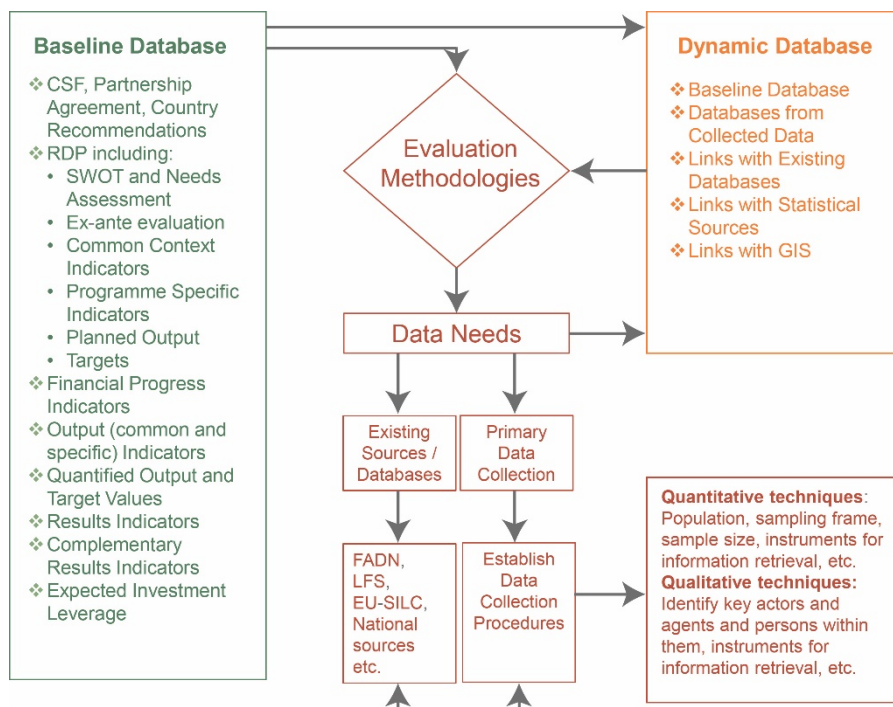
Source: European Evaluation Helpdesk for rural development, 2015

Responsibilities of stakeholders and the ToR

- Responsibilities of stakeholders are described in detail in the table above.

Relevance for ToR :

- This step is part of the ToR and reflects the responsibilities of evaluators.

Further reading

Commission Implementing Regulation (EU) No 808/2014, http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32014R0808
FADN Database (Eurostat),
http://ec.europa.eu/agriculture/ricaprod/database/database_en.cfm
National I/O Tables Database (Eurostat), http://ec.europa.eu/eurostat/web/esa-supply-use-input-tables
GTAP Database, Purdue University, https://www.gtap.agecon.purdue.edu/
EU SILC Database on subjective well-being (Eurostat), http://ec.europa.eu/eurostat/statistics-explained/index.php/Quality_of_life_in_Europe_-_facts_and_views_-_overall_life_satisfaction
OECD "How is Life" Database, http://www.oecd.org/statistics/how-s-life-23089679.htm
Gallup-Healthways Global Well-Being Index, **gallup-healthways global well-being index 2015**

8   CONDUCTING THE EVALUATION

While the preparation and structuring phase of the evaluation are primarily led by the Managing Authority (sometimes with the help of an evaluator) the conducting-phase is mainly carried out by the externally contracted evaluator. This chapter of the guidelines discusses the evaluation to be conducted in 2017 and onwards, in order to comply with evaluation tasks required by legal acts.

Figure 13. Conducting phase within evaluation process
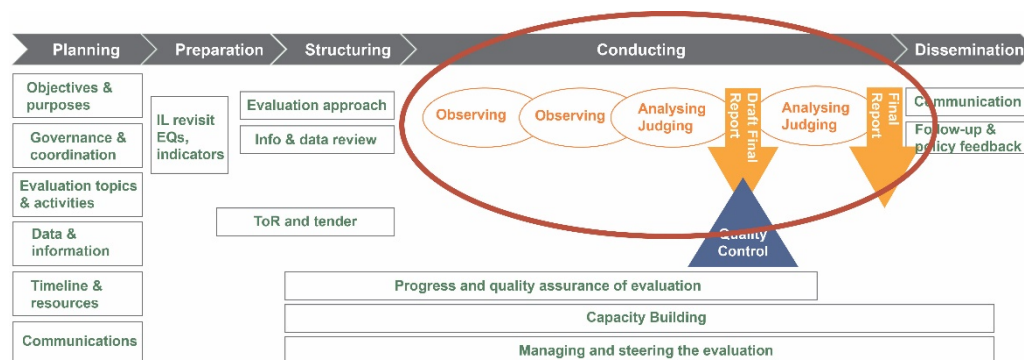
Source: European Evaluation Helpdesk for rural development, 2015

The conducting phase in this figure includes 3 of the 4 classical steps of an evaluation: observing, analysing and judging (whereas structuring has been already explained in the previous chapter). The evaluation steps can be further broken down into sequences, separated by different interim deliverables (Interim Report, Draft Final Report etc.). The ultimate aim is that the evaluation is carried out based on good quality data, appropriate observations, an effective analysis of the policy results and impacts (through sound answers to the evaluation questions), and accurate judgements, conclusions and recommendations.

Figure 14. Evaluation tasks for 2017 and beyond
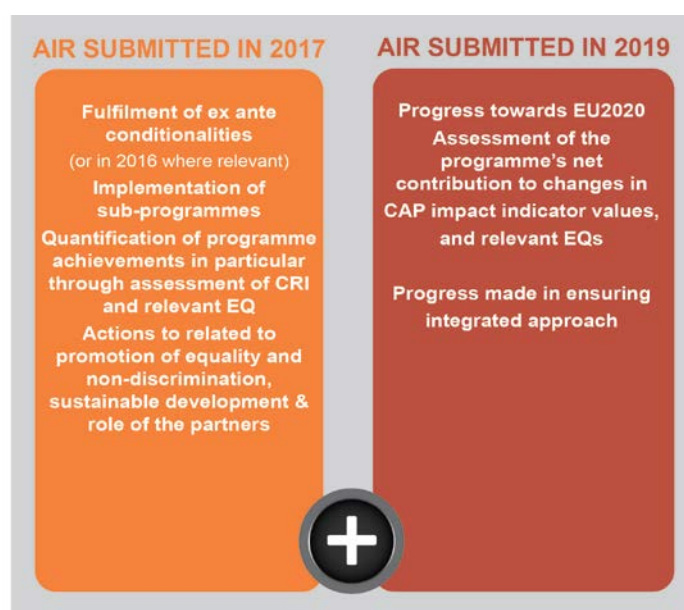
Source: European Evaluation Helpdesk for rural development, 2015

The setting up of the data management system, which allows to collect data on beneficiaries and non-beneficiaries, was described in previous chapters of PART II of the guidelines.

8.1.1 Observing

In this phase, the evaluators must proceed with the collection of relevant information and data. The critical point will be the definition of tools and approaches for gathering the information to bridge any identified data gaps. The data management system should provide the evaluators with the necessary data. Evaluators are responsible to link monitoring data, data from other sources[64] and primarily collected data, employing various tools and techniques.

Recommended working steps:

- **Create the tools needed for the quantitative and qualitative analysis**. Create the tools needed for the quantitative and qualitative analysis: interview guides, questionnaires, queries for extractions from databases, requests for maps, guidelines for case studies, and any other data collection instrument that the contractor deems appropriate
- **Collect data and qualitative information needed for answering evaluation question:** databases, studies, people to be interviewed, appropriate case study areas etc.

---

[64] E.g. Farm Bird Index (FBI), Farm Accountancy Data Network (FADN).

•	**Description of the process of programme implementation**, composition of programmes, priorities and target levels, budget
•	**Summary of the outputs:** Evaluators present the aggregated and quality-checked values for each of the common and programme specific indicators.

Expected outcome: Data and information for evaluation

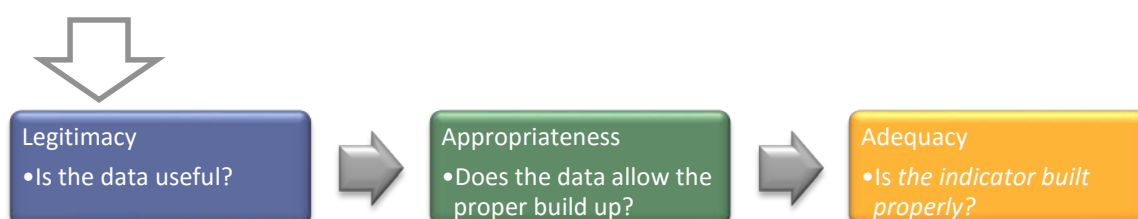### Create the tools needed for the quantitative and qualitative analysis.

During the observing phase, the evaluators use the data files to calculate values for common and programme specific indicators in line with the chosen evaluation methods. This process, however, can only happen if data on beneficiaries exists, in other words the observation phase cannot be completed fully without sufficient RDP uptake!

As the final objective is to answer all evaluation question evaluators should avoid to collect unnecessary data. Generally, the choice of data should be guided by the relevance of the variables for the assessment of achievements and the costs resulting from the planned investigation. For this reason it is necessary to coordinate planning of data-gathering along with data-processing.

Many Member States have recently established more coordinated procedures for data management for the new programmes. In these procedures, both MAs and, evaluation advisors have worked together on data management schemes. The workability of these databases should to be appraised before real data are integrated. Glitches should be addressed as a matter of urgency.

A separate document will contain templates for answering common evaluation questions no 1 – 21. These templates provide among others also the guidance with respect to types and volume of data to be collected for assessment of indicators and answering tee evaluation questions.

Evaluation objectives

Legitimacy	Appropriateness	Adequacy
•Is the data useful?	•Does the data allow the proper build up?	•Is *the indicator built properly?*

Quantitative data

In the observing phase data are collected for the following groups of indicators:

Context indicators (macro-data for RDP level)	• Member States should collect regularly data to quantify the common context indicators CCI (including impact indicators, which are part of the set, proxies agreed with the Commission and programme specific context indicators that have been used in the description of the programme area). • Data collected for CCI are macro-level data (RDP, regional or national levels) and are used in the assessment of programme impacts (this includes the netting out the respective impact indicators). • Data for context indicators are collected from EU level databases[65] or national/regional databases, on an annual basis, independently of the programme uptake.
Input and output indicators	• Micro-data, which is collected at the unit level (farm, community, business) for beneficiaries. It is automatically collected via the monitoring system (operations database). • Input indicators are linked to the RDP financial allocation and payments for individual measures, activities and operations). • Output indicators are linked to the basic description of the RDP implementation (number of beneficiaries, number of supported operations, …) • Also used to calculate target values for target indicators. • Application forms, payment request, standardised monitoring tables and other formats are used to collect these data and include them in the monitoring system. The structure of these forms and the decision on which data shall be collected (monitoring and evaluation) was part of the evaluation preparation and structuring.
Result indicators (micro-data at unit level)	• Data for result indicators are micro-data, which are collected at the unit level (farm, community, business) for beneficiaries and non-beneficiaries. This data, if properly standardized, can be aggregated to get gross values. • The operations database can be used to collect data for result indicators on beneficiaries, however there is always the need for additional data to be collected by the evaluator: data on non-beneficiaries and data that covers the period after the implementation of the project. • In line with requirements of specific evaluation methods, the data at micro level should be collected on the basis of secondary data (if available) and of surveys. • The earlier the collection of data starts in the RDP implementation, the better for ensuring the data timelines. • This includes also secondary effects and synergies. For practical reasons, it is important to automatize this system as much as possible, and ensure the electronic submission of above mentioned forms. In optimal case all of this information is within a dynamic database, described in the chapter 7.2

---

[65] Commission Implementing regulation No 808/2014, Annex IV, Working document. Context indicators fiches, Context indicator table and proposed list of common context indicators

Impact indicators (micro- and macro-data at RDP, regional, national levels)	• In addition to the micro-level data properly aggregated for beneficiaries and non-beneficiaries, there will exist the need to collect data for the macro-level analysis. This kind of data are to be collected from official statistics and other databases (NUTS 3-5 might be the lowest macro-level where effects of the RD support can be estimated (although this may vary according to the Member States) [66] • The operations database can be used to collect data for impact indicators on beneficiaries, however there is always the need for additional data to be collected by the evaluator:
Data for output, result and impact indicators on beneficiaries are collected on ongoing basis via the monitoring systems along with the application form and project finalisation processes and used on annual basis by MAs in the AIR. Data for non-beneficiaries are usually collected during the evaluation process (FADN, national/regional statistics, etc.)	

Qualitative data and information

Apart from quantitative data, also qualitative information shall be collected by the evaluator during the observing phase, using interviews, focus groups, case studies etc. For this purpose, tools to collect qualitative data and information are developed by the evaluator. As in case of quantitative data collection, it is important to:

- ensure the utility/relevance of the collected data to answer the evaluation questions (check judgement criteria); and
- ensure that qualitative information collected on beneficiaries can be paired with the information on non-beneficiaries to ensure the counterfactual and means to net out the programme effects.

It should be noted that qualitative data collection aims to provide the evaluation with empirical information around the vision of certain entities and leaders of opinion, about the objectives, [potential] results and effects of the programme.

Description of the process of programme implementation

Description of the process of programme implementation, composition of programmes, priorities and target levels, budget, financial and physical execution, major modifications to the strategy, structure and resource allocation. Evaluators shall ensure that the relevant information in this respect is collected.

For this step it is considered essential to conduct interviews already in the initial phase of the evaluation, including representatives of the MA and the respective technical teams.

Summary of the outputs

Evaluators present the aggregated and quality-checked values for each of the common and programme specific indicators and properly substantiated views on the progress observed or deviations occurring during the evaluation period. The aggregated values

---

[66] Guidelines for the ex-post evaluation of 2007-2013 RDPs, Part II, page 132.

must be in harmony with what is foreseen in the templates for common evaluation questions and fiches of complementary result indicators (to be published in separate document).

8.1.2 Analysing

The key task of this phase is to process and synthetize all available data and information in a systematic way, with the view of assessing the (net) effects of the RDP interventions and consider them at both - beneficiary and RDP territory levels. This means to analyse the collected quantitative and qualitative evidence in order to assess the effects, taking into consideration the baseline (CCI) and the trends over the programming period.

Recommended working steps:

- Introductory qualitative analysis
- Formulation of testable hypotheses regarding potential programme effects.
- Testing and verification of hypothesis
- Calculation of specific programme secondary effects at micro-level
- Identifying programme net-results

Expected outcome: calculated values of indicators (gross, net)

For 2017, the result indicators will be in the form of gross and/or net values. For 2019, all result indicators will be in form of net values, as well as impact indicators. This implies that there is sufficient data collected on beneficiaries in the system[67] . Only under this condition it will be possible to compare control groups and net out the values of result indicators and answer the evaluation questions foreseen for 2017 (focus area-related) and beyond (all). This will influence also the assessment of achievements and enable collecting enough evidence to answer all evaluation questions the same way.

Disentangling effects of the programme from effects of other intervening factors68

With the intervention logic the MA usually seeks to establish an unequivocal cause-effect chain for each measure. Strictly following this logic, environmental results/impacts would only be expected from environmental measures – or socio-economic results/impacts only from socio-economic or sectorial measures. However, the assessment of results/impacts is an integrative task. Transverse effects are most likely to occur, and the evaluation design has to take this into account. Furthermore, numerous exogenous factors may influence a given result indicator observed at the

---

[67] The question is if the beneficiaries with completed projects are representative of the total.
[68] Capturing the success of your RDP: Guidelines for the ex poste evaluation of 2007-2013 RDPs, PART II, Chapter 4, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

micro level. The task to separate the programme-borne effects from other intervening factors is specifically critical concerning environmental results, where there is still a lot of uncertainty about methods to determine the proper scale of appraisal (micro- or regional/local) and to derive appropriate counterfactual[69].

Given the above, there are three key methodological challenges:

- The requirement to assess, wherever possible, programme results against their counterfactual, i.e. calculating the changes that would have occurred without the specific programme intervention;
- The requirement to estimate the net effects of the programme, by netting out deadweight-, leverage-, substitution effects, etc.;
- The requirement to construct a data and information base which allows for the unbiased computation of the effects as stipulated above;

The evaluators' task is to establish a logical series of steps that connect the data derived from the database and other sources to the chosen analytical models in a way to respond to all effects expected in the evaluation process:

- direct and indirect effects: expected and unexpected (positive and negative)
- secondary effects: expected and unexpected (positive and negative);
- synergies and transverse effects[70].

This analysis will have to take into account the endogenous and exogenous factors (separation between the programme-borne effects from other intervening factors). (See CEQ-templates published in separate document)

The nature of those steps will be framed by the type and combination of measures and the availability of relevant context, input, output, and result data/information and their ability to appraise the programme results and impacts at later stages of RDP implementation. In some cases, quantitative data are available to estimate the RDP results in other cases, missing quantitative data must be replaced with qualitative evidence. Moreover the quantitative data should be triangulated with qualitative research, as it is proposed earlier in this guidelines.

The example of steps applicable in the assessment of RDP results is presented below:

Step 1: Introductory qualitative analysis

Introductory qualitative analysis (e.g. via application of theory of change) could be applied at the beginning of an evaluation process in order to reflect on various possible expected or unexpected, positive or negative effects of a RDP. If so, after linking it with the programme intervention logic, the evaluator may identify the magnitude and the scope of "observed" gross programme effects which at a later stage should be subject

---

[69] Due to these difficulties, the assessment of environmental effects was in many empirical studies restricted to the measurement aspect only. Indeed the assessment of environmental results/impacts is still in its infancy and a wide range of possible methods (applicable in the field of socio-economic results/impacts) still wait for their empirical testing.

[70] Also see chapter 5.1 of PART II in this document

to more rigorous quantitative verification using causal inference approaches. The main aim of this introductory qualitative analysis is the prioritization of the different potential programme effects by indicating and selecting those which, in view of the RD stakeholders and policy makers, are expected to make the biggest "change" in comparison with a situation "without" the programme (in AIR submitted in 2017 to be reported at a micro-level only). Clearly, at the first stage of evaluation, the preliminary qualitative assessments are essential because they can provide invaluable insiders' perspectives on a programme performance especially regarding its unintended results or direct and secondary effects. Stage 1 is also consistent with the approach recommended in order to develop PSEQs (see Chapter 6.2.3).

Step 2: Formulation of testable hypotheses regarding potential programme effects.

At this stage CEQ and judgement criteria linked to the assessment of the expected results at a micro-level should be complemented with programme-specific evaluation questions and respective judgement criteria and result indicators, and testable hypotheses derived from Step 1.

Step 3: Testing and verification of hypothesis

Under this step, the above-mentioned hypotheses have to be tested and/or verified using methodological approaches based on causal inference. At this stage it is important to use previously collected result indicators to enable a robust analysis of the most important expected or unexpected, positive or negative and secondary programme effects grouped under economic, social and environmental domains. Such testing can only be carried out under rigorous evaluation methodologies inter alia using credible counterfactuals.

The main components of Step 3 are as follows:

- Defining the outcome variables (in economic, social and environmental domains). The analysis can be conducted with as many outcome variables as there are data for. The analysis can be extended not only to cover direct effects, but also secondary effects, and positive, negative, expected unexpected effects.
- Defining the time dimension. For example, by comparing an average 2012-2013 (i.e. reflecting situation prior to the current programme) with 2016 (i.e., situation to be reported in AIR 2017).
- Applying suitable methodologies for finding credible control groups (a preference should be given to the application of both quantitative and qualitative methods). A suitable control group should be found for each type of effects analysed.
- Computing an average outcome effect for the group of programme beneficiaries.
- Computing an average outcome effect for a comparable control group.
- Calculating the expected or unexpected, positive or negative effect of the programme (for each outcome variable separately). Average Treatment Indicators (see definition below) can be applied to assess both programme direct as well as secondary effects – positive and negative.

Step 4. Calculation of specific programme secondary effects at micro-level[71]

Secondary effects are usually a derived causal consequence of programme direct effects at the micro-level (e.g. deadweight loss, leverage effects, or substitution effects) and at the macro or regional level (e.g. multiplier effect, displacement effect, etc.). Both types of secondary effects can be intended or unintended.

Due to the focus of AIR submitted in 2017, the analysis of programme results covering period 2014-2016 requires the calculation of programme secondary effects occurring at a micro-level only.

In general, a well-conceived evaluation design should be based on some verifiable hypothesis about whether programme secondary effects are present and why they exist. A typical feature of these kind of effects is that they can make an overall economic, environmental and social effect substantially larger than direct effects alone. Yet, estimation and especially quantification of secondary effects may not be a trivial task (see: ex-post evaluation guidelines, 2014). Methodologies recommended for estimating deadweight, leverage, and substitution effects are described and illustrated in several other publications.

Obviously abundant micro-economic data and advanced evaluation methodologies are crucial elements in a robust quantitative analysis of programme secondary effects. Additionally, it is recommended to complement quantitative estimates with qualitative information, e.g. by carrying out additional surveys focusing on specific issues (e.g. environmental), or interviewing programme non-beneficiaries that feel affected by the RDP implementation in order to find out a correct reference for a quantitative analysis.

Step 5: Identifying programme net-results

Programme results should be expressed in "net" terms, which means after subtracting the effects that cannot be attributed to the intervention, and by taking into account indirect effects (deadweight loss, leverage, substitution, etc.). The evaluator should specify in detail what elements have been accounted for when calculating programme "net" effects.

Evaluation of indirect effects is complex and requires advanced methods that can capture them. Most important secondary effects are: multiplier effect, substitution effect and displacement effect.

---

[71] Here, we focus on reporting in AIR 2017, i.e. on effects occurring at micro-level only. In the real world, secondary effects occur also at regional- and macro-level and point out how positive or negative effects originated from a given programme are transmitted through and/or across a given territory.

8.1.3 Judging

The task of the judging phase is to interpret the evaluation findings, formulate answers to evaluation questions on the basis of judgement criteria and indicators. The conclusions and recommendations relate to the effects of single focus areas as well as the programme as a whole.

Recommended working steps:

- Answer the evaluation questions foreseen in the 2017 enhanced AIR,
- Judge on expected results and identify the internal and external factors which contribute to the success or failure of the programme objectives.
- Draft conclusions as well as recommendations.

Expected outcome: Answers to EQs, conclusions, recommendations

Answers to evaluation questions must be based on sound evidence and accompanied by a critical discussion of the evidence. In this sense, evaluators shall discuss and interpret the qualitative information and data values for common and programme specific indicators obtained from the analysis. Moreover, the evaluators need to consider the context (e.g. socioeconomic situation, capacity of beneficiaries to co-finance projects). For example where net values of indicators are very low or negative, a proper explanation shall be provided. Or, if certain focus areas or a part of the programme has not delivered the expected results and impacts, an analysis of the reasons for this unexpected effect is necessary. In this respect, the evaluator will have to:

- Answer the evaluation questions foreseen in the 2017 enhanced AIR, taking into account the context. In the answer to the evaluation questions the efficacy and efficiency principles (obtained results and resources spent) should always be present.
- While answering, the evaluator has to guarantee a well thought judgment on the level to which the programme contributes in achieving the expected results and identify the internal and external factors which contribute to the success or failure of the programme objectives.
- Draft conclusions and recommendations appropriately substantiated by the findings.

In case the findings have limitations in their validity, (e.g. in case of insufficient RDP uptake, or using naïve evaluation methods), the resulting judgement should be critically reflected.

During the judging phase, the evaluator also draws the conclusions and recommendations to improve the strategy (design and implementation of RDP). The conclusions and recommendation should be based on the collected evidence and its

robust interpretation, as well as they should be rooted in the answers to evaluation questions.

The answers to evaluation questions, together with the set of conclusions and recommendations are the core messages to be integrated in the enhanced AIR.

Responsibilities

- Evaluators

Relevance for ToR:

- This task is fully part of the ToR, among others:
  - Focus the work on quantification of result indicators (ideally net values), including the programme specific results.
  - Create additional result indicators if needed to answer CEQ
  - Prepare the evaluation system for 2017 and beyond (e.g. resolve data gaps by modelling, extrapolations or other means).
  - Use specific methods & tools & techniques for quantifying and assessing the effectiveness and efficiency of the RDP, and for answering the evaluation questions.
  - Reflect about the need on the CCI update.
  - Provide conclusions and recommendations which are strictly based on evidence of the qualitative and quantitative assessment.

Further reading

Capturing the success of your RDP: Guidelines for ex post evaluation of 2007-2013 RDPs, PART II, Chapters 2, 3, 4 and 5, http://enrd.ec.europa.eu/enrd-static/evaluation/library/evaluation-helpdesk-publications/en/evaluation-helpdesk-publications_en.html

Handbook of the Common monitoring and evaluation framework 2007-2013 and Annexes, Annex 1, Guidance note B – Evaluation guidelines, http://ec.europa.eu/agriculture/rurdev/eval/index_en.htm

PART III - ANNEXES

Available as separate document

European Evaluation Helpdesk
Boulevard Saint-Michel 77-79
B - 1040 BRUSSELS
T: +32 2 737 51 30
Email: info@ruralevaluation.eu
http://enrd.ec.europa.eu