

EUROPEAN
EVALUATION
HELPDESK
FOR RURAL DEVELOPMENT



REPORT

HOW TO IMPROVE EVALUATIONS OF THE CAP BASED ON LESSONS LEARNED FROM 2014-2020

GOOD PRACTICE WORKSHOP
ONLINE, 25 NOVEMBER 2021

Copyright notice

© European Union, 2021

Reproduction is authorised provided the source is acknowledged.

Recommended citation:

EUROPEAN COMMISSION – Directorate-General for Agriculture and Rural Development – Unit C.4 (2021): New tools for monitoring and evaluation: insights from the Evaluation Knowledge Bank. Report of the Good Practice Workshop 20-21 October 2021. Online.

Disclaimer:

The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this report. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.



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 C.4 'Monitoring and Evaluation' 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>).

REPORT

HOW TO IMPROVE EVALUATIONS OF THE CAP BASED ON LESSONS LEARNED FROM 2014-2020

GOOD PRACTICE WORKSHOP
ONLINE, 25 NOVEMBER 2021

CONTENT

Content	1
Tables, Figures and Maps	2
LIST OF ACRONYMS	3
EXECUTIVE SUMMARY	4
1 SETTING THE FRAME	5
1.1. Introduction	5
2 SHARING LESSONS	6
2.1 Objectives and scope of evaluation	6
2.2 Lessons related to the types and topics of evaluations and evaluation criteria.....	6
2.3 Lessons from using common evaluation questions and judgment criteria	12
2.4 Lessons related to data availability and netting out of indicators	14
2.5 Support needs for evaluation	15
3 CONCLUDING REMARKS	17

TABLES, FIGURES AND MAPS

Figure 1. Participants of the Good Practice Workshop per role and Member State5

Figure 2. What were your main difficulties in terms of the evaluation of your RDP?.....6

Figure 3. Roadmap: What type and topic of evaluation do you consider most relevant, at which point in time and why?.....8

Figure 4. Outcomes of Menti poll on the benefits of having a common evaluation system.....13

LIST OF ACRONYMS

AIR	Annual Implementation Report
AKIS	Agriculture Knowledge and Innovation Systems
ANC	Areas with Natural Constraints
CAP	Common Agricultural Policy
CEQ	Common evaluation question
CMES	Common Monitoring and Evaluation System
CPR	Common Provisions Regulation
EAFRD	European Agricultural Fund for Rural Development
EIP	European Innovation Partnership
ENRD	European Network for Rural Development
EU	European Union
FADN	Farm Accountancy Data Network
FBI	Farmland Bird Index
GAEC	Good Agricultural and Environmental Conditions
GDPR	General Data Protection Regulation
GHG	Greenhouse Gas Emissions
HNV	High Nature Value
IACS	Integrated Administration and Control System
JC	Judgment criteria
LAGs	Local Action Groups
MA	Managing Authority
MC	Monitoring Committee
MS	Member State
NRN	National Rural Network
PA	Paying Agency
RDP	Rural Development Programme

EXECUTIVE SUMMARY

The 19th Good Practice Workshop of the Evaluation Helpdesk, 'How to improve evaluations of the CAP based on lessons learned from 2014-2020', took place online on 25 November 2021. It brought together 74 participants from 24 different EU Member States. These participants included RDP Managing Authorities, Paying agencies, evaluators, European Commission representatives, researchers, National Rural Networks, and other evaluation stakeholders. The overall objective of the workshop was to work towards identifying a set of common guiding principles for CAP evaluations.

The workshop was launched with a general reflection on the scope and achievements of evaluations, stressing the general awareness raising of the benefits of evaluations and the contribution to a constantly stronger evaluation culture.

Lessons from the different types and topics of evaluations and the evaluation criteria used fed into a discussion about what should be evaluated, when and why. An indicative roadmap comprising ideas and suggestions from Member States concluded that process-related evaluations are more pertinent at the beginning of the programming period, with achievements and impacts to be evaluated in later stages. However, short-term impacts may also be assessed in the middle of the period in order to justify previous choices and prepare the next programming period. In addition some environment/climate related interventions or other interventions which are high in the policy agenda may need to be assessed already from the middle of the programming period.

Lessons from the use of a common monitoring and evaluation system, focusing in particular on common evaluation questions, judgement criteria and indicators fed into discussions related to the improvement of the scope and content of these common elements. An overarching message is that the virtues of a common system, notably the possibility to exchange good practices and aggregate findings at EU level and ensure accountability at EU level, should be combined with flexibility for Member States to assess their specific needs and carry out evaluations to address these needs. Concerning indicators, while they are relevant for assessing the performance of the policy, there are still unresolved issues in relation to data availability which hinders their netting out. Netting out only when certain conditions are in place, are amongst the ideas for better using the indicators to assess the contributions of the policy.

Finally, Member States highlighted needs for support to improve evaluations of the CAP through capacity building, knowledge sharing events, networking and guidance. Quantitative approaches for assessing net effects, assessment of environmental impacts, the use of cost-effective evaluation methods, early assessment of data needed, methodological approaches for assessing eco-schemes and AKIS are the some of the main areas where support can be provided. There is also a need for a common learning process through collaborative approaches, such as multi-actor working groups, involving MAs, PAs, evaluators and the scientific community.

1 SETTING THE FRAME

1.1. Introduction

As the current programming period is coming to an end, Member States can look back at the accumulated experience on evaluation to identify lessons for improving evaluations in the future. Good quality, evidence based, well-conceived and structured evaluations contribute to better policy design and implementation, while addressing wider policy needs.

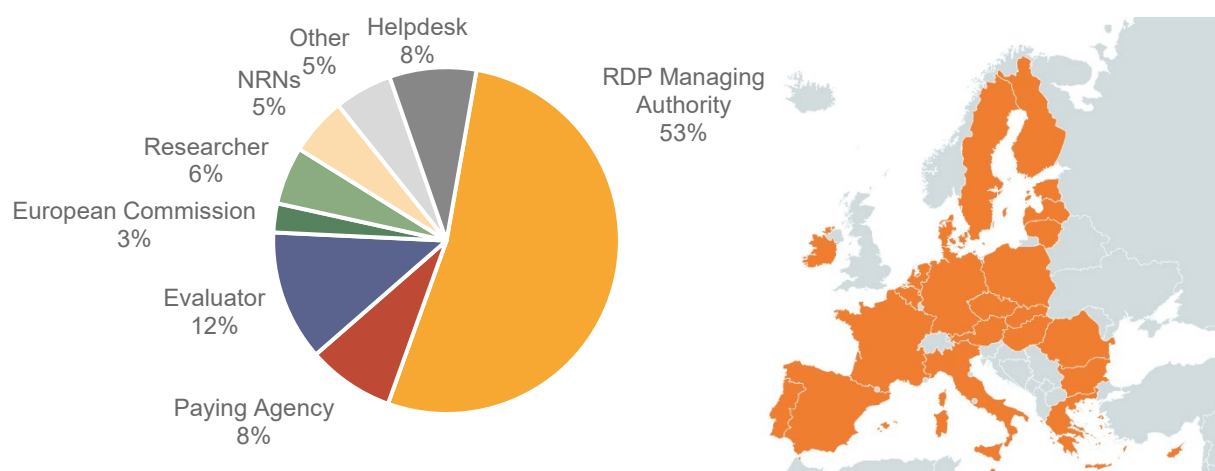
The evaluations conducted under the current common monitoring and evaluation system (CMES) and reported in the Annual Implementation Reports (AIRs) for instance, can offer useful insights into elements that worked well and aspects that could improve (e.g. what is evaluated, with what criteria, what gaps exist and how they could be addressed). Such lessons may help Member States in establishing a future framework for reporting, monitoring and evaluation of the CAP and therefore be better prepared to respond to the requirements of a more performance-oriented policy.

In order to capitalise on existing experience and within a context of a common learning process, Member States have worked together towards a common understanding of key concepts and elements of monitoring and evaluation and their practical application. This was the main focus of the Good Practice Workshop 'How to improve evaluations of the CAP based on lessons learned from 2014-2020'. Key issues considered in discussions include: the role of evaluation, the scope of different types of evaluations, the wider policy needs, the lessons learned from the current programming period, and the role of evaluation planning in the future.

Against this background, the overall objective of the workshop was to **work towards identifying a set of common guiding principles for CAP evaluations**. To achieve this overall objective, specific objectives included: analysing and exchanging knowledge on the lessons learned from the current period in relation to the CMES elements; identifying how these elements can be used/improved in the future so as to contribute to better CAP evaluations (with a view to improving the performance of the policy); discussing which of these elements could be common for everyone and which might require specific support to be provided to Member States.

74 participants from 24 different EU Member States attended the online event, including RDP Managing Authorities, evaluators, European Commission representatives, researchers, National Rural Networks, and other evaluation stakeholders.

Figure 1. Participants of the Good Practice Workshop per role and Member State



2 SHARING LESSONS

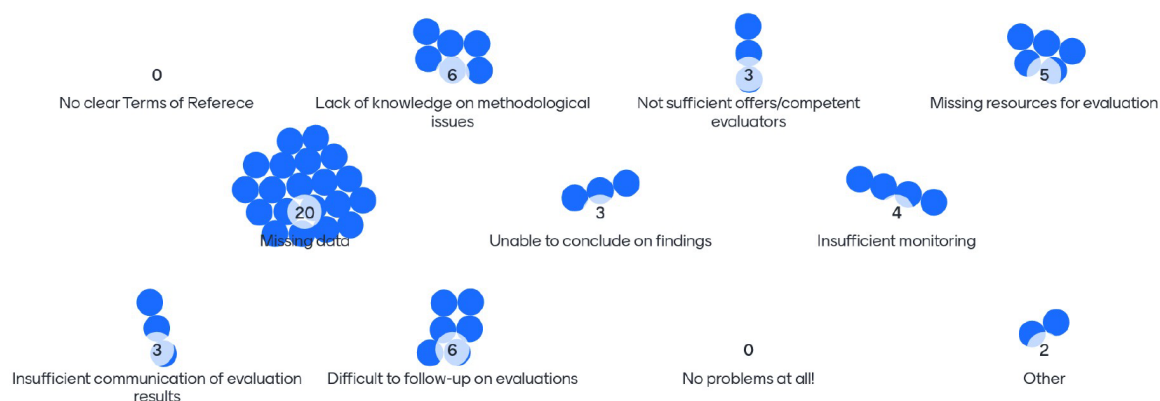
2.1 Objectives and scope of evaluation

The starting point of the workshop was a general reflection on the main achievements in terms of evaluations of RDPs. Participants brainstormed on the main achievements, concluding that:

- RDP evaluations offer an ongoing approach to evaluation and contribute to create an evaluation culture.
- Awareness of the benefits of evaluation has grown by convincing the wider audience of the importance of evaluation, promoting the participation of evaluators to the wider evaluation network at EU level and increasing the interest of partners.
- The quality of evaluations has improved, including the provision of good recommendations. A good understanding of what additional data is available also contributes to quality evaluations.
- Data availability has improved, there is for instance, good additional data and better data collection. This is supported by good dialogue between data providers and evaluators when processing data requests.
- Evaluations contribute to better structure and design of the RDP and improvement of measures, contributing thus to a more efficient and effective policy.
- Other more concrete achievements include improved impact calculations and the focus on targets.

Following the brainstorming on achievements, participants voted for the main difficulties in terms of evaluations of RDPs (Figure 2). Missing data remains by far the most pressing issue for evaluation, followed by lack of knowledge on methodological issues and difficulties to follow up on evaluations. In some cases, there are budgetary resources missing for conducting evaluations, which may be combined with lack of sufficient or competent evaluators.

Figure 2. What were your main difficulties in terms of the evaluation of your RDP?



2.2 Lessons related to the types and topics of evaluations and evaluation criteria

The Helpdesk offered an overview of how Member States have reported on completed evaluations in their AIRs. The analysis shows that the absolute number of completed evaluations reported per year has steadily increased between the AIRs submitted in 2016 and 2020, reflecting the overall increase of Member States' evaluation outputs in line with the RDP uptake and approaching the major evaluation milestones in 2017 and 2019. In the last available AIRs submitted in 2021 the number of evaluations reported by the Member States has dropped by more than 20%, evidently as no major evaluation milestone was related to this reporting year. In terms of evaluation topics addressed throughout the

reporting years it becomes evident, that most completed evaluations were reported in relation to the CAP objective 2 (258 evaluations), compared to CAP objective 3 (103 evaluations) and CAP objective 1 (76 evaluations), while however the number of evaluations addressing multiple RD priorities was quite high (110 evaluations). Additionally, at the beginning of the programming period also the assessment of the RDP delivery system, including the assessment of the functioning of the monitoring and evaluation system was quite important. Other RDP aspects (TA, NRN, synergies) are mainly addressed in the evaluations reported in 2018 and 2020. The horizontal RD Priority knowledge transfer and innovation had most completed evaluations in 2020 (19 evaluations).

A snapshot of the most common types of evaluations reported in the AIRs in 2021 shows that evaluations of achievements account for 36% of all reported evaluations. This is followed by impact evaluations with / or without counterfactual approach (18%). Impact evaluations with counterfactual approach concern only 3% of the total. Process evaluations of the implementation system have a share of 7% of the completed evaluations reported in 2021. Around 2% of the evaluations relate to the new programming period, e.g., ex-ante evaluation of the CAP Strategic Plan. Studies that are reported, but are not evaluations in the strict sense, account for about 27% (categories 'other than evaluation' and 'research study other than evaluation'). These are studies on an RD topic which are not examining the RDP contribution as such e.g., studies on context indicators such as the farmland bird index, publications with good practice examples. In about 10% of the evaluations, the available information does not allow to specify the type.

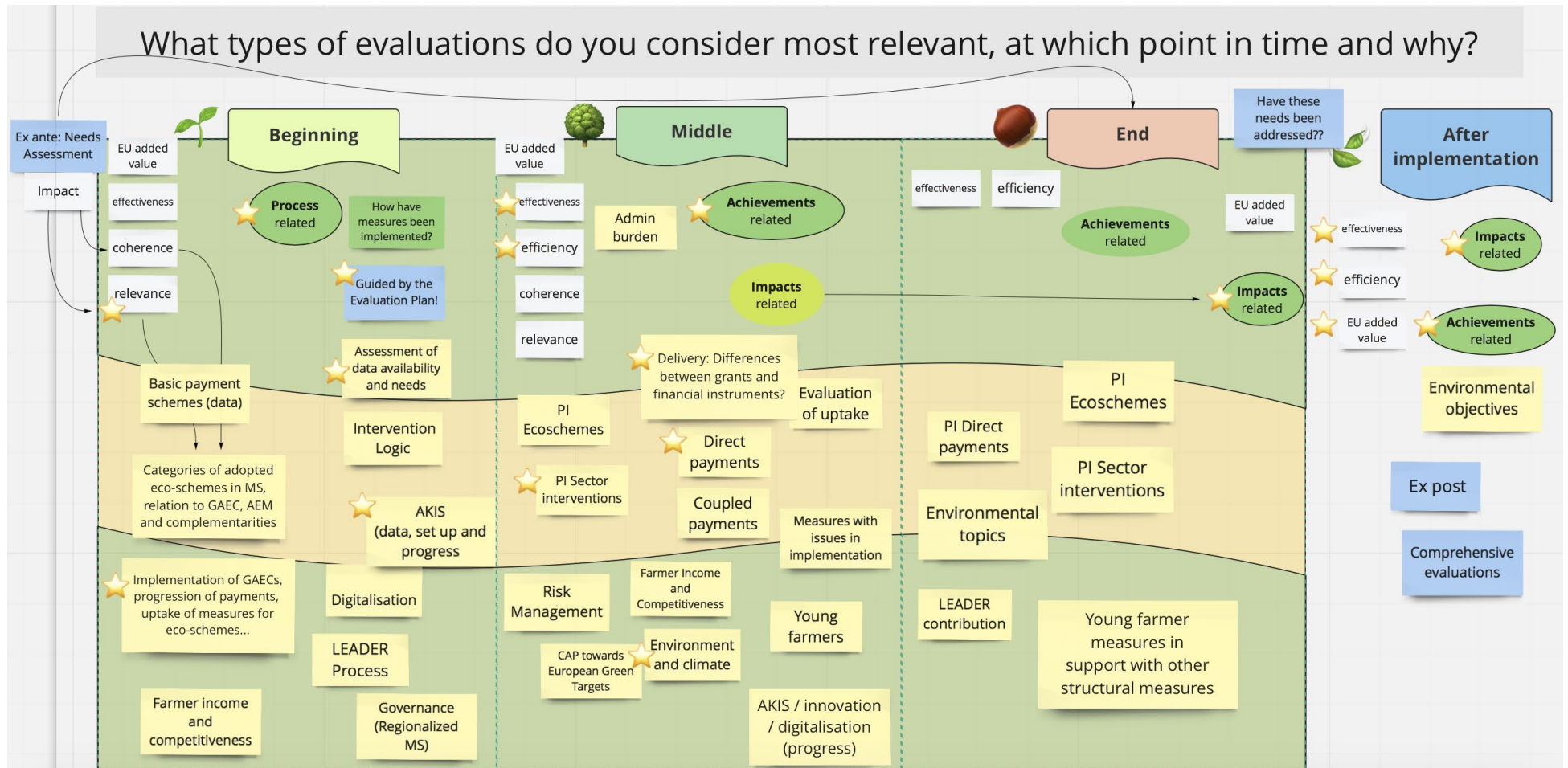
Within each CAP Objective, between 2019 and 2021, there have been thematic evaluations focusing on specific topics, in particular:

- In relation to competitiveness, most thematic evaluations focused on productivity and young farmers, while a few evaluations were carried out on animal welfare, job creation and diversification.
- In relation to the environment, the most common thematic evaluations were on biodiversity, followed by evaluations on soil and land management and then HNV and disadvantaged areas. Some thematic evaluations were related to water and a few to energy and emissions.
- In relation to territorial balance, some evaluations focused on diversification and local development, with LEADER being specifically assessed in thematic evaluations in 2021. Other topics include gender, social inclusion and infrastructures.
- In relation to the horizontal priority on knowledge and innovation, thematic evaluations covered innovation, the EIP and training and advisory services.

This overview was followed by group discussions, where participants developed a roadmap depicting what types and topics of evaluations are more relevant and at which stage of the programming cycle, as well as the rationale/justification for the choice. The relevant evaluation criteria were also discussed and placed in the roadmap.

The following roadmap (Figure 3) summarises the common points amongst the workshop participants on the types and topics of evaluations that are most relevant and at which point in time. A star represents the commonly agreed points amongst the different groups.

Figure 3. Roadmap: What type and topic of evaluation do you consider most relevant, at which point in time and why?



Colour codes: Green and blue: types of evaluations, Yellow: evaluation topics, White: evaluation criteria. Star: shared views amongst different groups

A summary of the main discussion points explaining the reasons behind this roadmap follows.

The beginning of the programming period

Types of evaluations:

There is a shared view that process related evaluations are most relevant at the beginning of the programming period, guided by the Evaluation Plan. The rationale for this is:

- The evaluation process for the whole period is built at this stage, through evaluation planning, including types of evaluations and how to spend the money. Lessons from the previous programming period can feed into the development of the evaluation plan.
- Process evaluations should involve checking the monitoring systems to collect data: who can collect what and when, make sure all the data is in place and is being collected to facilitate evaluations. It is essential to know what data is needed for evaluation early on, as well as ensure balance, while avoiding too much burden.
- Process evaluations early on can address the need to facilitate new ways to get the existing data FADN, IACS, etc. as well as mechanisms for combination of this data. Data protection and ownership issues should also be considered.
- When conducting process evaluation, it is important to get the contribution of different stakeholders and develop governance mechanisms.

Evaluation topics

Most Member States consider that an assessment of data needs and data availability should take place at an early phase of the programming period. Specific topics that can be assessed during the beginning phase of the programming period include:

- Governance, notably in regionalised Member States and linked to the new delivery model.
- Specific topics can be defined based on the intervention logic, according also to budget allocation and/or specific national needs (political, territorial, etc.). Evaluation questions can then be developed according to the intervention logic and help focus the evaluations.
- AKIS: an assessment at this stage would focus on data needed for the assessment of interventions¹ in the context of AKIS, how AKIS it is set up and how it is progressing, while building on an example of AKIS in the context of Horizon 2020 may also be useful.
- Eco-schemes: evaluation of uptake is particularly important for such schemes, i.e. what kind of eco schemes have been adopted in the different Member States (categorise them) and relate them with GAEC and AEM – complementarities (comparative analysis). If an assessment of uptake is not done early on, there is a risk of not being able to deliver what is needed for the green architecture and eco-schemes.
- LEADER: at this initial stage, the evaluation should focus on the LEADER process. LEADER evaluation has to be dynamic, i.e. a continuous process throughout the programming period.
- Basic payment schemes: Assessment of the data needed for direct payments and organisation of its collection is important from the beginning of the programming period.

According to Article 102 of the PRC, AKIS in CAP Strategic Plans shall contain: 1) their contribution to the cross-cutting general objective related to fostering and sharing of knowledge, innovation and digitalisation, notably by describing: the organisational set-up of the (*existing*)

¹ These include knowledge, exchange, advice and information; cooperation; EIP and operational groups.

AKIS, how advisors, researchers and CAP networks will work together within the framework of the (*future*) AKIS, and how advice and innovation support services are provided; 2) a description of the strategy for the development of digital technologies in agriculture and rural areas.

Evaluation criteria

There is a shared view that the assessment of relevance and coherence are relevant at the beginning of the programming period, especially when defining impact targets. This exercise can be linked to the ex ante. For instance, check the ex ante needs assessment, in particular, look at the needs assessment of previous programming periods and see if the programme actually addressed them. When assessing relevance, evaluation topics should be prioritised according to where most of the funding goes.

The middle of the programming period

Types of evaluations:

The assessing achievements can take place at this stage of the programming period. Although it requires a lot of efforts, if there is data availability, this could make it easier. This is why data needs, availability and collection should be assessed in the previous stage (beginning of the programming period). Evaluation of achievements can:

- be a preliminary assessment at this stage;
- consider the implementation progress;
- help understand the functioning and achievements of green architecture (CAP specific objectives 4, 5 and 6).

Impact evaluations may take place at this stage, but there are different views in relation to their scope (not mutually exclusive options below):

- One option is to evaluate short term impacts in order to validate previous choices and better define the next CAP Strategic Plan. For instance, a mid-term impact assessment in the current period helped identify whether ANC measures should also be applied in the CAP Strategic Plan.
- Another option is to focus on certain topics instead of doing an impact evaluation of all the programme.
- A third option is to define expected impacts including similar interventions implemented in the previous programme and eventually adjust the CAP strategic plan.
- Finally, it should be borne in mind that impact evaluations are often used to inform future programming periods, not necessarily the next one (e.g. impact evaluations of 2014-2020 may inform the post 2027 programming period).

Evaluation topics

There is a shared view that direct payments and sectoral interventions (previously Pillar I) should be evaluated at this stage of the programming period. This was not done at Member State level before. Most Member States also consider that environment and climate is an area that needs to be evaluated from this stage onwards, given its importance in the post 2020 programming period.

A variety of other topics are proposed to be evaluated in the middle of the programming period:

- Young farmers topic is important, as it is linked to ringfencing in the CAP Strategic Plan.

- Innovation: Innovation can be transferred very widely and can have major impacts on sectors.
- CAP contribution to EU Green Deal targets is important, but challenging to decide at which stage of the programming period. There is a need to clarify what can be the focus of topics related to Green Deal targets for evaluations taking place at this stage.
- Evaluations of measures with issues in implementation. These may help identify areas for improvement and provide recommendations to overcome any issues.
- Other topics include risk management, digitalisation, farmer income and competitiveness.

Evaluation criteria

There is a shared view that effectiveness and efficiency are the criteria to be applied in evaluations taking place in the middle of the programming period. For this, it is required to also have relevance and coherence assessed (already from the start of the period).

Efficiency may include the assessment of administrative burden (including administrative costs and time). It may not be possible to assess efficiency for all parts of the CAP Strategic Plan. It depends on the Managing Authority to suggest topics to be assessed that are important/relevant for reducing administrative burden. Efficiency may also include a comparison with other Funds or other forms of support.

Open questions include:

- whether efficiency can be assessed without measuring effectiveness;
- whether all Member States can assess effectiveness at the same time, given that implementation progress differs amongst Member States.

The end of the programming period

Types of evaluations:

Most Member States consider that impacts can be assessed at this stage, together with achievements (results) of interventions. In relation to the latter, evaluations should complement the preliminary assessment of achievements that may have taken place in the middle of the period.

Evaluation topics

The key question here is whether to evaluate the contribution of interventions to environmental objectives earlier than the ex post stage, given these are high in the policy agenda. There is a shared view that given that environmental effects take time to appear, this may be a good starting point in time to evaluate them. The use of a counterfactual approach may be possible at this stage if enough data is available.

In addition there are other pertinent topics to evaluate at the end of the programming period:

- LEADER: LEADER contribution is better evaluated at the end of the programming period as the LEADER process is slow. Open questions include: what should be evaluated and at which level? how can MAs support LAGs in the evaluations of their local development strategies?²
- Young farmers, together with other structural support measures that need time to be evaluated, e.g. training, investments.

² For requirements related to LEADER evaluations and the responsibilities of LAGs to evaluate local development strategies, see the Common Provisions Regulation (CPR 2021/1060), Articles 32 and 33.

Evaluation criteria

There is a shared view that impact is the most relevant evaluation criterion at this stage, while effectiveness and efficiency would have started to be assessed even from the middle of the programming period.

The end of the programming period

Types of evaluations:

This is the stage where the ex post evaluation will take place (legislative binding requirement). There is however an open question given that programmes are still being funded until 2025: how can Member States deal with the ex post in terms of time and resources along with all the other evaluations?

Evaluation topics

This is the stage where comprehensive evaluations would take place. However, there is a limit of what can be evaluated if certain evaluations have already been done before. It is important to use evaluation resources effectively. Potentially thematic evaluations related to environmental objectives are relevant given the environment/climate emphasis of the policy.

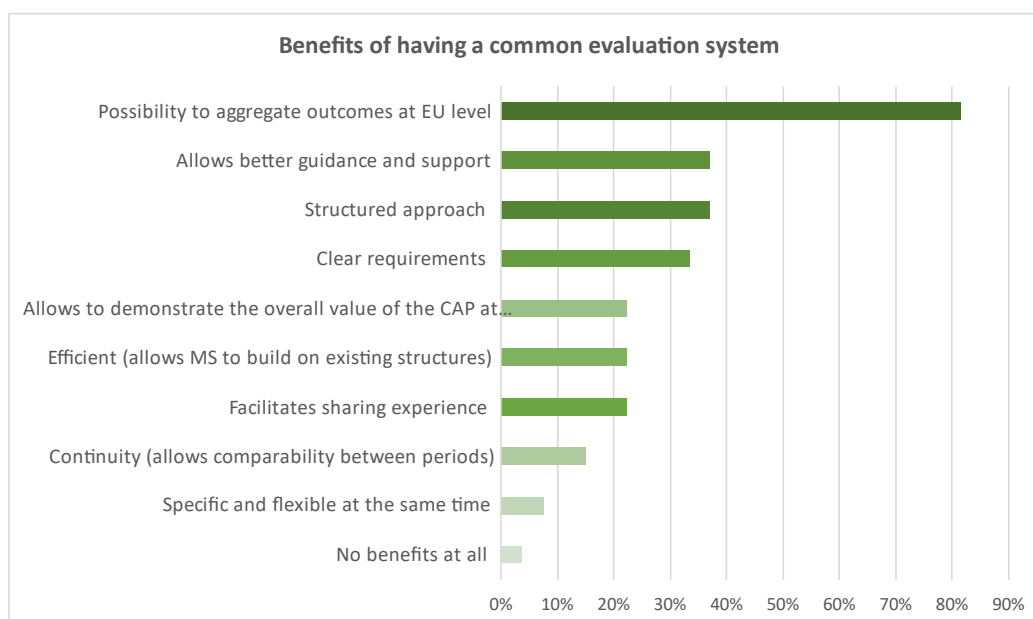
Evaluation criteria

Evidently the impact evaluation criterion applies here, together with effectiveness, efficiency and EU value added.

2.3 Lessons from using common evaluation questions and judgment criteria

The Helpdesk introduced the session related to the use of common evaluation elements and in particular of common evaluation questions (CEQs) and related judgment criteria (JC). The current CMES offers a structure for evaluations, consisting of 30 CEQs (18 related to results at Focus Area level, 3 related to other RDP areas and 9 related to the impacts of the programme). For each CEQ, it was recommended to define at least one JC, with the possibility to revise them or add further JC. CEQs were supported by common indicators (26 output, 24 target/result indicators, 6 complementary result indicator and 16 impact indicators, with the possibility to add proxy indicators. Finally, additional information could also be used to answer the CEQs. The objective of this session was to identify lessons from having a common evaluation system.

In order to further understand the benefits of having a common evaluation system for the Member States, a Menti poll was subsequently launched. The outcomes of the poll showed that the best perceived benefit by participants (81%) was the **possibility to aggregate/compare findings at EU level**. The CMES was also deemed relevant for providing a **structured approach** (to evaluation) and for allowing **better guidance and support** (37% each), as well as for setting **clear requirements** (33%). Few participants however believe that a CMES could enhance **continuity** (allowing comparability between periods) or that it is **specific and flexible** (to MS needs).

Figure 4. Outcomes of Menti poll on the benefits of having a common evaluation system

The discussion was later focused on the usefulness of the CEQs and JC for conducting evaluations. The Helpdesk provided an overview of the lessons learned on the subject, mainly about what worked well, issues, and ideas for improvement as they were collected from stakeholders through different sources. Following this overview, participants reflected further on the topic through discussions based on the following questions:

- To what extent were CEQs and JC helpful to guide and structure RDP evaluations?
- What should be considered when designing CEQs and JCs? (e.g. focus, level of specificity, clarity).

In general, the discussion stressed that CEQs (together with other common elements such as indicators, data, etc.) can be **useful to synthesize evaluations** at EU level and a correct answer to each CEQ seems challenging, both in terms of time and costs, given also the lack of data for counterfactual analyses.

When it comes specifically to the usefulness of CEQs for evaluation, it was remarked that while in the current programming period the CEQs reflected the objectives set in the Regulation 1305/2013, the new CAP Strategic Plan (the new Common Evaluation System related to it), it was pointed out that they were too broad to help Member States to steer their policy. **MS prefer assessing their more specific needs**. Such an approach seemed to be backed by the fact that several MS (e.g., DE, IT) have carried on thematic evaluations during 2014-2020 in order to address specific aspects which were more relevant for their programmes and their state of implementation.

However, the EC clarified that setting CEQs enables not only **accountability**, but also **the establishment of a communication channel** for policy steering between the EC and the MS providing common understanding on the content of evaluations. Moreover, it was also clarified that evaluations conducted at EU level (as in the example of DG Regio) would not double work which is conducted at MS level (subsidiarity), but still have to be designed based on a common ground with what is performed at MS level in order to understand and compare the findings (synthesis of policy achievements). DG

AGRI pointed out that need for a more systemic approach, is also dictated by other EU bodies and services, e.g., the European Court of Auditors.

Finally, some considerations for improvements were discussed such as introducing **more focused and better defined/specific CEQs** which would guide the evaluation throughout the period (potentially reflected in the Evaluation Plan) and enable the collection of relevant data, or foreseeing CEQs only at the ex post stage. Another suggested improvement was the elaboration of thematic reports which take into account the impact of the programme and are carried out for a specific sector or group of farms, or rural area, and have a higher added value.

Nevertheless, the role of the European Evaluation Helpdesk was deemed crucial for the sharing of knowledge and methods.

2.4 Lessons related to data availability and netting out of indicators

The Helpdesk introduced the session on the use of indicators for answering CEQs with the objective of identifying lessons from the current programming period. The overview presentation focused specifically on to what extent impact indicators have been quantified, showing that the indicators more difficult to calculate (data more difficult to collect) were those related to environment, namely soil organic matter (I.12), water quality (I.11), HNV (I.09), FBI (I.08), water abstraction (I.10), and soil erosion (I.13). Most net values were reported for competitiveness indicators (agricultural entrepreneurial income and agricultural factor income, with 29 and 28 MAs respectively reporting net values) and GHG emissions under the environment objective (30 MAs reported net values for GHG emissions).

The presentation also showed some information in relation to lessons on data collection for the calculation of common indicators, and particularly it indicated what worked well (e.g. good monitoring/information systems, use of qualitative methods to fill the gaps, long term collection of environmental related data), issues (e.g. lack of ongoing monitoring for some indicators, lack of regional data) and some ideas for improvements (e.g. targeted monitoring system on environment, use of all relevant data sources including Earth Observation).

In more detail, some lessons could be drawn with regards to the netting out of certain indicator values: a few MAs were in fact able to quantify the net contribution of the RDPs, especially for environmental effects for which the heterogeneity of the biophysical aspects makes difficult to apply a counterfactual approach. In relation to this some improvements could be suggested such as the use of qualitative methods and of surveys or thematic studies.

Following this overview, participants reflected further on the topic through discussions based on the following questions:

- To what extent were the common indicators sufficient to assess the effects of policy? (data availability, netting out, use of additional information)
- In which areas were the common indicators not sufficient to assess the effects of the policy? Would you suggest additional information for using the indicators?
- What is needed to better assess the contributions of the policy using common indicators?

As far as the first question is concerned, the participants underlined how the use of **common context indicators was important and relevant** for describing the situation in which the RDP was implemented.

It was also highlighted that in the current period the link between the RDP interventions/farm practices and the effects generated (impacts) are not always clear and there is therefore a need for a better understanding.

For the above, as far as environmental impacts are concerned, WikiCAP project of the JRC (https://marswiki.jrc.ec.europa.eu/wikicap/index.php/Main_Page based on all studies available could provide useful support.

Furthermore, with regards to what is needed to better assess the contributions of the policy, the participants suggested that:

- For socio-economic effects it could be relevant to use standardized and systematically collected data (e.g., FADN, I-O tables);
- For environmental effects it could be more realistic to focus on smaller territories, allowing the assessment of contributions to certain dimensions (e.g., soil erosion).

The improvement of the indicator fiches that has already taken place, taking into account also any additional information needed to better capture CAP effects, will help address Member State needs for more detailed information on the indicators.

Finally, it was agreed that **netting out** the contribution of CAP is important (e.g. employment income, GHG emissions, water quality) and should be done regardless of the robustness of the approach. However, the participants clarified that this may not be relevant or feasible for all indicators and it can be carried out only if certain conditions are in place (e.g. existing national database) or by focusing on the most important or relevant areas. In addition, national monitoring programmes are not always relevant for CAP evaluations, for instance some monitoring programmes of GHG emissions. In these cases, agreements with departments in charge of these programmes may help improve the relevance of the data collected for CAP evaluations.

2.5 Support needs for evaluation

A final session on further support for evaluations of the CAP needed by MS was introduced by the Helpdesk, based on a categorization of needs:

- Role of structures (e.g. Evaluation Helpdesk, Working Groups);
- Capacity building (e.g. legal framework, data collection);
- Knowledge sharing (e.g. examples of evaluation methods, networking);
- Guidance (guidelines for answering EQs, calculation of indicators).

Participants discussed their needs for support based on the following key question: What are your main needs for support for evaluations of the CAP and who would give this support?

The following table shows the main outcomes of the discussion:

SUPPORT NEED	TYPE OF ACTION
Capacity building	<ul style="list-style-type: none"> ❖ <i>Monitoring Committees should have more potential to enhance evaluations (e.g. data coming from organisations in the MC);</i> ❖ <i>Specific capacity building for Paying Agencies (i.e. in the context of both Pillars working together).</i> ❖ <i>Data gaps assessment</i>
Knowledge sharing events	<ul style="list-style-type: none"> ❖ <i>Pilot projects implemented in different MS (quantitative approaches, e.g. net effects);</i> ❖ <i>A platform on existing methodologically-oriented practices in MS targeted to the scientific community/evaluation;</i> ❖ <i>Involving the wider evaluator community, including Pillar I stakeholders (agriculture at large);</i> ❖ <i>Linkages with H2020 projects and the relevance of their outputs for evaluation (e.g. MonVia on biodiversity).</i> ❖ <i>Relate to new approaches (e.g. cooperative approach in NL for environment) and result-based approaches.</i>
Guidance	<ul style="list-style-type: none"> ❖ <i>Comparative analysis (especially of eco-schemes);</i> ❖ <i>Guidance documents to include examples from MS;</i> ❖ <i>Completed evaluations and study design in English in the Evaluation Helpdesk website and for specific cases, also available in the MS language.</i>
Networking	<ul style="list-style-type: none"> ❖ <i>Collaborative and innovative working groups linking MS needs (linking practical needs of evaluators with the scientific community)</i> ❖ <i>Interdisciplinary working groups with MA, PA, evaluators, researchers, etc.</i>

3 CONCLUDING REMARKS

The outcomes of the discussions of the workshop concluded with the following key messages in relation to improving evaluations of the CAP based on lessons learned from 2014-2020.

Scope and objectives of evaluations

The constantly growing evaluation culture is reflected in the improved quality of evaluations, backed up by improved data availability for assessing achievements and impacts and enhanced cooperation between data providers and evaluators. This may help overcome one of the most recurrent difficulties when evaluating RDPs, notably missing data.

The benefits of evaluations are evident in their contribution to a better structure and design of the policy and better targeting of the measures.

Types and topics of evaluations and evaluation criteria

Process related evaluations should take place at the beginning of the programming period, while the assessment of achievements can start in the middle of the period and definitely take place in the end and at the ex post stage. Impacts are mainly assessed at the ex post stage, however, for certain topics, notably environment and climate (including the contribution to Green Deal targets), a preliminary assessment may take place earlier (since the middle of the period), given the increased ambition of the post 2020 period in these areas.

Process evaluations should involve the assessment of the monitoring system, including the timing and availability of data, taking into account relevant data ownership and protection issues as well as administrative burden. Evaluations of achievements should take into account the progress of implementation, if conducted during the middle of the period. In addition to the complementary ex post assessment of impacts, other options for impact evaluations during the programming period, include the assessment of impacts of measures for which sufficient data exists or even identify lessons from the assessment of expected impacts of measures from the 2014-2020 period in order to improve the content and design of CAP Strategic Plans. Furthermore, the assessment of short-term impacts in the middle of the programming period on specific topics, can contribute to improve the future period (post 2027).

The intervention logic of the programmes and the level of uptake should be the starting point for deciding what to evaluate and at which point in time, reflected also in the Evaluation Plan. Following this, thematic evaluations could focus on aspects related to the new delivery model at the beginning of the programming period, as well as topics like LEADER from a process/delivery point of view. At the same time, specific topics merit attention early on, especially those for which the uptake is important in order to achieve the Green Deal targets, for instance eco-schemes and agri-environment measures. Certain topics may need to be assessed in the middle of the period, especially those that were not evaluated at Member State level before, like financial tools, direct payments and sectoral interventions. There is a breadth of other topics suggested for evaluations in this interim stage, such as young farmers, innovation, digitalisation, risk management, farmer incomes as well as ad hoc topics depending on which measures have issues with implementation. For the latter, evaluations can help produce recommendations for overcoming the issues and improving implementation.

Thematic evaluations of areas where effects/impacts take time to appear, would take place at the end of the programming period and ex post. This is the case for environmental topics (although as mentioned, some preliminary assessment is relevant earlier in order to ensure progress towards the ambitious targets), LEADER or young farmers.

In terms of evaluation criteria, the dominant view is that relevance and coherence are more relevant to be assessed at the beginning of the programming period, taking also into account links with the ex ante.

Effectiveness and efficiency are more pertinent from the middle of the programming period, with efficiency focusing, amongst others, on administrative burden. Finally, impacts are assessed ex post, but earlier interim impact evaluations may be pertinent for some topics, provided that the necessary data is already available.

Common evaluation questions and judgment criteria

The advantages of having a common evaluation system are the structured approach that this gives to evaluations, supported by common guidance and support and clear common expectations for everyone. A common evaluation system entails benefits in terms of aggregating and comparing findings at EU level, with common evaluation questions allowing a synthesis of policy achievements at EU level. A common system allows for accountability, however, there is a widespread need for more flexibility so that Member States can undertake evaluations depending on their specific needs. Common evaluation questions may become more focused and better defined to guide evaluations throughout the programming period. There is also a need to define what shall be assessed (with or without common evaluation questions) to give also an orientation for Member States' evaluation questions.

Data availability and netting out indicators

Data availability issues may be addressed through approaches like standard and systematic data collection for socio-economic effects, while for environmental effects, more focused assessments may be more realistic (e.g. smaller areas, certain dimensions like soil erosion).

Indicator fiches may improve to reflect with more clarity the link between CAP interventions and their effects or impacts.

Netting out is important for understanding the real effects of the policy, but it may not be relevant for all indicators and can only be carried out if evaluators have access to relevant data sources.

General support needs

Specific capacity building can improve the capacity of Monitoring Committee members or Paying Agencies in relation to data collection or to the specificities of the new CAP architecture. It can also contribute to improve their knowledge of evaluation concepts and to build evaluation culture.

Knowledge sharing events are claimed by participants as useful for learning from each other on evaluation approaches and methodologies, including from projects outside the agriculture/rural development field (e.g. Horizon 2020 projects).

The relevance of guidance was stressed, especially for assessing net effects, assessment of environmental impacts, the use of cost-effective evaluation methods, early assessment of data needed, methodological approaches for assessing eco-schemes and AKIS, amongst others. Guidance can also serve to share examples of evaluations from different Member States.

Finally, networking adds value to evaluations by bringing together the policy, evaluation and scientific communities to learn from each other and transfer experiences.

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>

