TECHNICAL REPORT

APPROVED: 06 April 2020

doi:10.2903/sp.efsa.2020.EN-1843

Draft framework for protocol development for EFSA's scientific assessments

European Food Safety Authority (EFSA),
Laura Martino, Elisa Aiassa, Þórhallur Ingi Halldórsson, Konstantinos Panagiotis Koutsoumanis, Hanspeter Naegeli, Katleen Baert, Francesca Baldinelli, Yann Devos, Federica Lodi, Alfonso Lostia, Paola Manini, Caroline Merten, Winy Messens, Valentina Rizzi, Jose Tarazona, Ariane Titz, Sybren Vos

Abstract

During 2014-2018, EFSA defined a series of principles for the scientific assessment process (impartiality, methodological rigour, transparency and engagement) and developed a 4-step approach (plan/do/verify/report) to facilitate their fulfilment. According to the approach, the methods for the scientific assessment must be planned upfront in a protocol to prevent data-driven decisions and to increase rigour and transparency of the process. Following the decision to gradually implement the 4-step approach in all EFSA's non-application scientific assessments, it was deemed necessary to set up recommendations for protocol development. This technical report provides these recommendations. The document is published as a draft because the framework for protocol development will be tested in EFSA's non-application assessments over a one-year period and revised accordingly.

Key words: Evidence use, methods, non-application scientific assessment, planning, problem formulation, protocol

Requestor: EFSA

Question number: EFSA-Q-2020-00274

Correspondence: AMU@efsa.europa.eu

Acknowledgements: EFSA wishes to thank the Scientific Committee members: Vasileios Bampidis, Diane Benford, Jos Boesten, Claude Bragard, Antonio F. Hernandez-Jerez, Susanne Hougaard Bennekou, Kyriaki Machera, Simon John More, Søren Saxmose Nielsen, Josef Rudolf Schlatter, Dieter Schrenk, Vittorio Silano, Dominique Turck, Maged Younes for the support provided to this scientific output and EFSA staff member Denise Candiani for the support provided to this scientific output. The report was endorsed by the Scientific Committee on 20.2.2020.

Suggested citation: EFSA (European Food Safety Authority), Martino L, Aiassa E, Halldórsson TI, Koutsoumanis PK; Naegeli H, Baert K, Baldinelli F, Devos Y, Lodi F, Lostia A, Manini P, Merten C, Messens W, Rizzi V, Tarazona J, Titz A, Vos S, 2020. Draft framework for protocol development for EFSA's scientific assessments. EFSA supporting publication 2020:EN-1843. 46 pp. doi:10.2903/sp.efsa.2020.EN-1843

ISSN: 2397-8325

Summary

This technical report provides a set of recommendations for developing protocols for EFSA's non-application scientific assessments, i.e. those classified in the EFSA Process Architecture (EPA) document (Figure 1) as Scientific Risk Assessment (E2.1); General Scientific and Technical Assistance (E2.2); Scientific Reports (E2.3); Emergency Responses (E3.2); Data Collection and Management (E4.1); Literature Management (E4.2); and Methodologies management (E5.1) including developing or updating methodological guidance documents in EFSA.

The recommendations have been defined by an EFSA's Scientific Committee Working Group and subsequently endorsed by the Scientific Committee. The framework is a draft because it will be refined and published after a trial phase in all EFSA's non-application scientific assessments over a year.

The framework for developing protocols illustrated in this document is intended for EFSA's staff and external experts. However, it may also be useful for contractors working on projects outsourced by EFSA or applicants submitting to EFSA dossiers on regulated products, when the methods for developing the dossier are not fully covered by the relevant legislative and guidance documents.

The report is divided into two parts. The first part describes the rationale for protocols and the second part provides steps and recommendations for developing protocols in EFSA's non-application assessments.

The report clarifies that protocol covers both step 1 the problem formulation (i.e. what the assessment aims to address) and step 2 which methods will be used for addressing the problem (i.e. how the assessment will be carried out). Problem formulation includes: (a) the clarification and acceptance of the mandate that takes place in dialogue with the requestor (not part of the protocol development); and (b) the translation of each mandate term of reference into a scientifically answerable assessment question, the related conceptual model (i.e. the sub-questions and their relationship) and the definition of the overall approach for the assessment (e.g. whether to apply a quantitative, qualitative, semi-quantitative approach or to adopt a tiered approach). Recommendations for phase (b) of problem formulation (Figure 2) are not given in this document being the topic of another EFSA project (EFSA-Q-2019-00256).

The advantages of defining the methods for the assessment upfront are listed in Section 3.1 and represent the motivation for planning prior to starting the assessment. They include guarding against confirmation bias, avoiding hypothesising after the results are known. This is achieved requiring assessors to articulate analytical decisions prior to acquiring knowledge about the available results. As clarified in Section 3.2, the protocol also represents a useful tool to better target the requestor needs and enhance the engagement with a wider community. The first part of the document also provides an illustration of the steps of the process (Figure 2) for EFSA's non-application scientific assessments. After problem

formulation, a method is chosen to address each sub-question. The various approaches adopted in EFSA can be: (1) using evidence extracted from the scientific literature or directly submitted to EFSA; (2) using data from databases other than literature (e.g. Eurostat database); (3) eliciting expert judgement; and (4) carrying out primary research studies. For each approach the steps of data collection, evidence appraisal/data validation, evidence synthesis/data analysis are followed by evidence integration across sub-questions in light of the identified uncertainties. The process is not topic-dependent and the report clarifies that, despite the variety of domains, the type of questions originating from EFSA's mandates and the resulting conceptual models are relatively limited and comparable (Table A.1). Acknowledging across-domain similarities of EFSA's assessment questions helps standardising the way the protocol is developed.

The second part of the document illustrates the actual steps and recommendations for developing and drafting protocols for EFSA's non-application scientific assessments. The elements to include in a protocol are not prescriptive. A central concept to allow flexibility in protocol development is the 'extent of planning', i.e., the degree of detail provided in the protocol for the methods that will be applied, which can be tailored to accommodate the characteristics of the mandate (e.g. the requestor's needs – including the deadline, and the available resources). It is also clarified that the extent of planning in the protocol is not related to the complexity of the methods that will be used in the assessment. For instance, for an urgent request for which the assessment methods are constrained by time limitation and fitness for purpose, a description of evidence needs, task allocation, evidence needs and methods for synthesis and integration should be prioritised in line with the EFSA procedures for responding to urgent advice needs. Box1 outlines the steps for drafting protocols in EFSA's non-application assessments. As part of the problem formulation (step 1), the translation of the mandate into the assessment question and the definition of the sub-questions and their relationship (conceptual model) take place. It is followed by a decision about the overall approach to take in terms of whether to apply a quantitative, qualitative, semi-quantitative approach; to adopt a tiered approach; to prioritise sub-questions over others. Step 2 of protocol development includes detailing: step 2.1 the evidence needs and the methods for collecting, assessing and synthesising evidence including uncertainty analysis within each sub-question; step 2.2 the methods for integrating evidence across sub-questions and addressing the remaining and overall uncertainty. Recommended content for a protocol, by approach, step and extent of planning for each sub-question (step 2.1) are detailed in Tables 1-4. Table 5 illustrates the recommended content for a protocol by extent of planning for integration across sub-questions. Appendix B gives an overview of methods for evidence synthesis and integration.

Table of contents

1. Introduction

The European Food Safety Authority (EFSA) contributes to the safety of the EU food chain primarily by providing scientific advice on food and feed safety matters and communicating on that advice.

EFSA operates in a "mandate-driven" environment and produces what in general terms is referred to as "regulatory science", i.e. scientific assessments and recommendations in support of managerial decisions. A fundamental requirement in this context is the production of fit-for-purpose scientific assessments that are timely and meet the requestors' needs.

During 2014-2018, EFSA defined a series of principles for the scientific assessment process and developed and tested a 4-step approach (plan/do/verify/report) to help fulfil those principles (EFSA, 2015). This approach, which is illustrated in Figure 2, puts an emphasis on the need to plan the methods for the scientific assessment in a protocol before starting the assessment. The result of the test period demonstrated that the approach is beneficial for increasing the methodological rigour, impartiality and transparency of EFSA's scientific assessments and at the same time ensures fitness for purpose (EFSA, 2015, 2016, 2018). The principles for the scientific assessment process and the 4-step approach also serve as a basis for revising EFSA's Quality Policy and Quality Management System.

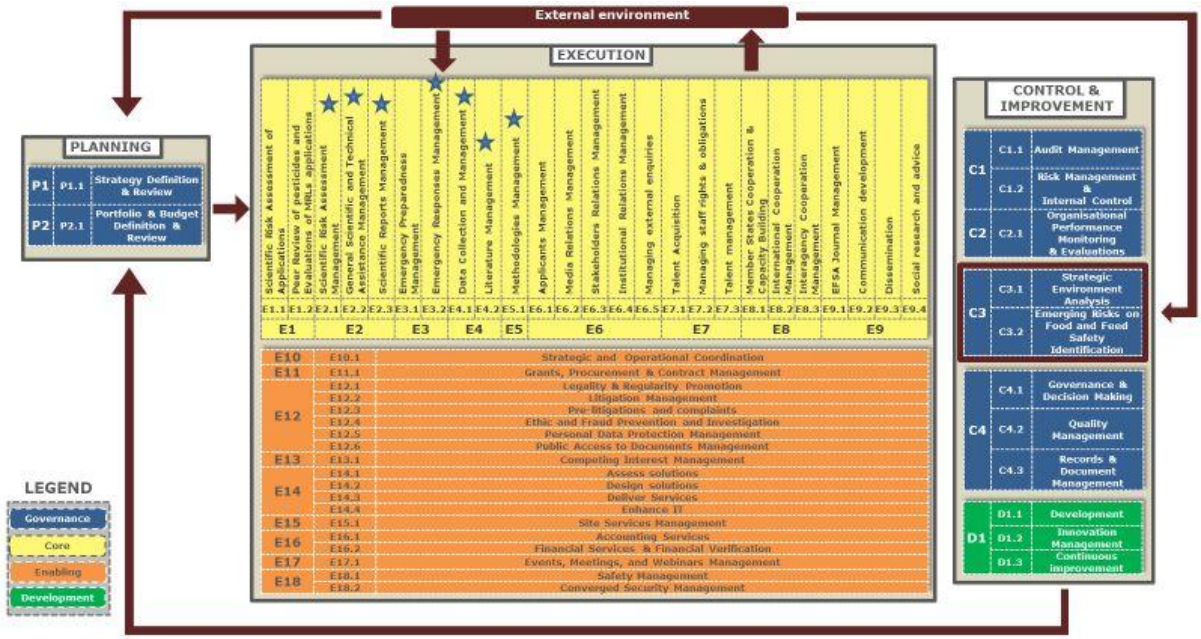
1.1. Background and Terms of Reference as provided by the requestor

In view of the benefits identified at the end of the PROMETHEUS testing phase, EFSA acknowledged the value of gradually implementing the 4-step approach as described in the Principles and process for dealing with data and evidence in scientific assessments (EFSA, 2015), starting from 2020, in its non-application scientific assessments. The non-application scientific assessments include those classified in the EFSA Process Architecture (EPA) document (Figure 1) as Scientific Risk Assessment (E2.1); General Scientific and Technical Assistance (E2.2); Scientific Reports (E2.3); Emergency Responses (E3.2); Data Collection and Management (E4.1); Literature Management (E4.2); and Methodologies management (E5.1). As for EFSA's assessments of regulated products, the trial phase of the 4-step approach (EFSA, 2018) showed that for these assessments the protocol is already described with different levels of detail, at least regarding the data requirements and methods for collecting those data, in the relevant EU legislation and technical documents (e.g. EFSA guidance document). The same applies to the types of generic scientific assessments where methods and data requirements are described in detail in EFSA guidance documents (e.g. EFSA PLH Panel, 2018) and/or legislation.[1] Nevertheless, recommendations for protocols were **not**

[1] e.g. for plant health see: Regulation (EU) 2016/2031 of the European Parliament of the Council of 26 October 2016 on protective measures against pests of plants, amending Regulations (EU) 228/2013, (EU) 652/2014 and (EU) 1143/2014 of the European Parliament and of the Council and repealing Council Directives 69/464/EEC, 74/647/EEC, 93/85/EEC, 98/57/EC, 2000/29/EC, 2006/91/EC and2007/33/EC. OJ L 317, 23.11.2016, pp. 4–104; Commission Implementing Regulation (EU) 2018/2019 of 18 December 2018 establishing a provisional list of high risk plants, plant products or other objects, within the meaning of Article 42 of Regulation (EU) 2016/2031 and a list of plants for which phytosanitary certificates are not required for introduction into the Union, within the meaning of Article 73 of that Regulation C/2018/8877. OJ L 323, 19.12.2018, pp. 10–15; Commission Implementing Regulation (EU) 2018/2018 of 18 December 2018 laying down specific rules concerning the procedure to be followed in order to carry out the risk assessment of high risk plants, plant products and other objects within the meaning of Article 42(1) of Regulation (EU) 2016/2031 of the European Parliament and of the Council. C/2018/8876, OJ L323, 19.12.2018, pp. 7–9.

deemed necessary for: (a) assessments of regulated products; and (b) generic assessments where methods and data requirements are already described in detail in relevant legislation and technical documents.

Figure 1: EFSA Process Architecture 2019/2020



Note: The blue star indicates the relevant processes

In light of the decision of implementing the 4-step approach in all non-application assessments, recommendations on how to develop and draft protocols were considered needed. To this scope, EFSA mandated the Scientific Committee Panel to establish a working group in collaboration with the Assessment and Methodological Support Unit (AMU) responsible for defining the main content and structure of EFSA's protocols for scientific assessments, including 'hints' on problem formulation (as this part will be further developed and completed in the project 'EFSA Framework for problem formulation').

1.2. Objectives, intended users, structure and next steps

This technical report is intended to provide **recommendations for** the **development of protocols for EFSA's non-application scientific assessments**, ensuring flexibility and enhancing harmonisation across EFSA's areas. After a one year trial phase in all EFSA's non-application assessments, during which recommendations will be tested and operationalised, the draft framework will be evaluated and finalised.

The recommendations are not prescriptive. They are intended for EFSA's staff and external experts. However, they may also be useful for other parties, such as contractors working on projects outsourced by EFSA, or those submitting dossiers to EFSA on regulated products, when the application process is not fully covered by the relevant legislative and guidance documents. The recommendations may also apply when developing or updating methodological guidance documents in EFSA.

The technical report is divided into two parts.

- The first part describes the rationale for protocol development in EFSA's non-application assessments. It clarifies that the planning can be applied to all domains within EFSA independent of the complexity of the scientific assessment being considered;
- The second part illustrates the steps and recommendations for developing and drafting protocols. Recommendations highlight that the protocol must state upfront: (a) the scientific question to address (WHAT); and (b) which methods will be used and how they will be applied in the scientific assessment process (HOW).

2. Data and Methodologies

To address the terms of reference of the mandate, EFSA established an *ad hoc* working group under its Scientific Committee composed by external experts and staff, that met regularly to share experiences gained by the individual members in using protocols in EFSA's scientific assessments.[2]

The methodology applied by the working group to develop this document consisted of discussions on the advantages, difficulties and best practices in designing upfront the methods for EFSA's non-application scientific assessments, taking into account the outcomes of a previous EFSA project aimed to define and test the use of protocols in EFSA's assessments (EFSA, 2015, 2016, 2018), along with relevant information reported in the scientific literature.

The staff members of the working group ensured proper reflection of the needs and view of their respective scientific units/teams. The technical report was endorsed by EFSA's Scientific Committee and is published as a draft because the framework for protocol development will be tested in EFSA's non-application assessments over a one-year period. Pending a decision of EFSA's management, dedicated workshops might be organised targeting each EFSA unit and the related panel to collect feedback after the testing phase on this draft framework on protocol development. A revised version is planned at the end of the testing phase also considering the outcome of the project on problem formulation.[3]

3. PART I: RATIONALE FOR PROTOCOLS in EFSA's non-application assessments

3.1. Scientific assessment process for EFSA's non-application assessments: planning as key phase

EFSA's non-application scientific assessments always start with a problem formulation phase that includes: (a) the clarification and acceptance of the mandate that takes place in dialogue with the requestor (not covered here, since it is not part of protocol development); and (b) the translation of each mandate term of reference (ToR) into a scientifically answerable assessment question, the definition of the related conceptual model and selection of the

[2] https://www.efsa.europa.eu/sites/default/files/wgs/cross-cutting-science/wg-protocol-development.pdf
[3] http://registerofquestions.efsa.europa.eu/roqFrontend/questionLoader?question=EFSA-Q-2019-00256

overall approach for the assessment (Figure 2). The description of the process of problem formulation is the topic of another EFSA project and will be detailed in this context (EFSA-Q-2019-00256)

The conceptual model illustrates all the sub-questions derived from breaking down the assessment question, along with a description (logical or mathematical) of their relationships (based on EFSA, 2015). A conceptual model can vary from being purely descriptive to mathematical with all the sub-questions expressed as parameters/variables. The pathway to harm (i.e. a causal chain of events that need to occur for a harm to be realised) represents an example of a conceptual model (e.g., Devos et al., 2019). A well-described conceptual model is essential for defining the evidence needs for each sub-question.

EFSA's mandates are very diverse and pertain to areas of food and feed safety, including animal health and welfare, plant protection, plant health, chemical and biological hazard, human and animal nutrition, the environment and emerging risks. Despite the large variety of topics, the type of questions originating from EFSA's mandates and the resulting conceptual models are relatively limited and comparable across domains (Table A.1). **Understanding the across-domain similarities of EFSA's assessment questions and related conceptual models helps standardise the problem formulation step during protocol development.**

EFSA recognised the value of complementing problem formulation (i.e. description of the *what*) with an upfront definition of methods for conducting the assessment (i.e. description of the *how* the assessment will be conducted), which include the methods for answering each sub-question and for integrating evidence across sub-questions, including uncertainty analysis (EFSA, 2015, 2016, 2018). Uncertainty has been defined by EFSA as "all types of limitations in available knowledge that affect the range and probability of possible answers to an assessment question". It can arise from limitations in the evidence (i.e. heterogeneity, degree of relevance, degree of internal validity and/or precision) and in the methods used throughout the assessment (EFSA Scientific Committee, 2018a-b).

Designing the methods upfront is a well-established practice in primary research and systematic reviews (Higgins et al, 2019). Moreover, there is a growing number of initiatives promoting and implementing this practice for other assessments in contexts similar to EFSA's (e.g. World Health Organisation, 2014; Woodruff and Sutton, 2014; OHAT-NTP, 2019; US-EPA IRIS[4]). Advantages of planning upfront the methods for the scientific assessment process include:

- Safeguarding against arbitrary decision making during the assessment process;
- Protecting from cognitive biases (Munafo et al., 2017; Shamseer et al., 2015) such as the confirmation bias, i.e. the tendency to focus on evidence that is in line with expectations or favoured explanation (Kerr, 1998);

[4] https://www.epa.gov/iris/iris-program-outlook

- Limiting methodological flaws like HARKing or data-contingent analysis decisions (P-hacking) by requiring assessors to articulate analytical decisions prior to acquiring knowledge about the available results so that these decisions remain data-independent (Munafo et al., 2017);
- Streamlining the implementation of the scientific assessment process. In summary, "*asking questions at the design stage can save headaches at the analysis stage: careful data collection can greatly simplify analysis and make it more rigorous*" (Kass, 2016).

Together, the description of the "what" (i.e. the part of the problem formulation that takes place after the mandate has been accepted) and the "how" represent the **protocol** for the assessment, which is developed during the planning phase of the scientific assessment process (Figure 2).

Protocol development is typically iterative and may take some time to be finalised, requiring input from all the expertise needed for the assessment itself and through scoping the literature or other types of evidence. Several consultations may be necessary within the expert group responsible for conducting the assessment to agree on a final protocol. A dialogue with the mandate requestor, especially during problem formulation, might also be needed to ensure the assessment question appropriately reflects the request. In some cases the draft protocol can be shared with a target audience (e.g. selected stakeholders) or made available for public consultation before finalization in order to collect feedback prior to the start of the assessment. This can be particularly useful when the topic of the mandate is sensitive and controversial. Once the protocol is finalised its implementation starts. During this phase there might be a need to revise the protocol in light of unforeseen elements. Deviations from the protocol can occur and are acceptable provided they are justified and documented.

The planning phase ("plan") is followed by the actual assessment, where the methods pre-defined in the protocol are implemented and conclusions are drawn in light of the identified uncertainties ("do"). Subsequently, compliance with the plan is checked and ensured ("verify") and the methods, assumptions, data, results and conclusions are reported and published ("report"). These phases of the assessment are detailed in Figure 2. During the "do-verify-report" phase, the different approaches to answer sub-questions in EFSA's non-application assessments are:

1. Collecting, appraising, synthesising and integrating evidence coming from the scientific literature or directly submitted to EFSA;[5]
2. Extracting, assessing and analysing data from databases other than literature (e.g. Eurostat database);[6]
3. Eliciting expert judgement, when evidence is scarce and/or of limited validity (EFSA, 2014);[7]

[5] Evidence is sometimes collected by EFSA launching calls for data that can yield heterogeneous types of evidence (e.g. papers published or not in literature databases, individual and aggregated data from unpublished studies). In addition, data for non-application assessments can be extracted from dossiers submitted to EFSA in regulatory applications, regular monitoring processes and data collections, EFSA networks, etc.

[6] https://ec.europa.eu/eurostat/data/database

[7] Expert Knowledge Elicitation (EKE) is also one of the methods that can be used for i) the problem formulation phase (e.g. for prioritising questions for more formal approaches), ii) synthesising and integrating evidence accounting for uncertainty (e.g.

4. Carrying out primary research studies.[8]

Combination of different approaches can be adopted for the same sub-question or, for broad assessments containing many sub-questions, for the various sub-questions.

EFSA's documents that provide guidance on how to conduct the scientific assessment process for all types of EFSA's non-application assessments are: the scientific report on principles and process for dealing with data and evidence (EFSA, 2015[9]); the Scientific Committee guidance documents on uncertainty (EFSA Scientific Committee, 2018b-c), weight of evidence (EFSA Scientific Committee, 2017a), and biological relevance (EFSA Scientific Committee, 2017b); and the EFSA guidance on expert knowledge elicitation (EKE) (EFSA, 2014).

[1] This is intended as the one defined throughout an iterative process and established at the end of it, before the start of the implementation phase

uncertainty arising from the evidence) and for iii) replacing data collections and analyses under pressing conditions. In addition, expert judgement is applied in any step if the scientific assessment process when subjective decisions are required.

[8] In EFSA, these are typically outsourced.

[9] This report describes EFSA's scientific assessment process, its guiding principles and the 4-step approach (plan/do/verify/report) to help fulfilling those principles.

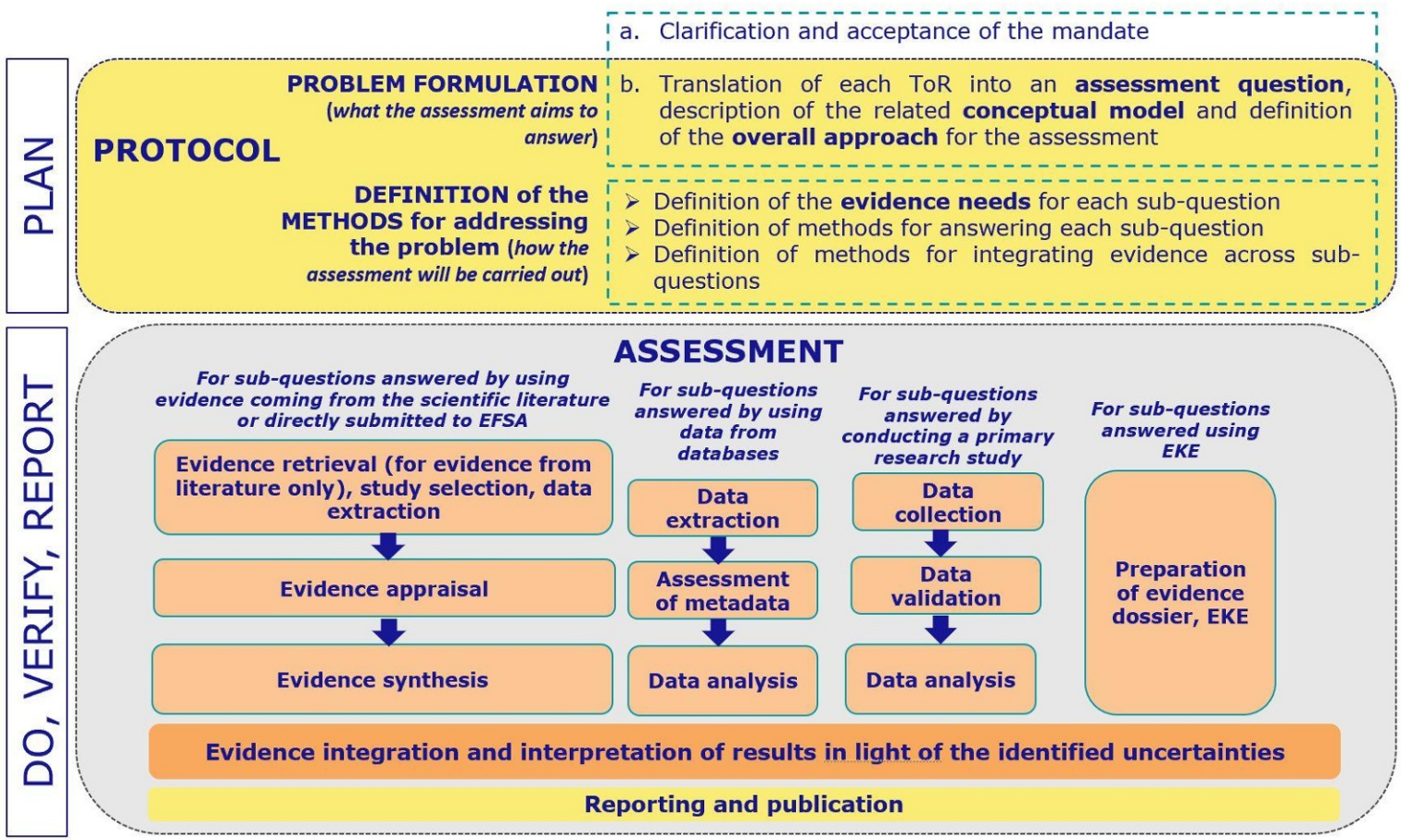


Figure 2: The scientific assessment process for EFSA non-application assessments

3.2. Planning and engaging upfront with the mandate requestor and the wider community

If protocols are exchanged with the requestor, they represent useful tools for ensuring, before starting the assessment, that the assessment will answer the question originally posed, and the plan would lead to an output that targets the requestor's needs.

In addition to the mandate requestor, draft protocols can be shared with other external parties to receive feedback and input on the methods to use for the scientific assessment and, if appropriate, refine them before starting the assessment. The extent of the consultation can vary depending on the context of the assessment (e.g. targeted consultation of relevant stakeholders or extensive public consultation).

4. PART II: STEPS and RECOMMENDATIONS for DEVELOPING and DRAFTING PROTOCOLS FOR EFSA'S NON-APPLICATION ASSESSMENTS

This section illustrates the actual steps and recommendations for developing and drafting protocols for EFSA's non-application scientific assessments (Box 1).

The recommendations are not prescriptive and allow flexibility.

It is always possible to revise the protocol once the implementation has started provided that the deviations are justified and documented.

The extent of planning upfront in a protocol depends on the context and characteristics of the assessment and the amount and degree of heterogeneity of the data that will be used. The more information one has in advance on the amount of data and its heterogeneity, the easier it is to develop ahead the most appropriate plan.

For all EFSA's domains, the **extent of planning in the protocol** (i.e. **the degree of detail provided in the protocol for the methods that will be applied in the assessment**) can be tailored to accommodate the characteristics of the mandate (e.g. the requestor's needs – including the deadline, and the available resources). For instance, the extent of planning can vary from being very comprehensive (e.g. thorough description of the final search strategies, tools for appraising evidence, statistical analysis plan) to generic outlines of the approach for performing the assessment and its rationale, limitations and related sources of uncertainty. The reasons for limited planning can be multiple: e.g. no resources, no time, or no need.

The **extent of planning in the protocol is not related to the complexity of the methods and tools that will be used in the assessment** (e.g. how extensively the search will be done, how complex will be the model to analyse the data and the related uncertainties). For instance, for an urgent request for which the assessment methods are constrained by time limitation and fitness for purpose, a description of evidence needs, task allocation and methods for synthesis and integration should be prioritised, in line with the EFSA procedures for responding to urgent advice needs (EFSA, 2019).

Box 1: Outline of the steps for developing and drafting protocols in EFSA's non-application assessments

Steps for protocol development

Step 1: Formulate the problem – phase (b)[1] of problem formulation (*what*)

Step1.1: Translate the mandate into assessment question(s);

Step1.2: Define the sub-questions of each assessment question and their relationship (conceptual model);

Step1.3: Select the approach to take i.e. whether:

- to apply a quantitative, qualitative, semi-quantitative approach;
- to adopt a tiered approach (from more conservative to more refined) for high-order sub-questions;
- to prioritise sub-questions over others.

The approach can refer to the scientific assessment overall or be detailed by individual sub-question. In the latter case a combination of approaches is possible.

Step 2: Plan the methods for conducting the assessment (*how*)

Step 2.1: Detail as appropriate the evidence needs and the methods for answering each sub-question, including uncertainty analysis;

Step 2.2: Detail as appropriate the methods for integrating evidence across sub-questions and addressing the remaining and overall uncertainty.

[1] Phase (a) of the problem formulation 'the clarification and acceptance of the mandate' that takes place in dialogue with the requestor is not covered in this document since it is not part of protocol development

Step 1: Formulate the problem

This section contains only some generic considerations on this step of protocol development, as EFSA has outsourced a project aimed at defining a detailed framework for problem formulation.[10] Once the project is completed, a summary of the final report will be included here.

After translating the mandate into as many assessment questions as needed (step 1.1 in Box 1) and defining the sub-questions and the related relationship (step 1.2 in Box 1), during step 1.3 of problem formulation (Box 1), the overall approach for the assessment is selected, including:

- The choice of a quantitative, qualitative or semi-quantitative approach. This choice will influence all the following decisions at protocol development level and during the conduct of the assessment;
- The application of a tiered approach. During problem formulation, it can be decided to adopt a tiered approach (from more conservative to more accurate) for high order sub-questions (Table A.1). For instance, for exposure assessment it can be decided to start

[10] http://registerofquestions.efsa.europa.eu/roqFrontend/questionLoader?question=EFSA-Q-2019-00256

from a conservative worst-case scenario including extreme centiles of the consumers' distribution to move, only if needed, to more accurate and less conservative estimates

- The prioritisation of the sub-questions, defined on the basis e.g. of the anticipated impact of the sub-question on the conceptual model and the scientific controversy (adapted from (EFSA, 2010)). Some criteria for prioritising sub-questions are illustrated in EFSA's systematic review guidance (EFSA, 2010) and a subsequent related project (O'Connor et al., 2012).

Step 2: Plan the methods for conducting the assessment

The recommendations on the elements to include in the protocol for the methods for conducting the assessment are illustrated in:

- Tables 1-4. Methods for collecting, assessing and synthesising evidence (including uncertainty analysis) within each sub-question (step 2.1 in Box 1);
- Table 5. Methods for integrating evidence across sub-questions and addressing the remaining uncertainty (step 2.2 in Box 1).

The 'recommended content for the protocol' in Tables 1-5 is intended as the one defined throughout an iterative process and established **at the end of it**, prior to the start of the implementation phase.

In each step of the scientific assessment process, the extent of planning can be "low" or "high" as highlighted in Tables 1-5. The low extent implies case-specific simplifications decided by the experts responsible for the assessment.

Step 2.1: Detail the evidence needs and the methods for answering each sub-question question including uncertainty analysis, as appropriate

Tables 1-4 illustrates the recommended elements to include in a protocol, tailored for the four possible types of approach to answering a sub-question (Figure 2), i.e.: (1) by using data extracted from literature or directly submitted to EFSA; (2) by using data extracted from databases other than literature; (3) by eliciting expert judgement; and (4) by conducting a primary research study).

For completeness, Tables 1-4 also reports the "formulation of the sub-question" step, though this step is conducted during phase (b) of Step 1 problem formulation.

As emphasised above, for each approach, the extent of planning can be low or high to allow flexibility in the level of the detail in the protocol that might be requested in each mandate depending on several factors.

In a broad assessment containing multiple sub-questions, the extent of planning of the methods can vary **by sub-questions**, depending on their relative "priority" (see problem formulation), but also **within the same sub-question and approach**, by step of the approach (first column in Tables 1-4). For instance, for the same sub-question in the protocol,

there can be an extensive plan of the methods for data collection and a generic description of the ones for evidence synthesis.

Call for data are also planned in the protocol and, depending on the data that will be submitted to EFSA (i.e. studies published in scientific papers or data from unpublished studies), the steps for planning will follow approach 1 and/or 2.

The type and extent of heterogeneity in the evidence can guide the planning of the methods for the synthesis within a sub-question. A stepwise approach can be indicated in the protocol for the evidence synthesis with alternative methods planned depending on whether the evidence is sufficiently homogeneous or not (e.g. quantitative methods such as meta-analysis can be foreseen in case of a body of evidence showing low heterogeneity, a qualitative synthesis when this condition is not met). The type of data (e.g. aggregated versus individual data) can also play a role when planning a method for the analysis including methods for addressing the uncertainty. An overview of methods for evidence synthesis is provided in Appendix B (together with methods for evidence integration).

As for the **uncertainty analysis**, the protocol illustrates, to the extent possible, the following elements with varying detail depending on the context of the assessment:

- Expected sources of uncertainty (which can arise from limitations in the evidence and in the methods in each step of the assessment process);
- Methods to assess the influence of the individual sources of uncertainty on the conclusions, in order to prioritise the most influential ones;
- Methods to analyse uncertainty sources individually and combined.

Step 2.2: Detail the methods for integrating evidence across sub-questions and addressing the remaining uncertainty, as appropriate

The plan covers also the methods that will be used for evidence integration across sub-questions and for accounting for the remaining uncertainty (Table 5).

A well-defined conceptual model is essential to plan and guide the process for evidence integration.

An overview of methods for evidence integration is provided in Appendix B.

Table 1: Recommended content for a protocol, by step and extent of planning when using **evidence from the scientific literature or directly submitted to EFSA for answering a sub-question**

Step of the approach	Recommended content for the protocol for answering each sub-question[1] Reference document: EFSA, 2010a	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate)[2]	**Extent of planning: HIGH**
Formulation of the sub-question	State the main objective(s) of the sub-question	• State the main objective(s), where appropriate in a single concise sentence • Formulate the sub-question into clearly defined key-elements (i.e. PICO, PECO, PIT, PO[3] • If applicable, define in advance which outcomes are primary outcomes and which are secondary outcomes
a. Definition of the eligibility criteria for study selection (i.e. evidence needs)	Briefly describe the evidence needs or justify why it is not possible or needed to plan	• Describe all eligibility criteria for study selection (i.e. the criteria related to study e.g. target population, intervention/exposure of interest, and the relevant outcomes and record characteristics e.g. time, language, publication type) • Provide the rationale for the choice of the eligibility criteria
b. Definition of the search strategy (only for literature-based approaches)	Briefly describe the search strategy or justify why it is not possible or needed to plan	• Describe the final search strategy in the protocol, i.e. search string(s) including planned limitations • Indicate the information sources (bibliographic databases and grey literature resources) that will be searched • Describe any other search approaches (e.g. citation indexes, hand-searching) • Indicate any software (e.g. for reference management) that will be used

Step of the approach	Recommended content for the protocol for answering each sub-question[1] Reference document: EFSA, 2010a	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate)[2]	**Extent of planning: HIGH**
c. Definition of the methods for selecting studies for inclusion/exclusion	Briefly describe the methods for study selection or justify why it is not possible or needed to plan	• Indicate the number of reviewers • Describe the method for study selection e.g. in parallel or not • If applicable, describe how conflicts will be solved, if and what Artificial Intelligence techniques will be used • Indicate the software that will be used for screening papers
d. Definition of the methods for extracting data from included studies	Briefly describe the data extraction process or justify why it is not possible or needed to plan	• Describe the main characteristics of data model (i.e. what data will be extracted from the included studies) • Indicate how data will be extracted (e.g. by two independent reviewers in parallel or one reviewer extracting and one validating the process) • Indicate the software that will be used for data extraction
e. Definition of the methods for appraising evidence (i.e. of the methods for identifying the uncertainty in the evidence)	Briefly describe the methods for evidence appraisal or justify why it is not possible or needed to plan	• Describe the finalised and customised version of the Critical Appraisal Tool (CAT) that will be used, for each study type/design • Indicate how overall conclusions on each individual study validity, by outcome, will be drawn • Indicate how study appraisal will be performed (e.g. in parallel for all studies - two independent reviewers) • Indicate how conflicts will be solved • Indicate software that will be used for evidence appraisal
f. Preliminary identification of the sources of uncertainty and definition of the methods for prioritising them[4]	Briefly describe the methods for prioritising uncertainty sources or justify why it is not possible or needed to plan	• List the main uncertainty sources • Describe the methods that will be used for prioritising the uncertainties

Step of the approach	Recommended content for the protocol for answering each sub-question[1] Reference document: EFSA, 2010a	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate)[2]	**Extent of planning: HIGH**
g. Definition of the methods for synthesising evidence within the sub-question	• Briefly describe the methods for evidence synthesis or justify why it is not possible or needed to plan • At least, indicate whether methods used for the synthesis are qualitative, quantitative or semi-quantitative	• Indicate whether the methods for the synthesis will be qualitative, quantitative or semi-quantitative • If applicable, describe the mathematical model that will be used (e.g. dose-response meta-analytic model) • Discuss biological relevance and plausibility of the possible results including level of the effect considered biologically relevant (if any) • Indicate how the results of evidence appraisal will be accounted for in the synthesis (e.g. sensitivity analysis, sub-group analysis, bias-adjusted meta-analysis) • Indicate the software that will be used for evidence synthesis
h. Definition of the methods for analysing uncertainties individually and combined[5]	• Briefly describe the methods for uncertainty analysis or justify why it is not possible or needed to plan • At least, indicate whether methods used for the uncertainty analysis are qualitative, quantitative or semi-quantitative	• Indicate whether the methods used for the uncertainty analysis are qualitative, quantitative or semi-quantitative • describe the methods to analyse uncertainty sources individually and combined • Specify whether variability and uncertainty will be addressed separately and in case how

[2] Case-specific simplifications agreed by the WG. If appropriate, refers to the case when all steps of the approach are addressed in the protocol with a low extent of planning. It doesn't apply to cases when different steps are addressed with different extents of planning

[3] Reference document: (EFSA, 2010a).

[4] Reference document: EFSA Guidance on Uncertainty Analysis in Scientific Assessments (EFSA, 2018b-c)
[5] Reference document: EFSA Guidance on Uncertainty Analysis in Scientific Assessments (EFSA, 2018b-c)

Table 2: Recommended content for a protocol, by step and extent of planning when using **data from databases other than literature for answering a sub-question**

Step of the approach	Recommended content for the protocol for answering each sub-question[1] Reference documents: DAMA-international, 2017; EFSA, 2010b	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate[2])	**Extent of planning: HIGH**
Formulation of the sub-question	State the main objective(s) of the sub-question	• State the main objective(s), where appropriate in a single concise sentence • Formulate the sub-question into clearly defined key-elements (i.e. PICO, PECO, PIT, PO[3]) • If applicable, define in advance which outcomes are primary outcomes and which are secondary outcomes
a. Definition of evidence needs based on the sub-question formulation	Briefly describe the evidence needs or justify why it is not possible or needed to plan	• Describe the evidence needs up to the level of variables, geographical and temporal coverage, levels of representativity etc. • Define the requested level of granularity of the evidence needs (e.g. species, sub-groups of the population). Describe how to re-cluster/re-group in case of need.
b. Identification of the adequate source of data - database	Briefly describe the criteria for selecting databases or justify why it is not possible or needed to plan	• Describe the criteria that will be used to evaluate relevance and representativeness of a source of data (database) and to decide whether to use it or not for the assessment
c. Definition of the data model to extract data from the selected databases	Briefly describe the characteristics of the data to be extracted or justify why it is not possible or needed to plan	• Describe the data models[4] detailing the list of variables, their format and related metadata (e.g. methods used to collect data, timeframe, reference population etc.) as required to address the specific sub-question • Develop/adapt/adopt (when available from EFSA Catalogues, OECD phrase lists, Global Agricultural Concept scheme etc.) classification schemes • Define the expected date for download from the database

Step of the approach	Recommended content for the protocol for answering each sub-question[1] Reference documents: DAMA-international, 2017; EFSA, 2010b	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate[2])	**Extent of planning: HIGH**
		• Define the export format considering transparency requirements for publication of data and supplementary materials
d. Plan for data check and validation (i.e. identification of the uncertainty in the evidence)	Briefly describe the data validation process or justify why it is not possible or needed to plan	• Define mandatory values, closed terminology lists and validation rules that will be used to validate the data • Describe how metadata will be assessed to evaluate validity and precision of the data; • Define pre-processing procedures (i.e. any data transformations), uniformity of statistical indicators and measurement units (only in case more than one database is used) that will be applied to the data
e. Preliminary identification of sources of uncertainty and definition of the methods for prioritizing them[5]	Briefly describe the methods for prioritising uncertainty sources or justify why it is not possible or needed to plan	• List the main uncertainty sources • Describe the methods that will be used for prioritising the uncertainties
f. Definition of the methods for evidence synthesis within the sub-question	• Briefly describe the methods for evidence synthesis or justify why it is not possible or needed to plan • At least, indicate whether methods used for the synthesis are qualitative, quantitative or semi-quantitative	• Indicate whether the methods for the synthesis will be qualitative, quantitative or semi-quantitative • If applicable, describe the mathematical model that will be used (e.g. dose-response meta-analytic model) • Discuss biological relevance and plausibility of the possible results including level of the effect considered biologically relevant (if any) • Indicate the software that will be used for evidence synthesis

Step of the approach	Recommended content for the protocol for answering each sub-question[1] Reference documents: DAMA-international, 2017; EFSA, 2010b	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate[2])	**Extent of planning: HIGH**
g. Definition of the methods to analyse uncertainties individually and combined[6]	• Briefly describe the methods for uncertainty analysis or justify why it is not possible or needed to plan • At least, indicate whether methods used for the uncertainty analysis are qualitative, quantitative or semi-quantitative	• Indicate whether the methods used for the uncertainty analysis are qualitative, quantitative or semi-quantitative • Describe the methods to analyse uncertainty sources individually and combined • Specify whether variability and uncertainty will be addressed separately and in case how

[1] This is intended as the one defined throughout an iterative process and established at the end of it, before the start of the implementation phase.

[2] Case-specific simplifications agreed by the WG. If appropriate, refers to the case when all steps of the approach are addressed in the protocol with a low extent of planning. It doesn't apply to cases when different steps are addressed with different extents of planning.

[3] Reference document: (EFSA, 2010a).

[4] Data models may be already available in EFSA or by other researchers working on the same topic or on data standard sharing sites e.g. https://fairsharing.org/search/?q=additives or https://joinup.ec.europa.eu/ or it might be needed to develop them from scratch.

[5] Reference document: EFSA Guidance on Uncertainty Analysis in Scientific Assessments (EFSA, 2018b-c)

[6] Reference document: EFSA Guidance on Uncertainty Analysis in Scientific Assessments (EFSA, 2018b-c)

Table 3: Recommended content for a protocol, by step and extent of planning **when using expert judgement for answering a sub-question**

Step of the approach	**Recommended content for the protocol for answering each sub-question**[1] Reference documents: EFSA, 2014; EFSA Scientific Committee, 2018a-b	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate[2])	**Extent of planning: HIGH**
Formulation of the sub-question	State the main objective(s) of the expert judgement	Translate the sub-question into a scientific question that can be addressed with an expert knowledge elicitation including the parameter/variable to be estimated by the expert and, if relevant, the reference population, the reference time, the geographical area, etc.
a. Definition of the approach	Briefly describe the approach that will be used for eliciting expert knowledge or justify why it is not possible or needed to plan	Specify: • The overall approach for the expert elicitation (e.g. level of formality including whether a Steering Group and an Elicitation Group will be set up, an external facilitator will be involved); • The method for eliciting the parameter including whether Sheffield, Cooke, Delphi or another approach will be used • the uncertainty in the method that can be identified at this stage
b. Identification of the experts	Briefly describe the criteria for experts selection or justify why it is not possible or needed to plan	Describe the expertise requirements and the methods to assess them
c. Preparation of the evidence dossier	Briefly describe the content of the evidence dossier or justify why it is not possible or needed to plan	Provide the table of content of the evidence dossier that will describe the evidence and the related uncertainties that will be used to support the elicitation process.

Step of the approach	Recommended content for the protocol for answering each sub-question[1] Reference documents: EFSA, 2014; EFSA Scientific Committee, 2018a-b	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate[2])	**Extent of planning: HIGH**
d. Definition of the methods for the synthesis of the individual expert estimates and their uncertainty	Briefly describe the methods for synthesising individual expert distributions/estimates or justify why it is not possible or needed to plan	• Describe the method that will be used to summarise the individual expert uncertainty distribution/estimates (e.g. mathematical synthesis, collegial discussion to achieve a consensus on the distribution centiles) • Indicate how many rounds will be used to achieve consensus

1 This is intended as the one defined throughout an iterative process and established at the end of it, before the start of the implementation phase.

2 Case-specific simplifications agreed by the WG. If appropriate, refers to the case when all steps of the approach are addressed in the protocol with a low extent of planning. It doesn't apply to cases when different steps are addressed with different extents of planning

Table 4: Recommended content for a protocol, by step and extent of planning **when conducting a primary research study for answering a sub-question**

Step of the approach	Recommended content for the protocol for answering each sub-question[1]	
	Extent of planning: LOW (justification for low extent of planning can be one overall for all steps if appropriate[2])	**Extent of planning: HIGH**
Formulation of the sub-question	• State the main objective(s) of the sub-question • Describe evidence needs in general terms without detailing all the aspects that might characterize the evidence. Provide a rationale for not providing details.	• Formulate the sub-question into clearly defined key-elements (e.g. PICO, PECO, PIT, PO[3]) • State the main and secondary objective(s), also those related to specific sub-populations • If applicable, define in advance which outcomes are primary outcomes and which are secondary outcomes
a. Design and conduct of the study	Refer to Manuals providing guidance on designing experimental and observational studies	

[1] This is intended as the one defined throughout an iterative process and established at the end of it, before the start of the implementation phase.

[2] Case-specific simplifications agreed by the WG. If appropriate, refers to the case when all steps of the approach are addressed in the protocol with a low extent of planning. It doesn't apply to cases when different steps are addressed with different extents of planning.

[3] EFSA guidance on systematic review (EFSA, 2010a).

Table 5: Recommended content for a protocol by extent of planning **for integration across sub-questions**

Steps	Recommended content for the protocol for integration across sub-questions[1]	
	Extent of planning: LOW[2]	**Extent of planning: HIGH**
a. Evidence integration across sub-questions	• Briefly describe the methods for evidence integration or justify why it is not possible or needed to plan • At least, indicate whether methods used for the integration are qualitative, quantitative or semi-quantitative	• Indicate whether the methods for the integration will be qualitative, quantitative or semi-quantitative • If applicable, describe the mathematical model that will be used; • If applicable, describe how indirect evidence will be extrapolated to the target population • Discuss biological relevance and plausibility of the possible results including level of the effect considered biologically relevant (if any)
b. Uncertainty assessment across sub-questions	• Briefly describe the methods for uncertainty and sensitivity analysis or justify why it is not possible or needed to plan	• Describe methods to analyse uncertainty sources collectively • If applicable, describe methods for sensitivity analysis to identify most influential sources of uncertainty

1 This is intended as the one defined throughout an iterative process and established at the end of it, before the start of the implementation phase

2 Case-specific simplifications agreed by the relevant Working Group.

References

DAMA-international, 2017: DAMA-DMBOK: Data Management Body of Knowledge: 2nd Edition. Technics Publications, Basking Ridge, New Jersey

EFSA (European Food Safety Authority), 2010a. Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal 2010; 8(6):1637, 90pp. doi:10.2903/j.efsa.2010.1637

EFSA (European Food Safety Authority), 2010b. Guidance on Data Exchange. EFSA Journal 2010; 8(11):1895, 50 pp. doi:10.2903/j.efsa.2010.1895

EFSA (European Food Safety Authority), 2014. Guidance on expert knowledge elicitation in food and feed safety risk assessment. EFSA Journal 2014;12(6):3734, 278 pp. doi:10.2903/j.efsa.2014.3734

EFSA (European Food Safety Authority), 2015. Principles and process for dealing with data and evidence in scientific assessments. EFSA Journal 2015;13(5):4121, 35 pp. doi:10.2903/j.efsa.2015.4121

EFSA (European Food Safety Authority), 2016. Analysis of EFSA methodological needs for evidence use in scientific assessments. EFSA supporting publication 2016:13(10):EN-1092, 55pp. doi:10.2903/ sp.efsa.2016.EN-1092

EFSA (European Food Safety Authority), Gundert-Remy U, Bodin J, Bosetti C, FitzGerald RE, Hanberg A, Hass U, Hooijmans C, Rooney AA, Rousselle C, van Loveren H, Wölfle D, Barizzone F, Croera C, Putzu C and Castoldi AF, 2017. Bisphenol A (BPA) hazard assessment protocol. EFSA supporting publication 2017:EN-1354. 76 pp. doi:10.2903/sp.efsa.2017.EN-1354

EFSA (European Food Safety Authority), Aiassa E, Martino L, Barizzone F, Ciccolallo L, Garcia A, Georgiadis M, Muñoz Guajardo I, Tomcikova D, Alexander J, Calistri P, Gundert-Remy U, Hart HD, Hoogenboom RL, Messean A, Naska A, Navajas Navarro M, Noerrung B, Ockleford C, Wallace RJ, Younes M, Abuntori B, Alvarez F, Aryeetey M, Baldinelli F, Barrucci F, Bau A, Binaglia M, Broglia A, Castoldi AF, Christoph E, De Sesmaisons-Lecarré A, Georgiadis N, Gervelmeyer A, Istace F, López-Gálvez G, Manini P, Maurici D, Merten C, Messens W, Mosbach-Schulz O, Putzu C, Ramos Bordajandi L, Smeraldi C, Tiramani M, Valtueña Martínez S, Vos S, Hardy AR, Hugas M, Kleiner J, De Seze G, Verhagen H and Verloo D, 2018. Implementation of PROMETHEUS 4-step approach for evidence use in EFSA scientific assessments: benefits, issues, needs and solutions. EFSA supporting publication 2018:EN1395. 31 pp. doi:10.2903/sp.efsa.2018.EN-1395

EFSA (European Food Safety Authority) (2019), EFSA procedures for responding to urgent advice needs, EFSA supporting publication 2019:EN-1228, 34 pp. doi:10.2903/sp.efsa.2017.EN-1228

EFSA ANS Panel (EFSA Panel on Food Additives and Nutrient Sources added to Food), 2015. Scientific opinion on the risk assessment for peri- and post-menopausal women taking food supplements containing isolated isoflavones. EFSA Journal 2015;13(10):4246, 342 pp. doi:10.2903/ j.efsa.2015.4246

EFSA PLH Panel (EFSA Panel on Plant Health), Jeger M, Bragard C, Caffier D,Candresse T, Chatzivassiliou E, Dehnen-Schmutz K, Gregoire J-C, Jaques Miret JA, MacLeod A, NavajasNavarro M, Niere B, Parnell S, Potting R, Rafoss T, Rossi V, Urek G, Van Bruggen A, Van Der Werf W,West J, Winter S, Hart A, Schans J, Schrader G, Suffert M, Kertesz V, Kozelska S, Mannino MR, Mosbach-Schulz O, Pautasso M, Stancanelli G, Tramontini S, Vos S and Gilioli G, 2018. Guidance on quantitative pest risk assessment. EFSA Journal 2018;16(8):5350, 86 pp. doi:10.2903/j.efsa.2018.5350

EFSA Scientific Committee, 2016. Guidance to develop specific protection goals options for environmental risk assessment at EFSA, in relation to biodiversity and ecosystem services. EFSA Journal 2016;14(6):4499, 50 pp. doi:10.2903/j.efsa.2016.449

EFSA Scientific Committee, Hardy A, Benford D, Halldorsson T, Jeger MJ, Knutsen HK, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Schlatter JR, Silano V, Solecki R, Turck D, Benfenati E, Chaudhry QM, Craig P, Frampton G, Greiner M, Hart A, Hogstrand C, Lambre C, Luttik R, Makowski D, Siani A, Wahlstroem H, Aguilera J, Dorne J-L, Fernandez Dumont A, Hempen M, Valtuena Mart ~ ınez S, Martino L, Smeraldi C, Terron A, Georgiadis N and Younes M, 2017a. Scientific Opinion on the guidance on the use of the weight of evidence approach in scientific assessments. EFSA Journal 2017;15(8):4971, 69 pp. doi:10.2903/j.efsa.2017.4971

EFSA Scientific Committee, Hardy A, Benford D, Halldorsson T, Jeger MJ, Knutsen HK, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Schlatter JR, Silano V, Solecki R, Turck D, Younes M, Bresson J-L, Griffin J, Hougaard Benekou S, van Loveren H, Luttik R, Messean A, Penninks A, Ru G, Stegeman JA, van der Werf W, Westendorf J, Woutersen RA, Barizzone F, Bottex B, Lanzoni A, Georgiadis N and Alexander J, 2017b. Guidance on the assessment of the biological relevance of data in scientific assessments. EFSA Journal 2017;15(8):4970, 73 pp. doi:10. 2903/j.efsa.2017.4970

EFSA (European Food Safety Authority) Scientific Committee, Benford D, Halldorsson T, Jeger MJ, Knutsen HK, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Schlatter JR, Silano V, Solecki R, Turck D, Younes M, Craig P, Hart A, Von Goetz N, Koutsoumanis K, Mortensen A, Ossendorp B, Martino L, Merten C, Mosbach-Schulz O and Hardy A, 2018a Guidance on Uncertainty Analysis in Scientific Assessments. EFSA Journal 2018;16(1):5123, 39 pp. doi: 10.2903/j.efsa.2018.5123

EFSA (European Food Safety Authority) Scientific Committee, Benford D, Halldorsson T, Jeger MJ, Knutsen HK, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Schlatter JR, Silano V, Solecki R, Turck D, Younes M, Craig P, Hart A, Von Goetz N, Koutsoumanis K, Mortensen A, Ossendorp B, Germini A, Martino L, Merten C, Mosbach-Schulz O, Smith A and Hardy A, 2018b. Scientific Opinion on the principles and methods behind EFSA's Guidance on Uncertainty Analysis in Scientific Assessment. EFSA Journal 2018;16(1):5122, 235 pp. https://doi.org/10.2903/j.efsa.2018.512

EFSA Scientific Committee, More SJ, Bampidis V, Benford D, Bennekou SH, Bragard C, Halldorsson TI, Hernandez-Jerez AF, Koutsoumanis K, Naegeli H, Schlatter JR, Silano V, Nielsen SS, Schrenk D, Turck D, Younes M, Benfenati E, Castle L, Cedergreen N, Hardy A, Laskowski R, Leblanc JC, Kortenkamp A, Ragas A, Posthuma L, Svendsen C, Solecki R, Testai E, Dujardin B, Kass GEN, Manini P, Jeddi MZ , Dorne J-LCM and Hogstrand C, 2019. Guidance on harmonised methodologies for human health, animal health and ecological risk assessment of combined exposure to multiple chemicals. EFSA Journal 2019;17(3):5634, 77 pp. doi:10.2903/j.efsa.2019.5634

EFSA (European Food Safety Authority) and EBTC (Evidence-Based Toxicology Collaboration) 2018. EFSA Scientific Colloquium 23: Evidence integration in risk assessment: the science of combining apples and oranges. EFSA Supporting publication 2018:16(3):EN-1396. 28 pp. doi:10.2903/sp.efsa.2018.EN-1396

Devos Y, Craig W, Devlin RH, Ippolito A, Leggatt RA, Romeis J, Shaw R, Svendsen C and Topping CJ, 2019. Using problem formulation for fit-for-purpose pre-market environmental risk assessments of regulated stressors. Proceedings of the EFSA Third Scientific Conference: Science, Food and Society, Parma, Italy. EFSA Journal 2019;17(S1):e170708, 31 pp. doi:10.2903/j.efsa.2019.e170708

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P and Schunemann HJ, 2011. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of Clinical Epidemiology, 64, 383-394. doi:10.1016/j.jclinepi.2010.04.026

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available online www.training.cochrane.org/handbook

Hill AB, 1965. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicines, 58, 295-300.

Kass RE, Caffo BS, Davidian M, Meng X-L, Yu B and Reid N, 2016. Ten Simple Rules for Effective Statistical Practice. PLoS Computatonal Biology 12(6). doi:10.1371/journal.pcbi.1004961

Kerr NL, 1998. HARKing: Hypothesizing After the Results are Known. Personality and Social Psychology Review 2(3), 196-217.

Lowell RB, Culp JM and Dube MG, 2000. A weight-of-evidence approach for Northern River risk assessment: Integrating the effects of multiple stressors. Environmental Toxicology and Chemistry, 19(4), 1182-1190. doi:10.1002/etc.5620190452.

Meek ME, Boobis A, Cote I, Dellarco V, Fotakis G, Munn S, Seed J and Vickers C, 2014. New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. Journal of Applied Toxicology, 34(1), 1-18. doi:10.1002/jat.2949

Munafo MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, du Sert NP, Simonsohn U, Wagenmakers EJ, Ware JJ and Ioannidis JPA, 2017. A manifesto for reproducible science. Nature Human Behaviour, 1, 0021. Doi:10.1038/s41562-016-0021

O'Connor AM, Lovei GL, Eales J, Frampton GK, Glanville J, Pullin AS and Sargeant J. Implementation of systematic reviews in EFSA scientific outputs workflow. EFSA Supporting Publication 2012;10(12) EN-367, 36pp. doi:10.2903/sp.efsa.2012.EN-367

OHAT-NTP (Office of Health Assessment and Translation-Division of the National Toxicology Program), 2019. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. Available online: https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookmarch2019_508.pdf

Pearl J, 2009. Causal inference in statistics: An overview. Statistics Surveys, 3, 96–146. doi:10.1214/09-SS057

Schünemann H, Hill S, Guyatt G, Akl EA and Ahmed F, 2011. The GRADE approach and Bradford Hill's criteria for causation. Journal of Epidemiology and Community Health, 65(5), 392-395. Doi:10.1136/jech.2010.119933

Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA and PRISMA-P Group, 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. The BMJ, 349:g7647. doi:10.1136/bmj.g7647

Swaen G and van Amelsvoort L, 2009. A Weight of Evidence Approach to Causal Inference. Journal of clinical epidemiology, 62(3), 270-277. doi: 10.1016/j.jclinepi.2008.06.013

U.S. EPA (United States Environmental Protection Agency), 2011. Expert Elicitation Task Force White Paper. Prepared for the Science and Technology Policy Council U.S. Environmental Protection Agency. Washington DC 20460, 142pp.

Woodruff TJ and Sutton P, 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. Environmental Health Perspectives, 122(10), 1007-1014. doi:10.1289/ehp.1307175

WHO (World Health Organisation), 2014. WHO Handbook for Guideline Development – 2nd edition. 179 pages.

Abbreviations

ADI	Acceptable Daily Intake
ADME	Absorption Distribution Metabolism and Excretion
BMDL	benchmark dose lower limit (lower bound of benchmark dose confidence interval)
CAT	Critical Appraisal Tool
DRV	Dietary Reference Values
EBTC	Evidence Based Toxicology Collaboration
EC_{50}	Half maximal effective concentration: the concentration of a test substance which results in 50% of the test organisms being adversely affected, i.e. both mortality and sublethal effects
EKE	Expert Knowledge Elicitation
EPA	EFSA Process Architecture
ErC_{50}	effective concentration (growth rate): the concentration of a test substance which results in a 50% of inhibition of the growth rate
GM	Genetically Modified
GMO	Genetically Modified Organism
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HARKing	Hypothesizing After the Results are Known
LC_{50}	Half population Lethal Concentration the concentration of a test substance which results in a 50% mortality of the test species
LOAEL	Lowest-Observed-Adverse-Effect Level
ML	Maximum Level
MoA	Mode of Action
MRL	Maximum Residue Limit or level
NOAEL	No-Observed-Adverse-Effect-Level
NOEC	No Observed Effect Concentration
NTO	Non Target Organism
NTP	National Toxicology Program
OECD	Organisation for Economic Co-operation and Development
OHAT	Office of Health Assessment and Translation
PEC	Predicted Environmental Concentration
PECO	Population (p), Exposure (e), Comparator (c) and Outcome (o) in a question about an exposure effect
PICO	Population (p), Intervention (i), Comparator (c) and Outcome (o) in a question about an intervention effect
PIT	Population (p), Index test (i), and Target population (t) in a question about test accuracy
PO	Population (p) and Outcome (o) in a descriptive question
QPS	Qualified Presumption of Safety

RA	Risk Assessment
RNAI	RNA interference
ROA	Rapid Outbreak Assessment
RoB	Risk of Bias
RPA	Reference Point for Action
SPG	Specific Protection Goal
TDI	Tolerable Daily Intake
TSE	Transmissible Spongiform Encephalopathies
UL	Upper Level of tolerable intake

Appendix A – EFSA's assessment questions

Table A.1 represents a first attempt to describe and classify the questions arising from EFSA's mandates into a common framework illustrating for each question possible split in common sub-questions.

The aim of this classification is to emphasise the similarities of EFSA's questions, irrespective of the type of assessment and the domain. **It is provisional** and it will be further refined based on the outcome of the testing phase of this framework, the discussion that will take place in the EFSA units and panels (possibly also in dedicated workshops) and the outcome of the project on problem formulation.

This table does not aim at harmonising terminology. Rather it offers a tool for reading-across different domains focusing on communality of concepts beyond terminological differences.

The classification does not discriminate between application and non-application mandates, since the objective here is to cluster assessments according to question types and not to the domain and regulatory context. The **recommendations listed in the core text of this document though apply only to non-application assessment**.

From Table A.1 it can be noted that, irrespective of the domain, EFSA's sub-questions at their lower order are classifiable into **two main classes**:

- Sub-questions **aiming to test a hypothesis** (e.g. sub-questions on safety/toxicity/pathogenicity and efficacy), aimed to assess the association between exposure(s) or intervention(s) and a specific outcome or comparative assessment of test accuracy);
- Sub-questions **aiming to estimate or predict a parameter in a descriptive way** (e.g. occurrence, prevalence or incidence).

DISCLAIMER to help reading Table A.1:

- The table provides the definition of questions and the related sub-questions with their 'hierarchical order'. For each type of question, the possible sub-questions are illustrated with hierarchical structure: 1. assessment pillars or higher-order sub-questions, 2. within each pillar, lower order sub-questions.
- The sub-questions in each question don't need to be followed from left to right;
- An assessment question does not necessarily include all sub-questions described in the table;
- The term 'dose/concentration-response' covers any kind of evidence synthesis, qualitative or quantitative, that infers association between 2 or more doses (or concentration levels) of a substance/agent and an outcome. The type of synthesis can range from a purely narrative summary of evidence to a continuous dose-response modelling (e.g. benchmark dose response model) covering the full spectrum of possible analyses in between (e.g. NOAEL). The outcome of the dose (concentration)-response is the definition of a reference dose that can be expressed as a single value or a distribution (in case the uncertainty is quantified);
- Prevalence also includes absence of a disease (zero prevalence);

- Exposure can include combined exposure and/or combined exposure via multiple pathways, i.e. aggregated exposure (definition provided in EFSA Scientific Committee, 2019);
- There might be multiple 'baseline scenarios' reflecting variability of conditions (e.g. worst and best case).

The question on 'Assessment of methods', although included in Table A.1, does not follow the pillar structure that is typical of the chemical risk assessment questions and applicable also to other EFSA's questions. Therefore it is included in the table only for completeness reasons. Questions on assessment of methods are about e.g. sensitivity and specificity of a diagnostic test/analytical method/outcome detection method, definition of a survey design e.g. to assess prevalence of a disease or to set up a surveillance plan, definition of welfare indicators.

Table A.1: EFSA's question types

ASSESSMENT QUESTION ARISING FROM EFSA MANDATE	ASSESSMENT PILLARS (or 'HIGHER ORDER SUB_QUESTIONS')					EFSA PANEL/UNIT or TEAM
	EFFECT IDENTIFICATION		EFFECT CHARACTERISATION	EXPOSURE ASSESSMENT	CHARACTERISATION of risk of adverse effect or of likelihood of beneficial effect	
	Preliminary phase	Actual phase				
	LOWER-ORDER SUB-QUESTIONS					
RA in humans and/or animals of: • **chemicals** (e.g. contaminants, food and feed additives, food contact materials, chemical active substances for PPPs) • **food and feed additives different from chemicals** (e.g. enzymes, microorganisms) • **active substances different from chemicals** • **novel foods** • **tolerable upper**	• **Identification of the agent[1] characteristics** (e.g. molecule or mixture, production of secondary metabolites, modification introduced in the genes) • **History of safe use and use patterns**	**Hazard identification** • **Inherent properties of the agent** (e.g. toxic effects, toxigenic or allergenic potential, genes of concern, AMR, virulence factors, antibiotics production) **Assessments of relationship between the agent and the adverse effect(s)**	• **Dose-response to establish a reference point/point of departure** (e.g. NOAEL, LOAEL, BMDL) • **Estimate of reference values for humans and animals applying uncertainty factors** to the established reference point (e.g. ADI, TDI) or deriving them from	• **Dietary exposure:** ✓ **Food Consumption** ✓ **Feed consumption** ✓ **Occurrence (prevalence and/or level of the agent** (it includes also viable cells, DNA, toxic metabolites, antibiotics) • **Non dietary exposure** (e.g. inhalation, dermal absorption)	**Risk characterisation** Comparison of exposure with: (i) reference points (e.g. NOAEL, BMDL) using the Margin of exposure or Margin of safety; (ii) reference values (e.g. ADI, TDI, UL) using the hazard quotient[3]	BIOMO[4] CONTAM DATA FEED FIP GMO NUTRI PRES PREV

ASSESSMENT QUESTION ARISING FROM EFSA MANDATE	ASSESSMENT PILLARS (or 'HIGHER ORDER SUB_QUESTIONS')					EFSA PANEL/UNIT or TEAM
	EFFECT IDENTIFICATION		EFFECT CHARACTERISATION	EXPOSURE ASSESSMENT	CHARACTERISATION of risk of adverse effect or of likelihood of beneficial effect	
	Preliminary phase	Actual phase				
	LOWER-ORDER SUB-QUESTIONS					
intake levels of nutrients • **GMOs** • **Antimicrobials**		*If needed, definition of the evidence streams (e.g. human, in vivo, in vitro, in silico* studies)	human studies (e.g. UL[2])			
	• ADME (it relates to both effect identification and characterisation) • MoA (it relates to both effect identification and characterisation)					
Nutritional assessments (e.g. DRVs, infant formula, exemptions from labelling)	• Properties of the nutrient/food constituent/food (all)	• Assessment of relationship between the nutrient/food constituent/food and adverse/beneficial effects (in human studies) (all)	Dose-response to establish: • reference values (DRVs) • minimal eliciting dose • threshold levels (exemptions from labelling)	Dietary exposure (DRVs, exemptions from labelling): • Food Consumption • Occurrence of the nutrient/food constituent	Comparison of exposure with: • minimal eliciting doses • threshold levels (exemptions from labelling)	NUTRI

ASSESSMENT QUESTION ARISING FROM EFSA MANDATE	ASSESSMENT PILLARS (or 'HIGHER ORDER SUB_QUESTIONS')					EFSA PANEL/UNIT or TEAM
	EFFECT IDENTIFICATION		EFFECT CHARACTERISATION	EXPOSURE ASSESSMENT	CHARACTERISATION of risk of adverse effect or of likelihood of beneficial effect	
	Preliminary phase	Actual phase				
	LOWER-ORDER SUB-QUESTIONS					
Monitoring and/or Surveillance[5] of: • **pesticides residues[6]** • **veterinary drug residues** • **pathogen/disease occurrence** and expected trends in food, feed and animals **(e.g. zoonosis, TSE, antimicrobial resistance)** • **contaminants**	*In case the monitoring data are used for risk assessment, this is addressed in the first question (RA on humans/animals) with some adaptation*	*In case the monitoring data are used for risk assessment, this is addressed in the first question (RA on humans/animals) with some adaptation*	*In case the monitoring data are used for risk assessment, this is addressed in the first question (RA on humans/animals) with some adaptation*	**Monitoring**: • Prevalence of exceedance of MRL/ML/RPA • Prevalence of diseases • Occurrence and/or level of potential hazard[7] in humans, animals and food and feed • Occurrence of foodborne outbreaks For **surveillance**, the above data are collected for assessing managerial/mitigation measures (see 'assessments of methods')	*In case the monitoring data are used for risk assessment, this is addressed in the first question (RA on humans/animals) with some adaptation*	AHAW BIOHAZ BIOMO DATA FEED PRES In future: PLH

ASSESSMENT QUESTION ARISING FROM EFSA MANDATE	ASSESSMENT PILLARS (or 'HIGHER ORDER SUB_QUESTIONS')					EFSA PANEL/UNIT or TEAM
	EFFECT IDENTIFICATION		EFFECT CHARACTERISATION	EXPOSURE ASSESSMENT	CHARACTERISATION of risk of adverse effect or of likelihood of beneficial effect	
	Preliminary phase	Actual phase				
	LOWER-ORDER SUB-QUESTIONS					
Animal welfare assessments	• Definition of animal population and system (e.g. on farm, during transport, at slaughter) • Identification of the hazards (e.g. lack of drinkers, lack of ventilation, rough handling from operators) • Identification of the welfare consequences (e.g. thirst, thermal stress, pain) including severity	Assessment of impact on animal welfare: • assessment of the relationship between the exposure to a system (and related hazards) and the welfare consequences • assessment of occurrence of welfare consequences in the system and related hazards The above includes the assessment of the duration of the consequences in the system	Outcome: Identification of the system-related hazards mainly contributing to the welfare consequences in the population	Typically, not done due to lack of data	Typically not done due to the lack of data on hazard characterisation and exposure assessment	AHAW

ASSESSMENT QUESTION ARISING FROM EFSA MANDATE	ASSESSMENT PILLARS (or 'HIGHER ORDER SUB_QUESTIONS')					EFSA PANEL/UNIT or TEAM
	EFFECT IDENTIFICATION		EFFECT CHARACTERISATION	EXPOSURE ASSESSMENT	CHARACTERISATION of risk of adverse effect or of likelihood of beneficial effect	
	Preliminary phase	Actual phase				
	LOWER-ORDER SUB-QUESTIONS					
Efficacy (e.g. of health claims, decontaminants, vaccines, control measures, feed additives, stunning methods)	• Inherent properties of the agent/measure of which efficacy is assessed	• Assessment of relationship agent/measure-beneficial effects	• Dose-response assessment to evaluate the efficacy of the agent (e.g. to estimate the log-reduction of the pathogen on a food item)	None	None	ALL
Emerging risks Identification	• Inherent properties of the microorganism, chemical substance, drug, additive	**Hazard identification** • Identification of a new hazard • Identification of new adverse effect of known hazard		**Exposure assessment** Assessment of increased exposure of a known hazard in terms of: new susceptibility; new target groups	**None**	SCER
Identify the food vehicle of infection (RoA) in outbreak	•	**Hazard identification** • **Step2:** Identification of food -	**None**	**Exposure assessment** **Step1:** Consumption of	Identify vehicle of infection and source of contamination of the	BIOMO

ASSESSMENT QUESTION ARISING FROM EFSA MANDATE	ASSESSMENT PILLARS (or 'HIGHER ORDER SUB_QUESTIONS')					EFSA PANEL/UNIT or TEAM
	EFFECT IDENTIFICATION		EFFECT CHARACTERISATION	EXPOSURE ASSESSMENT	CHARACTERISATION of risk of adverse effect or of likelihood of beneficial effect	
	Preliminary phase	Actual phase				
	LOWER-ORDER SUB-QUESTIONS					
		pathogen relation and tracing across countries[8] • **Step3:** microbiological characterization in the human cases and in the food to identify similarities		specific food items in human cases[9] **Step4**: Traceability of the food distribution across food production and consumption chain[10]	food item in the food production process	
Plant pest risk assessment	Pest categorisation (only PLH)		Dose-response assessment	• **Occurrence** (prevalence and concentration) of the pathogen at one or more stages • Consumption of the food (frequency and serving size)	Public, animal and plant pest health impact at baseline OR under alternative risk reduction options/managerial options	PLH AHAW BIOHAZ FEED
Microbial (risk) assessment **Animal health risk assessment**	• Definition of the characteristics of the biological agent (BIOHAZ and animal disease) • Identification of the food/vehicle-	Relationship pathogen-adverse effect(s) (e.g. QPS, Listeria monocyto-genes, Schmallenberg) (BIOHAZ and AHAW)				

ASSESSMENT QUESTION ARISING FROM EFSA MANDATE	ASSESSMENT PILLARS (or 'HIGHER ORDER SUB_QUESTIONS')					EFSA PANEL/UNIT or TEAM
	EFFECT IDENTIFICATION		EFFECT CHARACTERISATION	EXPOSURE ASSESSMENT	CHARACTERISATION of risk of adverse effect or of likelihood of beneficial effect	
	Preliminary phase	Actual phase				
	LOWER-ORDER SUB-QUESTIONS					
	pathogen relation (BIOHAZ, AHAW)			• Probability of exposure		
Environmental risk assessment (ERA) of: • **active substance for PPPs** • **FEED additives** • **GMOs post-market monitoring** • **invasive Alien Species (plant pests)**	• Identification of characteristics of the potential stressor • Translation, according to the SC GD (EFSA Scientific Committee, 2016), of the General Protection Goal into Specific Protection Goals (SPGs): ✓ biological entity; ✓ attribute; ✓ magnitude of effect; ✓ temporal and geographical	**Hazard identification** • Assessment of **relationship stressor-adverse effects and factors influencing the relationship** (e.g. soil characteristics, wind, rain, temperature) • Identification of Pathway to Harm • Sequence homology with	• **Concentration –response to establish a reference concentration value** (e.g. selecting the NOEC, EC50, LC50, ErC50 to be used in the risk assessment for each group of organisms, Predicted No Effect Concentration) • Assess activity spectrum of the molecule to identify	• Environmental fate for exposure assessment: ✓ Presence, concentration and biological activity of the stressor in the Environment ✓ Predicted environmental concentration (PEC) • Probability of exposure	Comparative safety using for instance: ✓ Risk Quotient (lethal or sublethal dose/environmental exposure) ✓ Exposure/ reference concentration	FEED GMO PLH PREV AHAW

ASSESSMENT QUESTION ARISING FROM EFSA MANDATE	ASSESSMENT PILLARS (or 'HIGHER ORDER SUB_QUESTIONS')					EFSA PANEL/UNIT or TEAM
	EFFECT IDENTIFICATION		EFFECT CHARACTERISATION	EXPOSURE ASSESSMENT	CHARACTERISATION of risk of adverse effect or of likelihood of beneficial effect	
	Preliminary phase	Actual phase				
	LOWER-ORDER SUB-QUESTIONS					
	scale of the effect; ✓ Tolerable harm	toxicants (in silico studies) • Read across for metabolites	species that it can affect	• intake/consumption/ingestion/absorption[11]		
	• Identification of pathway to harm/ MoA (e.g. how the protein operates in the target organisms – target gene in case of RNAi that and help understand off-target effect) • Identification of specific MoA of concern or potential for accumulation			•		
Assessment of methods (e.g. sensitivity/specificity of a diagnostic test, an analytical method, an outcome detection method; definition of surveillance design, sampling scheme as in Norovirus and welfare indicators)	This type of questions doesn't follow the structure of this table. Therefore it is described in the text above					ALL

[1] By 'agent' it is meant here 'any biological agent/chemical substance/genetical modification/novel food'.
[2] For companion and farm animals RA, Uncertainty factors are applied to NOAEL/LOAEL/BMDL of laboratory animals to derive safe concentrations in feed based on feed consumption.
[3] Margin of exposure: ratio of (a) a reference point of (eco)toxicity to (b) the theoretical, predicted or estimated exposure dose or concentration. Hazard quotient: the ratio of the potential exposure to the substance and the level at which no adverse effects are expected (EFSA Scientific Committee, 2019).
[4] e.g. Risk ranking of chemical and microbiological hazards in food (EFSA, 2018).
[5] These types of assessments can be specifically mandated to EFSA (sometimes as continuous data collection activities) or be embodied in broader mandates.
[6] Post-market monitoring including post-market risk assessment.
[7] Pathogen/infection/disease/contaminants.
[8] Supporting evidence provided by EC RASFF system.
[9] ECDC task – EPIS and EWRS database.
[10] Supporting evidence provided by EC RASFF system.
[11] For GM plants the main exposure route is 'oral intake through consumption'. For a herbivore (NTO) exposure can occur through direct feeding on the plant. For pollinators it can be through the consumption of pollen. In the soil, exposure is possible via root exudates. For an earthworm exposure is occurs through the feeding on soil (the newly expressed protein of the GM plant may bound to soil particles). Another route is via plant material entering the water comportments (ditches, etc.).

Appendix B – Overview of methods for evidence synthesis and integration accounting for uncertainty

This Appendix illustrates some of the qualitative and quantitative methods that can be adopted for synthesising and integrating evidence within and across sub-questions. The list of methods is not exhaustive and is intended to provide a broad overview of the possible options to choose when planning evidence synthesis and integration at protocol level. Some considerations on semi-quantitative methods are also made.

Qualitative methods

Qualitative methods range from purely narrative and unstructured to more structured approaches.

Structured qualitative methods for synthesis include tabular summary and graphical display of the evidence possibly stratified by factors that can influence the outcome of the assessment.

With regard to uncertainty analysis, a structured qualitative method for health care related questions is the GRADE (Guyatt et al., 2011). This approach involves rating an initial certainty in the body of evidence underpinning the research question at hand based on study design and then either downgrading or upgrading it using predefined criteria (i.e. risk of bias, imprecision, inconsistency in results across studies, indirectness, publication bias, strong association, dose-response gradient and opposing residual plausible confounding).

GRADE[11] incorporates and further elaborates on most of Bradford Hill considerations for causation (Hill, 1965; Schünemann et al., 2011) and its application for addressing uncertainty when assessing causality has gradually extended from clinical research to public health, health policy, environmental and occupational health and some GRADE-based approaches developed for environmental health assessments such as the University of California Navigation Guide (Woodruff and Sutton, 2014) and the OHAT/NTP Handbook 2019 (OHAT-NTP, 2019). GRADE-based approaches also draw upon Hill criteria for assessing the 'confidence' (i.e. certainty) in the association between chemical exposure and (adverse) effects from each evidence stream in hazard identification. For each stream, an initial confidence rating is defined based on study design and then up- or down-graded according to specified features of the body of evidence within- and across-sub-questions, which are derived from Hill criteria. GRADE-based approaches have been applied also in some EFSA scientific assessments (e.g. EFSA ANS Panel, 2015; EFSA, 2017).

Modified Hill's considerations are also incorporated in another structured qualitative approach, i.e. the WHO/IPCS framework on mode of action/species concordance analysis, aimed to increase transparency and consistency in integrating evidence for testing hypotheses of modes of action (or adverse outcome pathways) and in turn inform the risk assessment process (Meek et al., 2014).

[11] http://www.gradeworkinggroup.org/

For environmental risk assessment, Lowell and colleagues (Lowell et al., 2000) have developed a similar tool for synthesising evidence and for evidence integration in environmental risk assessment, also based on Hill's criteria. In their approach, developed to assess the effects of Contaminants of Potential Concern (COPC) on rivers, the criteria include: spatial and temporal correlation, plausible explanation linking stressor and effect, experimental verification of stressor, cause-effect relationship under controlled conditions, strength of correlation, specificity of the effect to the COPC, evidence of COPC exposure in the body of the Receptor of Potential Concern (ROPC), and consistency of association across other studies within the region and in analogous studies in other regions.

Quantitative methods

Quantitative approaches for evidence synthesis and integration exert large variability in terms of type and complexity ranging from traditional and bias-adjusted meta-analyses (EFSA (European Food Safety Authority) and EBTC (Evidence-Based Toxicology Collaboration), 2018) up to Bayesian approaches. For assessments of chemicals in humans, an example is provided by Swaen and van Amelsvoort (2009) whose approach to estimating the probability of a causal association involves using scientific evidence to quantify the probability that each Hill criterion is met and then assigning relative weights to each of them, using discriminant analysis. A different approach to evaluating causality is based on the theory of causal diagrams (Pearl, 2009). These are graphical models set up by translating both the hypothesised relations among the study variables and their full probability distribution into a Directed Acyclic Graph (DAG). A DAG is composed of nodes representing variables and directed edges representing conditional independence among subsets of variables. Bayesian networks are among the causal diagrams most commonly used for causality.

Expert judgement recently started to be used by international organisations (e.g. U.S. EPA 2011; EFSA, 2014) as a method for integrating evidence, drawing conclusions and expressing their level of certainty. As its level of formality varies significantly, expert judgement can be used as a narrative, qualitative or quantitative approach. The highest standard in the use of the method is achieved when a formal expert knowledge elicitation is performed (EFSA, 2014) and the results of the integration are expressed quantitatively as uncertainty probability distributions (e.g. uncertainty probability distribution around the population mean exposure to a hazard of concern and probability that a threshold is exceeded, uncertainty probability distribution of the dietary reference value to be used to target population intake for nutrients in order to minimise risk of deficiency and of chronic diseases).

Figure B.1 provides an overview of broad class of quantitative methods to integrate evidence at the level of individual sub-questions and across sub-questions. Those methods allow to address increasing levels of heterogeneity. Examples can also be found in EFSA Scientific Committee (2017a)

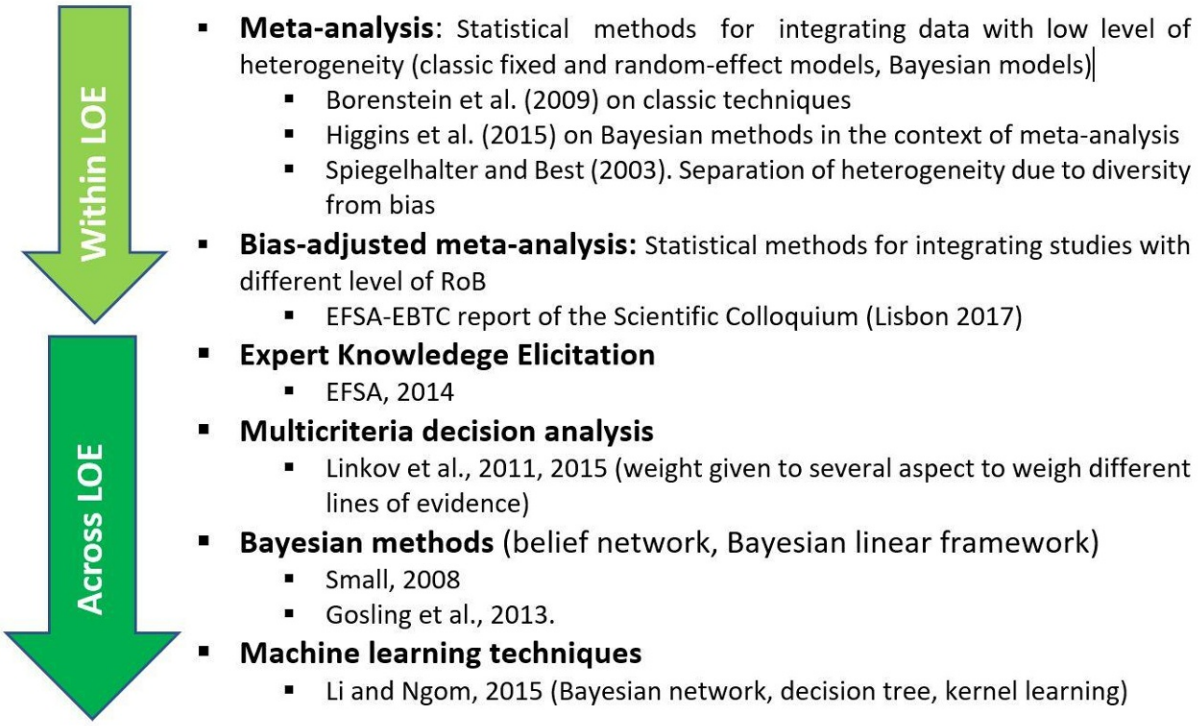


Figure B.1: Quantitative methods for evidence synthesis (within sub-question) and integration (across sub-question)

Semi-quantitative methods

Semi-quantitative approaches for evidence integration provide an intermediary level between the textual evaluation of existing evidence within stream in terms of risk and the numerical quantitative evaluation of risk.

Overall, semi-quantitative approaches offer a more consistent and rigorous approach to assess and compare risks and risk management strategies than qualitative risk assessment and avoid some of the ambiguities that a qualitative risk assessment may produce. They do not require the same mathematical skills as for quantitative risk assessment, nor the same amount of data, which means they can be applied more broadly.