

Concept Article

# EFSA's Framework for Evidence-Based Scientific Assessments: A Case Study on Uncertainty Analysis

Elisa Aiassa, Caroline Merten and Laura Martino

European Food Safety Authority (EFSA), Assessment and Methodological Support Unit, Parma, Italy

## Abstract

To provide a sound scientific advice in support of the European decision-making process in food and feed safety, the European Food Safety Authority (EFSA) defined the principles for producing 'evidence-based scientific assessments' (impartiality, methodological rigour, transparency and engagement) and, to help fulfilling them, developed cross-cutting methodological approaches. This paper focusses on two of these approaches: conducting scientific assessments in four steps - with an emphasis on developing a protocol for the assessment a priori - and analysing uncertainty. An overview of the 4-step approach and the methods for addressing uncertainty is given and a case-study on uncertainty analysis, developed in collaboration with the German Federal Institute for Risk Assessment, is illustrated. The main advantage related to the implementation of protocols and uncertainty analysis is the increase in the scientific value of the outputs. However, experience and further capacity-building is needed to better incorporate uncertainty analysis in the planning phase (protocol) of the scientific assessment process. The case-study is based on exposure in humans. Nonetheless it provides an example of a framework for evidence-based scientific assessments that is applicable also to other types of evidence, including the one arising from New Approach Methodologies. Adopting the proposed framework, that covers in the planning and implementation phase an analysis of the uncertainties, is also expected to foster the integration of multiple evidence sources, including alternative methods and testing strategies, in the regulatory scientific assessment process.

## 1 Background

The European Food Safety Authority (EFSA) is the European agency responsible for providing scientific advice in support to the European policies and legislations, and for communicating on that advice, in the area of food and feed safety, nutrition, animal health and welfare, plant health and protection and the possible impact of the food chain on the biodiversity of plant and animal habitats.

EFSA was established under the General Food Law - Regulation EC No 178/2002 (EC, 2002) which introduced a clear separation of responsibilities between risk assessment (science) and risk management (policy). The agency is funded by the European Union and functions in cooperation with the EU member States and independently from the European legislative and executive institutions (Commission, Council, Parliament). However, EFSA operates in a 'regulatory science' environment, meaning that the scientific advice it produces and the communications that accompany this advice are designed specifically to answer questions of — and support the decisions taken by — risk managers.

EFSA plays a fundamental role in collecting and analysing data to ensure that the European food and feed safety regulatory system is supported by impartial, methodologically rigorous and transparent advice. The recent adoption of the

---

This paper is part of the proceedings of the workshop on Development of an Evidence Based Risk Assessment Framework held at the University of Ottawa in December 2018.

**Disclaimer:** The authors are employed by the European Food Safety Authority (EFSA) in the Assessment and Methodological Support unit (Elisa Aiassa and Laura Martino) and the Scientific Committee and Emerging Risks unit (Caroline Merten). However, the present article is published under the sole responsibility of the authors and may not be considered as an EFSA output. The positions and opinions presented in this article are those of the authors alone and are not intended to represent the views or scientific works of EFSA. To know about the views or scientific outputs of EFSA, please consult its website at <http://www.efsa.europa.eu>

Received April 21, 2020; Accepted January 15, 2021;  
Epub June 22, 2021; © The Authors, 2021.

ALTEX 38(##), ###-###. doi:10.14573/altex.2004211

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.

Correspondence: Laura Martino, PhD  
EFSA  
Via Carlo Magno 1/a, 43126 Parma, Italy  
([laura.martino@efsa.europa.eu](mailto:laura.martino@efsa.europa.eu))

new legislation on transparency and sustainability of the EU risk assessment model in the food chain<sup>1</sup> has further emphasised the need for transparency and openness in the scientific assessment process in the fields covered under the EFSA remit.

## **2 Introduction**

To address the demand for evidence-based advice in support of policy-making, EFSA has implemented a series of activities aimed at further improving the soundness and transparency of its scientific assessments. These included the definition of a set of principles for ‘evidence-based’ scientific assessments and the development — or adaptation — and implementation of several cross-cutting methodological approaches that help fulfil those principles. These include a 4-step approach for conducting scientific assessment (plan/do/verify/report) (EFSA, 2015a and b), systematic review (EFSA, 2010), expert knowledge elicitation (EFSA, 2014), weight of evidence (EFSA Scientific Committee, 2017a), analysis of biological relevance (EFSA Scientific Committee, 2017b) and uncertainty analysis (EFSA Scientific Committee, 2018 a and b).

This paper focusses on how two of the approaches mentioned above contribute to set up a framework for evidence-based scientific assessments: (1) the 4-step approach for conducting an assessment (plan, do, verify, report) with particular focus on the planning phase (EFSA, 2015a), and (2) the analysis of uncertainty (EFSA Scientific Committee, 2018a,b). Over the time the two approaches have been tested and currently are being further refined and broadened in scope to fit new types of evidence such as those emerging from NAM and alternative testing methods as envisaged by EFSA’s strategy 2020 (EFSA, 2016). Moving to a framework for evidence-based scientific assessments where uncertainty analysis is inherent in the process can facilitate relying on novel methods in the regulatory context and making a better use of animal studies. This is in line with the effort done by other agencies to foster the use of sound practices to establish confidence in new methods (NRC, 2007, ICCVAM, 2017). Measuring the impact of the implementation of new testing approaches in terms of reduction of the uncertainty on hazard identification and characterisation conclusions is necessary to move from the prevalent use of animal models to a broader and more integrated set of evidence types.

Following an overall description of the four-step approach and uncertainty analysis and of their inherent relationship, a case study that was developed in collaboration with the German Federal Institute for Risk Assessment (BfR) is illustrated. The case study focuses on the identification and characterisation of uncertainty as an inherent part of the scientific assessment process and shows how uncertainty analysis can help to foster EFSA’s guiding principles for evidence-based scientific assessments. Although the case study is specific and makes use only of human data, its purpose is to provide a general illustration of the principles behind the framework and notably of the uncertainty assessment steps, highlighting that it is applicable to a more varied set of evidence.

## **3 EFSA’s guiding principles for evidence-based scientific assessments**

On the basis of its core values,<sup>2</sup> EFSA has defined four major principles to produce evidence-based scientific assessments: i) impartiality, ii) methodological rigour, iii) transparency, and iv) engagement. Overall, the extent to which these principles are fulfilled represents the ‘scientific value’ of an EFSA output (EFSA, 2018a; EFSA Quality Policy, in preparation).

Impartiality refers to the degree to which the scientific assessment process is free from preconceptions and bias due to prior knowledge of the results of the available studies or to any type of vested interest. It is promoted by planning the strategy for the scientific assessment *a priori* in a protocol (including anticipating as much as possible the possible sources of uncertainty), ensuring that the assessment is implemented in accordance with the protocol, and providing a mechanism to exclude those assessors for whom non-negligible conflicts of interest are identified (EFSA, 2015a, 2018a).

Methodological rigour concerns the extent to which systematic and random error are prevented, representativeness and generalisability are maximised, and uncertainty is accounted for in the scientific process (EFSA Quality Policy, in preparation). A structured and thorough analysis of the uncertainty arising from the methods applied and in the underpinning evidence is also fundamental to ensuring that an assessment is methodologically rigorous and sound.

An assessment is transparent when the supporting data, methods (including all assumptions), results and uncertainty are clearly reported, understandable and appraisable such that the process is as reproducible as possible. In regulatory science, the formal and transparent expression of the uncertainty in the results of an assessment is essential to providing the decision-maker with a basis for better-informed decisions (EFSA, 2019). Transparency also involves making all methods and data accessible (as appropriate considering the relevant legal constraints). Indeed, sometimes transparency is not fully achievable, given the confidentiality of certain data (EFSA, 2015a).

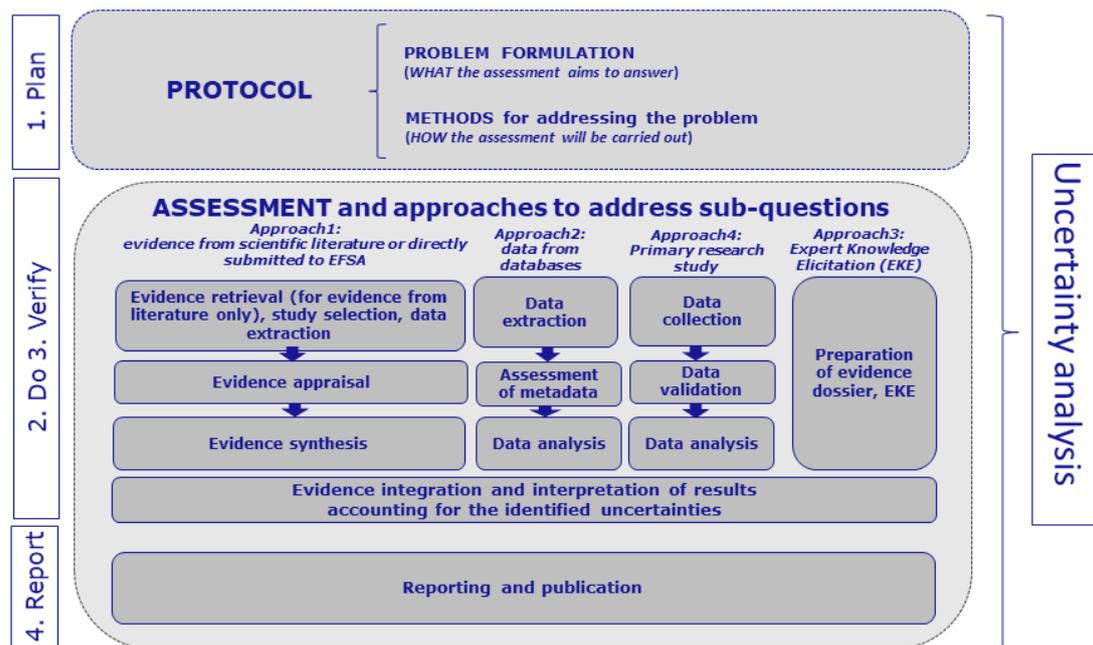
Engagement in a scientific assessment pertains to the ability to allow for, and benefit from, participation in the assessment from a wide range of parties — such as fellow institutional partners, individual professionals, national authorities, industrial or professional associations, academic operators, individual undertakings or Union citizens — who possess relevant knowledge and expertise (EFSA, 2018a).

In a mandate-driven environment like that of EFSA, assessments should also be targeted to the requestors’ needs. This is achieved by translating the terms of reference (ToRs) of the mandates into clearly formulated scientific problems that are clarified with the mandate requestors upfront, using methods and approaches that are functional within the context and requirements for the mandate, and meeting the expected deadlines. To this end, close consultation and on-going dialogue with the requestors throughout the assessment process is fundamental (EFSA, 2018a). This step, generally defined as

<sup>1</sup> <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R1381>. See also:

[https://ec.europa.eu/food/safety/general\\_food\\_law/transparency-and-sustainability-eu-risk-assessment-food-chain\\_en](https://ec.europa.eu/food/safety/general_food_law/transparency-and-sustainability-eu-risk-assessment-food-chain_en).

<sup>2</sup> <http://www.efsa.europa.eu/en/about/values>



**Fig. 1: The 4-step approach for EFSA's scientific assessments** based on (EFSA, 2015a, 2018a)

problem formulation, for mandates including a hazard identification or more generally a yes/no question, implies building a hypothesis with the purpose of falsifying it during the assessment process (e.g. H<sub>0</sub> - null hypothesis: 'there is no causal association between exposure to substance X and incidence of the disease Y'; H<sub>1</sub> - alternative hypothesis: 'there is causal association between exposure to substance X and incidence of the disease Y').

#### 4 Delivering 'evidence-based scientific assessments': the importance of planning upfront and of uncertainty analysis

Two central elements identified by EFSA for promoting the fulfilment of its principles for evidence-based assessments are the conduct of the scientific assessment process in four steps (plan/do/verify/report) and the thorough and transparent analysis of uncertainty throughout the process.

##### 4.1 EFSA's 4-step approach for conducting scientific assessments

EFSA's 4-step approach puts an emphasis on planning the strategy for the assessment in a protocol (step 1) developed prior to initiating any formal data collection, appraisal or synthesis (Fig. 1) (EFSA, 2015a, 2018a).

As a first step in protocol development, an explicit description of the mandate or question that must be answered is given (problem formulation). For broad questions, this involves defining as clearly as possible all potential underpinning sub-questions and their relationship. A well-defined (sub)-question refers to an outcome or quantity that could (in principle) be observed or measured without ambiguity in the real world or obtained from a defined scientific procedure. Each keyword requires an explicit definition and the population, region and time period of interest should be specified. For a variable quantity, the statistic(s) and/or quantities(s) are required (EFSA Scientific Committee, 2018a,b). A clear problem formulation is essential to defining the methods that will be applied for conducting the assessment, which include those for answering each sub-question (e.g., a primary data collection like an experiment or observational study; a review of existing data retrievable from the scientific literature or other sources; or — when evidence is scarce and/or of limited validity — expert knowledge elicitation)<sup>3</sup> and those for integrating evidence across sub-questions. Although a clear definition of the question is crucial to properly plan the methods and minimise ambiguity in the answer, uncertainties can arise from many other sources. The most recurrent ones are the limitations in the validity of the evidence and the methods to collect, analyse and integrate it, the heterogeneity of the results when it comes to the integration of several lines of evidence including NAMs, the difference in expertise and background of the people involved in the assessment. Also it must be acknowledged that any assessment is inherently dependent on the time and the place where it is performed. Both these aspects can limit the generalisability of the conclusions.

In the planning phase, the problem formulation is followed by the definition of the methods to address the (sub)questions arising from the mandate.

<sup>3</sup> Expert knowledge elicitation can also be used for i) supporting the protocol development phase, as e.g. a means to prioritise questions for more formal approaches, ii) replacing data collections and analyses under pressing conditions and iii), as explained in the next sections, assessing uncertainty (e.g. uncertainty arising from the evidence).

Uncertainty analysis is an inherent part of the scientific assessment process and methods to analyse it are also pre-defined in the protocol. Assessment methods are tailored according to the context, available resources and timelines. Therefore the 'extent of planning' can largely vary from one assessment to another.

The planning phase is followed by the actual assessment process, where the methods pre-defined in the protocol are implemented and conclusions are drawn in light of the identified uncertainties (step 2). During this step, the approach to answering each sub-question, including its degree of complexity and extensiveness, will vary depending on the decisions made in the protocol (Fig. 1): a) for sub-questions answered using data extracted from the scientific literature or submitted to EFSA via calls for data or application dossiers, the approach implies evidence retrieval/selection/data extraction, and evidence appraisal and synthesis and/or integration; b) when sub-questions are answered using data extracted from databases other than literature (e.g. Eurostat database),<sup>4</sup> the steps are those of data extraction, assessment of database meta-data and data analysis; c) when a primary research study (e.g. a survey) is carried out to answer a sub-question, the process is that of data collection, validation and analysis; while d) for those sub-questions answered by applying expert knowledge elicitation, an evidence dossier is prepared, followed by the actual elicitation process. Uncertainty analysis is inherent in all these steps. For instance, when using data extracted from the literature, evidence appraisal is the process of identifying and possibly quantifying the impact of the uncertainty arising from each individual study included in the assessment (owing to, for example, limited validity or precision). Then, such uncertainty is accounted for in the process for evidence synthesis and integration.

The last two steps in the 4-step approach for the scientific assessment process are those of checking and ensuring compliance with the plan (step 3) and of thoroughly reporting and publishing all methods, assumptions, data, results and related uncertainty, along with any deviations from the protocols (step 4) (Fig. 1) (modified from EFSA, 2015a, 2018a).

The use of protocols illustrating the design of a study is a well-established practice in primary research and several institutions have contributed to establishing the practice as a standard procedure also in literature-based assessments applying systematic review (e.g. The Cochrane Collaboration;<sup>5</sup> Campbell Collaboration;<sup>6</sup> Evidence-based Toxicology Collaboration;<sup>7</sup> The Collaboration for Environmental Evidence<sup>8</sup>). In addition, a growing number of initiatives have started to use protocols for broad assessments in contexts similar to EFSA's (e.g. OHAT-NTP, 2019; Woodruff & Sutton, 2014; WHO, 2012; or the US-EPA Integrated Risk Information System)<sup>9</sup>.

EFSA has tested and implemented the use of protocols in various types of assessment from different areas of food and feed safety and identified numerous benefits and some difficulties related to their application (EFSA, 2018a). Examples include:

- Risk assessment for peri- and post-menopausal women taking food supplements containing isolated isoflavones (EFSA Panel on Food Additives Nutrient Sources added to Food, 2015)
- Evaluation of the safety and efficacy of Listex™ P100 for reduction of pathogens on different ready-to-eat (RTE) food products (EFSA Panel on Biological Hazards, 2016)
- Risk for animal and human health related to the presence of dioxins and dioxin-like PCBs in feed and food (EFSA Panel on Contaminants in the Food Chain, 2018)
- Bluetongue: control, surveillance and safe movement of animals (EFSA Panel on Animal Health Welfare, 2017)
- Pest risk assessment of *Eotetranychus lewisii* for the EU territory (EFSA Panel on Plant Health, 2017)
- Bisphenol A (BPA) hazard assessment protocol (EFSA et al., 2017)
- Evaluation of the impact of glyphosate and its residues in feed on animal health (EFSA, 2018b)
- Protocol for the scientific opinion on the Tolerable Upper Intake Level of dietary sugars (EFSA, 2018c)
- Dietary reference values for sodium (EFSA Panel on Nutrition Novel Foods et al., 2019)<sup>10</sup>
- Draft protocol for the assessment of hazard identification and characterisation of sweeteners<sup>11</sup>

Currently the approach is being implemented in a project aimed at developing an informed Integrated Approach to Testing and Assessment (IATA) - Adverse Outcome Pathway (AOP) case study based on the risk assessment of Deltamethrin and Developmental Neurotoxicity outcomes<sup>12</sup>. In the case study, in addition to the traditional data from animal toxicity and human observational studies, a battery of invitro assays designed by EFSA and invitro studies from literature are being used. The focus of the project is measuring if and how uncertainty in the hazard identification of a class of chemicals can be reduced using novel types of evidence in combination with traditional ones. The EFSA framework for evidence-based scientific assessments is consistent with the principles of the IATA approach developed by the OECD<sup>13</sup> (described in Casati, 2018).

Planning and designing the methods for the assessment *a priori* acts as a guard against arbitrary decision-making during the assessment process and represents an effective way to protect from cognitive biases,<sup>14</sup> as the outcomes are not yet known when the methods are defined (Munafò et al., 2017; Shamseer et al., 2015). This increases impartiality in the

<sup>4</sup> <https://ec.europa.eu/eurostat/data/database>

<sup>5</sup> <https://www.cochrane.org/> See also (Higgins JPT, 2019)

<sup>6</sup> <https://campbellcollaboration.org/>

<sup>7</sup> <http://www.ebtox.org/>

<sup>8</sup> <http://www.environmentalevidence.org/>

<sup>9</sup> <https://www.epa.gov/iris>

<sup>10</sup> related protocol available at: <https://zenodo.org/record/1116290#.XZMWWEYzaUI>

<sup>11</sup> public consultation on draft protocol available at: <https://www.efsa.europa.eu/en/consultations/call/public-consultation-draft-protocol-assessment>

<sup>12</sup> <http://registerofquestions.efsa.europa.eu/roqFrontend/wicket/page?5>

<sup>13</sup> <http://www.oecd.org/chemicalsafety/risk-assessment/iata-integrated-approaches-to-testing-and-assessment.htm>

<sup>14</sup> E.g. confirmation bias: the tendency to focus on evidence that is in line with expectations or favoured explanation (Kerr, 1998).

assessment process. It also helps increase methodological rigour, as it reduces methodological flaws like HARKING (Hypothesising After the Results are Known) or data-contingent analysis decisions (P-hacking), by requiring assessors to articulate analytical decisions prior to acquiring knowledge about the available results, thereby ensuring that decisions remain data-independent (Munafò et al., 2017).

The time dedicated to planning a scientific assessment promotes a more efficient use of the resources in the subsequent implementation phase. This includes better-structured activities and tasks within the assessment group, making the evaluation easier, as well as more efficient discussions within the group of assessors (EFSA, 2018a).

If developed in consultation with the mandate requestors, protocols can also help ensure the subsequent assessment is tailored to their needs, as the question that will be answered — as well as the methods for addressing it — are agreed upon in advance.

Draft protocols can also be shared with external parties to receive feedback and input on the draft plan and refine it before starting the assessment, thereby encouraging engagement in the scientific assessment process. EFSA has tested this approach in some scientific assessments for which draft protocols were made available to the public and subsequently refined based on the outcomes of the consultation process (e.g. (EFSA Panel on Nutrition Novel Foods, 2019; EFSA, 2018c; EFSA et al., 2017)). EFSA's protocols are published also in the Agency's Knowledge Junction,<sup>15</sup> which is a curated, open repository for the exchange of evidence and supporting materials used in food and feed safety scientific assessments.

Among the difficulties of implementing protocols at EFSA, one is certainly the fact that, for broad assessments like the ones arising from EFSA's mandates, protocol development is resource-intensive and may require a long, iterative process supported by extensive literature scoping and continuous expert input prior to being finalized. This was particularly evident for those assessments where a protocol was implemented but extensive and complex methods were also required (e.g. systematic review) (e.g. (EFSA Panel on Contaminants in the Food Chain, 2018; EFSA Panel on Nutrition Novel Foods, 2019)).

## **4.2 Uncertainty analysis**

Uncertainty analysis is an integral and fundamental component of the scientific assessment process. Its role has been acknowledged by several organisations (e.g. EC, 2019; EU ANSA, 2018; IPCS, 2017; SAPEA, 2019), including EFSA and the BfR, which developed guidance aimed at supporting uncertainty analysis in their scientific assessments (EFSA Scientific Committee 2018a,b; Heinemeyer et al., 2015). EFSA defines uncertainty as all types of limitations in the knowledge available to the risk assessors at the time an assessment is conducted and within the time and resources available for the assessment. Uncertainty analysis is the process of identifying limitations in scientific knowledge and evaluating their implications for scientific conclusions, if possible, in terms of the possible range and probability of possible answers to the assessment question. The EFSA guidance on uncertainty is not prescriptive with regard to specific methods for uncertainty analysis and provides a flexible framework allowing for the selection of different quantitative and qualitative methods. The form and extent of uncertainty analysis, and how the conclusions are reported, vary depending on the nature and context of each assessment and the degree of uncertainty present.

The main elements of uncertainty analysis (EFSA Scientific Committee, 2018a,b) when applying the 4-step approach during the scientific assessment processes are described herein.

Uncertainty analysis starts at the planning stage of the EFSA 4-step approach — during the development of the assessment protocol — and ends at its final stage when conclusions and underlying uncertainties are reported and communicated.

When planning the data and methodologies for the scientific assessment during protocol development (step 1), any major uncertainties affecting the assessment need to be identified in a structured way. If possible, it should be clarified whether any of these uncertainties will be quantified individually within a particular assessment component or quantified collectively later (when assessing overall uncertainty). Sensitivity or influence analysis can help prioritise uncertainties in this step. In some assessments, it may be sufficient to characterise overall uncertainty for the whole assessment directly, by expert judgement. In other cases, it may be preferable to evaluate uncertainty for some or all parts of the assessment separately and then combine them, either by calculation or expert judgement.

EFSA's guidance on uncertainty offers examples of general types of uncertainties to support the identification of uncertainties in a scientific assessment. Although not strictly needed for conducting an uncertainty analysis, EFSA's classification of uncertainties into those associated with the evidence and those associated with the assessment methods can help identify uncertainties within an assessment process.

In a primary research study, examples of potential sources of uncertainty in the methods can be the non-random selection of the sampling units or the use of an imprecise scale or inadequate rules for data correction. They can also extend to the use of assumptions behind a mathematical model used to analyse data. When using existing data, other examples of uncertainty associated with methodological flaws can include the uncertainty in the conceptual model developed for the assessment (e.g. the probabilistic model used to estimate exposure to a chemical or the theoretical model used to describe the possible mode of action of a chemical) or the one related to the criteria for selecting studies for the assessment.

In an assessment based on existing data, uncertainty associated with the evidence can arise at two levels: a) from each piece of evidence, such as the uncertainty due to threats to the study validity and precision (e.g. limited sample size for a subgroup of the population in a food consumption survey, poor sensitivity of an analytical method used to measure the occurrence of a substance in a class of food items); or b) at the level of the overall body of evidence, like the uncertainty due to heterogeneity among different sources or — when using published data — uncertainty owing to publication bias.

---

<sup>15</sup> <https://zenodo.org/communities/efsa-kj/?page=1&size=20>

During the implementation and verification phases of the scientific assessment process (step 2 and 3 in EFSA's 4-step approach), additional sources of uncertainty may be identified and should be added to the initial list defined during protocol development. If the assessment involves calculations, preliminary results should be revised to consider whether it would be beneficial to quantify any additional uncertainties within the calculations or include them in a sensitivity analysis. When developing draft conclusions on the terms of reference, the risk assessor should ensure that these uncertainties are well-defined. The overall uncertainty of each conclusion should be quantified. If there are many conclusions, those where the uncertainty will have most impact on the conclusions, if this is known, should be prioritised. Otherwise, those conclusions that address the terms of reference most directly and/or are more uncertain should be prioritised. If no uncertainties were quantified in earlier steps, the combined impact on each conclusion of all the identified uncertainties should be assessed by expert judgement. If any uncertainties were quantified in earlier steps, these results should be considered together with all the other uncertainties, either by expert judgement or by a combination of expert judgement and calculation. Formal or semi-formal expert knowledge elicitation procedures can be used for the judgements required to assess overall uncertainty (EFSA Scientific Committee, 2018b). Finally, when reporting the scientific assessment process (step 4) methods and results of the uncertainty analysis should be reported in a transparent way. Any sources of uncertainty which were not included in the quantitative expression should be highlighted, and any assumptions about them should be reported. A recent EFSA guidance document on how to communicate uncertainty in scientific assessments (EFSA, 2019) contains specific guidance for assessors on how to best report the various expressions of uncertainty.

A challenge when addressing uncertainties, particularly in regulatory toxicology, is posed by possible resistances in embracing new methodological approaches by regulators. Historical, conventional agreements and regulatory, authoritative situations could make it difficult to discuss uncertainties of current practice, since political decisions could have been made already on their basis. Scientific evidence is currently lacking on the impact that assessing uncertainty can have on the decision-making process.

Reliability and relevance of animal tests for human or environmental risk assessments is also a field where a discussion of the critical uncertainties still faces hurdles. Data from NAMs and alternative testing strategies have a great potential in the regulatory context but they are also posing new challenges. The use of invitro assay battery for instance is currently hampered by limitations in the approaches to extrapolate in vitro effects to in vivo responses (Bell et al., 2018). Translating a nominal effective concentration into a toxic potency of a chemical in humans still represents a major uncertainty affecting invitro testing methods (Wetmore 2015, Groothuis et al., 2015). To improve the predictive power of the alternative methods, more knowledge is needed to understand possible differences between invitro and in-vivo systems in clearance, protein binding, bioavailability and other pharmacokinetics factors (Wilk-Zasadna et al, 2015).

Planning upfront, as proposed in the framework described in this paper, can represent an opportunity to engage in an early stage with regulators, stakeholders and general public also to clarify the importance to transparently assess the uncertainties surrounding the evidence and the methods in the scientific assessment process. Sharing the protocol can help agreeing, before conclusions are reached, on the approach to use to address the question, including methods for uncertainty analysis. It might represent also a tool to help producing a cultural shift in the regulatory field, focusing the discussion on the methods to reach the conclusions rather than solely on the conclusions.

## **5 A case-study: exposure assessment of aluminium in chocolate products**

In the context of their long-term collaboration, EFSA and BfR decided to test the applicability of their respective guidance on uncertainty analysis and to compare the related recommendations. The task was carried out by BfR based on a grant agreement with the EFSA (details can be found in (EFSA et al., 2018)). The exposure assessment for aluminium in cocoa and chocolate products was chosen as one of case study owing to the high level of aluminium detected in these food items and the many uncertainties identified by BfR in a previous assessment. This example is used in this paper to illustrate a) how uncertainty analysis represents an inherent element of all steps of the scientific assessment process; and b) the role of uncertainty analysis to fulfil EFSA's guiding principles for evidence-based assessments.

The implementation of the uncertainty analysis for the exposure to aluminium was preceded by a plan that covered all the steps recommended by the EFSA guidance on uncertainty (EFSA Scientific Committee, 2018b): the formulation of the question, the identification of the sources of uncertainty, the prioritisation of the most influential uncertainties, the analysis of the uncertainties and the overall characterisation of the uncertainty, and transparent communication (Fig. 2). The strategy for the analysis was not reported in a formal protocol since the implementation of the '4-step approach' was beyond the objective of the project.

During the planning phase, emphasis was put on the need to ensure that questions and quantities of interest were well-defined to reduce ambiguity and the risk of drawing conclusions that would not address the true question of interest. The scientific question was expressed as '*to estimate the 2017 average long-term aluminium intake (chronic toxicity) by consumption of chocolate and cocoa products in 2017 for infants from age 0.5 years to less than 5 years (which are not breastfed) in Germany for the 95th percentile of the population specified above (in µg/(kg bw)/week). Further stratification of the described population is not desired*'. The formulation of the problem included the identification of the target population, the exposure, the reference time and place and the level of risk that was considered acceptable (95th percentile). A tolerable weekly intake (TWI) of 1 mg/kg for aluminium was used as reference dose (EFSA Panel on Food Additives, 2008). In order to improve the precision of the estimates, the food items containing cocoa and chocolate were clustered in seven food categories: sugar panned chocolate; milk chocolate/baking chocolate; chocolate icing/chocolate sprinkles/chocolate coating; chocolate with fillings; dark chocolate; cocoa powder; beverages containing cocoa powder; chocolate-, nougat-, and cocoa-cream.

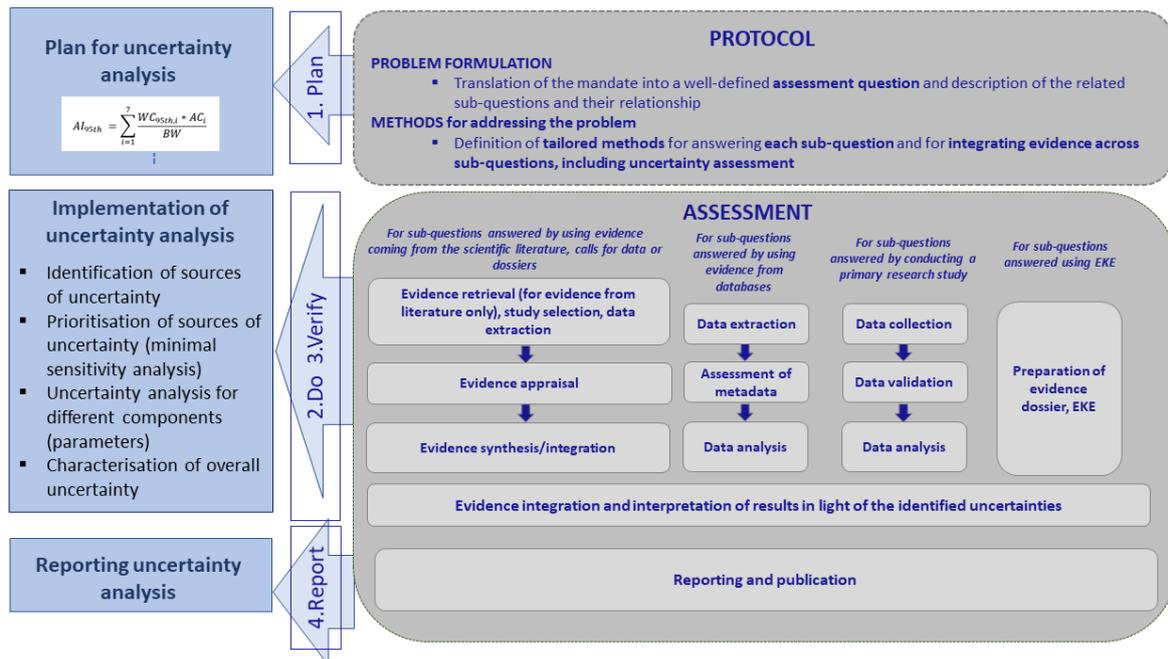


Fig. 2: The 4-step approach for EFSA's scientific assessments and the uncertainty analysis case study

A conceptual model, that took the form of a simple probabilistic model, was set up during the problem formulation (formula below):

The three components of the model were addressed as sub-questions requiring the quantification of the: a) 95th percentile of the weekly consumption of cocoa and chocolate products in the German population of children aged 0.5-5 years that were not breastfed ( $WC_{95th,i}$ ); b) aluminium content in each of the seven categories in which cocoa and chocolate products were aggregated ( $AC_i$ ); and c) children average body weight ( $BW$ ). Despite its simplicity, the model involved a series of

$$AI_{95th} = \sum_{i=1}^7 \frac{WC_{95th,i} \cdot AC_i}{BW}$$

assumptions and methodological choices that had the potential to influence the final results. For instance, a decision had to be made on the level of conservatism to take. Several options were available for the computation of the 95th percentile of the aluminium intake, including using the average aluminium concentration in the seven categories of products or using high aluminium concentration for one category and the average for the others. In addition, for the 'chocolate-, nougat, and cocoa-cream' category, the aluminium content was not available and it had to be estimated assuming that its average content of cocoa powder was 10%.

In parallel with the identification of the sources of evidence for each of the three components of the models (data collection in Fig. 2), the uncertainties stemming from lack of knowledge and other limitations in the evidence were discussed and listed (appraisal of the evidence in Fig. 2). No formal critical appraisal tools were used to assess the internal validity of the evidence sources, though considerations were made regarding the methods used to collect data that could have biased the results. The external validity of the evidence was also addressed in terms of the representativeness of the sample with respect to the target population (children 0.5-5 years old) and directness of the evidence to the target time of the estimate (year 2017) and quantity to be estimated (long-term 95th percentile of the weekly intake of aluminium from cocoa and chocolate products).

The dietary consumption data were derived from the VELs survey (Heseker et al., 2003) conducted in 2001/2002 on small children that were not breastfed (age 6 months to less than 5 years; 732 individuals in total). Several uncertainties were identified. It was not clear from the survey whether aluminium content is brand-dependent or whether consumers are brand-loyal. It is also possible that consumption of cocoa and chocolate products, the main target of the assessment, has changed between 2001/2002 and 2017. The sampling error associated with the estimate of the 95th percentile of the population is high, due to the limited number of participants in the survey. The long-term intake of chocolate and cocoa could not be directly assessed and had to be extrapolated from the 6-day dietary diaries survey. The portion size estimation might have been affected by measurement errors due to weighing. The weight of the children was also derived from the VELs survey. Evidence used for this variable was subject to uncertainty arising from sampling and measurement. The occurrence of aluminium in cocoa and chocolate was derived from 1,646 measurements taken from food products aggregated to the seven groups defined previously. The occurrence was not directly measured for the 'chocolate, nougat, and cocoa-cream' category and had to be derived using aluminium occurrence in the cocoa powder and assumptions about its content in the specific food category (10%). Estimates were affected by measurement and sampling errors. Food samples and aluminium content were collected prior to 2017 and the representativeness of the data for the target year is uncertain. The

analytical method was unable to detect content below a certain limit. Some measurements were reported as zero when below the limit of detection.

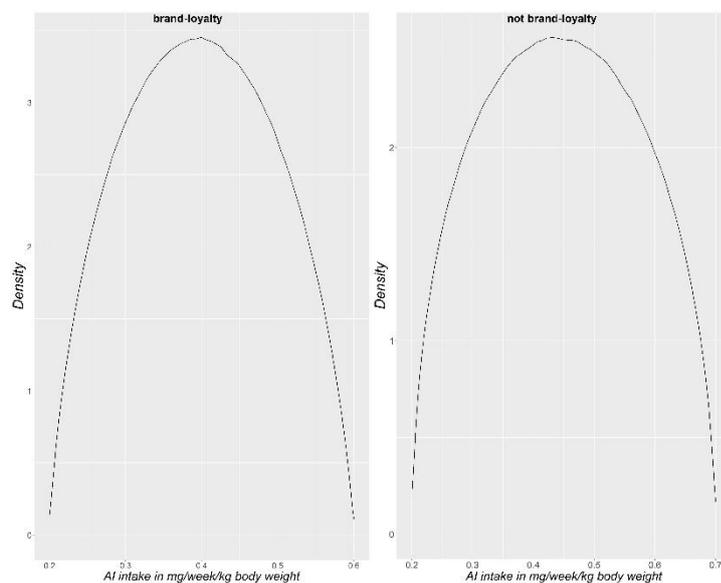
Following the recommendations within EFSA guidance, a simple influence analysis was performed to identify the sources of uncertainty that had the potential to be more influential on the estimate of the 95th percentile of aluminium intake in children. The 'one at a time' method was used, assessing the impact of each source of uncertainty one by one, while the other variables remain constant, assuming realistic alternative values for the variable affected by uncertainty. It is assumed that there is no dependency among variables. The results of the influence analysis indicated that four sources of uncertainty had a higher potential to influence the estimate of the aluminium intake of the 95th percentile of the target population: a) brand-loyalty of the consumers and dependency from brand of the aluminium content; b) actualization to 2017 of the consumption pattern from 2001/2002, especially for the cocoa powder (larger contributor to aluminium intake); c) sampling error of the estimated 95th percentile for the dietary intake of cocoa and chocolate products in the VELS survey; and d) sampling error in aluminium measurements.

For each of the most influential sources of uncertainty, an approach was planned to first quantify them and then to integrate them in the probabilistic model (synthesis and integration of the evidence in Fig. 2). Methods included scenario analysis (brand-loyalty of the consumers and brand dependency of the aluminium content), expert knowledge elicitation (change in consumption of cocoa powder), and bootstrapping methods (sampling error for dietary intake and aluminium measurement). Uncertainty distributions were derived for each of the three components/sub-question of the model and combined using Monte Carlo simulations in order to derive an uncertainty distribution for the aluminium intake.

The uncertainty distribution also embedded some uncertainties stemming from the methods used. The uncertainty in the estimate of the parameters used to describe the distributions of the aluminium content in the seven categories in the brand-loyal scenario were addressed using bootstrapping. Other uncertainties were not addressed individually. Different ways of grouping the food items and the assumptions made to extrapolate from short to long-term consumption might have meaningfully impacted results.

The uncertainties in the evidence and those stemming from the assumptions and methodological choices for the model that were not addressed individually were collectively assessed afterwards. These included: a) integrating all sub-questions (evidence integration across sub-questions in Fig. 2); b) deriving an initial uncertainty distribution from the probabilistic model; and c) integrating in the initial estimate the uncertainty in the possible change of cocoa power consumption between 2001 and 2017. The final step of accounting for the remaining sources and quantifying the overall uncertainty was achieved using an expert knowledge elicitation process (EFSA, 2014). Examples of methodological uncertainties addressed at this stage included the number of simulations used for the Monte Carlo algorithm, changes of cocoa power consumption being applied homogeneously to all individuals in the VELS survey, and the use of two-parameter distributions to model aluminium content. Remaining limitations in the evidence included the contribution to aluminium intake from foods containing cocoa/chocolate that were not included in the model. The final uncertainty distributions for the two scenarios (brand loyalty and not brand-loyalty) are depicted in Fig. 3.

Comparison of the above distributions and percentiles with the results of the original assessment performed by BfR led to the following conclusions (EFSA, 2018a). As for Scenario 1 (median aluminium concentration in each of the seven product groups), the BfR exposure assessment provided an estimate of 347  $\mu\text{g}/(\text{week}\cdot\text{kg bw})$  for the 95th percentile of population. The median estimate of the final distribution derived for Scenario 1 in this uncertainty analysis yielded 400  $\mu\text{g}/(\text{week}\cdot\text{kg bw})$  with an interquartile range of 320–470  $\mu\text{g}/(\text{week}\cdot\text{kg bw})$  and an overall range of 200–600  $\mu\text{g}/(\text{week}\cdot\text{kg bw})$  (Tab. 1). For Scenario 2, the opposite trend was observed in comparing the BfR estimate of between 488–565  $\mu\text{g}/(\text{week}\cdot\text{kg bw})$ , depending on the group of products for which 95th percentile of the content was assumed, to the median value of 440  $\mu\text{g}/(\text{week}\cdot\text{kg bw})$  derived from this analysis (interquartile range of 350–550  $\mu\text{g}/(\text{week}\cdot\text{kg bw})$  and an overall range of 200–700  $\mu\text{g}/(\text{week}\cdot\text{kg bw})$ ) (Tab. 2). In this second scenario, some differences in the methodological approaches taken by the two institutions that could go beyond the number and types of uncertainties that have been addressed must be acknowledged.



**Fig. 3: Uncertainty distributions of the 95<sup>th</sup> percentile of aluminum intake elicited for the two scenarios (brand loyalty on left, not-brand loyalty on the right).** Reprinted with permission from (EFSA et al., 2018), page 102, Fig. 9 and 10. The respective percentiles are reported in Tables 1 and 2.

Overall a large uncertainty was identified and quantified within the present analysis. Uncertainty quantification, expressed as the difference between the third and the first quartile, attained 38% of the median value for Scenario 1 and is closed to 45% in Scenario 2.

**Tab. 1: Scenario 1: 95<sup>th</sup> percentile of aluminium intake in µg/(week\*kg bw) among children aged 0.5–5.5 years who were not breastfed**

Percentile	1%	25%	50%	75%	99%
Initial model result	363	394	408	425	461
After one-parameter EKE	339	377	401	427	472
After overall EKE	200	320	400	470	600

Adapted from (EFSA et al., 2018), page 100.

**Tab. 2: Scenario 2: 95<sup>th</sup> percentile of aluminium intake in µg/(week\*kg bw) among children aged 0.5–5.5 years who were not breastfed**

Percentile	1%	25%	50%	75%	99%
initial model result	379	428	450	475	547
after one-parameter EKE	363	416	443	472	551
after overall EKE	200	350	440	550	700

Adapted from (EFSA et al., 2018), page 100

## 6 Considerations on uncertainty analysis

Overall, this example shows how a thorough application of uncertainty analysis throughout the entire assessment process can help to increase the scientific value of an output. Even though a protocol detailing the plan for uncertainty analysis was not published, an attempt was made to identify upfront as many sources of uncertainty as possible, with the aim of reducing data-driven decisions later in the assessment process and, in turn, making the assessment as impartial as possible. In addition, by addressing the main sources of uncertainty, the results and conclusions of the assessment are provided in a more rigorous and transparent way.

This conclusion was also reached at the end of a one-year trial phase at EFSA, during which the uncertainty guidance was tested in the various EFSA scientific panels (EFSA, 2017). Most of the participants involved in the trial phase (79%) agreed that applying the guidance improved transparency about uncertainties related to the risk assessment. The systematic identification of uncertainties provided a more structured and documented process, allowing for higher consistency in the formulation of scientific assessment conclusions. Analysing the uncertainties allowed the risk assessors to provide a more comprehensive scientific assessment, producing conclusions with which they felt more confident. Also, analysing uncertainty in a scientific assessment helped draw conclusions for which previously inconclusive opinions may have been the result (EFSA, 2017). In contrast, the risk assessors encountered several challenges during the trial phase. Experience in — and acceptance of — applying expert knowledge elicitation methodologies to assess uncertainty still require further development. Also, difficulties defining an acceptable level of quantified uncertainty were observed alongside challenges defining a threshold below which more data would be needed before concluding the scientific assessment. This would need close collaboration and agreement between the risk assessors and managers.

A few risk managers/policy-makers participating in EFSA's trial phase on the uncertainty guidance found it difficult to fully understand the scientific assessment's conclusion when uncertainty was presented, whereas, for others, their understanding of the conclusions improved as a result of uncertainty analysis aiding in providing the reasoning for the conclusions (EFSA, 2017).

The importance of uncertainty analysis for scientific assessments, the associated implications for decision making, and the need to communicate the most relevant uncertainties to decision makers and to the broad public was emphasised at a recent conference on Uncertainty in Risk Analysis in Berlin. Among other aspects, it was concluded that on the one hand, humans do not seem to have a good intuitive understanding of uncertainty, on the other hand they make decisions in the face of uncertainty all the time. It was stressed that understanding, processing, and accounting for uncertainty in decision making can be trained. It was recommended that for educational curricula, additional focus on statistics is desirable and that uncertainty analysis should be addressed in education of natural scientists (EFSA and BfR, 2019).

## 7 Discussion and concluding remarks

EFSA developed a methodological framework for producing evidence-based scientific assessments consisting of a set of guiding principles (impartiality, methodological rigour, transparency and engagement, which overall reflect the 'scientific value' of an output) and several cross-cutting methodological guidance documents that contribute to fulfilling those principles from different angles. Among these, the two approaches described in this paper (i.e. conducting the assessment in four steps – plan/do/verify/report - and analysing uncertainty throughout the process) were tested in several case studies which helped develop insights regarding the benefits and challenges related to their application. While the advantages are attributable to the increase in the scientific value of the final output, the related challenges require some considerations. First, from a methodological viewpoint, experience and further capacity-building is needed to more formally integrate the two approaches and better incorporate uncertainty analysis in the planning phase of the scientific assessment process. Second, it is observed that, in the short term, applying the 4-step approach (and, in particular, developing extensive protocols) and a structured uncertainty analysis in EFSA's assessments can be resource-intensive, especially with regard to the time and

expertise needed. In the longer term, however, as methods are developed further and regularly implemented and as experience in their application increases, the application of these approaches is expected to become progressively easier, with a likely decrease in the need for excessive methodological expertise, time and budget. Re-use and ad-hoc adaptation of existing protocols and approaches to analysing uncertainty can also facilitate regular implementation and, in turn, further increase the scientific value of EFSA's outputs.

Relying on an approach for evidence-based assessments that includes an analysis of the uncertainties can also contribute to foster the uptake of alternative methods and testing strategies in the regulatory context. Moving towards a more prominent use of NAMs is also envisaged by the EFSA's strategy 2020 (EFSA, 2016). The latter foresees the 'development and gradual integration in EFSA guidance of new approaches in prioritised chemical and biological risk assessment areas to strengthen EFSA's capability to deal with the absence of data, address complex questions and reduce uncertainty. Although the case study shown in this paper uses only human data, it provides a general illustration of the principles behind the framework and notably of the uncertainty assessment steps, highlighting that it is applicable to a much broader set and types of evidence, including those arising from NAMs.

## References

- Bell, S.M., Chang, X., Wambaugh, J.F. et al. (2018). In vitro to in vivo extrapolation for high throughput prioritization and decision making. *Toxicol in Vitro* 47, 213-227. doi:10.1016/j.tiv.2017.11.016
- Casati, S. (2018): Integrated Approaches to Testing and Assessment. *Basic & Clinical Pharmacology & Toxicology* 123, 51–55. doi:10.1111/bcpt.13018
- EC (2002). Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety. <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32002R0178>
- EC - European Commission, Chief Scientific Advisors, Scientific Advice Mechanism (SAM),. (2019). *Scientific Advice to European Policy in a Complex World*. doi:10.2777/80320
- EFSA - European Food Safety Authority (2010). *Application of systematic review methodology to food and feed safety assessments to support decision making*. Retrieved from <http://www.efsa.europa.eu/en/efsajournal/pub/1637.htm>
- EFSA (2014). *Guidance on expert knowledge elicitation in food and feed safety risk assessment* (1831-4732). Retrieved from <http://www.efsa.europa.eu/en/efsajournal/pub/3734>
- EFSA (2015a). *Principles and process for dealing with data and evidence in scientific assessments*. Retrieved from <http://www.efsa.europa.eu/en/efsajournal/pub/4121>
- EFSA (2015b). Editorial: Increasing robustness, transparency and openness of scientific assessments. *EFSA J* 13, e13031. doi:10.2903/j.efsa.2015.e13031
- EFSA (2016): EFSA Strategy 2020. Trusted science for safe food. Protecting consumers' health with independent scientific advice on the food chain. European Food Safety Authority. <http://bookshop.europa.eu/en/efsa-strategy-2020-pbTM0116289/>
- EFSA (2017). *Workshop on the trial of the EFSA guidance document on uncertainty analysis in scientific assessments 22 - 23 June 2017 Parma, Italy* (2397-8325). Retrieved from <https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/sp.efsa.2017.EN-1313>
- EFSA (2018a). *Implementation of PROMETHEUS 4-step approach for evidence use in EFSA scientific assessments: benefits, issues, needs and solutions* (2397-8325). Retrieved from <https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/sp.efsa.2018.EN-1395>
- EFSA (2018b). Evaluation of the impact of glyphosate and its residues in feed on animal health. *EFSA J* 16, e05283. doi:10.2903/j.efsa.2018.5283
- EFSA (2018c). Protocol for the scientific opinion on the Tolerable Upper Intake Level of dietary sugars. *EFSA J* 16, e05393. doi:10.2903/j.efsa.2018.5393
- EFSA (2019). Guidance on communication of uncertainty in scientific assessments. *EFSA J* 17, e05520. doi:10.2903/j.efsa.2019.5520
- EFSA and BfR. (2019). International Conference on Uncertainty in Risk Analysis. *EFSA Supporting Publication*, 1689. Retrieved from <https://www.efsa.europa.eu/en/supporting/pub/en-1689>
- EFSA, Gundert-Remy, U., Bodin, J. et al. (2017). Bisphenol A (BPA) hazard assessment protocol. *EFSA Supporting Publications* 14, 1354E. doi:10.2903/sp.efsa.2017.EN-1354
- EFSA, Schendel Thomas, J. C., Lindtner, O. et al. (2018). *Guidelines for Uncertainty Analysis: Application of the respective documents of EFSA and BfR for exposure assessments*. *EFSA supporting publication 2018: 15(7): EN-1472*. 113 pp. doi: 10.2903/sp.efsa.2018.EN-1472 (2397-8325). Retrieved from <https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/sp.efsa.2018.EN-1472>
- EFSA Panel on Animal Health Welfare (2017). Bluetongue: control, surveillance and safe movement of animals. *EFSA J* 15, e04698. doi:10.2903/j.efsa.2017.4698
- EFSA Panel on Biological Hazards (2016). Evaluation of the safety and efficacy of Listex™ P100 for reduction of pathogens on different ready-to-eat (RTE) food products. *EFSA J* 14, e04565. doi:10.2903/j.efsa.2016.4565
- EFSA Panel on Contaminants in the Food Chain (2018). Risk for animal and human health related to the presence of dioxins and dioxin-like PCBs in feed and food. *EFSA J* 16, e05333. doi:10.2903/j.efsa.2018.5333
- EFSA Panel on Food Additives, Flavoring, Processing Aids and Food Contact Materials. (2008). *Scientific Opinion on the Safety of aluminium from dietary intake - Scientific Opinion* (1831-4732). Retrieved from <https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2008.754>

- EFSA Panel on Food Additives and Nutrient Sources added to Food (2015). Risk assessment for peri- and post-menopausal women taking food supplements containing isolated isoflavones. *EFSA J* 13, 4246. doi:10.2903/j.efsa.2015.4246
- EFSA Panel on Nutrition Novel Foods (2019). Dietary reference values for sodium. *EFSA J* 17, e05778. doi:10.2903/j.efsa.2019.5778
- EFSA Panel on Plant Health (2017). Pest risk assessment of *Eotetranychus lewisi* for the EU territory. *EFSA Journal*, 15(10), e04878. doi:10.2903/j.efsa.2017.4878
- EFSA Scientific Committee (2018a). Guidance on Uncertainty Analysis in Scientific Assessments. *EFSA Journal*, 16(1), 39. doi:10.2903/j.efsa.2018.5123
- EFSA Scientific Committee (2018b). Scientific Opinion on the principles and methods behind EFSA's Guidance on Uncertainty Analysis in Scientific Assessment on The principles and methods behind EFSA's Guidance on Uncertainty Analysis in Scientific Assessment. *EFSA J* 16, e05122. doi:10.2903/j.efsa.2018.5122
- EFSA Scientific Committee (2017a). Guidance on the use of the weight of evidence approach in scientific assessments. *EFSA J* 15, 4971, 15, 69. doi:10.2903/j.efsa.2017.4971
- EFSA Scientific Committee (2017b). Guidance on the assessment of the biological relevance of data in scientific assessments. *EFSA Journal* 2017; 15,4970, 15, 73. doi:10.2903/j.efsa.2017.4970
- EU ANSA, (2018). *Approaches to assess and manage scientific uncertainty: examples from EU ANSA Agencies*. Retrieved from <https://publications.europa.eu/en/publication-detail/-/publication/9880c8bc-83eb-11e8-ac6a-01aa75ed71a1>.
- Groothuis, F. A., Heringa M. B., Nicol, B. et al. (2015); dose metric considerations in in vitro assays to improve quantitative in vitro- in vivo dose extrapolations. *Toxicology*, 332, 30-40. doi:10.1016/j.tox.2013.08.012
- Heinemeyer G., Kreienbrock, L., Schümann, M. et al. (2015). Guidelines on Uncertainty Analysis in Exposure Assessments. Recommendation of the Committee for Exposure Assessment and Standardisation of the Federal Institute for Risk Assessment (BfR). Series: BfR-Wissenschaft, Vol. 03/2015, 56 pp. Berlin, DE: Bundesinstitut für Risikobewertung (BfR). Retrieved from <https://www.bfr.bund.de/cm/350/guidelines-on-uncertainty-analysis-in-exposure-assessments.pdf>
- Heseker, H., Oeppining, A., and Vohmann, C. (2003). *Verzehrsstudie zur Ermittlung der Lebensmittelaufnahme von Säuglingen und Kleinkindern für die Abschätzung eines akuten Toxizitätsrisikos durch Rückstände von Pflanzenschutzmitteln (VELS)*. Paderborn DE: Universität Paderborn. doi.org/10.1007/s00003-006-0118-y
- Higgins J. P. T, Thomas, J., Chandler, J. et al. (editors). (2019). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). Retrieved from <https://training.cochrane.org/handbook>
- ICCVAM – Interagency coordinating Committee on the validation of alternative methods (2017): A Strategic Roadmap for Establishing New Approaches to Evaluate the Safety of Chemicals and Medical Products in the United States. [https://ntp.niehs.nih.gov/ntp/about\\_ntp/bsc/2017/december/meetingmaterials/iccvam\\_roadmap\\_508.pdf](https://ntp.niehs.nih.gov/ntp/about_ntp/bsc/2017/december/meetingmaterials/iccvam_roadmap_508.pdf)
- IPCS - International Programme on Chemical Safety (2017). *Guidance Document on Evaluating and Expressing Uncertainty in Hazard Assessment, second edition. IPCS Harmonization Project Document No. 11. World Health Organisation, Geneva*. Retrieved from <http://www.inchem.org/documents/harmproj/harmproj/harmproj11.pdf>.
- Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review* 2, 196-217. doi:10.1207/s15327957pspr0203\_4
- Munafo, M. R., Nosek, B. A., Bishop, D. V. M., et al. (2017). A manifesto for reproducible science. *Nature Human Behaviour* 1. doi:UNSP 002110.1038/s41562-016-0021
- NRC – National Research Council (2007): *Toxicity Testing in the 21<sup>st</sup> Century: A Vision and a Strategy*. Washington, DC: The National Academies Press. doi:10.17226/11970.
- OHAT-NTP. (2019). *Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration*. Available at: [https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookmarch2019\\_508.pdf](https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookmarch2019_508.pdf).
- SAPEA (Science Advice for Policy by European Academies). (2019). *Making sense of science for policy under conditions of complexity and uncertainty*. Berlin: SAPEA. doi:10.26356/MASOS.
- Shamseer, L., Moher, D., Clarke, M. et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*, 349, g7647. doi:10.1136/bmj.g7647
- Wetmore, B. A. (2015). Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment. *Toxicology*, 332, 94-101. doi:10.1016/j.tox.2014.05.012
- Wilk-Zasadna, I., Bernasconi, C., Pelkonen, O., and Coecke, S. (2015). Biotransformation in vitro: An essential consideration in the quantitative in vitro-to-in vivo extrapolation (QIVIVE) of toxicity data. *Toxicology* 332, 8–19. doi:10.1016/j.tox.2014.10.006
- Woodruff, T. J., and Sutton, P. (2014). The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environ Health Perspect*, 122(10), 1007-1014. doi:10.1289/ehp.1307175
- WHO - World Health Organisation (2012). *WHO Handbook for Guideline Development*. Retrieved from <http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441eng.pdf>

### Conflict of interest

The authors declare that they have no conflicts of interest.

### Acknowledgements

The authors wish to thank the German Federal Institute for Risk Assessment (BfR) and, in particular, T. Schendel, C. Jung, O. Lindtner and M. Greiner M, who were the major developers of the case study presented in this paper. They also extend their thanks to EFSA staff member Olaf Mosbach-Schulz for contributing to the case study.