# The European Commission's science and knowledge service

Joint Research Centre

# The assessment of uncertainty

**Stephanie Bopp**

***EURL ECVAM - SEURAT-1 – EU-ToxRisk workshop 10 November 2016***

European Commission

SCIENTIFIC OPINION

EFSA Journal

ADOPTED: dd mmmm yyyy PUBLISHED: dd mmmm yyyy AMENDED: dd mmmm yyyy

doi:10.2903/j.efsa.20YY.NNNN

*Revised Draft for Internal Testing*

# Guidance on Uncertainty in EFSA Scientific Assessment

**EFSA Scientific Committee**[1, 2]

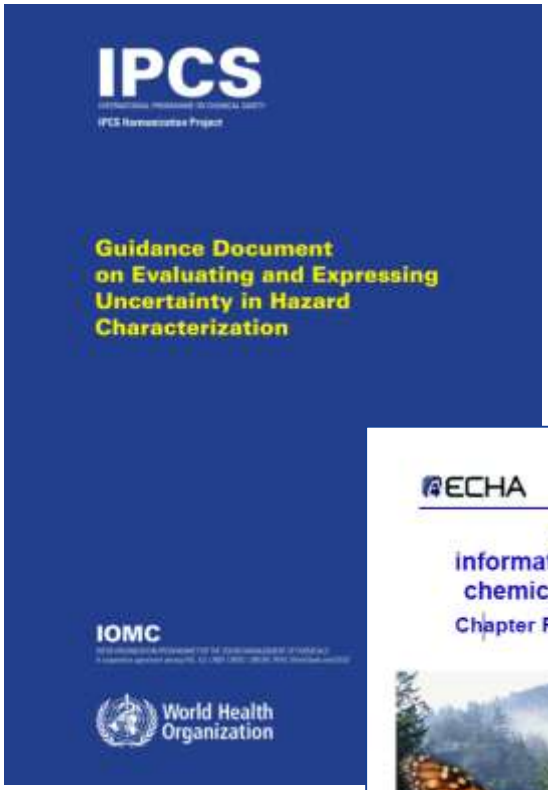
European Food Safety Authority (EFSA), Parma, Italy

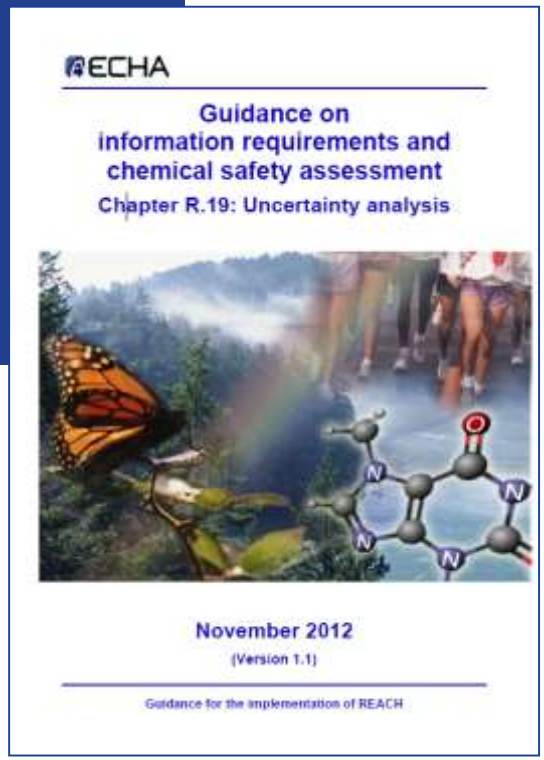
## Abstract

To meet the general requirement for transparency in EFSA's work, all its scientific assessments must include consideration of uncertainty. Assessments must say clearly and unambiguously what sources of uncertainty have been identified and what is their impact on the final assessment outcome: what range of outcomes is possible, and how probable they are. The Guidance is applicable to all areas of EFSA, all types of scientific assessment and all types of uncertainty affecting scientific assessment. It does not prescribe specific methods for uncertainty analysis but rather provides a harmonised and flexible framework within which different methods may be selected, according to the needs of each assessment. Worked examples are provided to illustrate different methods. Expert judgement plays a key role in uncertainty analysis, as in other aspects of scientific assessment. Assessors should be systematic in identifying sources of uncertainty, checking each part of their assessment to minimise the risk of overlooking important uncertainties. Uncertainty may be expressed qualitatively or quantitatively. It is not necessary or possible to quantify separately every individual source of uncertainty affecting an assessment. However, assessors should express in quantitative terms the combined effect of as many as possible of the identified sources of uncertainty. Practical approaches to facilitate this are described. Uncertainty analysis should be conducted in a flexible, iterative manner, starting at a level appropriate to the assessment in hand and then refining the analysis as far as is needed or possible within the time available. Some steps may be reduced or omitted in emergency situations and in routine assessments with standardised provision for uncertainty. Sensitivity analysis and other methods for investigating influence are used to target refinement on those sources of uncertainty where it will contribute most. The methods and results of all steps of the uncertainty analysis should be reported fully and transparently. Every EFSA Panel and EFSA Units that produce scientific outputs should apply the draft Guidance to at least one assessment during a trial period of one year, involving relevant decision-makers and supported by specialists in uncertainty analysis where needed. When the trial period is completed and any resulting improvements to the

[1] Scientific Committee members: Diane Benford, Thorhallur Ingi Halldorsson, Tony Hardy, Michael John Jeger, Helle Katrine Knutsen, , Simon More, Alicja Mortensen, Hanspeter Naegeli, Hubert Noteborn, Colin Ockleford, Guido Rychen, Josef Rudolf Schlatter, Vittorio Silano. Correspondence: scer@efsa.europa.eu

[2] Acknowledgement: The Panel wishes to thank the members of the Working Group on Uncertainty in Risk Assessment: Jan Alexander, Diane Benford, Peter Craig, Alessandro Di Domenico, Antony Hardy, Andrew Hart, Arie Havelaar, Michael Jeger, Robert Luttik, Ambroise Martin, Alicja Mortensen, Kaare Nielsen, Birgit Nørrung, Bernadette Ossendorp, Derek Renshaw, Josef Schlatter, Hans-Hermann Thulke, Natalie Von Goetz and Detlef Wölfle, for the preparatory work on this scientific opinion and the EFSA staff members Andrea Gennini, Laura Martino, Caroline Merten, Olaf Mosbach-Schulz and Anthony Smith for the support provided to this scientific opinion.






European Commission

# What is uncertainty?

- All types of **limitations in available knowledge** that affect
  - **the range and**
  - **probability**

  of possible answers to an assessment question.

Terminology from "**Guidance on Uncertainty in EFSA Scientific Assessment (Revised Draft for Internal Testing)** " EFSA, 2016

# What is uncertainty?

**Uncertainty**

*limitations in knowledge*

**Variability**

*observable differences in the real world*

- Variability is a property of the real world
- Uncertainty relates to our state of knowledge
- Uncertainty can be modified (e.g. by further research)
- Variability cannot be altered by additional research but by making changes to the real world e.g. by risk management (change exposures of specific population groups)

# Types of uncertainties (from draft EFSA Guidance 2016)

**Assessment inputs**

- Ambiguity
- Methodological quality of data sources
- Sampling uncertainty
- Assumptions and expert judgement
- Extrapolation uncertainty
- Distribution choice
- Other uncertainties

**Assessment structure/ model uncertainties**

- Ambiguity
- Excluded factors
- Use of fixed values
- Relationship between components
- Evidence for the structure of the assessment
- Calibration or validation with independent data
- Dependency between sources of uncertainty
- Other uncertainties

# Sources of Uncertainty (from ECHA Guidance R.19, 2012)

- **Scenario uncertainty**

level of accuracy of the scenario description/identified uses.

- **Model uncertainty**

adequacy of the (mathematical/statistical) model used with the scope and purpose of the assessment.

- **Parameter uncertainty**

  - Measurement errors
  - Sample uncertainty
  - Selection of the data used for assessing the risk
  - Extrapolation uncertainty

# Why do we need to address uncertainty?

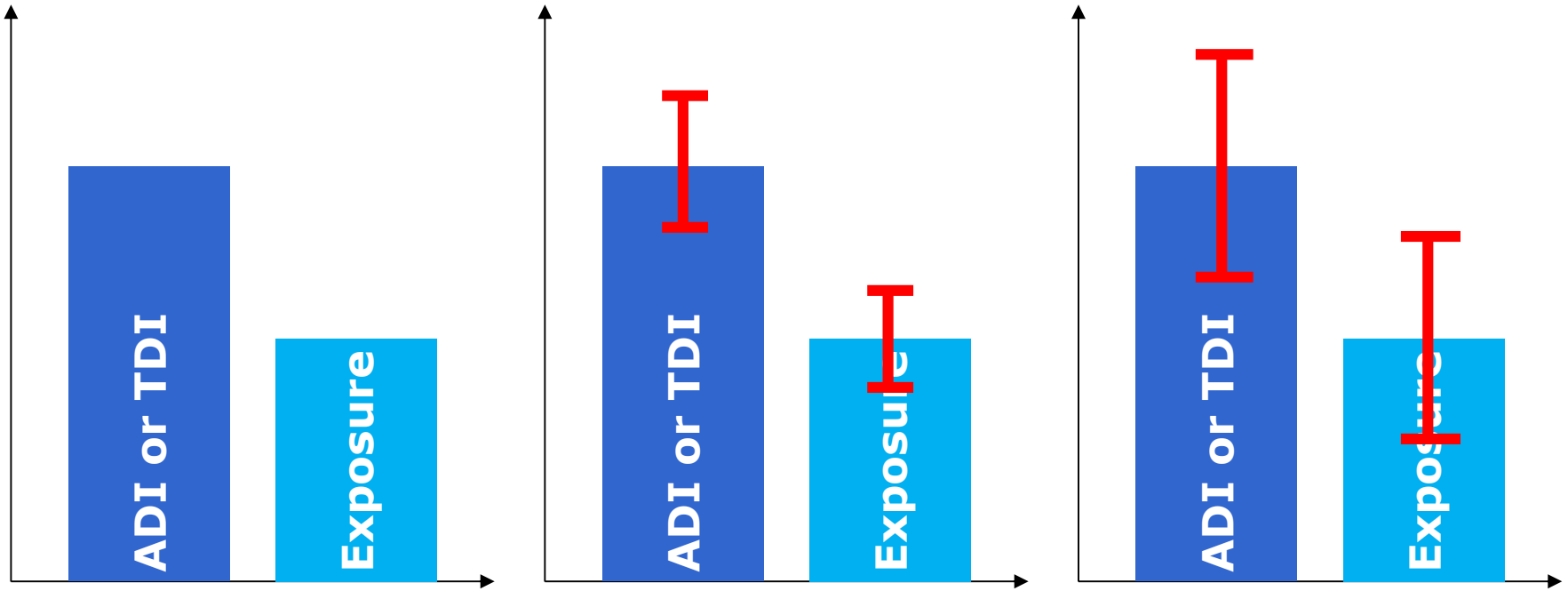
**Risk definition (IPCS 2004):**

*"The **probability** of an adverse effect in an organism system, or (sub)population caused under specified circumstances by exposure to an agent."*

IPCS, 2004. IPCS Risk Assessment Terminology. Harmonization Project Document No. 1. International Programme on Chemical Safety. WHO,Geneva.

**The degree of uncertainty determines the probability that an effect reaches unacceptable levels**

# Why do we need to address uncertainty?



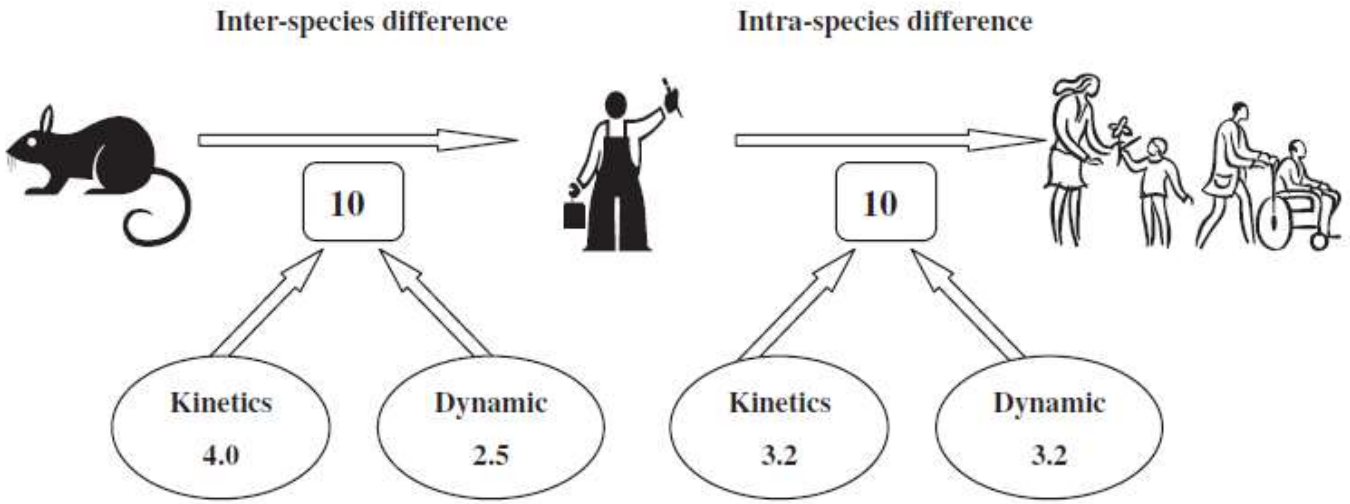

Decision makers need to know the related uncertainty around a result/prediction, i.e. the range and likelihood of possible outcomes

→ Avoid giving firm conclusions on weak basis

# How can we address uncertainties?

Incorporated in standardised procedures, i.e.

- Using predefined methods with e.g. default factors




Schroeder et al, Toxicology in Vitro 25 (2011) 589–604

- Using conservative default values (body weight etc.)

… but need to check always whether there are non-standard / uncovered uncertainties

# How can we address uncertainty?

**QUALITATIVE**
***describes the impact without quantification***

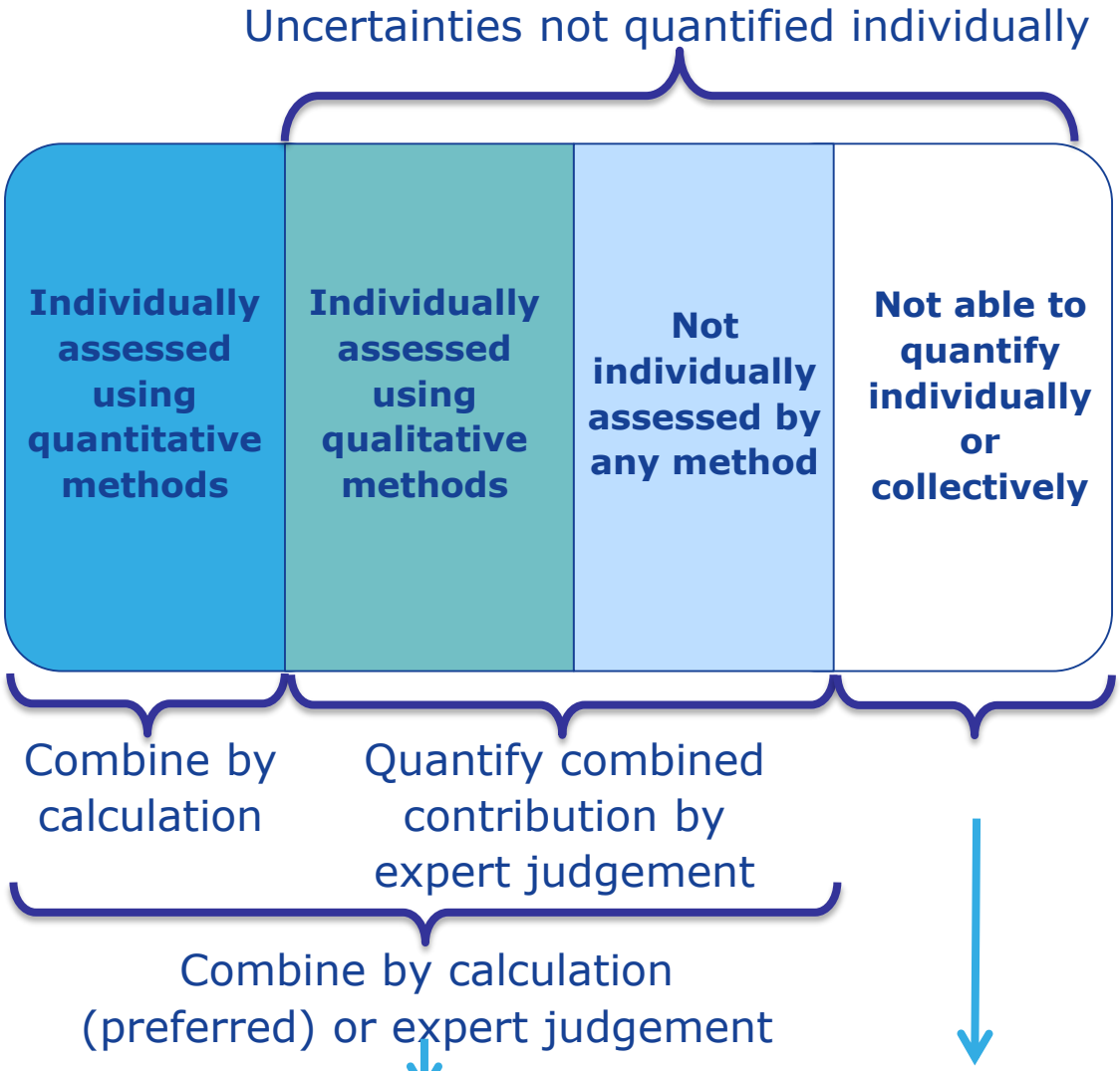
**DETERMINISTIC**
***range of possible outcomes without their probabilities***

**PROBABILISTIC**
***range and probability of possible outcomes***

- ➔ Choice of methods depends on needs
- ➔ Be quantitative wherever possible
- ➔ Words like negligible/likely are associated differently by each person
- ➔ Quantification allows also to combine several sources of uncertainty / multiple uncertainties

# How can we address uncertainty?

*From EFSA Draft Guidance on Uncertainty 2016*



# Qualitative uncertainty assessment

- Narrative text describing the direction and / or magnitude of the impact

  *"Children […] **could potentially** exceed the TDI by more than threefold […] **unknown** whether such high level exposure scenarios **may occur** in Europe."(EFSA Statement on Melamine 2008)*

- Ordinal scales/standardised terms like low/medium/high with some verbal definitions
- Scores or symbols with verbal definitions

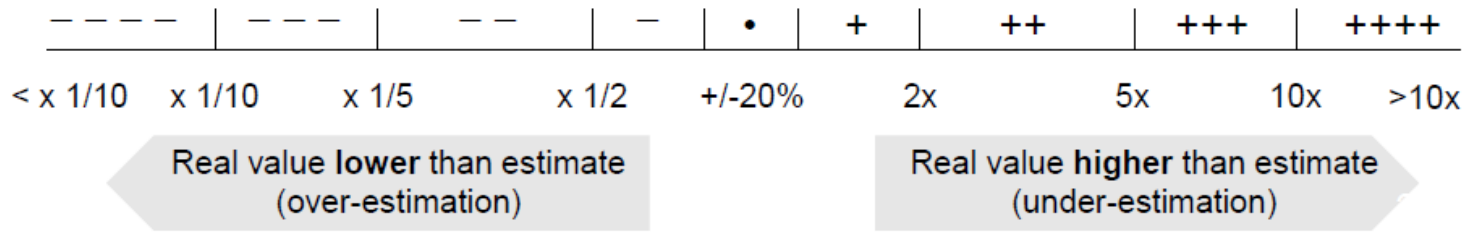
**Strengths**: simple to use, facilitate discussions, good for prioritisation and unquantifiable uncertainties

**Weaknesses**: Ambiguity, hamper combining uncertainties

# Qualitative example 1

| Parameter | Value in EFSA (2008) assessment | Range for uncertainty of individual parameters | Range for uncertainty of assessment output |
|---|---|---|---|
| TDI | 0.5 mg/kg bw/day | NQ/NQ or ---/++ | NQ/NQ or --/+++ |
| Highest concentration of melamine in milk powder | 2563 mg/kg | ---/+ | ---/+ |
| Highest consumption of Chinese chocolate by children | 0.044 kg | ---/++ | ---/++ |
| Assessment output: ratio of the calculated exposure to the TDI | 269% | | ----/NQ or ----/++ |

NQ = not quantified. See Figure B.5.1 for definition of scale for plus and minus symbols. See text for further explanation. Note that the results shown here differ from those in Annexes B.8 and B.9, as the latter were constructed as hypothetical examples and not elicited from experts.






From EFSA Draft Guidance on Uncertainty 2016

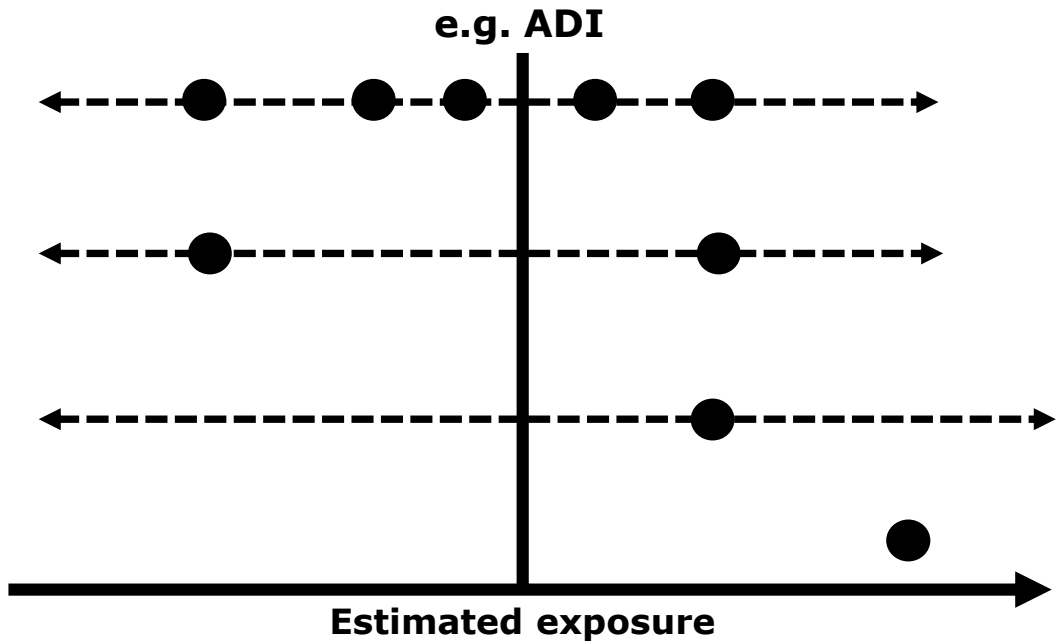
# Qualitative example 2

From Schultz et al. 2015,Regulatory Toxicology and Pharmacology 72, 586–601

| | Low uncertainty | Low-to-moderate uncertainty | Moderate uncertainty | High uncertainty |
|---|---|---|---|---|
| **Core structural similarity i.e., functional groups, extended fragments (especially those associated with chemical reactivity or its modification)** | Highly similar | Highly similar | similar | Difference in core structure and functional groups |
| **Physico-chemical and molecular** | Highly similar | Similar, having a consistent trend within values | Minor differences in values | Major differences in values |
| **Abiotic transformation and/or toxicokinetics, especially metabolism e.g. leading to a common metabolite** | Evidence demonstrating comparability in abiotic transformation and/or toxicokinetics and the same metabolites, similar bioavailability | Evidence demonstrating comparability in abiotic transformation and/or toxicokinetics and the same metabolites, similar bioavailability | No evidence that abiotic transformation and/or toxicokinetics, especially metabolism are dissimilar | Differences in abiotic transformation and/or toxicokinetics, especially in metabolism |
| **Mechanism of action and toxicological properties** | Evidence demonstrating comparability in mechanism supported by an AOP | Evidence demonstrating comparability in mechanism supported by an AOP | No evidence mechanism of action are dissimilar | Differences in mechanism of action and/or toxicological properties |

# Deterministic uncertainty assessment

- Impact of uncertainty on the outcome is expressed quantitatively, using one or more **point estimates**
- Use fixed numbers as input and provide fixed numbers as output




**range** of point estimates based on different combinations of assumptions

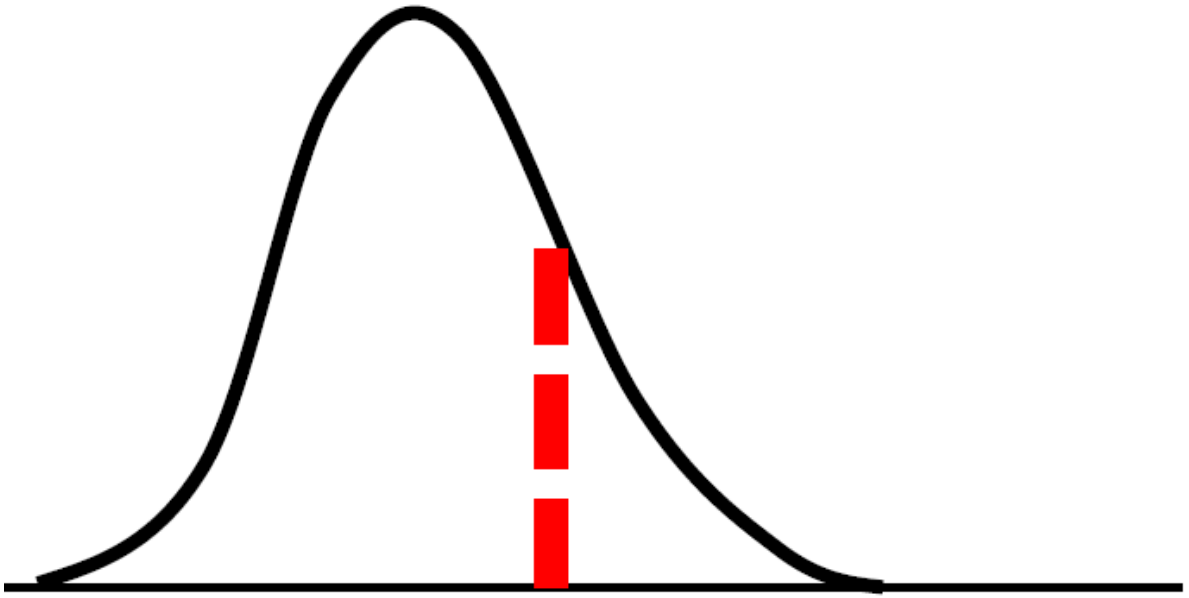
**optimistic and pessimistic** assumption point estimates

**refined** point estimate

**point estimate** with conservative assumptions and default values

# Deterministic uncertainty assessment

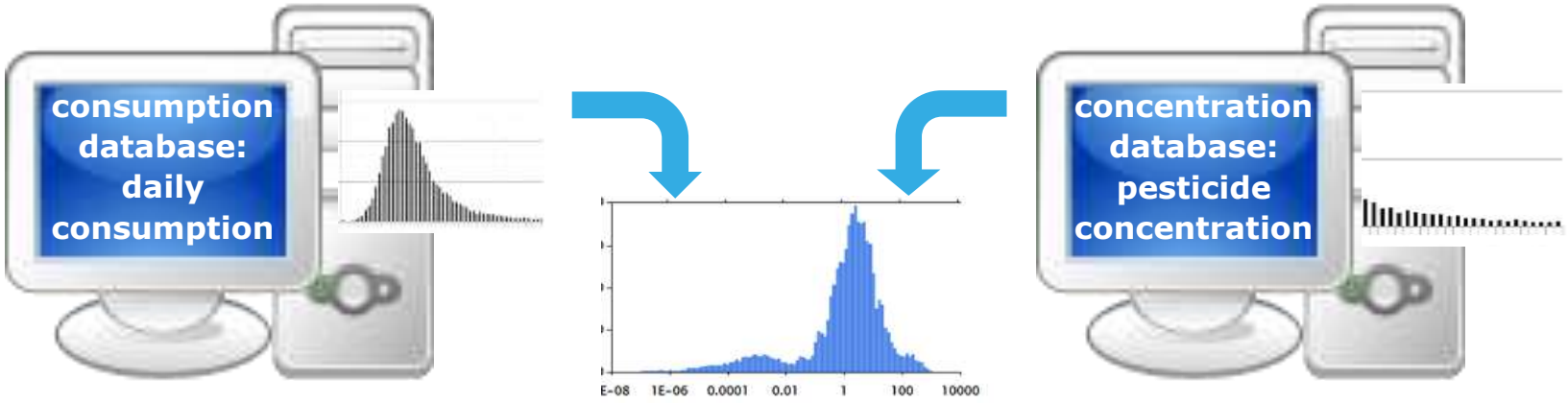
- Impact of uncertainty on the outcome is expressed quantitatively, using one or more **point estimates**



use specific percentile, often e.g. 95$^{th}$ percentile considered conservative

# Probabilistic uncertainty assessment

- Impact of uncertainty on the outcome is expressed probabilistically
  - Source of uncertainty is quantified using a probability distribution
  - Distribution is propagated through the assessment (e.g. by Monte Carlo simulation) to generate a distribution for the assessment output




**Distribution of daily intake**

# ECHA Read Across Assessment framework (RAAF 2015)

| SCORES | AOs | MEANING OF THE AOs |
|---|---|---|
| 5 | Acceptable with high confidence | Acceptance without reservations in the scientific explanation and documentation addressing the scientific aspects of the AE. |
| 4 | Acceptable with medium confidence | Acceptance with minor reservations about the scientific explanation and documentation addressing the scientific aspects of the AE. |
| 3 | Acceptable with just sufficient confidence | Acceptance with notable reservations. Minimum level of confidence in the scientific explanation provided in the documentation and addressing the scientific aspects of the AE. |
| 2 | Not acceptable in its current form | Acceptance for the AE under consideration may become possible if improved explanations and/or supporting evidence is made available by the registrant. |
| 1 | Not acceptable | A major flaw in the approach for the AE under consideration which is not expected to be resolved by the addition of supporting information. |

AO=Assessment Option; AE=Assessment Element

# Templates for Read Across

From Schultz et al. 2015, Regulatory Toxicology and Pharmacology 72, 586–601

| Factor | Uncertainty (low, medium, high) | Comment |
|---|---|---|
| The problem and premise of the read-across | | **Example:** The endpoint to be read across, developmental toxicity, for the category of branched carboxylic acids is well-studied and well-understood, The scenario of the read-across hinges on the inhibition of beta-oxidation of the acid and the subsequent build up of acid in the embryo leading to histone deacetylase inhibitors, increased cell adhesion and concomitant reduced cell motility, prevention of convergent extension during ontogenetic development. |
| **In vivo data read across** | | |
| Number of analogues in the source set | | **Example:** There are 3 suitable category members with *in vivo* apical endpoint data usable for read-across. |
| | | |
| Quality of the *in vivo* apical endpoint data read across | | **Example:** High quality empirical data from standard test guidelines for the stated regulatory endpoint exists for 1 category member. Similar non-standard test data of lower quality exists for 2 other category members. All these data [illegible] of effects [illegible] quality data. |
| | | |
| Severity of th[illegible] *vivo* hazard | | [illegible] endpoint (25 mg/kg/day) is limited to a single source substance. |
| Evidence to biologic[illegible] | | |
| Robustness of analogue data set | | **Example:** The available data from *in silico*, *in chemico* and *in vitro* studies for the category members were judged to be reliable and conducted under the appropriate conditions. |
| | | |
| Concordance with regard to the intermediate and apical effects and potency data | | **Example:** There is good agreement between the sequences of biochemical and physiological events leading to the *in vivo* apical outcome. There is consistency and high specificity for the association between the toxicophore and the structural domain of the category. There is general agreement among the dose-response relationships of the tested category members for mechanistically-relevant event(s) which may be assessed *in vitro*. |
| | | |
| Weight of Evidence | | **Example:** Overall the available information is generally consistent with the stated hypothesis. The sharp structural limitations of the category and narrow range of chemical properties strengthens the WoE. While the toxicokinetics data is limited, the lack of inconsistencies adds to the WoE. While the source substances data is limited, the fact that there is consistent relevant *in vitro* data for 50% of the category members, including the target chemical, strengthens the WoE. |
| | | |
| Overall uncertainty of the read across: (Low, Medium, High)<br><br>Uncertainty associated with the read-across is judged to be low. | | |

**Uncertainty associated with mechanistic relevance and completeness of the read-across**

| Similarity Parameter | Data Uncertainty [a] (empirical, modelled) (low, medium, high) | Strength of Evidence [b] (low, medium, high) | Comment |
|---|---|---|---|
| Substance Identification, Structure and Chemical Classifications | | | **Example:** All category members have CAS numbers, similar 2D structure and belong to the same chemical class/subclass. |
| | | | |
| Physio-Chemical & Molecular Properties | Empirical:<br><br>Modelled: | | **Example:** All category members are appropriately similar with respect to key physicochemical and molecular properties. There is a high degree of consistency between measured and model estimated values. |
| | | | |
| Substituents, Functional Groups, & Extended Structural Fragments | | | **Example:** Substituents, functional groups and extended structural fragments are consistent across all category members. |
| | | | |
| Transformation, Toxicokinetics and Metabolic Similarity | Empirical:<br>*In vivo*:<br>Simulated: | | **Example:** Based on *in vivo* [illegible] category members, there is evidence for similar toxicokinetics [illegible] metabolism among all category members. |
| | | | |
| Potential Metabolic Products | | | **Example:** Based on *in silico* metabolic simulations, potential metabolic products are similar among all category members. |
| | | | |
| Toxicophores /Mechanistic alerts | | | **Example:** Based on *in silico* profilers, all category members contain the same toxicophores. |
| | | | |
| Mechanistic plausibility and AOP-Related Events | | | **Example:** Although no AOP is currently available for the hypothesised toxicity pathway, many category members have been tested for what is generally accepted as a mechanistically-relevant event leading to the *in vivo* apical outcome of interest (*a citation could be provided*). |
| | | | |
| other relevant, *in vivo*, *in vitro* and *ex vivo* endpoints | | | **Example:** Although not part of the hypothesised toxicity pathway, many category members have been tested for rodent acute oral toxicity and there is general agreement among the reported LC50 values. |
| | | | |
| Overall uncertainty in similarity of category members: (Low, Moderate, High)<br><br>Summary: Key features of chemistry are similar within the category. Key features of transformation toxicokinetics and metabolism are common within the category. Category members are considered mechanistically similar. Category members exhibit a similar toxicological profile with respect to *in vivo* toxicity. | | | |

**Data uncertainty associated with the fundamentals of chemical, transformation, toxicokinetic and toxicological similarity**

# OECD GD on the Reporting of Defined Approaches

## Principle 6: *Consideration of known uncertainties*

a) the **relevance of the "model structure"** of the DIP, e.g. the extent of **coverage and weighting of AOP key events** and other **mechanistic considerations**;

b) the level of confidence (reliability of prediction) associated with the **application of the DIP to different chemicals**;

c) the **variability of the input data** used (generated by the information sources);

d) the **variability** of the output data associated with the gold standard (e.g. animal or human) data used as **benchmark data**, especially when these are used as the basis for regulatory decision making and;

e) **any other known source of uncertainty** (e.g. uncertainty in estimated exposure levels that are used in a safety assessment).

In each case, the magnitude and impact of the sources of uncertainty should be considered.



GUIDANCE DOCUMENT ON THE REPORTING OF DEFINED APPROACHES TO BE USED WITHIN INTEGRATED APPROACHES TO TESTING AND ASSESSMENT Series on Testing & Assessment No. 255, ENV/JM/MONO(2016)28

# Take home message

- **Do not neglect uncertainties**
- **Present uncertainties transparently**

## Start simple

- At least qualitative (tabular) description of individual uncertainties and their combined impact
- Use e.g. EFSA Uncertainty Toolbox

# Stay in touch

EU Science Hub: ***ec.europa.eu/jrc***

Twitter: ***@EU_ScienceHub***

Facebook: ***EU Science Hub - Joint Research Centre***

LinkedIn: ***Joint Research Centre***


YouTube: ***EU Science Hub***



European Commission