EUROPEAN COMMISSION

Community Research

SEVENTH FRAMEWORK PROGRAMME

DRAFT

Guidelines for Good Review Practice in GMO impact assessments

FP7 Collaborative Project
GRACE 311957

Author List:

Christian Kohl, Geoff Frampton, Christian Frølund Damgaard, Ralf Wilhelm, Joachim Schiemann

GRACE

GMO Risk Assessment and Communication of Evidence

GRACE
GMO Risk Assessment and Communication of Evidence

Table of contents

GRACE
GMO Risk Assessment and Communication of Evidence

Good practice for reviews in GMO impact assessment

1. Motivation

Evidence synthesis approaches represent powerful tools to collect, evaluate and summarize accessible research results to address a specific scientific issue in a transparent, reproducible and unbiased manner [1-5]. Their adaption to and implementation in the risk assessment process for genetically modified organisms (GMO) and products derived thereof, aims to increase the transparency of science based decision making processes regarding potential health, environmental and socio-economic impacts of GMOs. The "Good Review Practice" fosters the comprehensive, verifiable and inclusive evaluation of former and current studies.

Further, GRACE will contribute to a transparent risk assessment process by providing an open-access database in order to depict the actual evidence synthesis procedure and disseminate the results and the conclusions drawn thereof.

2. Background

The European regulatory framework for the market authorization of genetically modified organisms (GMO) and their products demands a comprehensive risk assessment (RA) in line with the concept of a knowledge based bio-economy that shall ensure safe access to the benefits of modern biotechnology.

Even though a lot of safety research has been conducted on GMO, it is very difficult for stakeholders, risk assessors and the general public, to fully overview existing outcomes which are hence not entirely considered in decision making.

This may partly be due to the fact that:

- recently published papers are more the focus of attention, even though other studies may also address the issue of concern
- the range of studies is extremely broad and rapidly increasing
- the use of different approaches and methods, the different ways to present data and results makes it hard to quantitatively and qualitatively compile/collate data from the published studies
- the validity of the conducted studies is not clear
- uncertainties and information gaps are sometimes over emphasized and often disregard the available evidence

- not all research results are readily available and conclusions are not always transparent or easily understood

In order to overcome these hurdles, available data dealing with potential impacts of GMOs on health, the environment and socio-economy will be collected, evaluated and analyzed in a standardized, transparent and reproducible manner within the course of GRACE. Doing so, two different evidence synthesis approaches (namely Systematic Reviews (SR) and Evidence Maps (EM)) will be adapted to and implemented in the impact assessment process for GMOs and their products. Both approaches aim to synthesis reliable outcomes by minimizing the introduction of bias originating from eventual subjectivity of the reviewers and from the included studies per se. The major methodological differences distinguishing them from traditional narrative reviews are shown in Table 1. The results and the conclusions drawn thereof will be discussed with stakeholders in a transparent manner and will further be made accessible via an open access database to the general public.

Table 1 : **Major methodological differences between systematic reviews, evidence maps and narrative reviews (adapted from [2])**

	Systematic Review	**Evidence Map**	**Traditional reviews**
Study question	Focused and explicit	Broad or focused and explicit	Often broad in scope
Eligibility criteria for inclusion or exclusion of studies	Pre-defined and documented; objectively applied	Pre-defined and documented; objectively applied	Not always explicitly stated
Description of the review method	Reported and also predefined in a protocol	Reported and also predefined in a protocol	Seldom reported
Literature search	Structured to identify as many relevant studies as possible	Structured to identify as many relevant studies as possible	Not always extensive
Critical appraisal of included studies	Included, typically using a critical appraisal tool	Variable	Variable
Reporting of study outcomes	Full reporting of relevant outcomes	Full reporting of relevant outcomes	Selective reporting; often of study author interpretation

Synthesis	Quantitative synthesis (meta-analysis) when possible	Usually narrative	Usually narrative, sometimes selective

3. Evidence synthesis for GMO impact assessment

3.1 Brief description of the approaches used for evidence synthesis

Two different evidence synthesis approaches, namely SR and EM, will be applied to integrate data from a set of relevant studies in a transparent, reproducible and case specific manner. Both approaches are based on a specific research question (see section 3.3) and employ the same methodological stringency for study identification and selection. However, an EM typically provides a less detailed synthesis of the evidence than a SR. The methodologies underlying both approaches will be briefly introduced in the following.

Systematic Review:

SRs are based on a focused question (see section 3.3) and give an overview of existing evidence using pre-specified and standardized methods to identify, critically appraise, collect, report, analyze, synthesize and if possible quantitatively (or qualitatively) combine data from the studies that are included in the reviews [2]. They are valuable for impartially synthesizing evidence relating to contentious topics where stakeholders may hold differing views and may inform the risk assessment process regarding a specific issue of concern.

The process of SR performance can be subdivided into eight core steps shown in Figure 1 which will be briefly summarized below.

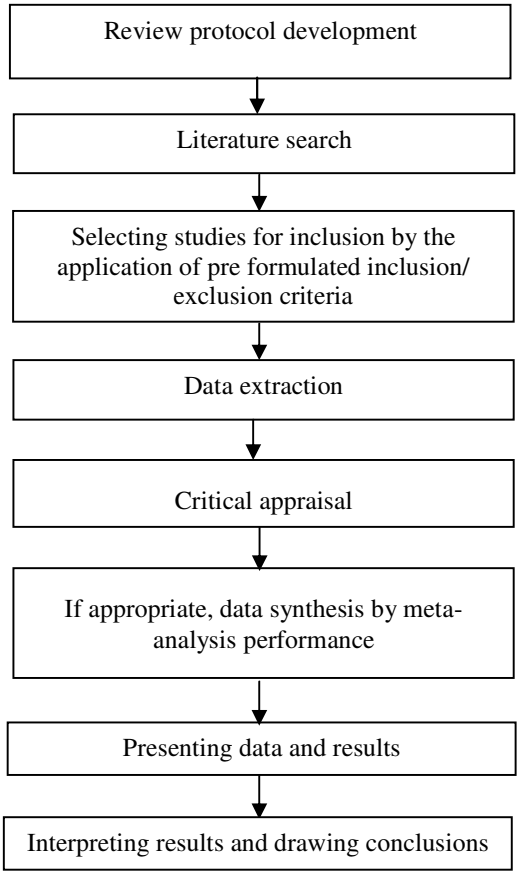

Figure 1 : Core steps of a SR (adapted from [2])

Review protocol development

The review protocol determines the detailed procedure used for evidence synthesis, as all methodical details envisaged to be applied throughout the entire reviewing process have to be made explicit *a priori*. The protocol has to be evaluated by stakeholder and thus aims to prevent the introduction of bias originating from methodological errors. It further includes the definition of the review question, setting the scope of the review.

Literature search

Relevant information sources, e.g. bibliographic databases, web-pages, are being screened for possibly eligible reports by the application of an extensive search strategy. The development of a proper search strategy is an iterative process aiming to retrieve a maximum amount of relevant reports from the information sources it is applied to [1]. It is a key step in the review process further aiming to minimize bias, originating from the

published data per se and from indexing (how and if e.g. an article is being listed by a bibliographic database) [2].

Study selection

A major difference between systematic and traditional narrative reviews is the use of specific eligibility criteria to determine the suitability of a certain report being included in the SR or not. These criteria have to be defined *a priori* in the review protocol, to ensure that the boundaries of the review question are clearly defined [6].

Data extraction

The process of data extraction aims to collect all the necessary information regarding study characteristics and outcomes of the included studies in order to give a sound answer to the review question. The extraction process will be standardized e.g. by the development of a data extraction form. The actual procedure for data extraction and a description of the information that will be collected from the included studies will have already been made explicit in the review protocol.

Critical appraisal

The internal validity of an identified study being included in the SR is strongly influenced by its associated methodological rigor. As the value of the conducted SR depends on the validity of included studies, a comprehensive critical appraisal has to be performed, determining the susceptibility of each included study to bias.

Doing so, critical appraisal criteria have to be developed and discussed by topic specialists, aiming to identify possible biases within included studies. The *a priori* specification of such criteria allows for a standardized and transparent study appraisal. Biases which can be introduced to scientific studies can be, beside others, caused by discrepancies in the selection of the study subjects, the performance of the study, the detection and reporting of outcomes, and the handling of missing data.

Data synthesis via meta-analysis (if appropriate)

A meta-analysis is a quantitative synthesis of the collective finding of several different empirical studies that have investigated the same phenomenon. Typically, outcome are summarized by a defined *effect size*, which is a descriptive statistic calculated from the data of each single experiment measuring the size of the effect.

Presenting/interpreting data and results and drawing conclusions

The results have to be objectively discussed and interpreted, taking the identified evidence base into account. A well structured and solely evidence based discussion sets the basis for drawing sound conclusions suitable to inform the decision making process regarding possible impacts of GMOs and their products on the environment, health and socio-economy.

Direct stakeholder involvement at this final step is important to ensure broader acceptance of the conclusions.

Evidence Map:

An EM is based on a specific research question and aims to identify all relevant literature within a research field in order to provide a comprehensive assessment of both what is known and where gaps in evidence exist [4].

There are no firm rules about when an EM should be used. EM may have two, more general functions:

- To provide an overview of the extent, range and nature of research activities in a particular field [4, 7].
- As a tool to inform stakeholder/risk assessors of the types of evidence available so that they may decide whether specific syntheses of outcomes would be valuable, in one or more SRs (see red arrow in Figure1). [5].

The decision for an EM may be based on the anticipated amount of available evidence. If a lot of evidence is available, mapping can help identifying areas where SR(s) may be valuable i.e. where a sufficient amount of robust evidence is available. In contrast, mapping the evidence can also help to identify research gaps when the evidence is rather scarce and can thus target further research in the future.

A comparison between EM and SR is given in Table 2.

Table 2: Key characteristics of Evidence Maps and Systematic Reviews

	Evidence Map	**Systematic Review**
Rationale	1) Provide an overview of existing evidence (e.g. which transgenes have been studied) 2) Inform stakeholder/risk assessors of the types of evidence available for further synthesis e.g. by SR	Directly inform the impact assessment process regarding a specific issue of concern
Nature of question	Broad or focused (see section 3.3)	Focused (see section 3.3)
Literature search	Thorough and reproducible	Thorough and reproducible
Data extraction	Limited extraction of data depending on the purpose of the map	Comprehensive data extraction
Critical appraisal	No in depth critical appraisal. Issues relating to the validity of included studies may be captured by extracting data like e.g. study design, sample size etc.	In depth appraisal
Data synthesis via meta-analysis	No	If possible

3.2 Overview of the general framework used for GMO impact assessment

The process of evidence synthesis being applied within GRACE can be subdivided into four distinct steps: i) the development of a specific research question, ii) the assessment of the need for evidence synthesis, iii) the selection of an appropriate evidence synthesis approach and iiii) its final performance (Figure 2).

The specific research question will be developed by the use of conceptual models (CM) (see section 3.3). At this stage, the availability evidence (quality and quantity of existing data) to give a sound answer to the addressed question is unknown. Thus, a scoping

exercise may help to estimate the actual evidence base regarding a specific issue of concern. Scoping results serve as a decision platform for a possible need to refine, specify or re-evaluate the addressed question and for the selection of the most adequate evidence synthesis approach. The preference for one or the other approach further depends on the intended purpose of the evidence synthesis (SR versus EM), already discussed above.

As the scope of this consultation is on "discussing, refining, and complementing both the conceptual models and the derived review questions", the question development process (Step1, Figure 2) will be explained in more detail in the following section.

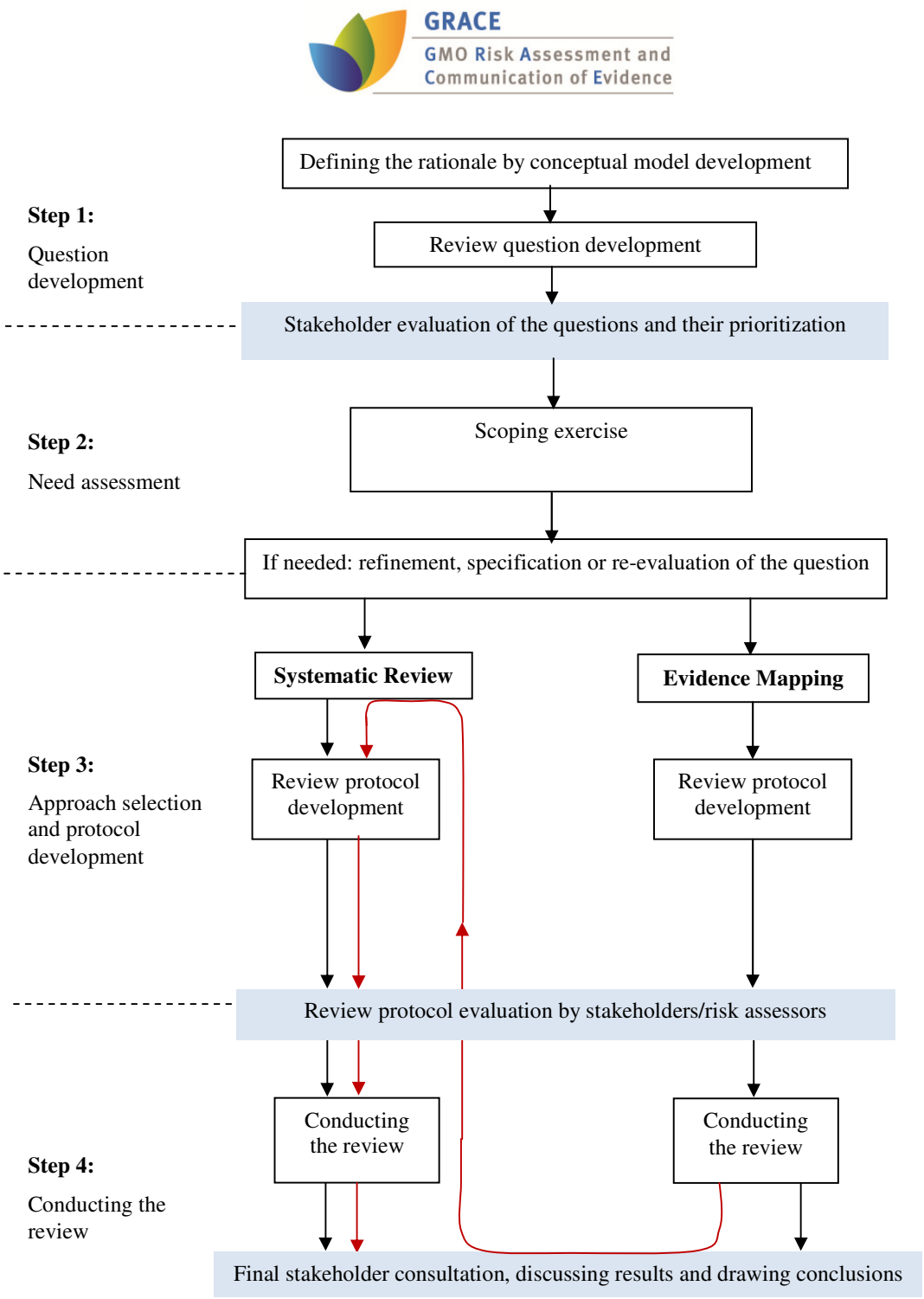

Figure 2 : General overview of the review strategy for GMO impact assessment (for more details see text).

3.3 Question development

As already mentioned, SRs and EMs are based on a specific research question setting the scope of the review. To guarantee transparent and traceable question development, both, the rationale for the assessment and the actual question development process have to be presented in a self-explanatory and reproducible way. Conceptual models will be used to illustrate and guide through the entire process of question development.

3.3.1 Define the rationale for GMO impact assessment

In environmental risk assessment (ERA), conceptual models (CM) are frequently used to facilitate problem formulation [8], leading to the development of risk hypotheses by establishing relationships between the stressor and the impacts of concern (assessment endpoints) [9]. The risk hypothesis is then translated into one or more experimental hypotheses (measurement endpoints) that can be used for testing and corroboration [10].

The CM-concept will be used to define the rationale of the impact assessment by illustrating the relationships between the valued entity, the stressor, pathways of exposure and potential impacts. There is flexibility in how many CMs are used and how broad they are. E.g. if a certain research area is of high complexity, it may be appropriate to use more than one CM. Further, a balance has to be found between the quantity of detailed information and the clarity of the CM. CM(s) allow to:

- Clearly define the rationale of the assessment
- Engage stakeholders from the outset
- Provide a logical source of structured information to develop review questions.

An example for a hypothetical CM setting the rationale for the environmental risk assessment of GMOs is given in Figure 3. The structure of the CM may vary depending on the different review topics (health, environment and socio-economy).

Possible concerns relating to the deliberate release of GMOs will be identified (primary concerns) and, if necessary, specified (secondary concerns). The proposed concerns may or may not lead to the identification of a secondary (specified) stressor, which differs from the GMO itself (e.g. in the case for HT crops, the secondary stressor could be the herbicide regimes used). Based on this information, broad protection/assessment goals are formulated and subsequently broken down into specific assessment endpoints. In a final step, the stressor is linked to the assessment endpoint by defining possible pathways of exposure.

.

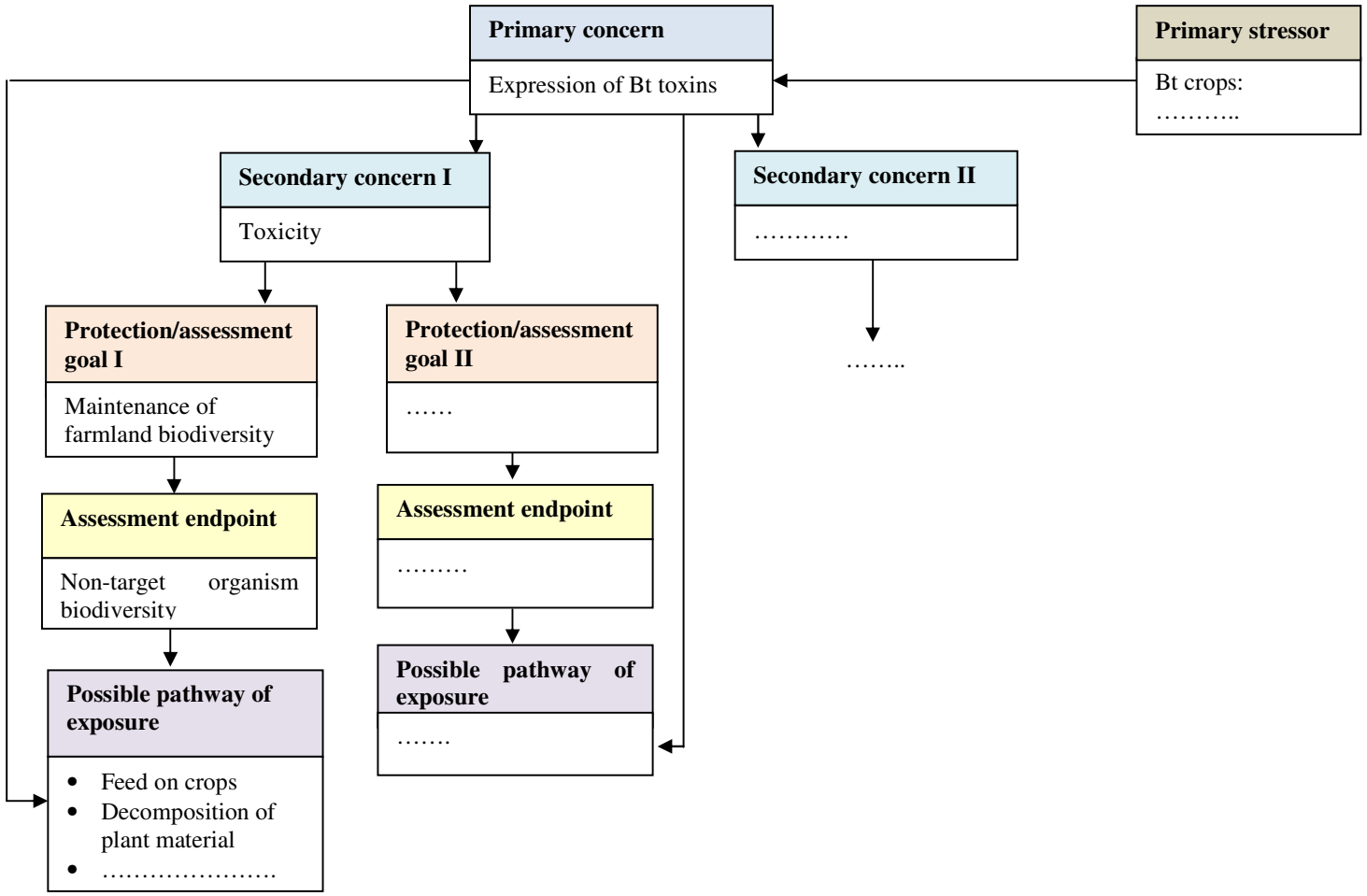

Figure 3 : Hypothetical example for a conceptual model defining the scope for the environmental impact assessment of GMO

3.3.2 Question formulation by conceptual model development

Based on the identified assessment endpoints, more focused CMs may be developed to facilitate the question formulation process and to guarantee for its accountability by stakeholders/risk assessors (a hypothetical example is given in Figure 4). Detailed research questions, which can be answered by a SR, can be developed based on important issues identified by the CMs.

This type of question is characterized by a precise description of its scope and is thus referred to as being focused. Key elements that define the scope of the research question can be specified e.g. in a "PICO" or "PECO" structure, where the key elements are the population (P) (equals the "Populations possibly affected" of the focused CM), Intervention (I) or exposure (E) (represented by the "stressor" in the focused CM), comparator (C) (relates to the "population" from the focused CM), and outcome (O) (represented by the "potential harm" in the focused CM) [1, 2].

The different key elements are defined as follows:

- Population of interest (P): Can be represented by a group of people, animals, plant species, a particular taxon or a sector of agriculture at a particular geographic scale[2].
- Intervention/Exposure (I/E): The factor to which the population is exposed [2].
- Comparator (C): The reference entity against which the intervention/exposure can be compared [2].
- Outcome (O): Measurable consequences of a certain intervention/exposure.

Possible questions derived from such a conceptual model could be e.g.:

- Does the cultivation of Bt maize (I) - when compared to their isogenic counterparts (conventional varieties) (C) - affect the biodiversity (O) of non-target arthropods (P)?

Questions following the PICO/PECO format do not represent the only type of question answerable by a SR or an EM. Questions seeking to determine the accuracy or the applicability of a certain technique to detect a certain parameter(s) of interest are called PIT question, where the key elements are defined as followed [2]:

- Population of interest (P): The population of interest can be represented by any kind of material (animal tissues, plant material) under investigation
- Index test (I): The technique/method under investigation which is needed to be evaluated
- Target condition (T): The variable intended to be detected (e.g. presence/absence of a certain compound/chemical)

Another type of question suitable for SR or EM aims to describe the prevalence, occurrence or incidence of a certain outcome (condition). This type of question is called PO question, where the key elements are defined as followed (taken from[2]):

- Population of interest (P): The population, organism or setting in which the condition of interest is measured.
- Outcome (O): What is assessed or measured in the population. For prevalence or incidence questions the condition of interest is often a disease; for occurrence questions it may be a possible outcrossing event, a chemical substance or pathogen.

If the review question lacks one or more of the above mentioned key characteristics, it is referred to as being broad. A broad research question cannot be answered by a SR but may be addressed by an EM.

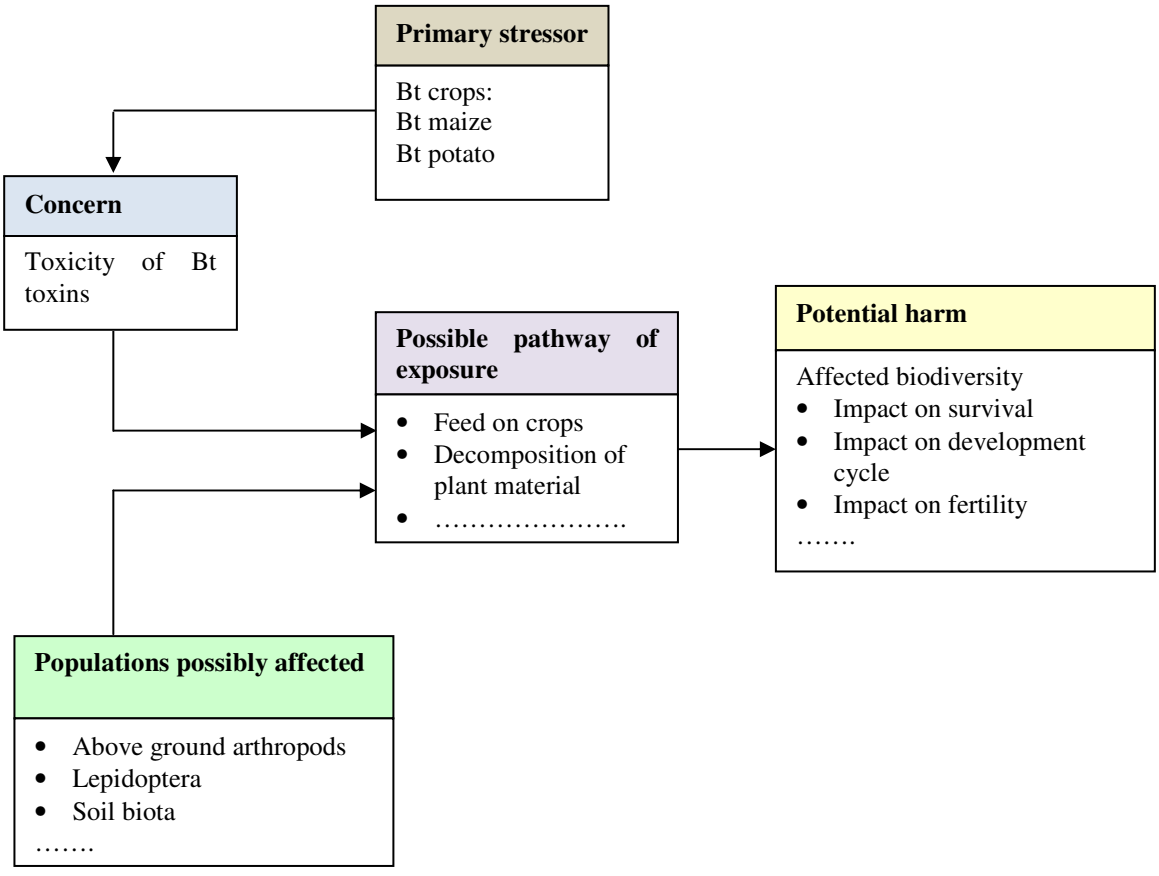

Figure 4 : Hypothetical focused CM describing potential impacts on the biodiversity of NTO caused by the deliberate release of Bt crops (adapted from [11]).

3.3.3 Question prioritization

The prioritization of the developed questions is crucial, as the conduct of a SR or EM is labor intensive and the work load has to be adapted to the availability of resources, like money, manpower and time. Review questions will be prioritized by ranking them against three domains (adapted from [12, 13]):

Domain 1: Relative importance for the impact assessment of GMOs (Is the review question of high importance for the impact assessment of GMOs?)

Domain 2: Disagreement/controversy among experts (Is there expert disagreement on the review question?)

Domain 3: High public scrutiny (Is the review question the subject of high public awareness?)

Stakeholders will be asked to rank each review question against those domains by applying a scoring system.

Taking suggestions made throughout the stakeholder workshop into account, the conceptual models and derived review questions will be revised and each review question envisaged to be addressed within GRACE will be prioritized by individual stakeholders and work package members by following the above mentioned approach via an online questionnaire. The results will be summarized and the final question prioritization will be transparently justified.

Reference list:

1. *Centre for Evidence-Based Conservation. 2010. Guidelines for Systematic Review in Environmental Management. Version 4.0. Environmental Evidence: www.environmentalevidence.org/Authors.htm.*
2. *European Food Safety Authority; Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal 2010; 8(6):1637. [90 pp.]. doi:10.2903/j.efsa.2010.1637. Available online: www.efsa.europa.eu.*
3. *Green S, Higgins JPT, Alderson P, Clarke M, Mulrow CD, Oxman AD. Chapter 1: Introduction. In: Higgins JPT, Green S (editors), Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.*
4. Hetrick, S.E., et al., *Evidence mapping: illustrating an emerging methodology to improve evidence-based practice in youth mental health.* J Eval Clin Pract, 2010. **16**(6): p. 1025-30.
5. Bragge, P., et al., *The Global Evidence Mapping Initiative: scoping research in broad topic areas.* BMC Med Res Methodol, 2011. **11**: p. 92.
6. Dissemination, C.f.R., *Systematic reviews: CRD's guidance for undertaking reviews in health care*. 2009: Centre for Reviews and Dissemination.
7. Arksey, H. and L. O'Malley, *Scoping studies: towards a methodological framework.* International journal of social research methodology, 2005. **8**(1): p. 19-32.
8. *Guidance on the environmental risk assessment of genetically modified plants.* EFSA Journal, 2010. **8**(11).
9. Romeis, J., et al., *Assessment of risk of insect-resistant transgenic crops to nontarget arthropods.* Nat Biotechnol, 2008. **26**(2): p. 203-8.
10. Wolt, J.D., et al., *Problem formulation in the environmental risk assessment for genetically modified plants.* Transgenic Res, 2010. **19**(3): p. 425-36.
11. Sears, M.K., et al., *Impact of Bt corn pollen on monarch butterfly populations: a risk assessment.* Proc Natl Acad Sci U S A, 2001. **98**(21): p. 11937-42.
12. *O'Connor AM, Lovei GL, Eales J, Frampton GK, Glanville J, Pullin AS, Sargeant J; Implementation of systematic reviews in EFSA scientific outputs workflow. Supporting Publications 2012:EN-367 [36 pp.]. Available online: www.efsa.europa.eu/publications.*
13. Clavisi, O., et al., *Effective stakeholder participation in setting research priorities using a Global Evidence Mapping approach.* Journal of Clinical Epidemiology, 2012.