Enhancements of pest risk analysis techniques*

Contribution to Work Package 3: enhancing techniques for standardising and summarising pest risk assessments – review of best practice in enhancing consistency

G. Schrader[1], A. MacLeod[2], M. Mittinty[3], S. Brunel[4], K. Kaminski[1], H. Kehlenbeck[1], F. Petter[4] and R. Baker[2]

[1]*Julius Kühn-Institut (JKI), Bundesforschungsinstitut für Kulturpflanzen, Messeweg 11/12, Braunschweig 38104 (Germany); e-mail: Gritta.Schrader@jki.bund.de*

[2]*Food and Environment Research Agency (Fera) Sand Hutton, York YO30 7BH (UK); e-mail: r.baker@csl.gov.uk*

[3]*School of Mathematical Sciences, Queensland University of Technology, 2 George Street, Brisbane (Australia)*

[4]*The European and Mediterranean Plant Protection Organization, 1 rue le Notre, Paris 7016 (France)*

International, regional and national standards for plant health pest risk assessment and examples of their use from countries around the globe were examined together with similar documents from related fields such as animal health, nature conservation and genetically modified organism (GMO) assessment to determine how the consistency of assessing risk, or components of risk, within and between assessments is addressed. A range of approaches was identified that could be adopted and adapted for use in a revised decision support scheme for quarantine pests by the European and Mediterranean Plant Protection Organization (EPPO) to aid consistency. No single scheme contained a mechanism to ensure or guarantee consistency and no single scheme contained all of the approaches identified to maximize consistency. If the approach of using a five division scale to describe individual components that contribute to evaluating the overall pest risk is to be maintained in the EPPO scheme, then the primary needs required in the EPPO scheme to enhance consistency are (i) the provision of examples that describe divisions within the scales, or alternative but equivalent descriptors that allow assessors to distinguish between divisions, and (ii) a mechanism to combine risk elements in a consistent and transparent way. Features that would help inexperienced assessors include a clear structure, clear rating guidance, questions posed unambiguously, provision of standardized answers and an easily applicable method to interpret and summarize risk ratings. Beyond improvements to the EPPO scheme, assessors using the scheme will need training. Providing links to information and suggesting data sources that would help assessors answer questions would also be helpful (this is being addressed in Work Package 1 of PRATIQUE).

Introduction

In the phytosanitary arena, pest risk analysis (PRA), 'the process of evaluating biological or other scientific and economic evidence to determine whether an organism is a pest, whether it should be regulated, and the strength of any phytosanitary measures to be taken against it' (FAO, 2008), is the science-based process that provides a systematic approach that helps to inform decision makers about whether and/or how to manage plant pests. PRAs provide the rationale for phytosanitary measures and, if they follow international standards, can be used to support the choices made by those responsible for regulatory decision-making in plant health. Consistency in the development of regulations enhances their credibility and reliability to stakeholders and is thus an important feature. With regard to any dispute settlement in unclear and challenged cases, the degree to which each stage in a PRA has been addressed correctly, clearly, transparently and consistently is very important. In contrast, inconsistent processes that result in similar situations being treated differently can result in the loss of trust and respect of those involved in the regulatory process and possible legal challenges.

Within plant health, consistency can be considered at different scales. The individual risk assessor is confronted with the challenge of making judgements in the presence of uncertainty, due to incomplete or contradictory data, but has nevertheless to strive to be consistent and transparent in the assessment of each risk element and consistent in the overall evaluation of pest risk, i.e. consistent in how elements contributing to overall risk are combined. At a higher level, using a PRA system that does not

*This work is part of the PRATIQUE Project. PRATIQUE is an EC-funded 7th Framework research project designed to address the major challenges for pest risk analysis (PRA) in Europe. It has three principal objectives: (i) to assemble the datasets required to construct PRAs valid for the whole of the EU, (ii) to conduct multi-disciplinary research that enhances the techniques used in PRA and (iii) to provide a decision support scheme for PRA that is efficient and user-friendly. For further information please visit the project website or e-mail the project office using the details provided below: Email: pratique@csl.gov.uk, Internet: http://www.pratiqueproject.eu.

provide consistent results or is unreliable is likely to result in inconsistent risk management measures being adopted. This breaches the key principle of minimal impact within the International Plant Protection Convention (IPPC), whereby phytosanitary measures must be consistent with the pest risk and use the least restrictive measures available (FAO, 2002, and Article VII.2g of the IPPC, 1997). It also breaches Article 5 of the WTO-SPS agreement on the application of sanitary and phytosanitary measures that states that each Member shall achieve consistency in the application of the concept of appropriate level of protection (WTO, 1995). Thus achieving consistency at all levels is not trivial and for a PRA scheme or protocol to be useful, reliable, widely adopted and its outputs trusted, it must be able to produce consistent results, that are not dependent on the individual risk assessor, i.e. the protocol and guidance is structured and applied in such a way that different assessors reach the same results when using the same information. However, developing a PRA scheme or protocol that delivers consistent results is recognised as a major challenge due to the lack of consistency in scoring responses to PRA questions (Baker *et al.*, 2009).

A primary challenge in improving consistency in risk assessment arises because qualitative assessments are generally made (Baker & MacLeod, 2005). The EPPO (2009) PRA scheme requires assessors to judge the likelihood or impact of an event from a limited predetermined choice of words arranged within a categorical ordinal scale, sometimes called a Likert scale, e.g. very likely, likely, moderately likely, unlikely, very unlikely (Anonymous, 2009). The current EPPO scheme uses such scales mostly with five divisions. However, no guidance is given on how to interpret each of the five divisions. Instead PRA assessors must make personal judgements about each risk element and support such judgements with detailed written evidence supported by references to the literature where available. Jablonowski (1994) and Theil (2002) showed that discrepancies, i.e. inconsistencies, arise when experts are asked to provide qualitative judgements (Baker *et al.*, 2005) and, whilst there are mechanisms to overcome this, e.g. using the Delphi technique (Ilbery *et al.*, 2004), which is a method applied by a group experts to find consensus by a series of intensive questionnaires that are combined with controlled opinion feedback (Dalkey & Helmer, 1963) such procedures can be expensive and time consuming. The Delphi technique is also increasingly becoming discredited (see e.g. Rowe & Wright, 1999).

This study aims to identify best practice for achieving and enhancing consistency in PRA by reviewing international, regional and national standards for risk analysis procedures and risk assessment protocols and examples from plant, public and animal health sectors around the world. By identifying factors contributing to ambiguity and inconsistency in other protocols, we aim to avoid such factors being part of a future EPPO scheme. In addition we aim to identify mechanisms that increase consistency and select those that can be examined further and perhaps taken forwards to enhance the existing EPPO scheme (OEPP/EPPO, 2007) specifically in relation to the protocol being able to produce consistent outputs between assessors examining the same pest, and between pests examined by the same assessor.

Methods

Most government risk assessment schemes are not generally available as published media and a conventional literature search did not provide any examples. However, web searches proved useful and examples were found as electronic media. Additionally, through professional colleagues around the globe that were willing to share schemes for this study, 43 documents, guidelines, protocols or examples of risk assessment schemes were eventually obtained and reviewed by six reviewers (five with experience as pest risk assessors, one with a more mathematical/statistical background) (Table 1). The documents were primarily examples of plant health pest risk assessment protocols although examples from other sectors, such as weed risk assessment, animal health, genetically modified organism (GMO) assessment, nature conservation, plant protection products and climate change[1] were also examined (Table 2).

Appendix 1 shows the standard template that was designed and used by the six reviewers to systematically examine each document. Unless after the first review it was clear that the document was not useful for the study, each document was reviewed at least twice by different reviewers. Reviewers focused on identifying whether standards and protocols explicitly identified the need for consistency and, if so, how consistency was addressed, e.g. to ensure a consistent approach to scoring/rating elements and other factors that contribute to risk.

In addition, documents[2]were examined to determine:

- their use of standard terminology, i.e. using definitions in ISPM No. 5;
- whether protocols provide guidance on risk assessment procedures;
- whether guidance is provided to assessors on how to select from a range of options on a categorical scale, i.e. is there guidance as to how 'unlikely' should be interpreted?
- how the degree of uncertainty, the combination of risk rating and uncertainty, and the summary of uncertainties are captured and analysed.

Results

Main findings

In most schemes and procedures, consistency is not directly addressed; consequently no clear suggestions are made on how to achieve or improve consistency within protocols. However, in an indirect way, some level of consistency is achieved or at least increased by, for example,

- using a clear and structured framework (e.g. New Zealand Animal health import risk analysis, # 54); such an approach

[1]Though the paper on climate change is not a guideline or standard for assessing risk, it nevertheless gives a guideline for risk rating, that was found useful to be included.

[2]For those documents that were complete risk analyses, i.e. those that included both risk assessment and risk management stages, only the risk assessment stage was reviewed. In one case, the title of the scheme referred to 'management' but included questions also relevant for risk assessment.

Table 1 Risk assessment documents (standards, guidelines, protocols and examples of their use) reviewed with regard to their approach to consistency, uncertainty, use of terminology and standardization

#	Name
Plant health: standards, schemes, guidelines	
1	ISPM2: IPPC International Standard for Phytosanitary Measures ISPM No. 2 (2007) Framework for pest risk analysis
2	ISPM11: IPPC International Standard for Phytosanitary Measures ISPM No. 11 (2004) Pest risk analysis for quarantine pests including analysis of Environmental risks and living modified organisms
8	EPPO DSS vers3: European and Mediterranean Plant Protection Organisation (EPPO) Pest Risk Analysis Decision support scheme 5/3 (3) version 2007
9	NAPPO RSPM 24: North American Plant Protection Organisation (NAPPO) Regional Standard on Phytosanitary Measures RSPM 24 (2005) Integrated Pest Risk Management Measures for the Importation of Plants for Planting into NAPPO Member Countries.
10	NAPPO RSPM 32: North American Plant Protection Organisation (NAPPO) Regional Standard on Phytosanitary Measures RSPM No. 32 (2008, draft) Pest Risk Analysis for Plants as Pests – Guidelines for Screening Plants for Planting Proposed for Import into NAPPO Member Countries
11	NAPPO RSPM 31: North American Plant Protection Organisation (NAPPO) Regional Standard on Phytosanitary Measures RSPM No. 31 (2008, draft) Guidelines for Conducting Pathway Risk Analysis
12	Canada Commodity: Canadian Food Inspection Agency (CFIA), Canada Plant Health Risk Assessment, Commodity Risk Assessment
13	Risk analysis NZ: New Zealand Biosecurity Risk Analysis Procedures (2006)
14	Australia IRA: Biosecurity Australia, Guidelines for Import Risk Analysis (draft 2001)
15	Turkey export PRA: Turkey Pest Risk Analysis Questionnaire (for export Agricultural Commodities, 2008)
18	Chile PRA: power point presentation of a regional workshop on Pest Risk Analysis, Santiago de Chile, 2007
PRA examples of use	
20	*Tilletia indica*: EC Fifth Framework Project QLK5-1999-01554: Risks associated with *Tilletia indica*, the newly-listed EU quarantine pathogen, the cause of Karnal bunt of wheat, Deliverable report
21	EPPO *Phytophtora lateralis*: Pest Risk Analysis for *Phytophthora lateralis* rev 1 Panel for *P. lateralis* following the EPPO Decision support Scheme for quarantine pests (draft 2006)
24	EPPO *Solanum elaeagnifolium*: Pest Risk Analysis for *Solanum elaeagnifolium*, EPPO Secretariat, 2006
25	Western Australia: Department of Agriculture, Government of Western Australia; Final State Pest Risk Analysis – Lettuce aphid 2004 (based on Australian national guidelines, Biosecurity Australia, 2001)
27	India PRA example: Phytosanitary importance of *Heterodera glycines* for South-east Asian countries, EPPO Bulletin 35, pp. 531–536, 2005
28	*Leucinodes orbonalis* NL: Plant Protection Service of the Netherlands, Pest Risk Analysis *Leucinodes orbonalis* (Guenée), 2005
29	*Erysiphe euphorbiicola* DE : Pest Risk Assessment *Erysiphe* cf. *euphorbiicola* (*Oidium sp*.), Germany, 2004
30	Stone fruit yellows CA: Canada (Plant Health Risk Assessment (PHRA) workshop Canada, Niagara Falls) *European stone fruit yellows phytoplasma*, 2005
31	Chile: PRA scheme for Chile (Plant Health Risk Assessment (PHRA) workshop Canada, Niagara Falls) *Marlene verdum*, 2005
32	Ghana: PRA scheme for Ghana (Plant Health Risk Assessment (PHRA) workshop Canada, Niagara Falls) *Vanilla tahitensis* and *V. fragrans*, 2005
34	EPPO *Pezothrips kellyanus* Sicily: Regione Siciliana, Assessorato Agricoltura e Foreste, Servizio Fitosanitario Regionale, Servizi allo Sviluppo, Pest Risk Assessment *Pezothrips kellyanus*, 2006
35	*Leucinodes orbonalis* UK: CSL Pest Risk Analysis for *Leucinodes orbonalis*, 2006
37	Australia IRA grapes from Chile: Australian Government, Biosecurity Australia, Final Report, Import Risk Analysis for Table Grapes from Chile, 2005
38	Australia IRA bulkmaize: Australian Government, Biosecurity Australia, Final IRA Report, Import Risk Analysis for the Importation of bulk maize (Zea mays L.) from the United States of America, 2002
55	Yams Jamaica: Ministry of Agriculture, Plant Quarantine/Produce inspection unit Jamaica, Pest risk assessment for the importation of *Dioscorea trifida* (yam, cush cush, yampee), and *Dioscorea alata/purpurea* (greater yam, water yam) from Gujarat, India into Jamaica
Weed risk analysis/assessment and invasive alien species	
39	Border control aquatic weeds NZ: New Zealand, Department of Conservation, Border control for potential aquatic weeds, stage 1. Weed risk model, 2000
40	Conservation WRA NZ: New Zealand Department of Conservation, A proposed conservation weed risk assessment system for the New Zealand border, 2002
41	WRA Canada: Canadian Food Inspection Agency (CFIA), Weed Risk Assessment
42	WRA South Africa: A proposed prioritization system for the management of invasive alien plants in South Africa, South African Journal of Science 99, 2003
43	Australian WRA: Australian Government, Biosecurity Australia, Weed Risk Assessment System (Pheloung), 2008
43a	South Australia WR manage: Department of Water Land & Biodiversity Conservation – South Australia, SA Weed Risk Management Guide, 2008
44	Exotic species ranking system USA: United States Department of the Interior, National Park Service, Midwest Regional Office, Nebraska, Handbook for Ranking Exotic Plants for Management and Control, 1993

Table 1 (Continued)

#	Name
45	Guidelines EIA Belgium: Belgian Biodiversity Platform, Guidelines for environmental impact assessment and list classification of non-native organisms in Belgium, 2007
46	UK non native: Department for Environment, Food, and Rural Affairs (DEFRA) United Kingdom, UK Non-native Organism Risk Assessment User Manual, Version 3.3, 2005 http://www.nonnativespecies.org/documents/uknnra%20user%20manual%20(v3.3).doc
56	Invasive species assessment protocol: NatureServe, Arlington, Virginia, USA, An Invasive Species Assessment Protocol: Evaluating Non-Native Plants for Their Impact on Biodiversity. Version 1, 2004.
Risk analysis/assessment schemes other sectors	
47	EFSA birds mammals: European Food Safety Authority (EFSA), Plant Protection Products, Risk Assessment for Birds and Mammals, The EFSA Journal 734: 1–181, 2008
49	EFSA GMPlants: European Food Safety Authority (EFSA), GMOs, Guidance Document of the Scientific Panel on Genetically Modified Organisms for the Risk Assessment of Genetically Modified Plants and derived Food and Feed (2006) The EFSA Journal 2004 – 99, pp. 1–94
50	EC GMPlants: European Commission, Health & Consumer Protection Directorate-General, GMOs, Guidance document for the risk assessment of genetically modified plants and derived food and feed
52	Canada Animal Health: Canadian Food Inspection Agency (CFIA), Animal Health: Protocol of the Animal Health & Production Division and Animal, Plant and Food Risk Analysis Network (APFRAN), Science Division, Animal Health and Production Risk Analysis Framework, 2000
53	Terrestrial animal health: World Organisation of Animal Health (OIE), Animal Health: Terrestrial Animal Health Code, Chapter 1.3.2 Guidelines for Risk Analysis, 2007
54	OIE handbook: New Zealand, Ministry of Agriculture and Forestry (MAF), Animal Health: Handbook on Import Risk Analysis for Animals and Animal Products, Volume 1 and 2, 2004
57	IPCC, 2005. Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties. Intergovernmental Panel on Climate Change. http://www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf (see footnote No. 2)

Table 2 Summary of risk assessment documents examined and reviewed for approaches to consistency. Numbers in the table are the reference numbers to the documents listed in Table 1

	Document		
Sector	International standards	National/regional guidelines	Examples of use
Plant health	1, 2	8, 9, 11, 12, 13, 14, 15	20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 34, 35, 37, 38, 55
Weed risk assessment		10, 39, 40, 44	41,42,43, 46a
Animal health	53	52, 54	
GMOs		48, 49, 50	
Nature conservation		45, 46	33, 56
Plant protection products		47	
Climate change	57		

ensures consistency in the format and structure of PRA between and within assessors using the scheme. However, in this example, it doesn't address consistency in how risk elements are assessed and combined

- clarity of questions (e.g. Proposed conservation weed risk assessment scheme, New Zealand Department of Conservation # 40); clarity of questions will help understanding but on its own does not ensure consistency in assessing risk elements or in how risk elements should be combined
- by requiring groups of assessors to conduct the assessment through the Delphi method, where different assessors do not work face to face/do not interact (e.g. South African Prioritization system for the management of invasive alien plants in South Africa, # 42); Through feeding back individuals' judgements over a number of rounds, such an approach leads to consistency within the group making the individual assessment consistent between assessors. However, this does not address consistency in assessing risk elements between different pests since different groups of experts would be used in each case
- the provision of examples and guidance (e.g. Canadian pest risk assessment, # 12). It was felt that this approach aids consistency in assessing individual risk elements on different pests by the same assessor and the same pest by different assessors. The system also provides a consistent mechanism for combining risk elements. Again this ensures consistency between pests within assessors and between assessors
- by providing a table that links qualitative descriptions to probabilities.

All examples reviewed were qualitative risk assessments. Qualitative assessment schemes are potentially inconsistent since they require inputs in the form of opinion and/or judgement of an assessor. In this case inconsistency can be a result of varying background knowledge of assessors, and their attitude to risk. This can be mitigated if a qualitative PRA system is supported with rating guidance and examples. It could stay qualitative but be consistent to a quite high extent. Poor protocol design, e.g. resulting from inaccurate definitions/unexplained terms or ratings, or no transparent method for combining ratings or scores could also produce inconsistencies. In the examples studied clear definitions of ratings were not provided. However, schemes 12, 14, 25, 37, 38, 41, 43, 43a, 45 and 57 (refer to Table 1 for details of schemes) did give clear guidance.

General sources of inconsistencies

Through reviewing the documents listed in Table 1, the reasons for inconsistencies are manifold. Those resulting predominantly from protocol design can be summarised as:

- Lack of clear structure of the scheme
- Lack of guidance on how to interpret questions, questions sometimes ambiguous
- Lack of rating guidance
- Lack of a standardized method to interpret ratings
- Lack of guidance on how to draw conclusions from many questions within a section examining one risk element, e.g. overall conclusion about likelihood of entry
- Lack of guidance in combining risk elements to make an overall conclusion of risk
- Lack of consideration in dealing with uncertainty, which is particularly important when data are lacking or conflicting
- Lack of mechanism to combine risk ratings and degree of uncertainty
- Lack of requirement to indicate where expert judgment has been used and clearly document and justify which rating/score has been chosen
- Lack of standardization of answers (using free text as the only way of expressing elements of risk does not help with consistency. A standardized set of 'defined' linguistic terms would help with transparency and consistency).

Within the plant health risk assessment procedures, IPPC terminology was generally used. However, no clear advice is given that the assessor should use such terminology. In some cases, terms are used differently (e.g. in Canada, 'Introduction' was used to refer to entry, and not entry and establishment, as defined in the IPPC Glossary ISPM No. 5). Strict use of agreed terminology and definitions increases consistency, e.g. by increasing a common understanding of technical terms. Animal health guidelines use OIE terminology.

Two major difficulties in developing a meaningful protocol for risk assessment that are generally valid were only noted specifically in a Weed Risk Assessment for aquatic plants (#39). These are: inadequate tools (i.e. models for determining potential weediness in this case) and unreliable data.

Other explanations for inconsistencies, though not shown through the review above, could reasonably be expected to arise from the requirement that assessors use subjective judgement and be summarized as:

- Scope of the PRA, e.g. the PRA area or the time scale for long term impacts to be considered, not being clearly specified
- PRAs being performed by specialists in specific taxa, as opposed to an assessment made by a generalist. This could be described as information asymmetry between assessors
- When a specialist in a specific taxon conducts a PRA on an organism outside his specialism, there may be a lack of consistency between PRAs performed by the assessor
- Different attitudes of assessors to risk (risk-averse or risk-willing) when guidance to questions and answers is not standardized
- Different background knowledge and experience (another example of information asymmetry between assessors)
- Misunderstandings of terminology.

Approaches to increase consistency

In several cases the protocols include features that would increase consistency, such as:

- Guidance to explain questions (e.g. notes provided in the EPPO scheme)
- Providing a fixed set of options to rate elements of risk. Depending on the protocol used, risk elements are divided into 4 (CFIA, 2008 #12), 5 (Baker *et al.*, 2008), 6 (Biosecurity Australia, 2001, #14) or 9 (OEPP/EPPO, 1997).[3] Due to cognitive limitations (limits of the human brain), a maximum of seven probability categories are the most that than can be handled efficiently (Beyth-Marom, 1982; Budescu & Wallsten, 1985). Indeed the EPPO (1997) risk assesment system has been revised (OEPP/EPPO, 2006, 2007, 2009) and now requires assessors to choose from 5 rather than 9 options when describing components of risk.
- Guidance and examples to help select answers when a categorical scale is used (e.g. when assessing pest impact, the Canadian risk assessment scheme (#12) has four options to choose from. The lowest impact is 'negligible'. This is intended to mean that there is 'no impact on yield, host longevity, production costs or storage'. This is further explained with an example 'Septoria leaf spot (*Septoria ampelina* Berk. and Curtis) infection results in leaf drop in grape that is premature by a few days only, with no treatment necessary and no economic losses.' Providing such descriptions for each division of a scale enables different risk assessors to have a common frame of reference when they are assigning risk ratings if sufficient detail is provided
- Linking qualitative descriptions to probabilities (#57)
- Separating questions that combine two issues, e.g. the volume and frequency of imports or the suitability of climate/hosts and the proportion of the PRA area which has suitable climate/hosts, or providing clear methods for combining the issues into one answer, e.g. by using a table or matrix
- Using closed questions, i.e. questions that can be answered only yes or no (e.g. #43) Whilst there is scope for debate, as

[3]The authors are aware that the US Guidelines for pathway-initiated pest risk assessments (USDA, 2000) uses 3 divisions. Unfortunately access to this document was not obtained during the period reviews were being conducted and so it was not included in the list of schemes originally reviewed.

this is only possible for a certain kind of questions, having such questions could improve consistency but perhaps at the expense of transparency

- Providing a method to combine answers to questions to determine an overall rating for each risk element
- Providing a method to combine risk elements to determine an assessment of overall risk
- Elaboration of generic criteria
- Explanation of procedural steps
- Training of assessors
- Review of draft PRAs
- Documentation of the different steps of the procedure
- Requirement for detailed, referenced documentation and justification for scoring/rating decisions
- Narrow range of clearly understandable options. There is a balance that needs to be struck between ensuring consistency and providing the ability to differentiate between risks. For example, in an absurd situation where assessors were not provided with a choice of answers to select from but were forced to select 'yes' then all answers would be the same, hence consistent between assessors. At the other extreme, questions that are answered using free text would allow different assessors to have many possible answers. With only a few options possible for each question then the scope for assessors to have different/alternative answers is more limited.

Uncertainty

In most schemes, the degree of uncertainty is not covered. Incorporation of uncertainty in overall risk assessment adds to transparency and its inclusion is good practice. See outputs from Task 3.2 for review of uncertainty within PRA and further discussion as to how it can be incorporated into the EPPO protocol.

Discussion of selected reviews

No one individual scheme/guideline fulfilled all conditions for 'best practice' but several had individual features that could increase consistency and be considered further for use in a revised EPPO PRA scheme. The following text highlights such features from a variety of the PRA documents reviewed.

ISPM 2 (1) *and 11* (2) provide a framework for PRA. They have a special status in this analysis, as they are the agreed international standards that contracting parties to the IPPC should comply with when preparing phytosanitary regulations. ISPM 2 gives an overview of the benefits of consistency, but is basically a document useful for 'pest categorization' and provides the general framework for PRA. However, ISPM 2 does not differentiate between variability, i.e. that which is due to natural variation, and uncertainty, that which is due to lack of information. ISPM 11 does not mention any methods or benefits regarding consistency. Uncertainty is discussed as being inherent to PRA, but again, no guidance or methods are provided to deal with uncertainty. Recommendation: as Baker & MacLeod (2005) recognized, there is a need for a guidance document showing how ISPMs can and should be used. With specific reference to uncertainty, the approach recommended by EFSA (2006) in the context of dietary exposure assessment, is equally suitable for expressing uncertainties in pest risk assessment. The method summarizes the uncertainties and their potential effects. By arranging an evaluation of the uncertainties in a table, the risk assessor can form a judgement about their combined effect on the assessment outcome. It provides risk managers and stakeholders with a transparent picture of the uncertainties affecting the assessment, and allows them to be taken into account in decision-making.

The EPPO PRA Decision Support Scheme (version 2007) (8) is a decision-support scheme to assess the potential risk of a particular pest for any clearly defined area. (However, there is no requirement for the time horizon to be specified). It provides guidance for some questions, but not for all questions. Rating guidance is usually not given in this scheme. The assessment stage of the scheme consists of questions concerning the temporal or spatial frequency of events or entities, the likelihood, the similarity of conditions, and the magnitude of impacts. The scheme is pragmatic in that it recognizes that very often there is a lack of empirical data to provide a qualitative or definitive answer to questions in the risk assessment scheme and the scheme allows expert judgement to be used to decide answers to questions. Over the approximately 15 years of its development, there have been continuous attempts to enhance the consistency of the EPPO scheme. Many problems have been resolved but some questions can only be answered with great difficulty and are the object of recurrent debates during EPPO Expert Working Groups for Performing PRA (F. Petter pers. comm.). Difficulties come either from (mis-) understanding and (mis-) interpreting the questions, or from a lack of rating guidance. Some PRA topics are inherently difficult to address especially when there is little or no information. The scheme already includes several good attempts to increase consistency, but also includes some elements decreasing or at least not adding to consistency:

Elements helping to increase consistency:

- notes explaining questions
- assessors are requested to provide details and references for ratings (nevertheless, such supporting data still requires interpretation and different assessors can interpret information in different ways)
- PRAs reviewed by peers (this is often done, but not requested in the scheme)
- ratings for uncertainty for every question
- clear and transparent structure.

Elements decreasing or at least not adding to consistency:

- in some cases, notes are not provided, not clear or are ambiguous
- no examples for what scores represent
- details and referencing are not requested in a standardised way
- provision of details and references does not guarantee that two assessors would come up with the same results but allows a reviewer to understand why there is a disparity in the assessment (as per note above)
- no guidance and examples for uncertainties

- no guidance on how to combine uncertainty ratings
- no advice on how to achieve consistency
- no mechanism for combining elements of risk
- geographic and temporal scope can vary between assessments of different pests.

The NAPPO RSPM 31 (11) (Regional Standard on Phytosanitary Measures No. 31 (2008, draft) Guidelines for Conducting Pathway Risk Analysis) provides an appendix (Appendix 2: Example of scoring system for section 2.2.1) with formulas for assessing pathways quantitatively, e.g. '*Probability of entry or spread of a pest or pests via the pathway = (ease of inspection for the pest factors) × (ease of detection factors) × (magnitude factors) x (survivability factors) × (environmental factors)*.'

The Canadian plant health risk assessment (12) (commodity PRA, but also applicable to pest specific PRA) is quite well advanced in respect of aiming to provide consistent outputs. Examples are provided – not for every detail, but for every section (likelihood of introduction, establishment potential, natural spread potential, potential economic impact, potential environmental impact). From these scores, a cumulative score is produced to summarize the consequences of introduction. No weighting of questions is done, the combination of scores is done by simple addition and the conclusions are drawn from a risk estimation matrix. However, as long as this combination of risk elements is done in a consistent manner, it may not really matter if answers are weighted or not. Risk ratings are not combined with a judgment on uncertainty. Having examples for questions or sections is a very useful approach that will be put forward to the EPPO scheme and is part of subtask 3.1.2.

The Risk Analysis Procedures of New Zealand (13) include descriptors for critical attributes of risk (Table 3). An example (from animal health) is used throughout the guidelines to illustrate the different steps and procedures. (An example from plant health used throughout the EPPO scheme would be a useful addition to a revised EPPO Protocol). The guidelines require a high level of detail.

The Australian Import Risk Analysis Guidelines and the State Pest Risk Analysis for Western Australia (14, 25) provide a description of what each score represents (e.g. 'High': the event would be very likely to occur; range = 0.7–1 or 'Negligible' the event would almost certainly not occur; range = $0–10^{-6}$) and a method for propagating uncertainty in the likelihoods (''The simulation-based approach provides a very simple and robust means by which the 'uncertainty' inherent in most import risk analyses can be represented and incorporated in the assessment process''). Of note is that the Australian system does not use a symmetrical score. There are more categories towards the lower end (negligible, very low, low) than at the higher end. Also, they include a procedure for involving stakeholders during risk communication. The guidelines provide a consistent procedure for combining the likelihoods and consequences by using risk estimation matrices. (WP 3.2 is exploring the use of risk matrices within a Bayesian Belief Network (BBN) framework).

The Border Control for Aquatic Weeds of New Zealand (39) provides guidance or examples for scores. The authors of this paper (Champion and Clayton) state, that 'in summary, the two major difficulties in developing a meaningful Weed Risk Assessment for aquatic plants have been inadequate tools (i.e. models for determining potential weediness) and unreliable data (i.e. identity of potential ecological weeds already in New Zealand and volume of traffic entering the border by illegal means).' (see General sources of inconsistencies above).

Table 3 Critical risk attributes and risk descriptors use in the Risk Analysis Procedures of New Zealand

Risk attributes	
Negligible	Not worth considering; insignificant
Non-negligible	Worth considering; significant
Risk descriptors (not all may be used)	
Very low	Close to insignificant
Low	Less than average, coming below the normal level
Medium	Around the normal or average level
High	Extending above the normal or average level
Very high	Well above the normal or average level

The Conservation Weed Risk Assessment (WRA) for the New Zealand Border (40) provides clear and well explained questions, examples and a useful background. As the scheme is only asking for measurable information and biological facts, judgements are not required and therefore the responses between different assessors should be consistent. However, for such an approach to be adopted and used in a new EPPO PRA protocol, there would have to be significant changes to the current EPPO PRA structure. More work would be needed to determine whether such an approach would be feasible for the EPPO scheme.

Prioritisation for management of invasive alien plants, South Africa (42) provides an interesting approach. Questions are clear and straightforward to answer. Many questions require quantitative data, thus judgement/interpretation of information is not needed (e.g. distribution has defined categories already). However, some questions do need information to be interpreted hence judgements are required (e.g. '*The options for mechanical control of the species are:(i) not available,(ii) impractical in most situations, (iii) partially successful, or (iv) effective and practical'*. When applied in a Delphi-technique style, several assessors are needed for each assessment and it is suggested that mean scores from assessors be used to determine the overall priority of an active management plan for each plant assessed. In plant health terms, an urgent need for a management plan would equate to a high-risk pest. In principle, avoiding the necessity of judgment by using quantitative data would be helpful in the EPPO scheme. However, this will only be possible for a few of the questions and species/situation where data are available – defined categories for distribution as mentioned above would be possible. More possibilities will be identified in the course of subtask 3.1.2.

The Australian Weed risk assessment (43) is a simple scheme, as most questions are closed (yes/no/don't know). Only two questions require rating (low/intermediate/high). Guidelines are provided to increase consistency, giving information about what the questions are asking and what information is required to answer them. If information on the plant's biology and its requirements is available, answering the questions is very easy and should be consistent between assessors having the same information. The WRA system can characterise species put through the scheme as useful, non-problematic plants or potential weeds of the environment and/or agriculture. It was calibrated and tested on 370 species, ranging from environmental and agricultural weeds to benign and beneficial plants. It was judged on its ability to correctly reject weeds, accept non-weeds and generate a low proportion of species requiring further evaluation (Pheloung, 1995). Being seen as a very useful tool, the WRA scheme has been adapted for use elsewhere (Gordon *et al.*, 2008a) i.e. in Hawaii (Daehler & Carino, 2000), Hawaii and Pacific Islands (Daehler *et al.*, 2004), the Czech Republic (Kriváne k & Pyšek, 2006), the Bonin (Ogasawara) Islands of Japan (Kato *et al.*, 2006), and Florida, USA (Gordon *et al.*, 2008b). Weed risk assessment type schemes to identify risky species based on their invasive attributes are especially helpful for plants and fish where many species are not invasive and it is important to have guidance for question 17 of the pest categorisation stage. This approach has been adopted by the UK Non-Native Risk Analysis Scheme (46) with adaptations of the Pheloung (1995) scheme for plants, freshwater fish etc. Such guidance is not really needed for plant pests, since it is generally already clear whether they are potentially damaging. Also studies by, e.g. Williamson (1999) have shown that once you try and find invasive attributes for broad taxonomic groups in areas with numerous and variable ecosystems the only evidence you can use is: 'is the pest invasive elsewhere?' and study propagule pressure. (This topic is being investigated further by deliverable 2.2).

South Australia Weed Risk Management Guide (43a) is concerned with the management of weeds, nevertheless it also includes points relevant for risk assessment (such as the probability of establishment, reproductive behaviour, long distance dispersal by natural or human means, impacts, potential distribution) and has a good approach to increase consistency by explaining questions clearly with examples and precise rating guidance. Answers are pre-formulated – see Table 4 as an example.

The Guidelines environmental impact assessment from Belgium (45) give general guidance for each question, rating guidance is provided with examples for each rating. Scores are summed without any weighting of questions (see comment regarding weighting above).

The Non-native organism risk assessment scheme from UK (46) is based on the EPPO scheme (former version) but it includes a rating scheme for interpreting the five level scores for impacts and likelihoods. It provides a table with a word scale of five divisions (minimal to massive) and equates each division to an economic value, with equivalent summary descriptors of health, environmental and social impacts. For example 'minor' impact equates to monetary loss or costs of responses of between 10 000 GBP and 100 000 GBP per year; mild short-term reversible effects on human health to identifiable groups at a local scale; some reversible local ecosystem changes; and significant social concern expressed at a local level. The scheme also has a table describing likelihoods quantitatively and qualitatively. For example a likelihood score of 2 (from within a scale of 1–5) means an event is 'unlikely' which is given to mean that 'this sort of event has not occurred anywhere in living memory' or is expected to happen only once in 1000 years. The magnitude and likelihood ratings were based on the Australia and New Zealand Risk Management Standard (AS/NZS 4360 Risk Management), but with some modification of the monetary values, and of the wording. The score is based on log (upper monetary value)-3. The Australia/New Zealand Standard uses seven categories of likelihood, including additional 1/3 year and 1/30 year frequencies, but these have been deleted to provide the 5 levels of risk. Such descriptors are useful because they provide a rating system for all questions requiring likelihood and impact responses, but no examples are given for each question. Three methods for summarizing risk were provided based on: the opinion of the assessors, simple averaging (but including, as in the EPPO scheme, the weighting given if key impact questions score major or massive) and conditional probabilities. This scheme is now used routinely in Great Britain providing support to non-native species policy. Consistency in scoring and conclusions is provided by a peer review system and a risk analysis panel that compares the 5 level scores for questions and conclusions with the detailed written justifications, pointing out discrepancies. Risk summaries are converted to graphs of accumulated risk.

The Invasive species assessment protocol, USA (56) has an interesting approach to scoring. 4 scores are possible and rank from A to D (A maximum score = 3; B = 2; C = 1; D = 0); the questions are weighted differently to reflect their relative contribution to the sub-rank. The maximum possible point total for each section is divided into four equal intervals representing sub-ranks of 'high', 'medium', 'low' and 'insignificant'. When a species is evaluated, the points for each answer in a section are tallied to yield a total that is used to determine the corresponding subrank; however, no concrete examples are given.

Table 4 Example of rating guidance in the South Australia Weed Risk Management Guide

Q1. What is the weed's ability to establish amongst existing plants?		Score
☐ Very high	'Seedlings' readily establish within dense vegetation, or amongst thick infestations of other weeds.	3
☐ High	'Seedlings' readily establish within more open vegetation, or amongst average infestations of other weeds.	2
☐ Medium	'Seedlings' mainly establish when there has been moderate disturbance to existing vegetation, which substantially reduces competition. This could include intensive grazing, mowing, raking, clearing of trees, temporary floods or summer droughts.	1
☐ Low	'Seedlings' mainly need bare ground to establish, including removal of stubble/leaf litter. This will occur after major disturbances such as cultivation, overgrazing, hot fires, grading, long-term floods or long droughts.	0
☐ Don't know		?

The risk assessment for birds and mammals by EFSA (47) is difficult to judge with regard to the fact that it is very different from the plant health sector. However, the approach whereby uncertainties are tabulated and each individual uncertainty is then described as either contributing to an increasing or decreasing risk would be useful in the EPPO PRA scheme. The approach would help highlight and focus on where resources could appropriately be used to reduce uncertainty and reduce risk (Table 5).

The Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties (57) use probability and impact language in a way that scoring systems can be consistent and provide the following table that links qualitative descriptions to probabilities (Table 6):

Discussion

All the assessment protocols reviewed were qualitative. When such an approach is taken, one of the main difficulties is ensuring consistency between assessors and between assessments. The reason for this is that qualitative methods rely on the subjective judgement of the assessor(s) using words to convey the magnitude of the each element of assessed risk. However, words can be interpreted or translated in a way that they lose their intended meaning, when no guidance and examples are provided. In contrast, quantitative techniques use measurable, numerical terms that are explicit and can be used to convey unambiguous meaning. Quantitative methods may overcome some of the limitations of qualitative methods, such as providing a more consistent interpretation, but they may also give a false sense of accuracy, depending on the quality of the source data. They are also hampered by time and data constraints (EFSA, 2008). As a consequence, quantitative assessments are rarely, if ever, used in plant health, and as noted previously none were included in this review.

Many of the risk assessment schemes require assessors to select descriptors from a categorical scale to summarize elements of risk (exceptions: ISPMs 2 and 11 (#1 and 2), #39, 42, 44, 56). At least three types or levels of inconsistency can be anticipated within such schemes,

- inconsistency in categorizing information, e.g. assessors using the same data/information but interpreting it in different ways and selecting different options to summarize it within the scale
- inconsistency in combining elements of risk, e.g. protocols could address likelihood of pest entry and likelihood of pest establishment separately then ask the assessor to reach a decision about pest introduction. Assessors could agree on entry and establishment but draw different conclusions about introduction
- inconsistency in the overall assessment of risk. This is similar to level 2 but is perhaps more important as it influences risk management decision-making.

There were a number of approaches used to describe components of risk/risk elements within the different schemes. Some used free text (EPPO scheme #8 and all related examples; #13, 15, 20), others used some form of categorical scale with either numbers or words used to indicate divisions within the scale. Schemes that provided more than one way of describing categories within the scale are more versatile and could be used according to the data available. The UK non-native scheme (2005) (#46) provides two such tables, one describing magnitude (of impacts) and one for likelihood. The Canadian pest risk assessment system (#12) also provides good examples and uses scales with four divisions. The scales are categorized using numbers (0–3), words and are accompanied with a realistic biological example to describe the overall risk element being assessed. The Australian system (#14) uses 6 divisions and provides descriptions for each division. MacLeod & Baker (2003) provided quantitative and verbal descriptions for a limited number of questions of an earlier version of the EPPO PRA scheme that used a scale of 1–9 ratings. Illustrating each division within a scale with examples or values more easily allows assessors to select from within the scale with greater consistency and enables the risks posed by different pests to be compared.

Examples of useful scales with alternative yet equivalent descriptors to rank/rate risk elements are seen in country specific systems (e.g. Australia, Canada, UK). The fundamental challenge to overcome within PRATIQUE is that the EPPO scheme covers 50 countries. It may help if, for the purposes of PRATIQUE, the geographic range of consideration could be limited to EU 27 countries. As an EU funded project, PRATIQUE focus is on EU 27, e.g. for dataset collection, modelling & mapping. Nevertheless if a rating system can be developed that is scale independent such problem can be overcome. To make this work, rating examples would have to be provided in terms of proportions, i.e. proportions of the PRA area, of the PRA area production, imports etc. To identify protected zones etc., the PRA area will often need to be much smaller than the EU.

Several different approaches are used to arrive at an overall assessment of risk. Some systems [e.g. No. 8 (EPPO, 2007); 13 (NZ Risk Analysis Procedures)] had no formal way of combining risk elements and assessors use judgement in reviewing previous components of risk to subjectively describe their perception of overall risk. Such an approach does not provide transparency and could be very inconsistent. Good practice was seen in other systems that had designed a mechanism for combining risk elements (e.g. #12, #14# and others). Devising a transparent mechanism to combine risk elements to arrive at a final overall evaluation of risk would dramatically improve consistency within the EPPO scheme, and is one of the aims of PRATIQUE (3.2).

For a pest risk assessment protocol to be described as being consistent, it should be demonstrated that assessors using the same data for the same pest should reach the same conclusions about PRAs given the same remit. Table 7 is a matrix illustrating how a PRA protocol can be judged when the criteria considered within a protocol (or the answers to questions within a protocol), and the conclusions of analysts that use the same information, are compared in terms of their variability/consistency.

Thus assessors that independently use a protocol to assess the same pest for the same PRA area over the same time scale and each judge elements of risk to be the same and subsequently draw the same conclusions would be using a useful and reliable

Table 5 Example of a table showing where uncertainties lie within an EFSA risk assessment, and how the uncertainties increase or decrease risk. (TER; toxicity-exposure-ratio)

Evaluation of uncertainties affecting comparison of the first-tier TER assessment of acute avian risks with data on mortality in field studies. The aim is to find the critical TER value above which any mortality will be unlikely (surrogate protection goal). Symbols are used to indicate the extent to which the true critical TER could be lower (−) or higher (+) than 10. The number of symbols provides a subjective evaluation of the approximate magnitude of the effect, e.g.+++ indicates a factor that could increase the critical value by a factor of 10.

Source of uncertainty	Potential to decrease critical TER	Explanation	Potential to decrease critical TER	Explanation
Uncertainties affecting the evaluation of the field studies				
Variable quality of field studies	−	Evaluators took account of study quality when assigning subjective probabilities. It is possible that in doing so they might have overstated the probability of effects for poorer studies.	+	It is possible that evaluators might have understated the probability of effects for poorer studies.
Matching field studies to TER scenarios		Two of the field study evaluators matched the field studies to the TER scenarios, which were then checked by a third person.		
Subjectivity of evaluation		Four evaluators gave similar results. Average values used for analysis.		
Relationship of results to actual effects			+	Probable that some field studies with TER > 10 caused some undetected (hence not visible) mortality (see discussion in text).
Uncertainties affecting the form of the relationship between TER and field effects and its extrapolation from the available studies to other pesticides and scenarios				
Toxicity		Most field study pesticides had mean LD50 < 100 mg kg^{-1} bw but relationship of TER to field effects is expected to be similar for less toxic pesticides		
Molecular weight		Larger substances with lower uptake may present less hazard, but this should be reflected in LD50.		
Other pesticide properties	−	Field study pesticides cover the general range for Kow, need for activation, and reversibility of effects. It is possible that these or other factors reduce critical TER substantially for some pesticides but they are not well enough understood to be included in the first-tier assessment (see text).	+	Field study pesticides cover the general range for Kow, need for activation, and reversibility of effects. It is possible that these or other factors increase critical TER to a limited extent for some pesticides (see text).
Application method		Many field studies used aerial applications but in a separate analysis, effect of application method was not significant ($P > 0.05$).		
Multiple applications	−	Few studies with multiple applications, no obvious difference. TER calculations include theoretically appropriate adjustment. However, possible that MAF factors over-represent the contribution of repeat applications, which would cause critical TER to be lower.		Possible that more studies with multiple applications might have indicated a higher critical TER. Possible that MAF factors under represent the contribution of repeat applications, which would cause critical TER to be higher.
Crops	−	Studies cover wide range of crops, no sign of consistent differences. Possible some crops not included in field studies may have lower critical TER.	+	Possible that some crops not included in field studies may have higher critical TER.

The EFSA Journal (2008) **734**, 41–181.

PRA protocol (top left of Table 7). In contrast, if assessors used a PRA protocol and given the same scope and the same information, made many different judgements about the risk elements/criteria considered by the protocol, and subsequent conclusions between many assessors were different, then the PRA protocol would not be very useful due to it being very

Table 6 Suggested descriptions of a likelihood scale by the IPCC (2005)

Terminology	Likelihood of the occurrence/outcome
Virtually certain	>99% probability of occurrence
Very likely	>90% probability
Likely	>66% probability
About as likely as not	33 to 66% probability
Unlikely	<33% probability
Very unlikely	<10% probability
Exceptionally unlikely	<1% probability

inconsistent (bottom right of Table 7). However, it is important to recognize that the current EPPO risk assessment method, which requires score ratings to be justified by a detailed written description does enable the scores to be reviewed independently and clear inconsistencies to be identified and corrected. This process is followed by EFSA when reviewing the pest risk assessments that are sent to the Plant Health Panel and has been formalized in Great Britain, where all non-native risk assessments (using a scheme that is based on the EPPO scheme) are subject to peer review and scrutiny by an expert panel. But even if this process exists, a consistent approach right from the beginning of conducting a PRA would facilitate procedures significantly.

Recommendations

In order to achieve best practice in consistency in the enhanced EPPO Decision Support scheme, the scheme should:

- be clearly structured
- include explicit advice on using IPPC terminology
- give clear guidance for all questions (questions should be unambiguous)
- provide a selection of standardized answers
- provide clear rating guidance with examples; comments on ratings should be provided in a way that allows assessors to clearly distinguish between, e.g. likely and very likely
- provide a standardized method to interpret and summarize ratings
- provide a mechanism to combine risk elements to provide an assessment of overall risk
- provide a method to combine ratings with judgement on uncertainty
- specify the elements that must be addressed in the text that justifies the ratings given.

Assessors using the scheme should indicate clearly where expert judgement by the assessor or in consultation with others (as, e.g. a personal communication) has been used so that score ratings can be reviewed by others and clear inconsistencies corrected.

As a consequence of implementing the recommendations above, the same assessor should, when assessing different pests representing similar risks, obtain the same level of risks for each pest. Different assessors assessing the same pests with the same amount information should produce similar assessments. In addition, when drawing a conclusion about risk elements, such as overall likelihood of establishment, or spread, or impacts etc. the answers to questions should be used in a consistent way to determine the overall likelihood for each risk element, and in the same way in respect of overall magnitude for impacts. Furthermore, each conclusion about the likelihood or magnitude for impacts for each risk element should be used in a consistent way to determine the final pest risk. Clear and explicit definitions improve the transparency of risk assessments and to increase the consistency of the ratings among risk assessors. A risk assessment procedure that is consistent and transparent should ensure fairness and rationality. If consistency in the EPPO scheme can be increased, e.g. by comparison of risks, this would also facilitate stakeholder engagement, risk communication and improve understanding by all stakeholders.

Table 7 A matrix to evaluate a PRA protocol by determining the variability of assessors' responses to criteria within the protocol and the overall conclusions

		Variability of conclusions		
When analysts use the same information…		All conclusions are the same	A few conclusions are different	Many conclusions are different
Variability of judgments about criteria/answers to questions in the PRA protocol	All answers are the same/all criteria are judged to be the same	A useful/reliable/consistent PRA protocol	Depending upon how different the conclusions are, this protocol may need more work to improve how conclusions are reached.	Unreliable mechanism for drawing conclusions/needs further work to improve the protocol.
	A few answers are different/few criteria are judged to be different	Depending upon how different the answers are, this is still a useful PRA protocol.	May need to check sensitivity of conclusions to particular questions. More work needed.	Unreliable mechanism for drawing conclusions/needs further work to improve the protocol.
	Many answers are different/many criteria are judged to be different	May need to check the usefulness of questions.	May need to check the usefulness of questions. Which questions actually make a difference to conclusions?	Very inconsistent, not useful.

Where this work contributes to pratique

This milestone directly feeds into deliverable 3.1, due in month 33 (November 2010), that will provide guidance with respect to answering questions within a modified EPPO decision support scheme. This text also supports WP2 (Task 2.3) and WP4 (Subtask 4.2.2) that will develop methods for enhancing consistency in the scoring of spread, impacts and entry potential. The deliverable resulting from Task 3.1 will be the construction of an enhanced EPPO protocol with decision rules, illustrated by examples, for scoring levels of risk in the EPPO PRA scheme. This milestone is the first step towards that aim.

Acknowledgements

This work was supported by the EU Framework 7 Research Project, Enhancements of Pest Risk Analysis Techniques (Grant Agreement No. 212459). This publication has been funded under the small collaborative project PRATIQUE, an EU 7th Framework Programme for Research, Technological Development and Demonstration addressing theme: (kbbe-2007-1-2-03: development of more efficient risk analysis techniques for pests and pathogens of phytosanitary concern call: fp7- kbbe-2007-1).

Améliorations des Techniques d'Analyse de risque phytosanitaire. Etude des bonnes pratiques pour améliorer la cohérence

Les normes internationales, régionales et nationales pour l'évaluation du risque phytosanitaire et des exemples de leur utilisation dans les pays autour du monde ont été examinés avec des documents similaires venant de domaines associés comme la santé animale, la préservation de l'environnement, l'évaluation des organismes génétiquement modifiés (OGM) et l'évaluation des risques des adventices pour déterminer comment est traité la cohérence de l'évaluation du risque, ou des composants du risque, au sein et entre les évaluations. Différentes approches ont été identifiées et pourraient être adoptées et adaptées pour être utilisées dans une révision du schéma d'aide à la décision pour les organismes de quarantaine de l'Organisation européenne et méditerranéenne de protection des plantes (OEPP) pour en améliorer la cohérence. Aucun schéma ne comporte de mécanisme pour assurer ou garantir la cohérence et aucun schéma ne contient toutes les approches identifiées pour maximiser la cohérence. Si l'utilisation d'une échelle à cinq niveaux pour décrire les composants individuels qui contribuent à évaluer le risque phytosanitaire global est une approche à conserver dans le schéma OEPP, il est avant tout nécessaire pour améliorer la cohérence dans le schéma OEPP de (i) donner des exemples qui décrivent les divisions dans les échelles, ou des descripteurs alternatifs mais équivalents qui permettent aux évaluateurs de faire la différence entre les divisions, et (ii) d'élaborer un mécanisme pour combiner les éléments du risque de façon cohérente et transparente. Les caractéristiques qui aideraient les évaluateurs inexpérimentés sont notamment une structure claire, un guide de notation clair, des questions posées sans ambigüité, la mise à disposition de réponses normalisées et une méthode facilement applicable pour interpréter et résumer les notes de risque. Au-delà de ces améliorations dans le schéma OEPP, les évaluateurs utilisant le schéma ont besoin de formation. Donner des liens vers des informations et suggérer des sources de données qui aideraient les évaluateurs à répondre aux questions serait aussi utile (ceci est traité dans le module WP 1 du projet PRATIQUE).

Совершенствование методов анализа фитосанитарного риска. Обзор лучших способов обеспечения последовательности в решениях

Международные, региональные и национальные стандарты по оценке фитосанитарного риска и примеры их использования в различных странах были исследованы совместно с аналогичной документацией, полученной из смежных областей, таких как здоровье животных, охрана природы и оценка генетически модифицированных организмов (ГМО), с тем, чтобы определить, каким образом обеспечивается последовательность оценки риска или его компонентов внутри этих оценок и между ними. Был определен целый спектр подходов, которые могут быть приняты и приспособлены к использованию в пересмотренной схеме поддержки принятия решений в отношении карантинных вредных организмов Европейской и Средиземноморской организации по карантину и защите растений (ЕОКЗР), позволяющий содействовать обеспечению этой последовательности. Ни одна схема не содержит в себе механизм, способный обеспечивать или гарантировать последовательность решений, и ни одна схема не содержит в себе все выявленные подходы, обеспечивающие максимальную последовательность. Если подход с использованием пятибалльной шкалы для описания индивидуальных компонентов, который способствуют оценке общего фитосанитарного риска, должен быть сохранен в схеме ЕОКЗР, главными потребностями этой схемы для увеличения последовательности являются: (i) предоставление примеров, которые описывают подразделения внутри шкалы или альтернативные, но эквивалентные дескрипторы, которые позволяли бы экспертам лучше различать подразделения, и (ii) механизм, позволяющий сочетать элементы риска постоянным и прозрачным образом. Факторы, которые помогли бы неопытным оценщикам, включают: ясную структуру, четкие указания по рейтингу, однозначно поставленные вопросы, обеспечение стандартизированных ответов, а также легко применимый метод, позволяющий интерпретировать и суммировать итоги по оценкам риска. Кроме совершенствования схемы ЕОКЗР оценщики, использующие эту схему, будут нуждаться в техническом обучении. Было бы полезно также обеспечить связь с данными и предложить возможные источники данных, которые помогли бы оценщикам отвечать на вопросы (эти проблемы были подняты на заседании WP1 PRATIQUE).

References

Anon (2009) Likert scales. http://www.statistics.com/resources/glossary/l/likertsc.php [Accessed on 7 February 2009].

Baker R & MacLeod A (2005) Pest risk assessments: tools, resources and key challenges. In: IPPC Secretariat *Identification of risks and management of invasive alien species using the IPPC framework*, Proceedings of the workshop on alien species and the International Plant Protection Convention, Braunschweig, Germany, 22-26 September 2003, pp. 106–109. FAO, Rome (IT), xii + 301 pp.

Baker R, Cannon R, Bartlett P & Barker I (2005) Novel strategies for assessing and managing the risks posed by invasive alien species to global crop production and biodiversity. *Annals of Applied Biology* **146**, 177–191.

Baker RHA, Black R, Copp GH, Haysom KA, Hulme PE, Thomas MB, Brown A, Brown M, Cannon RJC, Ellis J, Ellis M, Ferris R, Glaves P, Gozlan RE, Holt J, Howe L, Knight JD, MacLeod A, Moore NP, Mumford JD, Murphy ST, Parrott D, Sansford CE, Smith GC, St-Hilaire S & Ward NL (2008) The UK risk assessment scheme for all non-native species. *Neobiota Series* ISSN: 1681-5947.

Baker R, Battisti A, Bremmer J, Kenis M, Mumford J, Petter F, Schrader G, Bacher S, De Barro P, Hulme PE, Karadjova O, Lansink AO, Pruvost O, Pyšek P, Roques A, Baranchikov Y & Sun J-H (2009) PRATIQUE: a research project to enhance pest risk analysis techniques in the European Union. *Bulletin OEPP/EPPO Bulletin* **39**, 87–93.

Beyth-Marom R (1982) How probable is probable? A numeric translation of verbal probability expressions. *Journal of Forecasting* **1**, 257–269.

Budescu DV & Wallsten TS (1985) Consistency in interpretation of probabilsitic phrases. *Organizational Behaviour and Human Decision Processes* **36**, 391–405.

Daehler CC & Carino DA (2000) Predicting invasive plants: prospects for a general screening system based on current regional models. *Biological Invasions* **2**, 93–102.

Daehler CC, Denslow JS, Ansari S & Kuo H (2004) A risk assessment system for screening out invasive pest plants from Hawaii and other Pacific islands. *Conservation Biology*, **18**, 360–368.

Dalkey N & Helmer O (1963) An experimental application of the Delphi method to the use of experts. *Management Science* **9**, 458–467.

EFSA (2006) European Food Safety Authority (EFSA), Guidance Document of the Scientific Panel on Genetically Modified Organisms for the Risk Assessment of Genetically Modified Plants and derived Food and Feed. *The EFSA Journa*l 2004–99, pp. 1–94.

EFSA (2008) *Pest Risk Assessment: Science in Support of Phytosanitary Decision-making in the European Community*. Summary report of an EFSA scientific colloquium, No. **10**, 06-07 December 2007. EFSA, Parma (IT).

FAO (2002) *Guide to the International Plant Protection Convention*, 20 pp. FAO, Rome (IT).

FAO (2008) *International Standards for Phytosanitary Measures 1 to 31* (2008 edn). FAO, Rome (IT).

Gordon DR, Onderdonk DA, Fox AM & Stocker RK (2008a) Consistent accuracy of the Australian weed risk assessment system across varied geographies. *Diversity and Distributions* **14**, 234–242.

Gordon DR, Onderdonk DA, Fox AM, Stocker RK & Gantz C (2008b) Predicting invasive plants in Florida using the Australian Weed Risk Assessment. *Invasive Plant Science and Management* **1**, 178–195.

Ilbery B, Maye D, Kneafsey M, Jenkins T & Walkley C (2004) Forecasting food supply chain developments in lagging rural regions: evidence from the UK. *Journal of Rural Studies* **20**, 331–344.

IPPC (1997) International Plant Protection Convention, new revised text. FAO, Rome (IT). http://www.ippc.int [Accessed on 1 February 2010].

IPCC (2005) Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties. Intergovernmental Panel on Climate Change. http://ipcc-g1.ucar.edu/wg1/Report/AR4_UncertaintyGuidanceNote.pdf [Accessed on 1 February 2010].

Jablonowski M (1994) Communicating risk. Words or numbers? *Risk Management* **41**, 47–50.

Kato H, Hata K, Yamamoto H & Yoshioka T (2006) Effectiveness of the weed risk assessment system for the Bonin Islands. In: *Assessment and Control of Biological Invasion Risk* (ed. Koike F, Clout MN, Kawamichi M, De Poorter M & Iwatsuki K), pp. 65–72. Shoukadoh Book Sellers, Kyoto, Japan and IUCN, Gland (CH).

Krivánek M & PyÍek P (2006) Predicting invasions by woody species in a temperate zone: a test of three risk assessment schemes in the Czech Republic (Central Europe). *Diversity and Distributions* **12**, 319–327.

MacLeod A & Baker RHA (2003) The EPPO pest risk assessment scheme: assigning descriptions to scores for entry and establishment questions. *Bulletin OEPP/EPPO Bulletin* **33**, 313–320.

OEPP/EPPO (1997) Pest risk assessment scheme. *Bulletin OEPP/EPPO Bulletin* **27**, 281–305.

OEPP/EPPO (2007) Guidelines on Pest Risk Analysis: Decision-support scheme for quarantine pests EPPO Standard PM 5/3(3). EPPO, Paris (FR). http://archives.eppo.org/EPPOStandards/PM5_PRA/PRA_scheme_2007.doc [Accessed on 1 February 2010].

Pheloung P (1995) Determining the weed potential of new plant introductions to Australia. 26 pp. Agriculture Protection Board Report, West Australian Department of Agriculture, Perth (AU).

Rowe G & Wright G (1999) The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting* **15**, 353–375.

Theil M (2002) The role of translations of verbal into numerical probability expressions in risk management: a meta-analysis. *Journal of Risk Research* **5**, 177–186.

USDA (2000) Guidelines for pathway-initiated pest risk assessments. US Dept. of Agriculture, Animal and Plant Health Inspection Service, version 5.02, October 2000.

Williamson M (1999) Invasions. *Ecography* **22**, 5–12.

WTO (1995) Agreement on the application of sanitary and phytosanitary measures, (the SPS Agreement). In: *The Results of the Uruguay Round of Multilateral Trade Negotiations: The Legal Texts*. WTO Secretariat, Geneva (CH).

Appendix 1 – Template review on consistency of risk analysis/assessment schemes and other guidance for performing risk analysis

Version 27 June 2008

This template is intended to identify best practice worldwide for addressing consistency in pest risk analysis. Results will be used to enhance the EPPO decision support scheme in this regard.

Name of PRATIQUE scientist
1. Step: information on scheme/guidance
1.1 Name of scheme/guidance 1.2 Sector 1.3 Reference/link/attached documentation 1.4 Authority or scientific body using this scheme/guidance 1.5 Type of risk assessed

2. Step: analysis of schemes and guidance

Clarity of questions

2.1 Are questions posed clearly or are guidance or examples given to understand the question?

Rating method

2.2 How is the risk rating done: numerical or verbal?
2.3 How many choices of answers are available?
2.4 Are examples or descriptions given of what each score represents?
2.5 If not – is other assistance given for risk rating?
2.6 Is the risk assessor asked to comment/explain his rating, including the outlining and interpretation of available research and data, or the lack thereof?

Uncertainty

2.7 Is the degree of uncertainty discussed?
2.8 Are individual ratings combined with a judgement on uncertainty?

Compliance with standardization

2.9 Is a standardized terminology used in all parts of the risk assessment?
2.10 Which methods are used to standardize answers?
2.11 If it is a plant health risk assessment: does the scheme/guidance comply with ISPMs 2 and 11?
2.12 If it is a plant health risk assessment: does the scheme cover all aspects of ISPM 11?
2.13 If it is not a plant health risk assessment: is it based on a standard?

Consistency

2.14 Is the assessor advised to aim at consistency?
2.15 If yes, is this only generally mentioned or is concrete assistance given on how to reach consistency?
2.16 How are risk ratings interpreted at the end of the assessment? By expert judgement? With a statistical approach?
2.17 If a statistical approach is used, is it explained?
2.18 Has the scheme been tested for sensitivity and robustness, to ensure it can give a full range of outcomes consistent with reality?
2.19 What is your conclusion about the consistency approach in the scheme?
2.20 Is the process subject to peer review?

3. Step: interpretation of results

After finalizing step 1 and 2 for all schemes/guidelines available, a synthesis of the results will be done in step 3.

3.1 Which methods are applied to achieve consistency?
3.2 What are the pros and cons of the different methods applied?
3.3 Which methods allow for arriving at a logical, scientifically based and transparent decision?
3.4 Which methods are most effective (some may be more effective than others in different circumstances)?
3.5 Are the methods applicable or can they be modified for use within the EPPO scheme?
3.6 What kind of expertise does the risk analyst need for application of these other methods – expertise on the risk topic only or in addition e.g. statistical knowledge?

Review procedures

- Gathering of various schemes (to be referenced);
- analysis/review of consistency through different schemes;
- interpretation of results (as detailed in template);
- incorporation into EPPO scheme (if appropriate);
  (i) identification of questions which need to be rephrased or for which more guidance or examples are needed (here, we can circulate a questionnaire to our experts involved in the EWG, and they could explain which questions they had problems with).