SESSION 6B

ADVANCES IN RISK ASSESSMENT

Chairman & Session Organiser: Dr Helen M Thompson, *Central Science Laboratory, York, UK*

Platform Papers: 6B-1 to 6B-5

GM Nation? - Lessons learnt from the public debate

R S Hails
Centre for Ecology and Hydrology - Oxford, Mansfield Road, Oxford, OX1 3SR, UK
Email: rha@ceh.ac.uk

P J Dale
Emeritus Fellow, John Innes Centre, Colney, Norwich, NR4 7UH, UK

ABSTRACT

Over the summer of 2003, a public debate on the commercialisation of GM crops took place in the UK. This was the first large scale exercise to involve the public in a deliberative process about matters of science and policy, and represented a consequence of the government's emphasis on stakeholder involvement. Subsequent analysis of the debate process has concluded that it was flawed in a number of respects, an inevitable consequence, perhaps, of the novelty and scale of the exercise. Here, we outline the different components of the debate, and discuss three elements which have proved controversial: 1) provision of information to inform the debate; 2) representativeness of participants; 3) presentation of the debates results. We conclude that experts and those involved in the media should work together to play a key role in societies future debates.

INTRODUCTION

The recommendation for a national debate came from the Agriculture and Environment Biotechnology Commission (AEBC), and arose from an investigation into the controversy surrounding the farm-scale evaluations (FSEs) (AEBC, 2001). This series of experiments investigated the potential impact of growing four GM herbicide tolerant (GMHT) crops on farmland biodiversity, compared to the equivalent conventional crops. The GM crops involved in the experiments had passed through the regulatory system, and so had already been judged to pose no harm to human health or directly to the environment. The FSEs were designed to compare potential indirect impacts on farmland biodiversity due to the associated herbicide regimes, so contributing one of the final pieces of evidence on the potential ecological consequences of commercialising these crops. However, the experiments themselves became something of a focus for general concerns surrounding the technology. This highlighted the lack of a framework within which members of the public could debate the issues surrounding GM, and it was in this context that the AEBC recommended a debate. This resonated well with government, being in line with their desire to engage more effectively with stakeholders. Consequently, it was announced in the autumn of 2002 that there was to be a national dialogue on genetic modification.

THE DESIGN OF THE DEBATE

The debate consisted of multiple components summarised in Figure 1 and outlined below.
Three interacting strands. When the intention to hold a public debate was announced, the government also initiated two other studies: one investigating the potential economic costs and

benefits of GM technology, and another a science review covering the current state of knowledge on GM risk assessment, highlighting gaps and uncertainties (UK government response to AEBC advice July 2002). The intention was that these three strands would interact over the course of the following year, with the science and economics reviews informing the public debate.

Management. The debate was funded by UK government with a budget of £500,000, but managed by an independent steering board. The board consisted of seven members of the AEBC and four others.

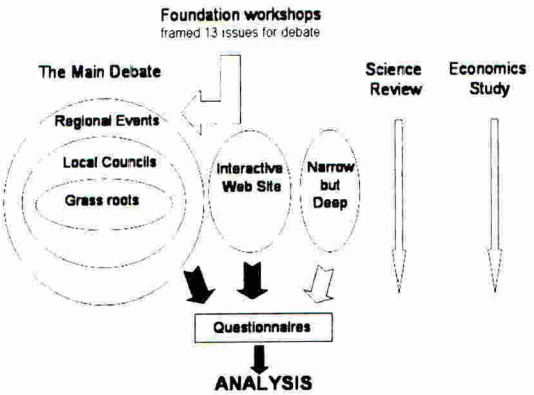

Figure 1: The construction of the debate.

Foundation discussion workshops. A series of workshops involving a representative cross section of the public was held in November 2002, to allow the public themselves to frame the issues for debate. These individuals were selected in such a way that anyone who had already been actively involved in GM issues was excluded from eight of the nine workshops, with one specifically involving those who already had an active interest in GM issues.

Stimulus material. A variety of outputs were produced to inform and stimulate the debate, including a film, a CD-ROM, printed workbooks and an interactive web site.

The main debate. Three tiers of public discussion meetings were held, starting with six regional meetings (three in England and one in each of Scotland, Wales and Northern Ireland) resourced and facilitated by the main contractors. This led to a cascade of 'second tier' (around 40) meetings typically at the level of local county councils, followed by local third tier (around 600), grass roots events, usually staged by a local organisation. Those that participated in these meetings were self-selected.

Narrow but Deep focus groups. In parallel with the main debate, a relatively small group of people (78) were selected to be a representative cross-section of the public. These people participated in two discussion groups, with a gap of two weeks in between which allowed them to think over the issues and collect their own evidence. The results of these focus groups were to be compared with the outputs from the main debate, to detect if the main debate was dominated by people with a particular perspective.

Questionnaires. Thirteen issues were raised at the foundation workshops, and these were used to design a questionnaire. This was answered by participants in the main debate (either sent in by mail or via the *GM Nation?* website, amounting to 36,000 in total) and the Narrow But Deep (NBD) focus groups – in the latter case, both before and after the two weeks of deliberation. These provided the raw data for much of the analysis.

The outcome. A final report was drafted by a professional writer, edited by the steering board and presented to government in September 2003 (Public Debate Steering Board , PDSB 2003)
Independent evaluation. An independent team of researchers from the University of East Anglia, Cardiff and Brunel Universities and the Institute of Food Research, all part of the research initiative 'Understanding Risk', undertook the evaluation. This was funded by the Economic and Social Research Council (Horlick-Jones *et al.,* 2004).

We now focus on four aspects of the debate's construction and execution, to discuss some of the lessons we may draw from this exercise.

SETTING THE OBJECTIVES

The steering board laid out nine objectives for the debate [see Box 1 – (PDSB 2003)].

Box 1: Objectives for *GM Nation?* set by the debate steering board

The public debate will aim to:

1. allow the public to frame the issues for debate so that the programme of debate focuses on what the public sees as the relevant issues;
2. focus on getting people at the grass roots level whose voice has not yet been heard to participate in the programme;
3. create new and effective opportunities for deliberative participation about the issues;
4. enable (through dialogue with experts and other activities) access to the evidence and other balanced and substantiated information the public may want and need to debate the issues;
5. create widespread awareness among the UK population of the programme of debate, even if people do not wish to participate directly in events; and give widespread opportunities to register views;
6. provide occasions within the programme of debate for interactions between members of the public in debate, and mutual learning between the public and experts;
7. seek to complement and inform the economic and science strands and in turn, as appropriate, utilise their outputs;
8. calibrate the views of organisations who have already made their views known by contrasting their views with other participants in the debate;
9. provide intelligent, qualitative information about public views emerging from the debate in a report to government by end June 2003.

Components of the debate were designed to meet specific objectives: for example, the foundation workshops were clearly designed to allow the public to 'frame the issues' (objective 1). The website provided 'widespread opportunities to register views' (objective 5). In the following sections we focus on the quality of the information provided in the debate (objectives 4 & 6); drawing in people who had not previously engaged in these issues (objective 2); and finally objective 9, which emphasises that the information emerging will be qualitative.

INFORMING THE DEBATE

The foundation workshops, organised and run by a contractor (CorrWillbourn Research and Development), were designed to elicit from the public the terms and frame of reference in which they wished to discuss GM issues. One general conclusion from the workshops is that the public knows that it lacks facts and information about GM, and there is a desire to learn more, preferably through documentary style programmes on the television. This is hardly surprising given the popularity of the high quality science documentaries often produced by television companies. Unfortunately this was beyond the budget and time frame of the debate.

Examining the report on the foundation workshops (CorrWillbourn, 2003) reveals that much emphasis is placed on statements such as 'more information alone will not be enough', with the social and ethical implications also being important. This is undoubtedly true, and this underlies the shift in the culture of science communication that has taken place in the last decade. Until recently, science research councils referred to the 'Public Understanding of Science', which implies a one way flow of information from the expert to the public. It is now widely recognised that this flow of information should be a two-way engagement – with the public participating in a debate on the social and ethical implications – moving from the 'deficit model' to the 'dialogue and debate' model of science communication (House of Commons, 2000).

We suggest the new dialogue and debate model should go one step further, so that the information flow is a full two-way process. Science gives birth to social and ethical implications, and it is the latter which form the basis of much of the debate. Rarely does the debate return to discuss the science itself, as if only the implications, and not the content, of the scientific aspects are of concern. Just as scientists need to learn more about societies responses to new technology, so would the public benefit from learning the language of scientific discourse, and a better understanding of the scientific process. Science can provide facts, which act as the starting point for debate, without displacing the social and ethical considerations. However, there is a perception, even in the House of Commons, that there is considerable disagreement between scientists about the safety and environmental impact of GM crops (House of Commons 2002). This should now have been clarified by the Science Review, one of the three parallel strands (GM Science Review First Report, 2003), which collated the current state of scientific knowledge on the safety of GM food and crops, highlighting any remaining areas of uncertainty. Even though this was not available to act as stimulus material for the debate, the source material (principally the peer reviewed literature) was. However, for the sake of equality and inclusivity of different viewpoints, a different approach was adopted.

Seven areas emerged from the foundation workshops, and for each of those areas between four and thirteen questions were posed. These questions defined the information that the members of the public required in order to debate and deliberate over the issues. The difficulty then comes in how to answer these questions – many of which would have been very similar if the topic of concern was nanotechnology; for example, trust in governments, profit motive of industry, and the vested interests of scientists. It fell to the steering board to provide balanced answers, which they inevitably could not agree on – so two answers would be provided under the headings 'views for' and 'views against', perpetuating the polarised nature of the debate. One example is used to illustrate the difficulties of this process.
One question was '*Is GM a natural process?*' The answer to this question for the 'views against' begins 'It is unnatural to bring together genes from quite different species that would

never mix'. It has been proposed that this view originates from a Platonic view of the world, which has been an established way of thinking for more than 2000 years until the arrival of Darwin in the mid-nineteenth century (Davies, 2001). The Platonic view of species is as eternal, ideal forms, of which individuals are imperfect approximations. With this philosophy, the transfer of genes between species would seem unnatural and objectionable. Empirical biologists, in contrast, would see all individuals as unique, with species acting as approximate groupings with similar characteristics. Species are therefore on a phylogenetic continuum, linked by evolutionary history to other species. Events such as hybridisation and horizontal gene transfer cause gene flow between closely and distantly related species, sometimes providing the raw material for new species to be formed. Thus the answer to this question touches upon two quite different schools of thought that would be almost impossible to explore and summarize in a paragraph. This is an example where the debate would benefit from an exploration of the language of scientific discourse. True deliberation requires time to access and consider different viewpoints.

The production of stimulus material will always be problematic in controversial areas. In this case, the scientists involved found this a very unsatisfactory process. How is it possible to inform a debate, integrating all viewpoints, yet give the appropriate weight to the results of several years of scientific research compared to relatively unsubstantiated opinion? It is important that in moving from the deficit model to the dialogue and debate model of science and communication, the quality of information flow from scientists to the public is not sacrificed. The provision of good quality information is an absolute requirement for a truly deliberative process.

PARTICIPATION AND REPRESENTATION

One of the main criticisms of the debate has revolved around the extent to which the open debate was representative of UK public opinion as a whole. The public discussion meetings were populated by self-selecting participants, whereas the "Narrow But Deep" focus groups were made up of a stratified random sample of participants, chosen to be representative of the UK population. Yet the open meetings that involved most people (around 36,000 as opposed to 78) consumed most of the resources, and figured most prominently in the outputs. How did these two groups compare?

Campbell and Townsend drew direct comparisons between the two groups in their responses to specific questions (Campbell *et al.,* 2003). One example is the response to the question 'I would be happy to eat GM food'. Only 8% of the open debate respondents agreed compared to 35% of those in the NBD focus groups. Campbell and Townsend argued that the gap between the two sets of responses calls into question the legitimacy of the debate. However, the sample size for the focus groups is so small as to make quantitative comparisons difficult. To overcome this, a statistically representative sample of the British public (1,363) was asked the same questions after the close of the debate by the independent evaluators (Pidgeon *et al.,* 2005). An analysis of their responses can be used to divide participants into one of four attitudes: a) indifferent when perceived risks and benefits are low; b) positive if perceived risks are low but benefits high; c) negative if perceived risks are high and benefits low; d) ambivalent if perceived risks and benefits are high. Their analysis illustrated a very clear mismatch between the debate participants and the general public. Specifically, participants of the open debate were overwhelmingly negative, consistent in their perception of the high risks

involved with the technology, and consistent in their rejection of any perceived benefits (Pidgeon *et al.*, 2005). The general public were more ambivalent – accepting that there may be risks, but there may also be benefits – with only a small proportion in the negative category. The open meetings had therefore attracted a very distinct subset of the public.

This does not completely devalue the main debate, as the views of the self-selecting participants are the views of an engaged sector of society. However, it does illustrate problems with adopting an open meeting approach, as it has clearly distorted the kind of data collected.

THE MEDIA AND THE DEBATE

The House of Lords heard evidence that once they had left school, 74% of people obtain their information about scientific issues from national newspapers (House of Lords, 2000). Newspapers have certainly been influential in setting the context for the debate, with a few of them running an anti-GM campaign in 1999 (POST, 2000). They interacted with *GM Nation?* at two points. The first of these was during the two-week period when the NBD participants were deliberating between discussion meetings. This coincided with a cabinet reshuffle, which displaced Michael Meacher as environment minister, and during the following week, there was extensive newspaper and television coverage of his criticisms of the government over GM foods. Newspaper articles alluded to 'toxins that could damage embryos in the womb' being used in conjunction with GM crops, and quoted Mr Meacher as saying 'it was really extraordinary that there had so far been virtually no independent studies of the health effects of GM' (2003). It is, of course, impossible to assess the extent to which this coverage influenced the participants, or to control for such events occurring during the course of a debate. However, in the 'evidence' collected by the NBD participants in the course of their two-week research period, material involving Mr Meacher constituted a significant proportion of the total material assembled (Dale *pers. comm.*). The attitudes of the NBD participants 'hardened with time' over this period in their response to the 13 questions (PDSB, 2003). In other words, the number of 'don't knows' declined significantly, and while there was some movement to agree with positive benefits (e.g. cheaper food, medical benefits and help for developing countries), there was more movement to agree with potential risks (e.g. GM will have negative impacts on the environment, we don't know enough about health, and GM represents unacceptable interference with nature).

The second point of interaction with the media, and newspapers in particular, was when the final report was published in September 2003. Throughout the management of the debate, one of the justifications for the methodology used was that it was not designed to be a referendum on GM crops, but a qualitative exploration of people's attitudes. Indeed in the context of EU legislation, a referendum would be inappropriate. However, perhaps inevitably, the results were summarised in a quantitative manner in the final report – for example key messages included 'People are generally uneasy about GM' and 'There is little support for early commercialisation'. This quantitative flavour was echoed and exaggerated by the newspaper headlines, and this will have become the main take home message that many of the public will have received. In this light, the issue of the unrepresentative nature of those participating in the main debate becomes much more important.

CONCLUSIONS: LESSONS LEARNT FROM *GM NATION?*

This was the first attempt to conduct a public debate on a contentious and technical issue. One of the very positive aspects of the design was the multi-strand approach. This three-pronged approach, including the science and economics reviews, was a potential strength, although in the event it may not have been fully exploited.

It was recognised by the board that quality information is required for meaningful deliberation. This is arguably where the greatest problem lay. The time-frame within which the debate was conducted, with the main debate occupying just six weeks in the summer of 2003, restricted the preparation of the stimulus material and the extent to which the science and economic strands could contribute to this process. But as we have noted, there is a much greater underlying problem. The dialogue and debate model has not yet been developed into a fully two way process. Scientific advances lead to social and ethical implications which fuel the debate – but the science itself is largely absent. Methods to draw the public in to the language of scientific discourse and a better understanding of the scientific process are required.

Objectives included reaching people who have not yet participated in the debate, and providing new and effective opportunities for deliberative debate. *GM Nation?* had limited success on both of these counts, as the evidence suggests that those who attended the open meetings were those who had already engaged in the issues, with very few altering their opinion over the course of the debate (Pidgeon *et al.*, 2005). The shortcomings of the stimulus material also reduced the opportunity for real deliberation. A number of methods have been developed over recent years to engage with the public on matters of policy, including focus groups, deliberative polling, stakeholder dialogues, citizens' juries and consensus conferences (POST, 2001). Each method is more suited to a particular objective. Given the stated aims, the emphasis on open meetings with self-selected participants and relatively little opportunity to interact with experts was inappropriate. It is only by the application of appropriate deliberative methods that individuals have the opportunity to become better informed and to engage in deliberation. The disadvantage is that these methods are relatively expensive, and so limit the number of people who can be involved. In conclusion, the original motivation behind *GM Nation?* was to engender a debate amongst the British public. In the event, public opinion was not respected, as what emerged was not an accurate flavour of what 'the many' think. Thus public opinion was both the *raison d'être* and the casualty. This suggests that attempting to interpret the data qualitatively rather than taking a more quantitative approach was not appropriate in this case.

The press often relies on emotive headlines and campaigning in order to sell newspapers, so communication through the press in a balanced and informed way is often difficult. However, proactively and positively engaging with certain sections of the media could bring significant benefits to such engagement exercises. The participants of the foundation discussion workshops were keen to learn more about the issues, particularly through television. This was perhaps a missed opportunity, but beyond the budget and timescale of *GM Nation?* In the future, we should look to our television companies to play a key role in society's future debates by the provision of quality information. The challenge is for scientists to fully collaborate and participate in this process, and perhaps for research councils to channel significant resources to this purpose.

REFERENCES

AEBC (2001). *Crops on Trial.* Agriculture and Environment Biotechnology Commission: London.

Campbell S; Townsend E (2003). "Flaws undermine results of UK biotech debate." *Nature* **425**, 559.

CorrWillbourn (2003). *A report on the foundation discussion workshops conducted to inform the GM public debate.*

Davies K G (2001). "What makes genetically modified organisms so distasteful?" *Trends in Biotechnology* **19**, 424-427.

GM Science Review First Report (2003). *An open review of the science relevant to GM crops and food based on interests and concerns of the public.* UK Government: London.

Horlick-Jones T; Walls; Rowe G; Pidgeon N F; Poortinga W; O'Riordan T (2004). *A Deliberative Future? An independent evaluation of the GM Nation? public debate about the possible commercialisation of transgenic crops in Britain, 2003,* Working Paper 04-02. Centre for Environmental Risk, University of East Anglia: Norwich.

House of Commons (2002). Genetically modified organisms: fifth report of session 2001-2002. The Stationary Office Ltd.: London

House of Lords Select Committee on Science and Technology (2000). *Science and Society: 3rd report of session 1999-2000.*

Parliamentary Office of Science and Technology (2000). *The great GM food debate: a survey of media food coverage in the first half of 1999.* Report no. 138.

Parliamentary Office of Science and Technology (2001). *Open channels: public dialogue in science and technology.* Report 153.

Pidgeon N F; Poortinga W; Rowe G; Horlick-Jones T; Walls J; O'Riordan T (2005). "Using surveys in public participation processes for risk decision making: The case of the 2003 British GM Nation?" *Risk Analysis* **25**(2), 467-479.

PDSB (2003). GM Nation? *The Findings of the Public Debate.* (Report by the Public Debate Steering Board). Department of Trade and Industry: London.

Probabilistic approaches to assessing ecological risks of pesticides

A Hart
Central Science Laboratory, Sand Hutton, York, YO41 1LZ, UK
Email: a.hart@csl.gov.uk

ABSTRACT

Directive 91/414/EEC requires that if a plant protection product fails to pass preliminary, first-tier assessment criteria for environmental risk, then it may not be authorised for use unless an "appropriate risk assessment" shows that it will cause no unacceptable impact. One option for a refined, higher-tier risk assessment is to use probabilistic approaches. The defining feature of probabilistic risk assessments is that they quantify variability and/or uncertainty. Potential benefits of quantifying variability include increased realism, representing real-world variation in factors that influence risk, the opportunity to replace or refine worst-case assumptions and they provide an alternative to conducting higher-tier laboratory or field studies by making more use of the available data.

INTRODUCTION

Directive 91/414/EEC requires that if a plant protection product fails to pass preliminary, first-tier assessment criteria for environmental risk, then it may not be authorised for use unless an "appropriate risk assessment" shows that it will cause no unacceptable impact. 'Plant protection product' is the formal term for pesticides, safeners and plant growth regulators that are within the scope of Directive 91/414/EEC. For simplicity, the word "pesticide" is used in this document to cover all types of plant protection product. Various options for these refined, higher-tier risk assessments are identified in existing EU Guidance Documents, including probabilistic approaches. Until now, however, probabilistic approaches have gained only limited acceptance, partly due to a lack of guidance on how to implement and evaluate them.

The objective of EUFRAM is to provide a framework of basic concepts, principles and methods that will help users to conduct, report, evaluate and communicate probabilistic assessments in appropriate ways. It is aimed primarily at risk assessors in government, industry and consultancy companies. EUFRAM aims to assist the implementation of probabilistic methods for assessing the environmental risks of plant protection products in Europe. The main outputs are being developed as a report in 3 volumes: (1) a framework of basic principles and methods for probabilistic assessment, together with summaries of selected examples; (2) eight chapters providing more detail on selected aspects of probabilistic approaches; and (3) detailed case studies developed or evaluated during the project.

The framework is being developed by the EUFRAM project, an EU-funded concerted action involving 29 organisations including regulatory authorities, government research institutes, agro-chemical companies, consultancy companies and universities. It will be refined and revised to account for feedback from all interested parties, including two workshops scheduled for October 2005 and July 2006. We invite people to visit the website (www.eufram.com) to provide feedback on the framework report and/or to get involved in refining the document.

DETERMINISTIC METHODS

Deterministic methods are defined as:

> ***Methods that use point estimates to represent one or more factors in a risk assessment and treat them as if they were fixed and precisely known.***

Point estimates represent a measured or estimated quantity by a single number, e.g. the minimum, mean or maximum value, rather than a distribution. Most risk assessments include at least some deterministic elements, but few are wholly deterministic. This is because many of the point estimates that are used for exposure and/or toxicity have in fact been derived from probabilistic calculations. Of the first tier assessments for ecological risk under Directive 91/414/EEC, perhaps only the approach for soil micro-organisms is entirely deterministic. Thus almost all assessments under 91/414/EEC already incorporate at least some probabilistic elements. On the other hand, it is practically impossible to quantify absolutely every source of variability and uncertainty affecting an assessment. In fact, most assessments are neither fully deterministic nor fully probabilistic, but somewhere in between (Figure 1). So the question is not whether to start doing probabilistic assessments, but whether it may be helpful to include more probabilistic elements than we already do and – if so – when and how to do it.

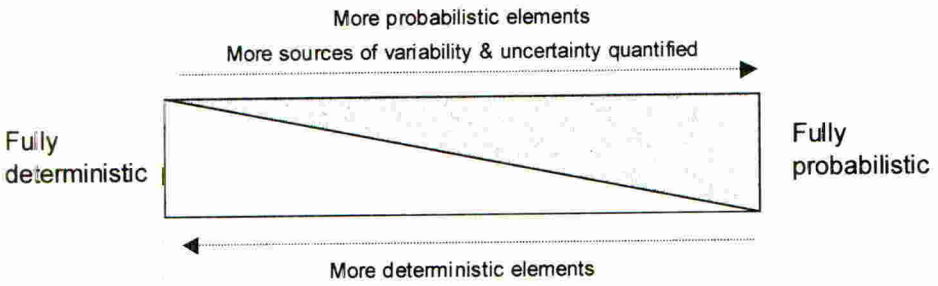

Figure 1. Continuum between wholly deterministic and wholly probabilistic risk assessments

PROBABILISTIC RISK ASSESSMENT

EUFRAM defines probabilistic risk assessment as:

> ***Risk assessments that use probabilities or probability distributions to quantify one or more sources of variability and/or uncertainty in exposure and/or effects and the resulting risk.***

This definition itself contains several terms that require definition. Variability is defined as *Real variation in factors that influence risk.* For example, toxicity varies between species, and exposure varies in time and space. Variability matters because risk assessment usually needs to address a range of relevant species and exposures, not just one particular species and one exposure. Uncertainty is defined as: *Limitations in knowledge about factors that influence risk.* For example, there is uncertainty when we extrapolate toxicity from a small number of tested species to other, untested species, and uncertainty when we extrapolate from mathematical models of exposure to the real world. Uncertainty matters because decision-makers and stakeholders need to know how different the real impacts might be from the scientists' best estimates. An important practical difference between uncertainty and variability is that:

uncertainty can often be reduced by obtaining further data, whereas variability can be better quantified but not reduced by further data.

Probabilities can be used to quantify variability and/or uncertainty. The probability of something can be defined as its frequency in repeated independent trials, or as the degree of belief that it will occur, expressed as a proportion (i.e. a number between 0 and 1). For example, when tossing an unbiased coin, it is expected to land on "heads" on 50% of occasions, so the probability of obtaining "heads" is 0.5. Similarly, if you know or believe that half the people in a population are taller than 1.8m, then you can express the probability of heights above this value as 0.5.

The most familiar form for a probability distribution is a graph showing the relative probabilities of different values of a variable such as height. For example, Figure 2 shows the variation in height of participants at the first EUFRAM workshop in two ways: as a histogram showing the raw frequencies (the number of participants in particular height ranges), and as a curve showing a "Normal" distribution fitted to the raw data. This curve is a "probability density function" (PDF) and shows the relative frequency of occurrence for each point on the x-axis. In principle, distributions can be used to quantify variability and uncertainty for any number of inputs to a risk assessment, and for any output of the assessment that may be of interest to the decision-maker.

The defining feature of probabilistic risk assessments is that they quantify variability and/or uncertainty. Potential benefits of quantifying variability include:

- Increased realism, representing real-world variation in factors that influence risk
- The opportunity to replace or refine worst-case assumptions
- Provides an alternative to conducting higher-tier laboratory or field studies
- Makes more use of the available data.

Potential benefits of quantifying uncertainty include:

- Provides an objective basis for discussions about reducing uncertainty factors (e.g. the TER thresholds of 10 and 100) when additional data is provided
- Indicates the influence of quantified uncertainties on the assessment outcome
- May help increase the cost-effectiveness of higher-tier studies by targeting them on major sources of uncertainty.

COMPATIBILITY WITH EXISTING LEGISLATION AND GUIDANCE

Probabilistic risk assessment is not mentioned in Directive 91/414/EEC or its Annexes. However, Annex VI states that a plant protection product that fails the preliminary, first-tier assessment shall not be authorised "unless it is clearly established through an appropriate risk assessment that under field conditions no unacceptable impact on the viability of exposed species occurs". This opens the way for probabilistic assessments at higher-tiers of the assessment process if they are considered "appropriate" by the responsible authorities. Furthermore, using distributions to represent variation in exposure and toxicity can be regarded as one way to take account of "field conditions" and "exposed species".

Current EU Guidance Documents for both aquatic and terrestrial ecotoxicology state that traditional deterministic assessment methods have limitations that could be overcome by probabilistic approaches. The Aquatic Guidance Document (European Commission, 2002a)

states that probabilistic risk assessment is usually a tool for higher-tier assessments and hence its suitability needs to be considered case-by-case. The Terrestrial Guidance Document (European Commission, 2002b) states that probabilistic methods are promising tools and already now there may be situations where their use could be envisaged.

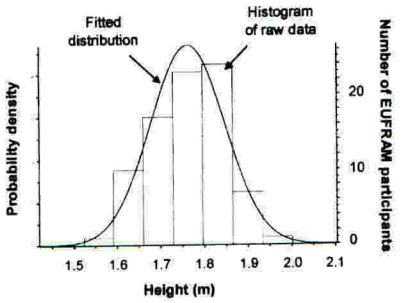

Figure 2. An example of a probability distribution: variability in the height of participants at the first EUFRAM workshop, in March 2005

CONDUCTING A PROBABILISTIC RISK ASSESSMENT

The main steps of a probabilistic assessment are as follows:

1. **Define the assessment objectives**. The objectives should reflect as closely as possible the information needs and protection goals of decision-makers. They should define the pesticide use, non-target organisms and types of effects to be considered, what types of variation are of interest, and whether confidence intervals are required with the outputs.
2. **Define the assessment endpoint**, that is, the primary output of the probabilistic assessment. In particular, define what measure to use for the magnitude of hazard, exposure or risk, for what ensemble or group of entities (e.g. species or community) it should be estimated. It is essential to ensure that these choices provide an output that will be meaningful and relevant to decision-makers.
3. **Identify the key factors and mechanisms** that influence the effect that is to be assessed, and how they interact, and develop an assessment model to represent them.
4. Consider each part of the model in turn, and **decide which factors and mechanisms might contribute significantly to variability** in the assessment endpoint. These will be represented by distributions in the assessment model.
5. **For each input variable, identify the appropriate ensemble**. This depends on the ensemble of the assessment endpoint and the model structure. For example, if the output is frequency of effects in an ensemble of water bodies, then the appropriate ensemble for exposure is concentrations for different water bodies.
6. **Identify what data are available** that can help in quantifying each factor and mechanism, including distributions for those that will be treated as variables.
7. **Decide whether any extrapolations or adjustments are needed** to model the key factors and mechanisms from the available data, e.g. to account for lab-to-field extrapolation or non-random sampling. If adjustment or extrapolation factors are required, then they should be identified as part of the assessment model.
8. Consider each part of the model in turn to **identify possible sources of uncertainty**. Decide which uncertainties might contribute significantly to uncertainty in the assessment endpoint, and which of these will be quantified using distributions (if any). Decide what to

do about important uncertainties that cannot or will not be quantified (e.g. use conservative values).

9. Consider carefully for each input distribution whether it should be treated as contributing uncertainty or variability to the assessment output.
10. **Identify potential dependencies** affecting the assessment. Dependencies occur where the value of one variable depends upon the value of another variable (e.g. food intake may be positively related to body weight) and can have a major impact on the assessment outcome.
11. Express the entire assessment model as a set of mathematical equations, so that they can be used for calculations.
12. **Select appropriate computational or graphical methods** for combining the input distributions to obtain the assessment output. This choice will depend on various factors including the number of input distributions, whether a confidence interval is required for the output, the approaches that will be used to handle uncertainties and dependencies, and ease of use. More specific recommendations will be developed during the remainder of the project.
13. **Specify distributions for the input variables**. This requires expertise in statistics as well as expert knowledge of each variable and how it relates to the assessment output.
14. **Consider conducting sensitivity analyses**. These may help in various ways, e.g. in deciding which sources of variation and uncertainty should be quantified.
15. **Select appropriate software** to carry out the computations and generate outputs. EUFRAM has defined desirable characteristics for probabilistic software and databases. Existing tools meet these criteria to a limited extent and can be used for some types of pesticide assessment, but further development is highly desirable.
16. **Check the outputs of the probabilistic assessment and consider their implications for decision-making**. This will have to be done case-by-case, as for other refined assessments, unless standard criteria for decision-making are established.

OUTPUTS

Probabilistic methods can produce many different types of output. Consultations with EUFRAM end-users indicate a desire to facilitate communication by adopting a single format as standard. **EUFRAM provisionally recommends cumulative distributions (with confidence intervals if required) as a preferred choice for the primary output of probabilistic assessments.**

Cumulative distributions can be used to represent three basic dimensions of risk: magnitude, frequency and uncertainty (e.g. Figure 3). The x-axis shows the magnitude of hazard, exposure or risk (how bad), the y-axis shows the frequency (how often), and the confidence intervals show uncertainty (how sure). It can be used to read off the frequency of effects below (or above) any given magnitude, together with confidence intervals if required. Different measures of magnitude and frequency can be used, according to the needs of the assessment. In addition, quantitative results should always be accompanied with a written statement, expressing and interpreting them in language that can be understood by non-specialists.

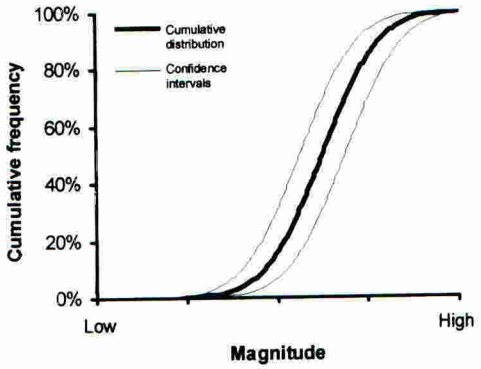

Figure 3. Generalised example of a cumulative distribution. This is provisionally recommended by EUFRAM as the preferred format for graphical output of probabilistic assessments

EUFRAM envisages that probabilistic assessments for pesticides will generate four main types of output using cumulative distributions, depending on the type of question posed by the assessment objectives:

- A CDF for hazard (effects), of which the most familiar is a species sensitivity distribution (SSD), Posthuma *et al.*, 2002), showing variation in sensitivity between species. This can be compared to a point estimate for exposure (e.g. Figure 4a).
- A CDF for exposure, showing variation of exposure, e.g. between water bodies. This can be compared to a point estimate for effects (Figure 4b).
- Two CDFs, one for exposure and one for effects, so that they can be compared with one another (Figure 4c).
- A single CDF for "risk" or impacts, which takes account of both exposure and toxicity (Figure 4d). This can be used by decision-makers to make judgements about the acceptability of the risk, or compared to standard decision criteria for risk, if these have been established.

An important limitation of probabilistic outputs is that they cannot quantify all sources of variability and uncertainty. This is because (a) there are simply too many sources of variability and uncertainty for it to be practical to quantify them all, and (b) many uncertainties can only be assessed subjectively (e.g. by expert judgement) and are difficult to quantify. Consequently there will always be some sources of variability and uncertainty that are not accounted for in the quantitative output of a probabilistic assessment. Therefore, the outputs of probabilistic assessments should include a list of unquantified sources of variability and uncertainty, and a qualitative evaluation of how they might affect the assessment outcome.

In addition, the results of probabilistic assessments should be considered together with conventional deterministic results and other lines of evidence (e.g. field studies or monitoring), to arrive at overall conclusions. This may include consideration of the wider ecological consequences of predicted impacts (e.g. extrapolation from effects on individual organisms to consequences for the wider population). The formal report of a probabilistic assessment should document and justify the methods, results and conclusions clearly and concisely, but in sufficient detail to enable critical evaluation of all stages by other specialists (e.g. peer reviewers). It should include an executive summary, communicating the main points required for decision-making. Sufficient background information for other specialists to duplicate the assessment should be provided in appendices.

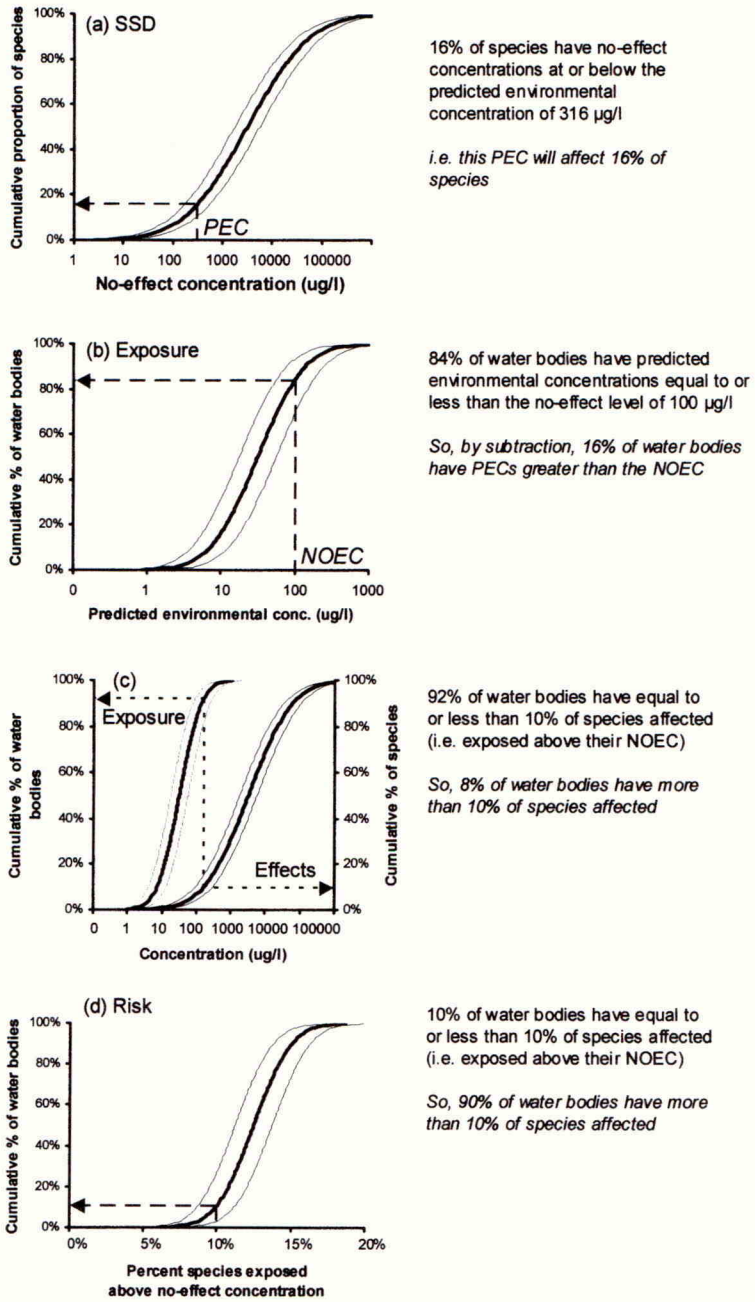

Figure 4. Four main types of graphical output proposed for probabilistic assessments of ecological risks of pesticides, together with examples of risk statements that they could be used to make. The measures of magnitude and frequency (x and y axes) will vary according to the needs of the assessment (Note: these examples are illustrations and do not relate to real data)

Validation is possible only to a limited extent for both deterministic and probabilistic approaches. Consequently, thorough peer review by relevant technical experts will be a key requirement for the acceptance of new approaches, and for the acceptance of individual probabilistic assessments.

Effective communication is essential if probabilistic approaches are to be accepted. Different approaches are required when communicating to different audiences, e.g. technical specialists, decision-makers and the public. The format of graphical, tabular and textual outputs needs careful and detailed consideration, to maximise their effectiveness.

Probabilistic approaches are still evolving and it would be premature to attempt harmonisation at the present state of the art. Nevertheless, if a probabilistic assessment has been conducted for one pesticide and accepted by the relevant authorities, then much of the approach may be directly transferable to other, similar pesticides with similar use patterns. This transferability of approaches opens the possibility of establishing generic, peer-reviewed probabilistic approaches and tools for scenarios that frequently require refined assessment under 91/414/EEC, analogous to the FOCUS models and scenarios for exposure assessment. This would increase efficiency both for people conducting probabilistic assessments, and also for people evaluating them. However, flexibility is important and it should remain open for assessors and decision-makers to select other approaches where appropriate.

CONCLUSIONS

After considering their strengths and weaknesses, EUFRAM concludes that there is scope for deploying probabilistic approaches to a greater extent, as one of several alternatives for higher-tier assessment. They can be applied either to exposure assessment, or effects assessment, or both. EUFRAM proposes a framework for probabilistic risk assessment, which is summarised in this paper. The basic steps of a probabilistic assessment are (1) define the objective of the assessment and decide what form of probability or distribution is required for the assessment output; (2) identify one or more inputs to the risk assessment, for which variability and/or uncertainty is to be considered, and quantify them using appropriate probabilities or distributions; (3) use appropriate methods to combine the different input distributions and produce the distribution for the assessment output, showing the variability and uncertainty of the predicted impacts; and (4) interpret and communicate the results.

ACKNOWLEDGEMENTS

We would like to thank the EUFRAM partners (www.eufram.com) for their invaluable contributions to the framework.

REFERENCES

European Commission (2002a). *Guidance document on aquatic ecotoxicology in the context of Directive 91/414/EEC.* Sanco/3268/2001 rev.4 (final), 17 October 2002: Brussels.

European Commission (2002b). *Guidance document on terrestrial ecotoxicology under Directive 91/414/EEC.* SANCO/10329/2002 rev 2 final, 17 October 2002: Brussels.

Posthuma L; Suter G W III; Traas T P (2002). *Species sensitivity distributions in ecotoxicology.* Lewis Publishers: Boca Raton, FL.

FOCUS Landscape and Mitigation: Mitigating risk posed by pesticides to the aquatic compartment

C D Brown
Environment Department, University of York, Heslington, York, YO10 5DD
Email: colin.brown@csl.gov.uk

M Streloke, J-J Gril, S Loutseti, R Stadler, J van de Zande, A Alix, J L Alonso-Prados, D Auteri, R Hiederer, C Holmes, A Huber, F de Jonge, M Liess, N Mackay, S Maund, C Pais, W Reinert, M Russell, T Schad, M Styczen
On behalf of the FOCUS Working Group on Landscape and Mitigation (under the auspices of the European Commission's DG SANCO)

ABSTRACT

The FOCUS Landscape and Mitigation group has reviewed potential approaches to undertake higher-tier exposure assessments for pesticides in European surface waters and possibilities for the use of mitigation measures to reduce risks to non-target aquatic organisms. This paper summarises draft recommendations from the working group with respect to the use of mitigation within European regulatory procedures for pesticides. This includes a review of current practice, suggestions for how mitigation measures could be included within risk assessment, and proposals for measures suitable to reduce risk from exposure via spray drift, surface runoff and drainflow. The views expressed are those of the working group.

INTRODUCTION

FOCUS (FOrum for the Coordination of pesticide fate models and their USe) was established under the auspices of DG SANCO to develop approaches to environmental exposure assessment issues under Directive 91/414/EEC. The aim of FOCUS is to develop guidance for notifiers and Member States concerning appropriate methods for calculating exposure concentrations for EU dossiers on plant protection products (Annex I). Over recent years, significant advances have been made in the development of exposure assessments for surface waters through the activities of FOCUS working groups on this topic, most recently with the release of the FOCUS surface water scenarios (FOCUS, 2002). Whilst these approaches have led to the development of harmonised approaches for conducting lower-tier exposure assessments, to date little guidance has been available on the topics of higher-tier exposure assessments and the implementation of mitigation measures suitable for managing risk. Consequently, a working group was established in June 2002 to review potential approaches to higher-tier surface water exposure assessments and mitigation measures. The remit of the group was to review the current state-of-the-art, recommend approaches that could be implemented forthwith, and identify requirements for further work. The group considered approaches suitable for supporting the assessment necessary for authorisation on a Community level, but also those that could be applied in risk assessments to support national registration.

The material in this paper draws on a draft report submitted to the European Food Safety Agency (EFSA) in June 2005. The paper concerns the conclusions regarding mitigation of risk posed by pesticides to the aquatic compartment. Some of the recommendations of the working

group reflect the current technical state of the art rather than what may be possible to implement in regulatory practice. Some of the issues may need to be further discussed and decided upon by policy makers at the European Community level and at Member State level. The views expressed are those of the FOCUS Landscape and Mitigation Working Group and do not necessarily reflect the views of DG SANCO, EFSA or regulatory authorities in individual Member States.

REVIEW OF CURRENT PRACTICE

A review was undertaken of risk mitigation measures currently used in EU Member States to protect aquatic life within the authorisation procedure of plant protection products. Representatives of the different Member States were surveyed and literature sources were collated. The results of the review are summarised in Table 1.

Table 1. Examples of risk mitigation measures currently used in European Member States (includes examples of post-authorisation and voluntary measures)

Member State	Spray drift			Surface runoff	Drainflow
	No-spray buffer zone	Drift-reducing techniques	Other		
Austria	Up to 50 m	Yes	Bankside veg.; applic. type	-	-
Denmark	By crop (up to 20-50 m)	-	-	-	Application window
Finland	10-25 m	-	-	-	-
France	Mitigation devised/ implemented based on local conditions	-	-	Mitigation devised/ implemented based on local conditions	Mitigation devised/ implemented based on local conditions
Germany	Up to 20 m	Yes	-	Grassed buffer zones; min. tillage; detention ponds	Application window; soil type
Greece	Up to 20 m	-	-	-	-
Ireland	By crop (up to 5-50 m)	-	Dry ditch	-	-
Italy	Up to 50 m	Yes	-	-	-
Netherlands	0.25 - 14 m	Yes	Windbreak	-	-
Portugal	By crop (up to 5 - 40 m)	Yes	-	Grassed buffer zones; min. tillage	-
Spain	Up to 5 - 50 m	Yes	-	Applic. window; grassed buffer zones	-
Sweden	By water body (1 - 10 m)	Yes	Wind speed/ direction; field size; temp.	-	-
UK	By crop (up to 5 - 50 m)	Yes	Water body type and size; windbreak	-	Application window

A broad view of risk mitigation measures is taken and these are defined as all measures and conditions that mitigate risk compared with the standard use situation considered during risk assessment in accordance with the Uniform Principles. This means that not only active mitigation such as implementation of a no-spray (or no-crop) buffer zone, but also the absence of a vulnerable situation (e.g. large and or flowing water bodies with large dilution potential) is considered at this stage.

The current position on the stipulation of mitigation measures during authorisation is variable, although mitigation options for all potential exposure routes are already considered by several Member States. Whereas risk mitigation during authorisation is routine in some Member States, measures are only applied post-authorisation at the regional or local scale in others. Mitigation of risk arising from spray drift is much further developed than that for exposure via surface runoff or drainflow. No-spray buffer zones are the most widely used mitigation measure although maximum widths and local conditions considered in their adoption vary considerably. Drift-reducing techniques are also considered in several Member States.

GENERAL PRINCIPLES FOR IMPLEMENTING RISK MITIGATION MEASURES UNDER 91/414/EC

The Work Group reviewed current practice in risk mitigation for pesticides across Europe and examined the literature investigating the efficacy of individual mitigation measures. It concluded that there is already sufficient evidence to implement certain measures into ecological risk assessment. Authorisations of products that present unacceptable ecological risk under standard use conditions can be made subject to the application of suitable restrictions ensuring mitigation of the risk. These mitigation measures should be grouped by the extent to which they reduce exposure in the following categories: 50, 75, 90, 95 and 99%. The Work Group has adopted a reasonable worst-case approach in assigning measures to different categories (e.g. exposure reductions based on larger datasets are assigned as a nominal 10^{th} percentile of the actual range of efficacy).

The topic of risk mitigation measures needs to span both regulatory procedures at European level leading to the listing of plant protection products on Annex 1 and authorisation procedures at Member State level where individual measures must be implemented. It is unlikely that the implementation of mitigation measures can be harmonised at Member State level in the short-term because: (i) there is large variability in the status of measures currently implemented to mitigate risk; (ii) there are different legal frameworks and enforcement possibilities in the different Member States; (iii) individual measures may be particularly suited to specific use conditions; and (iv) there are differences in agricultural practice and regulatory assessment procedures at Member State level (e.g. the use of locally collected data to calculate exposure via spray drift). It is thus suggested that a sequential procedure is adopted for incorporating mitigation measures into ecological risk assessment where it is considered that mitigation is required to protect non-target organisms:

[1] There should be a harmonised listing of the level of mitigation afforded by different mitigation measures. The purpose of the list would be to support the authorisations of plant protection products at the European level. For the reasons outlined above, it is not intended that the listing should be considered as mandatory, as different Member States will have varying use conditions and differing potential to implement a specific measure.

[2] The notifier will need to demonstrate the efficacy of one or more measures through a suitable refinement of the risk assessment.

[3] Within the EU registration process, the actual measure to be applied to mitigate risk should not be specified. Rather, the listing on Annex I should state that the decision to authorise the active substance was made on the basis of a mitigated risk and the level of mitigation that must be achieved for a particular input route in the different scenarios to assure safe use. It is possible to establish the level of mitigation that can be achieved for different measures with proved efficiency; such a listing implies that the maximum level of mitigation specified during Annex I listing is capped and that it will not be possible to authorise products where an unrealistic mitigation of risk would be required.

[4] Individual Member States must decide on national authorisations for products subject to appropriate risk mitigation. In so doing, they should consult the harmonised listing of mitigation measures for those approaches that are both appropriate and practicable to implement. Some mitigation measures will vary between Member States (e.g. the mitigation afforded by drift-reducing techniques may vary with standard machinery set-up and/or environmental conditions). Here, it would be appropriate to supplement the standard listing with alternative classifications at Member State level.

MITIGATION MEASURES SUITABLE FOR INCORPORATION INTO ECOLOGICAL RISK ASSESSMENT

It was agreed that several criteria must be met before a specific mitigation measure can be recommended for inclusion in the risk assessment and for subsequent implementation into risk management: (i) the measure must be practicable with a reasonable possibility of enforcement; it was beyond the scope of the working group to review the enforceability of a measure in individual Member States, but clearly successful mitigation depends on the measure being operational, controllable and backed by suitable enforcement; (ii) there must be a weight of evidence to demonstrate the efficacy of the measure under European (or directly correlated) conditions; the evidence must be quantitative so that the effect of the mitigation can be described numerically; and (iii) the risk assessment based on FOCUS guidance is complex and considers multiple routes of exposure (spray drift and either drainflow or runoff); methods must be available to include mitigation against a single route of entry into exposure assessments so that the total reduction in risk can be calculated.

The Work Group examined measures aimed at mitigating exposure via spray drift, surface runoff and drainflow. There are examples of measures for each exposure route where there is sufficient evidence of efficacy to recommend immediate inclusion within the risk assessment and these are discussed in turn below. Mitigating influences operating at the landscape level were also considered, though in less detail. For example, the high external recovery potential associated with inter-connected water bodies in differentially contaminated landscapes will mitigate risk from specific contaminants. It was agreed that the impact of such influences may be significant and that further work is required to develop and evaluate such approaches.

RISK MITIGATION FOR SPRAY DRIFT

The science of mitigation for pesticide exposure via spray drift is better developed than that for exposure via surface runoff or drainflow. Spray drift has been considered as a main route of entry to surface waters within risk assessments both at European level and within national procedures in all Member States surveyed. Many Member States have existing procedures for enforcing mitigation of spray drift during authorisation, although the complexity of the restriction possibilities and the range of mitigation approaches vary significantly. Three types of mitigation measure are proposed for immediate implementation into the risk assessment. These are the use of no-spray buffer zones, the application of drift-reducing technology and the reduction of exposure using windbreaks. Wind direction and wind speed will significantly affect spray drift, but the potential for control and policing is low so these factors were not considered as viable mitigation options.
No-spray buffer zones are widely implemented at present and have been successfully incorporated into the risk assessment over several years. Enforcement of the mitigation may be simpler where no-spray buffers are legislated as no-crop buffers, as in the Netherlands.

Technical solutions to reduce spray drift have advanced significantly over the last 10 years. Drift-reducing nozzles are widely adopted by farmers in some Member States and have been incorporated into the risk assessment. It is suggested that the use of this technology is incorporated into risk assessment at the European as well as Member State level. Specific technologies that are recommended for use include drift-reducing nozzles, air assistance, tunnel sprayer, shielded spraying, and band spraying. Classification systems for drift-reducing techniques already exist in several Member States. Windbreaks comprising trees or vegetation of at least 1 m higher than the crop have been successfully implemented as a mitigation measure in the Netherlands. Similar approaches are also used in the UK. The approach is suitable for incorporation into ecological risk assessment, but applies only to a windbreak planted immediately adjacent to the water body.

RISK MITIGATION FOR SURFACE RUNOFF AND EROSION

The potential for runoff entry can be separated into two main components: (i) the portion of pesticide transported in association with particulate, eroded material in the runoff; this is likely to be the major contributor for low solubility, sorptive compounds; and (ii) the portion of pesticide transported in association with the water phase of the runoff; this is likely to be the major contributor for high solubility, mobile compounds. For the former case, interception of the transported soil particles will provide the greatest mitigation benefit, whereas for the latter, water transport (and hence infiltration capacity) will be more important. Appropriate mitigation measures for runoff entry should therefore take into account the mobility properties of the compound in question. In experimental studies, Koc is rarely identified as a determining factor for buffer efficacy. In agricultural settings subject to excessive soil erosion, various soil management practices (such as conservation tillage or contour ploughing) can limit the transport of pesticides by eroded sediment.

Because pesticide transfer in runoff varies considerably in relation to climatic conditions and numerous local parameters, the effects of mitigation measures in reducing pesticide transport in surface runoff can be variable. The most effective implementation of mitigation will take place through the application of pesticide management at the local scale, and for Member State

registrations, it is important that local climatic, soil and agronomic practices are taken into account when determining suitable levels of mitigation. Nevertheless it is recommended that runoff mitigation approaches can now be broadly implemented into the regulatory risk assessment for Annex I registration in the EU. Three mitigation options that are suited to regulatory assessments are: (i) a reduction in the application rate, giving a similar reduction in losses to surface waters via surface runoff or erosion; (ii) a restriction in the application window, normally to avoid application during or immediately before periods when the risk of runoff is greatest; and (iii) the application of a vegetated buffer zone (or filter strip) to intercept runoff prior to entry into surface water. For the first two options, the principles are similar to approaches applied in many Member States to mitigate the risk of leaching to groundwater. For the third option, there are already good examples of such approaches being successfully applied at Member State level, where label restrictions are applied to limit runoff input at the point of entry (i.e., next to the water body).

RISK MITIGATION FOR DRAINFLOW

The number of effective mitigation options for reducing exposure via drainflow is limited. This is partly because drainflow has only recently been considered as a primary route of exposure both at Annex I and in many national registration procedures. However, losses via drainflow are also very difficult to control through intervention other than to limit the amount of pesticide applied, the timing of treatment or the types of soil treated. The only regulatory mitigation options at present are: (i) a reduction in the application rate, giving a similar reduction in losses to surface waters via drainflow; and (ii) a restriction in the application window, normally to avoid application just before the onset of winter drainage. Although these options have only been used in practice in two Member States (Germany and the UK), the principles are similar to approaches applied in many Member States to mitigate the risk of leaching to groundwater.

A number of mitigation options will be suited to local management of pesticides and/or product stewardship. These include management of soil structure, avoiding application to very dry or very wet soil, and discouraging the practice of "over-draining" (installing more efficient drains than required for a particular soil type). However, none of these approaches is suitable for inclusion in ecological risk assessment as impact on pesticide transport is unpredictable and none can be rigorously controlled or policed. In the absence of further mitigation based on site management, the only additional option to mitigate risk within ecological risk assessment appears to be a restriction in the soil to which a product may be applied. This could be applied at Member State level according to the level of risk and could take two forms: (i) a blanket restriction from use on any drained land; this mitigation measure is simple to communicate and should reduce transport of a pesticide in drainflow to zero; and (ii) a restriction based on soil vulnerability prohibiting use of a product on soils associated with unacceptable risk. Such an approach would be more flexible and it mimics label restrictions imposed by risk managers in the United States. Differentiation of use by soil type is already implemented in Germany and the Netherlands to protect groundwater.

REFERENCE

FOCUS (2002). *FOCUS Surface Water Scenarios in the EU Evaluation Process under 91/414/EEC.* EC Document Reference SANCO/4802/2001-rev.2. pp.245.

A new empirical approach to estimate the short-range transport and dry deposition of volatilised pesticides

G Fent, M Moendel, R Kubiak
Institute for AgroEcology, RLP AgroScience, Breitenweg 71, 67435 Neustadt/W., Germany
Email: gunnar.fent@agroscience.rlp.de

ABSTRACT

As a continued effort to reduce uncertainties in pesticide non-target exposure assessment, the short-range transport and deposition of volatilised pesticides was identified as a relevant entry pathway. Consequently, this volatilisation pathway should also be considered in the overall exposure assessment. The results from 15 outdoor wind tunnel studies with 10 pesticides covering a wide range of vapour pressures (v.p.) formed the basis of an empirical assessment tool. The variance analysis identified three subgroups of pesticides with different deposition behaviour. Because the v.p. was the most influential factor, three v.p.-dependent deposition classes were established. Within each deposition class, the pesticide that represented the highest deposition was selected and the 90th percentile 1-m distance deposition of several trials of this pesticide was determined. The decrease of deposition was found to be a function of the distance from the treated crop. This could be described by an exponential function normalised to the 90th percentile deposition at the 1-m distance. A comparison with independent field results indicated that the new assessment tool can serve as a suitable conservative approach to estimate the short-range transport and dry deposition of volatilised pesticides to surface waters under field conditions.

INTRODUCTION

Plant protection products are needed by farmers to protect their crops from the attack of pests, fungal diseases and weeds. A further heightened understanding of the behaviour of these products in the environment also helps to increase the safety of the populations of non-target organisms in off-crop areas close to treated fields.

Volatilisation, short-range transport and deposition of pesticides from the air have recently been identified as an additional potentially relevant entry path into the habitats of aquatic and terrestrial non-target organisms. In 2002, the German Federal Consumer Protection and Food Safety Agency (BVL, formerly BBA) published a guideline for the assessment of the volatilisation, short-range transport and deposition of plant protection products from the air (Winkler *et al.*, 2002). It was intended to include this path into the exposure assessment for the authorisation of plant protection products. The guideline describes a tiered assessment scheme consisting of model calculations and field experiments. In principle, the volatilisation of pesticides and subsequent off-site deposition from the air can be determined in outdoor field trials. However, because meteorological conditions are highly variable and less reproducible over the test period of at least one day, outdoor field studies may be appropriate to identify possible pathways of entry. Here, however, systematic investigations about the order of magnitude and influencing factors are limited. Both assessment methods (model calculations

and field experiments) showed some uncertainties in their predicted estimated concentrations. This was due to the fact that the volatilisation, transport and deposition of volatilised pesticides are a highly complex process consisting of several mass transfer and transport phenomena. A need for additional research in this area was concluded.

The main objective of this research project was the development of a large-scale outdoor wind tunnel test system for reproducible measurements of the atmospheric short-range transport and dry deposition of volatilised pesticides in order to develop and improve models and test systems. The volatilisation, short-range transport and deposition behaviour of ten pesticides, representing a wide range of physical and chemical properties, was investigated in fifteen semi-outdoor wind tunnel experiments.

MATERIALS AND METHODS

The semi-outdoor wind tunnel test system was introduced in detail by Fent (2004) and was developed to conduct volatilisation/deposition studies under realistic outdoor conditions independent of outdoor wind speed, wind direction and rainfall. Two semicircular film greenhouses with an open front and rear walls were constructed (length 55 m, width 7 m, height 3 m) that served as outdoor wind tunnels. At one end an assembly of 26 synchronously working axial flow fans was established to produce a constant air-flow ranging from 1 m/s to 5 m/s. The size of the field area in the wind tunnel was 4 m wide and 25 m long (0.01 ha). It was cropped alternately with sugar beet and winter wheat, and had an almost complete crop canopy at the time of spray treatment. Artificial water bodies (steel bowls, 0.5 m^2 filled with tap water) served as model surface waters in the non-target area in downwind direction at defined distances (1, 3, 5, 10, 15 and 20 m). These water bodies were put into place 5 min after the spray treatment in order to prevent sampling of spray drift. Following the spray application of formulated pesticides the substances were allowed to volatilise in a constant air stream and to depose to the downwind water bodies for a period of up to 24 h after spraying. Then, the deposition was determined by analysing water samples by means of GC and HPLC. The Plant Protection Products (PPP) used in the experiments were applied as spray mixtures in various compositions, covering a wide range of v.p. (from 1.2×10^{-2} to 1.7×10^{-6} Pa) and water solubilities (from 1.4 to 10^6 mg/litre at 20°C). The pesticides investigated in this project were named in an anonymous form (PPP-A to PPP-I), with the exception of the active ingredient lindane, in order to protect the data protection rights of the manufacturers.

Spraying was carried out according to agricultural practice at application rates of between 200 and 1000 g a.i./ha. The wind tunnel experiments were carried out under the following conditions:

- The wind tunnel trials covered a wide range of weather conditions with respect to air temperature, relative air humidity, solar radiation, wind speed and leaf wetness duration.
- Wind speed and direction were fixed and reproducible.
- The use of two crops with different habits, leaf morphologies and leaf area indexes provides information about the influence of crops on volatilisation.
- Processes competing with volatilisation, such as wash-off by rainfall or photo-degradation, did not take place in the wind tunnel test system. Volatilisation observed in the wind tunnel can therefore be classified as the worst-case situation compared with outdoor field studies.

RESULTS AND DISCUSSION

The results were reported in detail by Fent (2004). In the following, selected results are reported which formed the empirical basis of the estimation tool. To obtain an initial overview of the magnitude and limits of variation, the relative non-target deposition (mean of measurements of all distances) for all trials (mean and standard variation) are given in Figure 1. Using these data as input data for an analysis of variance (LSD: Least Significant Differences), three subsets with significantly different deposition ranges (95%-level) were identified. Subset "A" was only represented by lindane (0.59% relative deposition; mean of measurements of all distances), subset "B" was represented by PPP-A, PPP-B, PPP-C and PPP-E (mean of measurements of all distances ranged from 0.07 to 0.11% relative deposition) and subset "C" was represented by PPP-D, PPP-F, PPP-G, PPP-H and PPP-I (mean of measurements of all distances ranged from <0.001 to 0.023% relative deposition).

In spite of the variation caused by changing experimental boundary conditions (e.g. wind speed 2 or 4 m/s, winter wheat or sugar beet, mean air temperatures between 4.4 and 24.5°C), it was possible to identify subsets of pesticides with different deposition behaviour. It can therefore be stated that the process of short-range transport and deposition after volatilisation:

- was confirmed as a potential entry path for non-target areas.
- could be quantified for 10 different pesticides
- depends not only on experimental boundary conditions but also on pesticide properties.

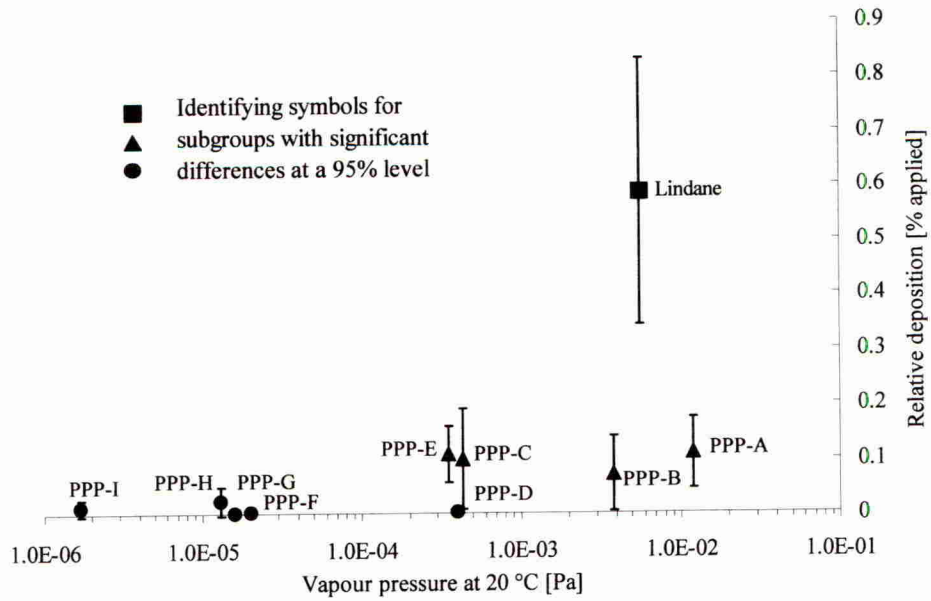

Figure 1. Relative deposition (mean of measurements of all distances) observed in the wind tunnel as a function of v.p. (mean of all experiments and standard deviation)

The decrease of deposition after volatilisation was found to be a function of distance from the treated crop and could be described by an exponential function (r^2 = 0.9758) normalised to 100% deposition at the 1-m distance (with x = distance in metres):

$$y = 100 \cdot e^{-0.05446(x-1)}$$

In order to identify a relationship between deposition after volatilisation and potentially influencing parameters, correlation analysis and t-tests were performed. The results are summarised in the following:

- The v.p. is the most influential factor of the pesticide-specific properties, but fitting a continuous equation to observed deposition data was statistically not relevant.
- A comprehensive effect of air temperature, air humidity and wind speed on the concentration in the air and the deposition after volatilisation was not observed.
- The kind of crop did not affect the aerial concentration nor the deposition after volatilisation.

Despite the fact that controlled boundary conditions in the wind tunnel are in contrast to field experiments, it appeared that deposition after volatilisation is highly complex, non-linear and multi-factorial. Prospects for the development and improvement of physically based and process-oriented models are therefore limited. Nevertheless, the wind tunnel results can be used for an empirical approach of an estimation model.

The variance analysis identified three subgroups of pesticides with different deposition behaviours. Because the v.p. was the most influential factor, three deposition classes were established, taking into account the v.p. of the pesticides and the results of the variance analysis.

Within each deposition class the pesticide that represented the highest deposition (mean values 1-m distance) was selected and the 90th percentile of several volatilisation/deposition trials was determined.

The only pesticide with a v.p. < 10^{-5} Pa (PPP-I) showed a 90th percentile deposition at 1-m distance of $\leq$ 0.05% of the applied amount. The percentage of deposition after volatilisation was < 5% of the total deposition from the aerial route [deposition after volatilisation + spray drift deposition amounts taken from the literature (Rautmann *et al.*, 2001)]. On the basis of these findings the following proposal can be made:

- Dry deposition and short-range transport after volatilisation is negligible for pesticides with a v.p. < 10^{-5} Pa in relation to spray drift deposition. An estimation of dry deposition after volatilisation for pesticides with a v.p. < 10^{-5} Pa is not necessary.

Lindane applications resulted in by far the highest deposition observed in the wind tunnel experiments (mean of all trials 0.94% of applied amount at a 1 m distance from the field). Deposition was higher by a factor of about 5 than that of the next volatile pesticide. These findings are in line with the field results published by Siebers *et al.* (2003) and Gottesbüren *et al.* (2003). In these field studies lindane also showed by far the highest deposition after volatilisation. Therefore, the following proposal can be made:

- Due to the highest deposition generally observed with lindane, the 90th percentile of lindane deposition from a large number (n = 15) of wind tunnel trials can be used as a conservative value for the uppermost deposition class, valid for pesticides with a v.p. > 5 x 10^{-3} Pa .

The other deposition classes are shown in Table 1.

Table 1. Deposition classes, corresponding pesticides and 90th percentiles of the pesticides with the highest deposition (mean values 1-m distance) as a basis of a deposition class model

Deposition class v.p. at 20°C	Pesticide code[1)] [Number of trials]	90th percentile relative deposition (% applied at 1-m distance)
< 10^{-5}	PPP-I [15]	≤ 0.05
10^{-5} – 10^{-4}	PPP-H [15], PPP-G [3], PPP- F [12]	0.089
10^{-4} – 5 x 10^{-3}	PPP-B [14], PPP-C [3], PPP-D [12], PPP-E [5]	0.22
> 5 x 10^{-3}	Lindane [15], PPP-A [15]	1.56

[1)] Pesticides representing the highest deposition in their class are underlined (guidance substances)

In order to describe the deposition as a function of distance from the treated crop, the 90th percentiles of the guidance substances were used as an input parameter in the distance function as already mentioned above, resulting in deposition curves shown in Figure 2.

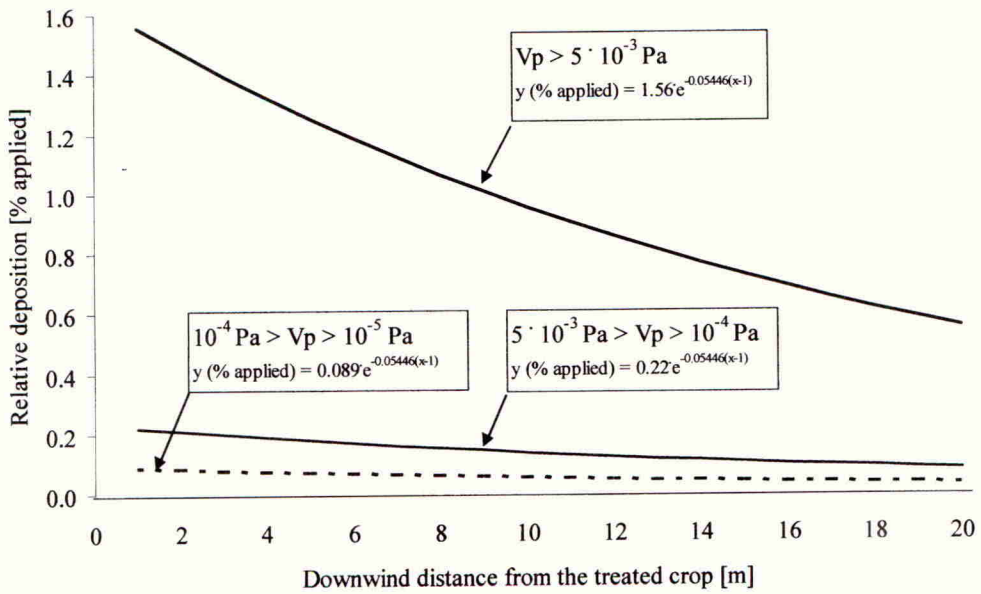

Figure 2. Dry deposition of volatilised pesticides. Deposition as a step-wise function of v.p. and distance from the treated crop

The predictions of this proposed deposition model for each of the three deposition classes were compared with the experimental wind tunnel results. The minimum over-estimation of the model for the pesticides representing the highest deposition was by a factor of 1.8 (lindane), a factor of 1.4 (PPP-E) and a factor of 2.7 (PPP-H) when compared to the wind tunnel measurements. The deviations for the other measured pesticides were consequentially higher. The estimation functions were developed on the basis of the wind tunnel results. It can therefore be expected that the estimations reflect the results observed in the wind tunnel. It is, however, more important to validate the model with independent field results. So far only the results of three field experiments are available (Gottesbüren *et al.*, 2003; Siebers *et al.*, 2003). Lindane applied to cereals in two field trials showed that field results were over-estimated by the estimation model by a factor of up to 2. The results of field experiments with parathion and pendimethalin were also in agreement with the corresponding model calculations.

CONCLUSIONS

The experimentally derived v.p.-dependent deposition classes and the decrease of deposition as a function from the treated crop were used to establish an empirical deposition model that can serve as a suitable conservative approach to estimate the short-range transport (1 - 20 m) and dry deposition of volatilised pesticides to surface waters under field conditions.

A model assumption is that pesticides are applied to arable crops with an interception larger than 90%. If larger pesticide amounts reach the soil, volatilisation and ensuing deposition might be less than applied on only plant surfaces.

ACKNOWLEDGEMENTS

We express our sincere thanks to the "Ecochemistry Team" of the Institute for Agroecology and we would like to thank the German IVA for their financial support of this research project.

REFERENCES

Fent G (2004). *Short-range transport and deposition of volatilised pesticides*. ISBN 3-8322-3568-X. Shaker publication house: Aachen.

Gottesbüren B; Grote C; Hassink J; Pestemer W; Richter T; Schmidt H; Stadler R; Zangmeister W (2003). Deposition of volatilised pesticides on non-target areas. *Proceedings of the XIIth Symposium Pesticide Chemistry*, 247-256.

Rautmann D; Streloke M; Winkler R (2001). New basic drift values in the authorisation procedure for plant protection products. In: *Workshop on risk assessment and risk mitigation measures in the context of authorisation of plant protection (WORMM)*, ed. R Forster & M Streloke, Mitteilungen der BBA, **383**, 133–141.

Siebers J; Binner R; Wittich K P (2003). Investigation on downwind short-range transport of pesticides after application in agricultural crops. *Chemosphere*, **51**, 397-407.

Winkler R; Binner R; Gottschild D; Koch W; Siebers J (2002). *Bewertungskonzept zum Nahtransport von Pflanzenschutzmitteln infolge Exposition über den Luftpfad (Abtrift, Verflüchtigung und Deposition)*, Berichte der BBA, **110**.

Honeybee testing and risk assessment in ecotoxicology – current approaches and novel aspects

Ch Maus
Institute for Ecotoxicology, Bayer CropScience, Alfred-Nobel-Strasse 50, D – 40789 Monheim, Germany
Email: christian.maus@bayercropscience.com

ABSTRACT

Current approaches and requirements for ecotoxicological testing and risk assessment on bees in Europe are reviewed and discussed, in particular in terms of their suitability to cover all possible aspects of risk which may be posed by pesticides to bees. Moreover, concerns on appropriateness of the existing systems are described, and selected novel approaches for testing and risk assessment are presented and discussed. Crucial points are that all testing methods suggested for implementation should be appropriately validated, that sublethal endpoints only should be considered if their ecological significance is clearly proven, and that all testing and risk assessment approaches should be in line with the European Directive 91/414/EEC. An important principle of this directive is that higher tier tests will in any case of doubt always provide information superior and more reliable compared to laboratory studies. Given that tests performed are appropriately designed, the range of studies as referred to in Directive 91/414/EEC will fully cover any possible risk of pesticides to bees.

INTRODUCTION

One of the aspects of ecotoxicology which were most intensively and most emotionally discussed during the last years in the scientific community as well as in the public are the potential side effects of plant protection products to honeybees. In terms of these issues, there was also some discussion about the requirements for ecotoxicological testing and risk assessment, and in how far requirements at the current state would guarantee a reliable protection of bees. In this article, a review of the current European testing and risk assessment schemes as put down in the European Council Directive 91/414/EEC is presented as well as an outline of the new approaches discussed, followed by a discussion whether and in how far the novel approaches are really necessary to fill potential gaps in the current scheme, or in how far they are already covered by the existing scheme, when appropriately applied.

Honeybees exhibit some very specific characteristics, which make them different from other organisms treated in ecotoxicology, and which give them a kind of special position in ecotoxicology in terms of several aspects. First, honeybees, like other eusocial hymenopterans, have a virtually unique social system which is based on particular genetic features. Bees are haplo-diploid with haploid drones and diploid workers and queens, whereby worker bees normally not reproduce themselves, but act as nurses for the queen's offspring, which are their sisters. Due to their haplo-diploid genetic constellation, sisters are more closely related to each other than they would be to their own offspring – thus, it is more attractive to them to foster their sisters than to have own offspring, from an evolutionary genetic perspective. Therefore, the individual does not play such an important role in the socio-ecological system of this kind

of organisms, and a bee colony has thereby to be seen and treated differentially compared to a population or an individual of other organisms, also under ecotoxicological aspects. Second, different from other test organisms in ecotoxicology which are mainly considered to be model organisms to cover a broad range of diversity with regulatory testing, honeybees are of importance in terms of three additional aspects: apart from their function as model organism covering related species, they also have a key position in terrestrial ecosystems as pollinators of plants, and play thus an important role in maintenance of plant and thereby of insect diversity. Furthermore, they are not only an ecological but also an economic factor, as domestic animals producing honey, wax, and other commercial products, and pollinating agricultural crops. And third, honeybees play an important role in human culture since long times, and traditionally are associated with positive values. This can readily be seen by the emotional dimension which bee incidents can easily gain in public perception. Accordingly, particular attention is paid to the honeybee in the European ecotoxicological testing and risk assessment systems. It is one of the few standard test species in ecotoxicology which is not only tested as a model organism, but which is also considered in testing and risk assessment for the sake of protection of one individual species.

RESULTS AND DISCUSSION

Current ecotoxicological testing requirements and risk assessment procedures in Europe

In the EU, testing requirements for honeybees are fixed in Council Directive 91/414/EEC (Council of the European Communities, 1991). Further guidance for testing and risk assessment on bees is given in the Guidance Document on Terrestrial Ecotoxicology (European Commission 2002), furthermore in the EPPO Standards (EPPO, 2003). Relevant testing guidelines for regulatory bee testing are for instance EPPO 170 (EPPO, 2001) and OECD 213 and 214 (OECD, 1998a, b). In addition to this, there are methods for special testing designs available as well as special testing guidelines on national level (see below).

The routinely required and basic acute test is the laboratory test on oral and contact toxicity according to OECD 213/214 or EPPO 170. It has to be carried out for all products to be registered. The endpoint is the LD_{50}, unless a product is nontoxic to bees (LD_{50} > 100 µg a.s./bee), then a limit test is sufficient. With the LD_{50} found, the Hazard Quotient (HQ) can be calculated by dividing the highest single application rate (g a.s./ha) by the contact or oral LD_{50} in µg a.s./bee. If the HQ is below 50, a substance is considered a priori nontoxic to bees, and further testing is not required. If the HQ is greater than 50, a more in-depth assessment of a potential risk to bees will be required. For example, an aged-residues test on natural leaf substrate can then optionally be carried out, in particular when significant residual traces of the substance of concern are likely to remain on crops after application and could affect foraging bees. However, this test was never very common in Europe (rather in USA), and may be considered no more in a revised Directive 91/414/EEC. Normally, compounds with an intrinsic toxicity potential to bees are as a next step tested under semifield (tunnel or tent) conditions in European practice (this would also be the next step in case of effects in an aged residues test). The studies are usually conducted according to the guideline EPPO 170, which is, however, not very strict in its provisions concerning study design. In France, the national guideline CEB 230 (Malet, 2003) is available for semifield testing, which is slightly different from- and more specific than EPPO 170. If there are adverse effects seen in the semifield study, a field test can be carried out according to EPPO 170. The EPPO guideline, as well as the Terrestrial Guidance Document, provides a considerable degree of freedom in setting up such field study design.

The aforementioned study cascade is mainly designed to detect effects of a compound to adult bees, yet the higher tier study designs could also detect effects to bee brood. To cover compounds which are not toxic to adults, but potentially to bee brood, an additional line of testing is provided under the Directive 91/414/EEC: if a compound may act as an insect growth regulator or has larvicidal properties, or if other hints to a brood-toxic potential are existing, a brood-feeding test has to be conducted, e.g. according to Oomen *et al.* (1992), irrespective of the results of the acute laboratory test. In case of effects in this test, a semifield test can follow (since a few years, there is a validated method available for a specific semifield test on brood effects (Schur *et al.*, 2003)), whereas in cases where a potential for brood effects clearly exists, the semifield test can be directly conducted, without preceeding brood-feeding test. Next and last step would be a field study also in this cascade.

Peculiar features of honeybees in terms of ecotoxicology testing and risk assessment

Considering the particular position of the honeybee in the system of ecotoxicology, it is not surprising that there are several special features in ecotoxicological bee testing and risk assessment, compared to other domains of ecotoxicology.

Bees are routinely tested for effects caused by contact as well as by oral exposure to the same compound, which is done in separate tests. This is exceptional in ecotoxicology, in most other test organisms, there is only one main route of exposure considered. Moreover, there are different testing cascades for different developmental stages of one organism existing for the bees; this is likewise rather uncommon in ecotoxicological testing. In contrast to Non-Target Arthropod testing, extended laboratory approaches are less important in bee testing.

In the risk assessment, the HQ approach is applied instead of a TER approach as in many other sub-domains of ecotoxicology. The HQ for bees is an empirically derived value which does not directly include PEC data.

There are in addition further points in which risk assessment for bees differs from risk assessment for other organism groups, in particular Non-Target Arthropods. Whereas in the risk assessment on Non-Target Arthropods recovery or potential for recovery play an important role, this is not considered very relevant in the current risk assessment on bees. Then, in bee risk assessment, the differentiation between in-crop and off-crop is not as pronounced as it is in the arthropod risk assessment. This due to the facts that bees are very mobile and will not differentiate between in-crop and off-crop, and that bees will not have separate populations in-crop and off-crop. A further difference between bees and other non-target organisms is the goal of protection: for other organisms this is the population (or in exceptional cases the individual), for bees, however, it is the colony.

All these peculiar features underline the special position of the honeybee in the ecotoxicological testing and risk assessment system.

Concerns regarding the existing testing and risk assessment schemes

Although the European risk assessment approach proved largely reliable and effective in the last years, as can be seen by the relatively low number of bee and bumblebee poisoning incidents reported (see e.g Oomen, 2001; Thompson & Hunt, 1999; Fletcher & Barnett, 2003; Brasse, 2003), there were still occasionally concerns brought up that this approach would not appropriately address certain sources of risk (e.g. Cluzeau, 2002; Tasei *et al.*, 2003).

One of the concerns sometimes expressed was that sublethal and behavioural effects were not sufficiently covered in the current testing and risk assessment schemes; this was especially related to supposed effects on foraging and homing behaviour, orientation and learning capacity, and behavioural patterns associated with broodcare. Another complex of aspects which some assumed not to be fully covered by existing regulations is chronic toxicity on individual and colony level, including longevity of individuals, cryptic brood effects, colony overall vitality, and colony overwintering (see e.g. Thompson & Brobyn, 2003; Tasei *et al.*, 2003).

A further critical point discussed concerns systemic compounds used for seed dressing. The HQ approach is difficult to apply to such compounds since the empirical correlation between a hectare application rate of a sprayed compound and its intrinsic bee toxicity cannot directly be extrapolated to soil-systemic products, and a validated HQ for seed-dressing compounds has not been established so far (e.g. Tasei *et al.*, 2003). Likewise, the concern was brought up that other hymenopteran pollinator species like wild bees or bumblebees might not be sufficiently covered by the studies conducted with the honeybee (see e.g. Thompson, 2001).

The aforementioned concerns were partly driven or enhanced by severe colony losses which occurred in many countries in Europe and overseas in the last years. Despite the majority of experts agrees that this phenomenon is caused by a complex of multiple factors interacting, among which pesticides play, if any, only a minor role (e.g. Otten, 2003; Ritter, 2003; von der Ohe, 2003a, b), the hive losses stirred up a discussion about possible cryptic effects of pesticides to bees. This happened most pronounced in France where the discussion attracted considerable public attention and focussed on systemic insecticidal seed-dressings. The dispute culminated in the issue of the "CST report" which was corroborated in 2003 by the Comité Scientifique et Technique de l' Etude Multifactorielle des Troubles des Abeilles on behalf of French Ministry of Agriculture (Doucet-Personeni *et al.*, 2003). In this document, most of the concerns expressed were taken up, and a novel kind of risk assessment was generated which seemingly covered open points raised before. This new approach apparently strongly influenced the new French Guidance Document (version 6.5, SSM 2004), which was released shortly after.

Several, but not all of the approaches reviewed and discussed below are also referred to in the current French Guidance Document.

Novel approaches for bee testing and risk assessment

Proposed new tests included a chronic laboratory study with dietary exposure. Likewise, in-vitro laboratory tests with feeding of larvae were suggested. Furthermore, new testing requirements discussed were on behaviour of the bees; orientation, foraging and homing behaviour, as well as learning capacity (e.g. Proboscis Extension Reflex test) (see for instance the French Guidance Document 6.5). The need to implement testing requirements with non-Apis bees (like bumblebees, wild bees) was occasionally claimed, for instance in a recent draft version of a Dutch National Guidance Document (unpublished).

Another novel approach for bee risk assessment is the PEC/PNEC approach. It is suggested that it is used in particular for systemic applications of pesticides to which bees may be exposed through nectar or pollen, and for which the HQ approach cannot readily be applied (e.g. French Guidance Document 6.5). In this approach, the PEC, here defined as consumption (mg food/bee) x residue level (g a.s./kg food), divided by the PNEC, which is NOAEL (or LD_{50}) / uncertainty factor. Risk is considered acceptable if PEC/PNEC values are smaller than

1. The comparison of a PEC with a NOEC in the risk assessment for bees is not new per se (see for instance Maus *et al.*, 2003), however, the new approach also includes the feed consumption of a bee as an input parameter and, in contrast to a TER approach, an uncertainty factor. New in this, but not necessarily restricted to the PEC/PNEC approach, is also the option of implementation of uncertainty factors even for highest tier study data.

Comments on novel approaches

A chronic laboratory feeding study on worker bees is on principle feasible yet relatively high control mortalities may occasionally cause technical problems influencing the robustness of the study design and the statistical power to detect effects. However, all the information that can be gained from a chronic laboratory study is already covered in the European study scheme with a semifield test. As such, the chronic laboratory test could reasonably be conducted to replace a semifield test when the concern is limited to intrinsic chronic adult toxicity. In this case, however, it should be defined how this study is triggered and under which circumstances it could replace a semifield study. It should in no case be required in addition to a semifield test, because it would only create redundant data.

The usefulness of the implementation of *in-vitro* larvae tests in the laboratory is questionable. Although several methods have been published (e.g. Wittmann, 1981; Malone *et al.*, 2002; Brødsgaard *et al.*, 2003), they have not been validated through appropriate ring-testing as required under international methodology schemes (i.e. OECD). So it is not known whether they will allow for generating reproducible results when tested in different testing facilities. In addition, all relevant aspects of larval testing in the laboratory are likewise covered by semifield tests with an appropriate design.

Similar concerns apply to the implementation of lower-tier tests on sublethal effects and the use of their results for the risk assessment. One important point is that many of the sublethal effects proposed to be tested are unclear, or even doubtful, with respect to their ecological relevance. For instance, if measurable differences in learning behaviour are found in an artificial study design, this does not necessarily mean that a colony or even an individual would be adversely affected under field conditions by this effect (see also Thompson, 2003). Moreover, for testing sublethal effects on bees, there are at present hardly any validated methods available. Study designs proposed are mostly derived from publications based on one or a few trials frequently carried out in only one laboratory, and they have not been validated by appropriate ring-testing so far (e.g. Lambin *et al.*, 2001; Guez *et al.*, 2001; Dechaume Moncharmont, 2003; Decourtye *et al.*, 2004). In addition, any sublethal effects that are ecologically relevant, should also show up under field or tunnel conditions, otherwise they would be artifacts of laboratory conditions and could thus not be relevant for the final risk assessment. Thus, also here, sublethal studies would be redundant if appropriate higher tier semifield or field data are at hand.

Inclusion of tests of additional hymenopteran pollinator species into the testing and risk assessment scheme does not appear necessary: In the European risk assessment schemes, the honeybee is separately tested (in addition to sensitive NTA testing) due to its economic and ecological importance and to its highly complex social system and colony structure. Both features may also be relevant for other hymenopteran pollinators, but most pronounced they are in the honeybee. So this species represents an extreme case of a eusocial hymenopteran pollinator with great relevance in ecosystems. Thus, there is no reason why wild bees and bumblebees should not be covered by bee testing in terms of their special position in ecosystems, their sociobiological properties, and their economic importance. Likewise, the

possibility that there may be more sensitive pollinator species than the honeybee is already covered by the European risk assessment scheme as well: the Non-Target Arthropods *Typhlodromus pyri* and *Aphidius rhopalosiphi* have been identified as the most sensitive out of a broad range of arthropods. These two species are regularly tested and considered in the risk assessment for any pesticide to be registered. Thus, the aspect of possibly existing more sensitive apoid species is fully covered by the Non-Target Arthropod risk assessment.

In conclusion, it can be stated that all aspects of pesticide side effects to bees can be covered in the test types which are currently referred to in the current European study schemes. Additional study types, if properly validated, can give insight in intrinsic effect potentials of compounds, and they may in case be suitable to address certain aspects on a lower tier. However, they should not routinely be required for regulatory risk assessment. Semifield and field studies with a realistic design will still provide most reliable highest-tier evidence, and any kind of concern can appropriately be addressed in a semifield or field trial with a suitable design, if neccessary by adapting a standard design accordingly in order to account for particular concerns, e.g. on behavioural or cryptic long-term effects. Thus, semifield or field data should always override lower-tier data of any source, and no additional lower-tier data should be required, if higher-tier data are available. This principle is complied with in all other sub-domains of ecotoxicology in EU Directive 91/414/EEC, for example aquatic or soil ecotoxicology, or Non-Target Arthropods. There is no obvious reason not to treat the bees according to the same basic principle. It may in fact be the case that certain effects have only be detected in special-design studies, however, effects which do not occur or have no further impact under field conditions are not of ecological relevance and do thus not need to be considered in ecotoxicological risk assessment.

The strength of the current HQ approach for bee risk assessment lies in the fact that it is an empirically derived correlation between intrinsic toxicity of a compound and its field application rate. It does as such not need arbitrary uncertainty factors. It is mainly designed for spray-applied compounds, and it cannot be directly extrapolated to systemic seed dressing compounds. With the PEC/PNEC approach, there is a new alternative implemented to the HQ approach. The PEC/PNEC approach resembles a TER approach in so far as that both compare a PEC with effect data. However, both are still different approaches and should not be mixed up. For a sound risk assessment on bees, a TER approach should, in cases where an HQ approach is not applicable or not validated, be applied rather than a PEC/PNEC approach. Ecotox risk assessment as such should be consistent and be conducted according to homogenous principles. The risk assessment for bees should thus stay within the concept of the EU Directive 91/414/EEC, and the tool considered in this concept is the TER approach, not the PEC/PNEC approach. There is no legitimate reason to switch to another approach for single sub-domains unless there is a compulsory necessity to do so. However, this is not the case for the bees, since a TER approach would be fully sufficient for a sound risk assessment. Apart from this, as far as it includes an uncertainty factor, the PEC/PNEC approach differs from the TER approach in a decisive point. In the latter approach, the uncertainty factor remains separate from the TER calculation, and can thus be adapted to the particular circumstances of each case; e.g. it can be lowered to account for an extensive data basis on the endpoint in question. In the PEC/PNEC approach, an uncertainty factor would be fixed for every type of test and of resulting endpoint, which would make the system less flexible. Furthermore, it should be generally questioned in how far it is legitimate to consider uncertainty factors for endpoints that have been generated in higher-tier studies in a realistic design. Uncertainty factors are usually applied to correct for uncertainties in extrapolating from an artificial

laboratory design to field conditions, or for possible sensitivity differences between the test organism and other species to be covered by this model organism. The first aspect does not need to be considered when a test design already reflects field conditions. The sensitivity aspect is, as mentioned above, already covered by the Non-Target Arthropod risk assessment so that the honeybee stands as such for itself in terms of the sensitivity aspect and not as a model organism chosen to represent particularly sensitive arthropods in the ecotoxicological risk assessment. Moreover, there is no evident need to generally assess all systemic compounds according to such an approach. It would rather be reasonable to restrict the requirement of a TER calculation to systemic compounds which have been shown to exhibit significant toxicity to bees in the laboratory test, or which have an intrinsic potential for brood effects.

REFERENCES

Brasse D (2003). Poisoning incidents involving honeybees in Germany (1999-2002) and new problems for beekeeping. *Bulletin of Insectology* **56**, 199.

Brødsgaard HF; Brødsgaard C J; Hansen H; Lövei G L (2003). Environmental risk assessment of transgene products using honeybee (*Apis mellifera*) larvae. *Apidologie* **34**, 139-145.

Cluzeau S (2002). Risk assessment of plant protection products on honeybees. Regulatory aspects. In: *Honeybees: Estimating the Environmental Impact of Chemicals*, eds J Devillers & M H Pham-Delègue, pp. 42-55, Taylor & Francis: London, New York.

Council of the European Communities (1991). Council Directive of 15 July 1991 concerning the placing of plant protection products on the market (91/414/EEC). *Official Journal of the European Communities* No L 230/1 (19.8.91), pp. 32.

Dechaume Moncharmont F X; Decourtye A; Hennequet-Hantier; Pons O; Pham-Delègue M H (2003). Statistical Analysis of Honeybee Survival After Chronic Exposure to Insecticides. *Environmental Toxicology and Chemistry* **22**, 3088-3094.

Decourtye A; Devillers J; Cluzeau S; Charreton M; Pham-Delègue M H (2004). Effects of imidacloprid and deltamethrin on associative learning in honeybees under semi-field and laboratory conditions. *Ecotoxicology and Environmental Safety* **57**, 410-419.

Doucet-Personeni C; Halm M P; Touffet F; Rortais A; Arnold G (2003). Imidaclopride utilisé en enrobage de semences (Gaucho®) et troubles des abeilles. Rapport final. *Comité Scientifique et Technique de l' Etude Multifactorielle des Troubles des Abeilles*, pp. 221.

EPPO (European and Mediterranean Plant Protection Organization) (2001). Guidelines for the efficacy evaluation of plant protection products: Side-effects on honeybees (PP 1/170 (3)). *EPPO Bulletin* **31**, 323-330.

EPPO (European and Mediterranean Plant Protection Organization) (2003). Environmental risk assessment scheme for plant protection products. Chapter 10: Honeybees (PP 3/10(2)). *EPPO Bulletin* **33**, 141-145.

European Commission, Health & Consumer Protection Directorate – General (2002). Guidance Document on Terrestrial Ecotoxicology Under Council Directive 91/414/EEC. *European Commission, Health & Consumer Protection*. SANCO/10329/2002 rev 2 final, pp. 39.

Fletcher M; Barnett L (2003). Bee pesticide poisoning incidents in the United Kingdom. *Bulletin of Insectology* **56**, 141-145.

Guez D; Suchail S; Gauthier M; Maleszka R; Belzunces L P (2001). Contrasting effects of imidacloprid on habituation of 7- and 8-day-old Honeybees (*Apis mellifera*). *Neurobiology of Learning and Memory* **76**, 183-191.

Lambin M; Armengaud C; Raymond S; Gauthier M (2001). Imidacloprid-induced facilitation of the proboscis extension reflex habituation in the honeybee. *Archives of Insect Biochemistry and Physiology* **48**, 129-134.

Malet C (2003). Méthode d' Evaluation des Effets de Toxicité Aiguë et a Court Terme des Préparations Phytopharmaceutiques sur l' Abeille Domestique *Apis mellifera* L. Methode N° 230. 1ére Edition. *Commission des Essais Biologiques. Association Française de Protection des Plantes*, pp. 21.

Malon L A; Tregidga E L; Todd J H; Burgess E P J; Philip B; Markwick N P; Poulton J; Christeller J T; Lester M T; Gatehouse H S (2002). Effects of a biotin-binding protein on adult and larval honeybees. *Apidologie* **33**, 447-458.

Maus Ch; Curée G; Schmuck R (2003). Safety of imidacloprid seed dressings to honeybees: a comprehensive overview and compilation of the current state of knowledge. *Bulletin of Insectology* **56**, 51-57.

OECD (1998a). Honeybees, Acute Oral Toxicity Test. *OECD Guidelines for the Testing of Chemicals*, TG No. 213, pp. 8.

OECD (1998b). Honeybees, Acute Contact Toxicity Test. OECD *Guidelines for the Testing of Chemicals*, TG No. 214, pp. 7.

Ohe W v d (2003a). Völkerverluste im Sommer und Herbst 2002. *Deutsches Bienen Journal* 1/2003, 4-6.

Ohe W v d (2003b). Was sind die Ursachen von Völkerverlusten? *ADIZ/db/IF* 9/2003, 8-10

Oomen P A; de Ruijter A; van der Steen J (1992). Method for honeybee brood feeding tests with insect growth-regulating insecticides. *Bulletin OEPP/EPPO Bulletin* **22**, 613-616.

Oomen P A (2001). Honeybee poisoning incidents over the last ten years, as reported by bee keepers in the Netherlands. In: *Hazards of pesticides to bees.*, Avignon 1999. INRA Paris, Les Colloques 98, 129-135.

Otten Ch (2003). Das Völkersterben: Daten und Fakten.*Deutsches Bienen Journal* 8/2003, 1-3.

Ritter W (2003). Warum sterben unsere Bienenvölker? *Deutsches Bienen Journal* 2/2003, 4.

Schur A; Tornier I; Brasse D.; Mühlen W; von der Ohe W; Wallner K; Wehling M (2003). Honeybee brood ring-test in 2002: method for the assessment of side effects of plant protection products on the honeybee brood under semi-field conditions. *Bulletin of Insectology* **56**, 91-96.

SSM (2004). Évaluation du risque environnemental des produits phytopharmaceutiques, revision 6.5, octobre 2004. *Commission d' étude de la toxicité des produits antiparasitaires à usage agricole et des produits assimilés des matières fertilisantes et des supports de culture.* Versailles, pp. 145.

Tasei J N; Pham-Delègue M H; Belzunces L (2003). Registration of systemic insecticides and EPPO guidelines. *Bulletin of Insectology* **56**, 189-191

Thompson H (2001). Assessing the exposure and toxicity of pesticides to bumble bees. In: *Hazards of pesticides to bees*, Avignon 1999. INRA Paris, Les Colloques 98, 197-205.

Thompson H (2003). Behavioural effects of pesticides in bees – their potential for use in risk assessment. *Ecotoxicology* **12**, 317-330

Thompson H M; Brobyn T (2003). Sub-lethal effects to honeybees: their significance and use in pesticide risk assessment. *Bulletin of Insectology* **56**, 198.

Thompson H M; Hunt L V (1999). Extrapolating from honeybees to bumblebees in pesticide risk assessment. *Ecotoxicology* 8, 147-166.

Wittmann D (1981). Bestimmung der LC_{50} von Dimilin 25 WP für Bienenbrut mit einem neuen *Apis*-Larven-Test. *Zeitschrift für angewandte Entomologie* **92**, 165-172.