UK case studies on quantitative risk assessment

Report to:
Ingrid Meakin
Arable Crop Sciences and Pesticide Safety
DEFRA
Cromwell House
Dean Stanley Street
Westminster
London SW1 3JH

Paul Adamson
Pesticides Safety Directorate
Mallard House
3, Peasholme Green
York YO1 2PX

Mark Clook
Pesticides Safety Directorate
Mallard House
3, Peasholme Green
York YO1 2PX

CSL Project Leader:
Dr A D M Hart
Central Science Laboratory
Sand Hutton
York
YO41 1LZ

PROJECT PN0920: UK CASE STUDIES ON QUANTITATIVE RISK ASSESSMENT
Final report

Andy Hart
Central Science Laboratory
July 2002
EXECUTIVE SUMMARY

1. Probabilistic methods of risk assessment take account of the variability and uncertainty that exists in the real world. Current methods are deterministic and take only limited account of variability and uncertainty.

2. This project used case studies to evaluate the feasibility and benefits of applying probabilistic methods to regulatory assessments for pesticides.

3. It was concluded that probabilistic methods can be accommodated within EU regulations and could provide a significantly better basis for decision-making, if they are used appropriately.

4. Basic forms of probabilistic assessment require no more data than current approaches. More sophisticated forms require more data, just as refined deterministic assessments do.

5. Further work is needed to develop and implement probabilistic methods and to achieve a consensus on how they should be used.
LIST OF CONTENTS

EXECUTIVE SUMMARY ............................................................................................................... 3
LIST OF CONTENTS .......................................................................................................................... 4
Background AND OBJECTIVES ...................................................................................................... 5
CASE STUDIES .................................................................................................................................. 5
  Summary of case study 1 – birds in orchards .................................................................................. 5
  Summary of case study 2 – Geese on cereals .................................................................................. 6
USEFULNESS OF PROBABILISTIC RISK METHODS FOR THE UK ............................................. 7
  Advantages relative to current approaches ...................................................................................... 7
  Disadvantages relative to current approaches .................................................................................. 8
  Practicality and cost .......................................................................................................................... 8
  Acceptability to other European Member States .............................................................................. 9
Conclusion ......................................................................................................................................... 9
RECOMMENDATIONS FOR FURTHER WORK .................................................................................. 9
Technology transfer .......................................................................................................................... 11
Acknowledgements ........................................................................................................................... 11
References ........................................................................................................................................ 11
APPENDICES - detailed publications on the case studies:
BACKGROUND AND OBJECTIVES

Current methods for assessing pesticide risks are deterministic – they use fixed values for toxicity and exposure, and produce a single measure of risk (e.g. a toxicity-exposure ratio). In the real world, toxicity and exposure are not fixed, but variable. Furthermore, many aspects of risk assessment involve uncertainty – for example, when extrapolating toxicity from test species to humans or wildlife. Consequently, the effects of pesticides are both variable and uncertain.

Deterministic methods cannot incorporate variability and uncertainty directly. Instead, uncertain or variable factors are fixed to worst-case values, or dealt with subjectively using expert judgement, or simply ignored.

Probabilistic methods can incorporate variability and uncertainty in toxicity and exposure, because they use probability distributions instead of fixed values. Distributions for toxicity and exposure can then be combined, to estimate a distribution for the measure of risk. This provides a much more complete description of the range of risks, which can be very helpful for decision-making. For example, instead of producing a single value for the toxicity-exposure ratio, probabilistic methods can estimate how often the ratio will exceed a regulatory trigger.

Probabilistic methods have been developed over many years and are actively used in other fields such as engineering, insurance, finance, and chemical contamination. There is widespread interest in applying them to pesticides, especially in the USA where the US Environmental Protection Agency (USEPA) has recently committed to implementing them.

CSL was the sole European participant in a major USEPA project (ECOFRAM) that developed plans for introducing probabilistic methods in the USA (ECOFRAM 1999, US EPA 2000). At the same time, CSL carried out this project to explore the potential of these methods for DEFRA. Specifically, the objective of this project was to assess the applicability and feasibility of the methods emerging from ECOFRAM for UK needs and conditions, by developing case studies using existing data.

CASE STUDIES

Two case studies were completed: one concerning exposure of small insectivorous birds to pesticide sprays in orchards, and the other concerning exposure of large herbivorous birds to pesticide sprays on cereals. Detailed accounts of the case studies are presented in Appendices 1-3.

Summary of case study 1 – birds in orchards

The case study focussed on acute risks to birds from the use of the organophosphorus insecticide chlorpyrifos in apple orchards in the United Kingdom, applied by air-blast sprayer at 0.96 kg/ha. The focal species selected for the example was the blue tit (Parus caeruleus), a small insectivorous bird which is common in orchards in the UK.

The study began with a deterministic assessment, conducted in a manner consistent with standard European regulatory practice. Exposure was estimated using a simplified version of the calculation proposed by ECOFRAM.
The probabilistic assessment explored some of the approaches proposed by ECOFRAM, in particular: (a) methods for extrapolating acute toxicity between species, (b) a method for using toxicity data to calculate the tolerances of individual birds, (c) the use of Monte Carlo simulations for propagating uncertainty, (d) progressive refinement of the assessment, by incorporating additional distributions in place of worst-case assumptions, (e) the use of generic field data to estimate distributions of pesticide residues on invertebrates, (f) the use of radio-tracking to estimate empirical distributions for the proportion of food obtained by birds from treated areas.

Progressive refinement was achieved by developing a series of three probabilistic models. Model 1 used distributions to represent uncertainty in the estimate of toxicity, but used a deterministic worst case estimate for exposure. Models 2 and 3 introduced distributions for exposure by replacing fixed values with distributions, first for residues and then bird foraging behaviour. The calculations were executed using the computer programs @Risk (Version 3.5.2 for Excel, ©Palisade Inc.) and Microsoft Excel 97 (©Microsoft Corporation).

As expected, the level of risk decreased markedly from Model 1 to Model 3, because worst-case assumptions regarding residues and bird behaviour were being replaced by distributions based on more realistic data. A similar decrease could be shown by using point estimates for residues and bird behaviour (e.g. means or specified percentiles) in a refined deterministic assessment, but this would not reflect the full range of variability in these parameters nor the uncertainty regarding toxicity.

The detailed results of case study 1 are reported in Appendix 1, which has also been accepted for publication in a forthcoming book (Hart, in press). An earlier version was published in the ECOFRAM report (Anon. 1999).

**Summary of case study 2 – Geese on cereals**

This case study was designed to assist the Environmental Panel of the Scientific Committee on Pesticides in discussing the relative advantages and disadvantages of deterministic and probabilistic approaches. It presents a direct comparison of probabilistic and deterministic approaches to the same worked example.

The example involves a fictitious insecticide, to be used at the rate of 150 g a.s./ha on cereals between growth stage 11 and 32 (early growth, in the spring). For the purposes of the case study, only the acute risk to birds was considered. There is a single avian acute toxicity study with an LD50 of 38 mg a.s./kg bw.

The deterministic assessment started with a tier 1 assessment and then proceeded with the exploration of several refinement options, applying deterministic approaches recommended in a draft Guidance Document currently under development by an EU expert group. All of the information used in refining the deterministic assessment, and the way in which it was used, were taken without alteration from the worked example in the draft Guidance Document.

The probabilistic risk assessment used the same starting point and refinement steps as the deterministic one. It used precisely the same data and assumptions as the deterministic assessment, except that the assumptions were expressed in more quantitative forms. The probabilistic assessment produced toxicity-exposure ratios (TER’s), calculated using exactly the same equations as the deterministic TER’s.

The probabilistic assessment used distributions instead of fixed values for one or more of the inputs. In each step of the assessment, the TER was calculated many times using different values selected at random from the input distributions. Hence the initial output for each
probabilistic assessment was a distribution of TER’s. Summary statistics were taken from these distributions and tabulated for comparison with the deterministic TER’s. The calculations were carried out with the spreadsheet program ‘Excel’, which is installed on most office computers, together with a readily-available add-on probabilistic modelling program called ‘Crystal Ball’ (© Decisioneering Inc.).

The detailed results of case study 1 are reported in Appendices 2 and 3. Appendix 2 is a slightly edited version of the paper submitted to the Environmental Panel (Hart, 2001a). Appendix 3 reports additional results and is in the form of a short poster paper that was presented at the EUPRA workshop in June 2001 and published in Hart (2001b).

USEFULNESS OF PROBABILISTIC RISK METHODS FOR THE UK

The primary objective of this project was to assess the usefulness of probabilistic methods for pesticide risk assessment in the UK. The following sections present conclusions drawn from the case studies 1 and 2, and from other experience gained during the course of the project. Additional discussion of the advantages and disadvantages of probabilistic methods may be found in the report of the EUPRA workshop (Hart, 2001c), together with an extensive collection of summary papers on relevant research.

Advantages relative to current approaches

1. Probabilistic risk assessment can predict how likely impacts are, how often they will occur, and how large they will be, as illustrated in case study 1. These outputs have much more ecological meaning than a TER. However, if the legislation requires outputs to be expressed as TERs, this can also be done within a probabilistic approach as shown in case study 2. In fact, both types of output can be produced side by side, and this will probably be helpful in familiarising users with the new approaches.

2. Other current approaches such as semi-field or field studies provide a measure of actual impact but only under the conditions they are performed in, whereas probabilistic methods can predict impacts over a range of conditions. Data from semi-field or field studies can be incorporated into probabilistic assessments, although this may require additional methods such as Bayesian updating and was not explored in the present study.

3. Probabilistic methods provide a means of quantifying variability and uncertainty in the factors which influence risk. This is much more objective than the current approach of applying uncertainty factors to the TER.

4. Currently there is no consensus on how to relax conservative worst-case assumptions when moving beyond the screening assessment. Because it quantifies uncertainty and variability, probabilistic methods provide an objective way of doing this.

5. Probabilistic methods make it clear which sources of uncertainty and variability are accounted for, whereas this has not been defined for the TER thresholds. Because of this, there is no objective way to adjust the TER thresholds when additional data are provided for the assessment (to reflect the decrease in uncertainty). This problem disappears with probabilistic methods: when additional data are provided the model output quantifies their effect both on the central estimate of risk and on the uncertainty around it.

6. Probabilistic methods facilitate sensitivity analysis which can identify which sources of variability and uncertainty are most influential. They therefore provide an objective way of prioritising data requirements to refine the assessment – a frequent source of controversy with current approaches.
7. Because of the advantages mentioned above, probabilistic methods can provide the ‘appropriate risk assessment’ required by Directive 91/414/EEC whereas this is proving difficult to achieve with current methods.

Disadvantages relative to current approaches

1. It is frequently suggested that probabilistic methods require more data than current approaches. This is not necessarily true. As demonstrated by both Case Study 1 and 2, a simple probabilistic risk assessment can be conducted with the same minimum dataset which is currently used for the initial TER. Furthermore, probabilistic methods get more value out of the current studies because they use more of the data than current methods (e.g. the slope of the dose-response curve, and whole distributions rather than just selected values). Monte Carlo methods (as used in this study) may give misleading results if inappropriate distributions are used, or if important dependencies are overlooked. However, other methods (such as probability bounds analysis) are available which, in theory, are less affected by these problems. The suitability of these methods for pesticide assessments needs to be evaluated (Hart, 2001c).

2. It is possible to carry out very sophisticated probabilistic risk assessments which would require very large amounts of data. However, probabilistic methods provide the means to determine when this sophistication is necessary to answer the regulatory question, so large quantities of data would only be requested when they were actually needed. It would be for the users to decide, whether the increased precision justified the additional cost.

Practicality and cost

1. As already noted, in theory probabilistic methods should not require more data than current approaches, except when this is genuinely needed to enable a regulatory decision. Probabilistic methods will get more value out of those data which are produced.

2. Modern probabilistic software such as @Risk and Crystal Ball is inexpensive. People who are familiar with current risk assessment approaches and with computer spreadsheet packages such as Excel can begin to use @Risk or Crystal Ball in a few hours at their desks, with no formal training. A one-week course provides a good general grounding and awareness of pitfalls etc. More expertise (e.g. a professional statistician) is required for more sophisticated analyses, and to advise on the choice of distributions for new data, although the frequency with which such assistance is required would decrease over time as standard approaches become established and experience grows. It would be advisable to arrange for assessments to be peer-reviewed by suitable experts (e.g. by appointing them to the relevant regulatory committees).

3. Significant research investment is required to develop probabilistic methods and apply them in suitable ways for pesticides. Significant work is also required to develop generic data for key inputs to probabilistic risk assessment, for example data on exposure variables and species sensitivity distributions, although these data are needed anyway for refining deterministic assessments. However, the costs of both types of research could potentially be greatly reduced by sharing them with other countries and/or with industry.

4. Non-probabilistic solutions are unlikely to be a cost-effective alternative. They would incur many of the same costs (because the same technical problems need addressing) but are unlikely to provide effective solutions (for the reasons stated earlier).
5. It has occasionally been suggested that it would be better to resort to a simple hazard assessment with large safety factors. It is true that this would maximise simplicity and transparency for users. However, (a) some type of uncertainty analysis would still be required at the outset, to determine how large the safety factors should be, and (b) to provide protection in all circumstances would probably require such large safety factors that few pesticides would pass.

Acceptability to other European Member States
1. Naturally, there are concerns about the introduction of new methods, partly because of the increased complexity of probabilistic methods and the perception that they may require more data than current approaches.

2. Increasing acceptance of probabilistic methods should be facilitated by the support which already exists in the Commission, the Scientific Committee on Plants, in several EU Member States, in industry, in some of the EPPO sub-panels on risk assessment, and in other countries (especially the USA).

3. Further support has been provided by the European workshop on probabilistic risk assessment for the environmental impacts of pesticides (EUPRA, June 2001) which concluded that probabilistic methods would be useful provided they were implemented and used appropriately (Hart, 2001c).

4. The EUPRA workshop has been followed by a proposal for a 4-year collaborative EU project with 29 partners, aimed at establishing a framework of basic principles for appropriate use of probabilistic methods and providing end-user training and testing for representatives from all EU member states and industry. If approved by the European Commission, the project will proceed in late 2002 or early 2003, and should help to establish wider acceptance of probabilistic methods at European level.

CONCLUSION
1. Probabilistic methods appear to provide the most promising approach for dealing with uncertainty and variability in risk assessment, and could be implemented in ways that are compatible with current UK and EU regulations.

2. Implementation of probabilistic approaches requires significant investment and needs to be done in a way that facilitates acceptance by other Member States. For both reasons it may be attractive to pursue this in collaboration with other countries and stakeholders on a shared-cost basis.

RECOMMENDATIONS FOR FURTHER WORK

The key components required to develop and implement probabilistic methods are:

1. **Identification of assessment endpoints** (what the assessment is to predict) – initially these would probably be very similar to current practice (short, medium and long-term impacts on survival and reproduction) but in the longer term, higher level endpoints might be considered (e.g. at population level).

2. **Understanding mechanisms of exposure and effects** – e.g. models of spray drift, drainflow, deposition on food materials, food intake rates, dose-response relationships and how they vary between species, to name but a few. Some of these things are common
to several taxonomic groups (e.g. fate in soil is relevant to both soil organisms and aquatic organisms), others are specific to one group (e.g. taxonomic patterns in toxicity).

3. **Generation of generic input distributions** which can be used for all pesticides – e.g. distributions of spray deposition, food intakes, toxicity etc. Again some of these are common to several taxa, others are specific. In some cases, distributions can be derived by collating and analysing existing data, but in other cases new data will be required.

4. **Methods of analysing variability and uncertainty** – there are many methods for analysing variability and uncertainty (frequentist vs. Bayesian probability, and Monte Carlo vs. other methods of propagating variability and uncertainty), many ways of using them, and many ways of presenting the results. Inappropriate methods give misleading results, and experts currently differ on what is appropriate. There is a need to establish which methods are appropriate for which purposes, and develop guidance on appropriate ways of using them. In particular, there is a need to evaluate methods which can be used with the levels of data normally available in routine pesticide assessments, and methods for separating variability and uncertainty.

5. **Methods for presenting and interpreting the output** - the outputs of probabilistic approaches express risk in terms of the probability, frequency and magnitude of effects. This makes it easier to assess the ecological significance and acceptability of impacts, but will require new approaches to interpretation. At least in the short term, it will be desirable to present new types of output together with current types of output, to help users become familiarised.

6. **Efficient strategies for refining the assessment** – e.g. some form of tiered approach with clear guidance on how to decide whether refinement is required, and which aspect of the assessment to refine. A key issue here is achieving a good balance between flexibility and predictability.

7. **Risk communication** – it is essential to develop effective ways of communicating risk to decision-makers, to stakeholder organisations and the public. It is not yet clear whether probabilistic methods will help this (because people have some familiarity with the concept of odds, from betting) or hinder it (if the outputs cannot be simplified enough without losing their meaning).

8. **Case studies** – case studies are a vital component of the development process. They keep it focussed on the practical regulatory need, and they provide immediate feedback on the practicality of the methods and the reasonableness of the outputs. Case studies should be carried out for each taxonomic group and assessment endpoint (e.g. acute and chronic effects) for which methods are developed.

9. **Validation** – clearly it is essential to test whether the emerging approaches give appropriate results. This can include testing the validity of individual components (e.g. predicted environmental concentrations) or of the final outputs. It can be done relatively cheaply by comparing case study results with existing data from field studies or monitoring, but such data are limited in quantity, quality and relevance so new, more definitive tests may also be needed.

10. **User support tools** – when consensus approaches begin to take shape it would be useful to consider developing computer-based tools for probabilistic risk assessment, e.g. standard distributions for key input variables, standard exposure scenarios, and standard models for estimating and combining exposure and effects. This would be helpful to users and would also greatly facilitate consistency. At the very least, there will be need for effective written guidance for regulators and industry alike.
TECHNOLOGY TRANSFER

The results of this project have been disseminated in many ways at various stages of the work, to assist DEFRA and other stakeholders in assessing the potential of probabilistic methods. They have been:

- presented in two seminars given to PSD staff at different stages of the project.
- presented at two meetings of the Environmental Panel of the Scientific Committee on Pesticides.
- published in the MAFF Environment R&D Newsletter.
- discussed at ECOFRAM meetings and published in the ECOFRAM report (Anon. 1999).
- presented at SETAC conferences and training courses in Europe and America.
- presented at the EUPRA workshop and published in the workshop report.

Also, the example of blue tits in orchards developed in this project has been adopted for case studies at a SETAC Pellston workshop and for a forthcoming book on probabilistic methods for general chemicals (a preliminary account of this has been published by Moore et al., 2001).

ACKNOWLEDGEMENTS

The author is grateful to ECOFRAM project members, Mark Clook, Dwayne Moore, Scott Ferson and many other individuals for ideas and discussion; to Pierre Mineau, Alain Baril, Dave Fischer and colleagues at CSL for data used in Case Study 1; and to the UK Department for Environment, Food and Rural Affairs for funding.

REFERENCES


Probabilistic assessment of pesticide risks to birds

Andy Hart

Central Science Laboratory, Sand Hutton, York, YO41 1LZ, UK

Probabilistic risk assessment methods offer substantial advantages for assessing the impacts of pesticides on birds. These were explored using a simple example, involving insectivorous birds exposed to chlorpyrifos in apple orchards. Unlike current approaches for assessing pesticide risks, probabilistic methods do not rely on subjective ‘worst-case assumptions’ and safety margins. Instead, probabilistic methods allow the risk assessor to quantify the effects of variability and uncertainty on risk. They estimate the probabilities that different levels of effect will occur. This provides a more realistic and meaningful measure of risk. However, substantial further work is required to implement probabilistic methods and achieve a consensus on how they should be used.

Current methods for assessing pesticide risks are deterministic, in that they use single, fixed values to estimate toxicity and exposure, and produce a single measure of risk (e.g. a risk quotient or toxicity-exposure ratio). In the real world, toxicity and exposure are not fixed, but variable. Furthermore, many aspects of risk assessment involve uncertainty – for example, when extrapolating toxicity from test species to humans or wildlife. Consequently, the effects of pesticides are both variable and uncertain.

Deterministic methods cannot incorporate variability and uncertainty directly. Instead, uncertain or variable parameters are fixed to worst-case values, or dealt with by applying assessment factors (sometimes called safety factors or uncertainty factors) based on expert judgement, or simply ignored.

Probabilistic methods can incorporate variability and uncertainty directly, by using probability distributions instead of fixed values for uncertain or variable parameters. These distributions can then be combined, to estimate a distribution for the measure of risk. This provides a much more complete description of the range of risks, which can be very helpful for decision-making. For example, instead of producing a single value for the toxicity-exposure ratio, probabilistic methods can estimate how often the ratio will exceed a regulatory trigger.

Probabilistic methods have been developed over many years and are actively used in other fields such as engineering, insurance, finance, and chemical contamination. There is widespread interest in applying them to the ecological risks of pesticides, both in North America (1,2,3) and Europe (4).

This paper reports a case study exploring the application of probabilistic methods to assessing pesticide risks to birds. The aim of the case study was to explore some of the potential advantages and disadvantages of probabilistic methods. It was not intended to define how probabilistic methods should be used: this will require further work.

This case study was initiated in conjunction with the US Environmental Protection Agency’s ECOFRAM project (1). It therefore focussed on methods proposed by ECOFRAM, but the approach and conclusions are of general relevance for the implementation of probabilistic risk assessment for pesticides.

**Case study scope and scenario**

The case study focussed on risks to birds from the use of the organophosphorus insecticide chlorpyrifos in apple orchards in the United Kingdom, applied by air-blast sprayer at 0.96 kg/ha. The focal species selected for the example was the blue tit (*Parus caeruleus*), a small insectivorous bird which is common in orchards in the UK.

The scope of the case study was limited to considering acute lethal effects, exposure via the dietary route only, and an exposure period of one day. These limitations were intended to provide a simple illustration of the approaches. A risk assessment for regulatory purposes should take account of additional effects and timescales.
Appendix 1

Deterministic assessment

The study began with a deterministic assessment, conducted in a manner consistent with standard European regulatory practice. Exposure was estimated using the following equation, which is a simplified version of the calculation proposed by ECOFRAM (1):

\[
\text{One day dietary dose (mg chlorpyrifos/kg bodyweight)} = \frac{TFIR \times PT \times FDR \times C}{W}
\]

(1)

where
- \( TFIR \) = Total Food Intake Rate (kg dry weight/day).
- \( PT \) = Proportion of food obtained from Treated area (unitless).
- \( FDR \) = Fresh to Dry weight Ratio (unitless).
- \( C \) = Concentration of chlorpyrifos on food (mg chlorpyrifos/kg wet weight of food).
- \( W \) = Weight of bird (kg).

As is normal in a preliminary assessment, it was assumed that birds obtain all their food in the treated area, so \( PT \) was set to 1. Blue tits are assumed to feed entirely on small insects, so the term \( PD \) that is used by ECOFRAM (1) for ‘proportion of diet’ can be set to 1 and is omitted from the equation. Body weight \( W \) was set to 13.3g, an average value for blue tits (5). Daily total food intake \( TFIR \) was estimated as 3.3g dry weight/day using a standard equation published by Nagy (6). Fresh to dry ratio \( FDR \) was set to an approximate value of 5, and is necessary to convert \( TFIR \) to wet weight. \( C \) was set to 27.8 mg chlorpyrifos/kg food (wet weight), based on the approach suggested by Kenaga (7) as applied in Europe (8) (i.e. assuming that initial residues on the small insects eaten by blue tits are similar to those on plant material of similar surface area / volume ratio, and taking the ‘typical mean’ value from Kenaga’s analysis). These estimates resulted in an estimated one day dietary dose of 34.5 mg/kg.

The one day dose was compared to 32 mg/kg, the median lethal dose (LD50) from a study with the bobwhite quail, a standard test species (9). In Europe this comparison is done by calculating the ratio of toxicity (the LD50) to exposure (the daily dose). This results in a TER of 0.9. This is well below the threshold of 10 which is set down in EU regulations, so the pesticide could not be authorised for sale unless a more refined assessment showed the risk to be acceptable (10). In North America this comparison would be done by calculating the inverse ratio (i.e. exposure/toxicity), which is called the ‘risk quotient’, and would lead to a similar conclusion (further assessment required).
Appendix 1

Probabilistic assessment

The probabilistic assessment explored some of the approaches proposed by ECOFRAM (1), in particular: (a) methods for extrapolating acute toxicity between species, (b) a method for using toxicity data to calculate the tolerances of individual birds, (c) the use of Monte Carlo simulations for propagating uncertainty, (d) progressive refinement of the assessment, by incorporating additional distributions in place of worst-case assumptions, (e) the use of generic field data to estimate distributions of pesticide residues on invertebrates, (f) the use of radio-tracking to estimate empirical distributions for the proportion of food obtained by birds from treated areas.

Progressive refinement was achieved by developing a series of three models. Model 1 used distributions for toxicity, but a worst case estimate for exposure. Models 2 and 3 introduced distributions for exposure by replacing fixed values with distributions, first for $C$ and then $PT$. A summary of the data used in the three models is given in Table I.

The calculations were executed using the computer programs @Risk (Version 3.5.2 for Excel, ©Palisade Inc.) and Microsoft Excel 97 (©Microsoft Corporation).

Model 1 – a basic assessment

Model 1 differed from the deterministic assessment by (a) using a probabilistic approach for estimating toxicity and (b) generating a distribution for the percentage mortality of exposed individuals as output, rather than a fixed estimate of the toxicity-exposure ratio. Model 1 used the same method for estimating exposure as the deterministic assessment, described earlier, resulting in a fixed daily dose of 34.5 mg/kg.

Representation of uncertainty about the LD50

Acute toxicity varies between species for the same pesticide, and is usually only measured for a small number of standard test species. The LD50 for our focal species, the blue tit, is unknown, so the probabilistic assessment used a distribution to represent this uncertainty. Although the toxicity of chlorpyrifos has been tested for a large number of species this study used only the single result for bobwhite, as quoted earlier. This was done to explore the applicability of probabilistic methods to the more normal situation, especially for new pesticides, where toxicity data are available only for one or two species.
Table I. Summary of data used in case study.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Deterministic assessment</th>
<th>Probabilistic Model 1</th>
<th>Probabilistic Model 2</th>
<th>Probabilistic Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD50</td>
<td>32 mg/kg</td>
<td>Distributions estimated from one toxicity study for bobwhite quail and historic data on many pesticides (see text)</td>
<td>Distributions, as Model 1</td>
<td>Distributions, as Model 1</td>
</tr>
<tr>
<td>Probit slope of LD50</td>
<td>Not used in deterministic assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C – Concentration of pesticide on food</td>
<td>27.8 mg/kg</td>
<td>27.8 mg/kg</td>
<td>Distribution based on field data on concentrations in insects (see text)</td>
<td>Distribution, as Model 2</td>
</tr>
<tr>
<td>PT – Proportion of food from treated area</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Distribution based on radio-tracking blue tits in orchards (see text)</td>
</tr>
<tr>
<td>TFIR – Total food ingestion rate</td>
<td>3.3 g/day</td>
<td>3.3 g /day</td>
<td>3.3 g /day</td>
<td>3.3 g /day</td>
</tr>
<tr>
<td>FDR – Fresh to dry weight ratio for food</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>W – body weight of bird</td>
<td>13.3 g</td>
<td>13.3 g</td>
<td>13.3 g</td>
<td>13.3 g</td>
</tr>
</tbody>
</table>
Appendix 1

A distribution for the LD50 was estimated in two steps using methods proposed by the ECOFRAM project (1). First, the mean of the distribution of log LD50s between species was estimated by adding an extrapolation factor to the LD50 for the bobwhite quail:

\[
\text{Mean (log LD50)} = \log \text{LD50}_{\text{bobwhite}} + \text{Extrapolation factor}
\]  

ECOFRAM estimated this extrapolation factor by calculating the ratio between the log of the bobwhite quail LD50 and the geometric mean LD50 for 56 cholinesterase-inhibiting pesticides. The mean value of this ratio for the 56 pesticides was –0.0177, with a standard deviation of 0.38 (1). A normal distribution with this mean and standard deviation was therefore used in Model 1, to take account of uncertainty due to the variation of the extrapolation factor between pesticides.

Second, the standard deviation of the distribution of log LD50s between species was set to 0.428, this being the pooled standard deviation calculated by ECOFRAM (1) for the same set of 56 pesticides using the formula proposed by Luttik and Aldenberg (11). Model 1 therefore used a standard deviation of 0.428, together with means calculated using equation (2), to define a normal distribution representing uncertainty about the LD50 of chlorpyrifos for blue tits.

**Representation of uncertainty about the slope of the dose-response**

The slope of the probit dose-response also varies between species. Unfortunately most of the toxicity studies done with non-standard species do not report slopes, so variation in the slope cannot be analysed satisfactorily with the method used for the LD50. However ECOFRAM (1) presented evidence that the overall variation between species is not much greater than the variation due to other sources (variation within and between tests with the same species). Therefore, as a simple approximation, Model 1 defined a normal distribution to represent uncertainty about the slope using the reported slope (4.6) and its standard deviation (1.2) from the same bobwhite LD50 study for chlorpyrifos, referred to earlier.

**Estimation of percentage mortality**

ECOFRAM (1) identified a range of options for combining exposure and toxicity, including the simple option of generating a distribution of toxicity-exposure ratios in place of the single ratio produced by the deterministic
assessment. However, this study examined another approach recommended by ECOFRAM, which enables risk to be estimated in terms of the predicted mortality, because this was thought to give a more interpretable representation of the expected impact in the field. The method proposed by ECOFRAM (1) for this purpose involved calculating the dose required to kill a random individual bird (its tolerance), estimated as:

\[
\text{Individual tolerance (mg/kg)} = LD50 \times 10^{\left(\frac{z}{\text{slope}}\right)}
\]

where

- \(LD50\) = median lethal dose for species
- \(\text{slope}\) = slope of probit curve for LD50
- \(z\) = standard normal deviate.

This calculation was necessary because the tolerances of individual animals are distributed around the median lethal dose (LD50). Assuming the probit model, the distribution of the base 10 logarithms of the tolerances is Normal with mean = \(\log LD50\) and standard deviation = 1/slope (ECOFRAM report Appendix D1 (1)). The standard normal deviate (\(z\) in equation 3) is a number taken at random from a normal distribution with a mean of zero and standard deviation of 1, and is used to model the distances of randomly chosen individuals from the median tolerance. The effect of this calculation is that each individual is equally likely to fall in any percentile of the distribution of tolerances.

The tolerance for each individual from equation (3) was compared with its daily dose, which in Model 1 was fixed at 34.5 mg/kg (as stated earlier). If the dose exceeded the tolerance, the individual was assumed to have died; if the dose was less than the tolerance, the individual was assumed to survive. This operation was repeated for 1000 individuals, each using a different value for \(z\), and the percentage of individuals dying was calculated.

In order to examine the effects of uncertainty about the LD50 and slope of the dose-response relationship, the calculation for percentage mortality was repeated in 1000 simulations. Each simulation took one value for the LD50 and one value for the slope at random from the distributions defined above, used them to calculate tolerances for 1000 individuals with equation (3), compared the tolerances to the daily dose of 34.5 mg/kg and determined the percentage of mortalities. Altogether the 1000 simulations generated 1000 estimates of percentage mortality, forming a distribution for percentage mortality that represented the effect of uncertainty about the true values of the LD50 and slope for the blue tit.
Appendix 1

Results

An ‘exceedance curve’ was plotted to show what proportion of the 1000 simulations exceeded any given level of mortality (Figure 1). For example, the lines drawn on Figure 1 show that mortality exceeded 90% in 33% of the simulations, and exceeded 56% in 50% of the simulations.

In effect, the 1000 simulations represent 1000 hypothetical species of small insectivorous birds with different LD50s and dose-response slopes. The exceedance curve can be interpreted in two ways. Most simply, it shows the proportion of small insectivorous bird species that will exceed any given level of mortality. Alternatively, if one is interested in a particular species (such as the blue tit) for which the LD50 is unknown, the exceedance curve shows the chance that its mortality exceeds any given level. For example, there is a 1 in 3 chance of mortality over 90%, and a 1 in 2 chance of mortality over 56%.

It is important to note that Model 1 did not include all possible sources of uncertainty regarding toxicity; for example it did not take account of possible differences between the dose response relationship in laboratory studies and the field. Nevertheless, the results are likely to substantially over-estimate the true level of risk because Model 1 used the same exposure estimate as the deterministic assessment, including some potentially conservative ‘worst-case’ assumptions. Models 2 and 3 examined the effects of replacing two of these assumptions with more realistic data.
Figure 1. Probability of exceeding any given level of mortality for blue tits exposed to chlorpyrifos. Results of Model 1, incorporating uncertainties regarding toxicity.

**Model 2 – incorporating variability in pesticide residues**

Model 2 explored the effect of using more realistic estimates of the concentrations of pesticide on the small insects eaten by blue tits (C in equation (1)). The assumed residue of 27.8 mg/kg used in Model 1 was derived from measurements of residues on plant material, as explained earlier. This was replaced in Model 2 by a distribution based on measurements of residues on insects collected within 24 hours after spray applications of pesticides in the USA (see chapter 3.10.6.3 of the ECOFRAM report (1)). Only data for applications to orchards (citrus and apples) were used. These data derived from 4 studies with unnamed pesticides applied at rates between 1 and 6 kg/ha. Measured residues were divided by the corresponding application rate to obtain a ‘residue per unit dose’ (RUD). RUDs ranged from 0.04 to 10.99 mg/kg with a mean of 2.5 mg/kg and standard deviation of 2.8 mg/kg. The RUDs fitted reasonably well to a lognormal distribution, but it was decided instead to use the
non-parametric ‘General’ distribution fitted by RiskView (© Palisade Corporation), which in effect adopts the shape of the empirical distribution. Model 2 therefore sampled 1000 values at random from the distribution of RUDs and then multiplied each value by 0.96 kg/ha, the application rate for chlorpyrifos in the model scenario. The resulting concentrations (mg/kg) were then simply substituted for \( C \) in equation (1), generating 1000 estimates for the daily dose. As the RUD distribution represents variation between orchard plots, the 1000 estimates of daily dose can be regarded as representing 1000 individual blue tits foraging in different plots. Tolerances were estimated for the 1000 individuals and combined with their daily doses to produce an estimate of percentage mortality, using the same method as for Model 1. Finally the whole procedure was repeated 1000 times, with different values for the LD50 and probit slope, to obtain a distribution of percentage mortalities, again using the same method as for Model 1.

**Results**

The results for Model 2 show much lower mortalities than Model 1, because the range of residues is much lower. For example, the exceedance curve for Model 2 shows roughly a 20% chance of mortality exceeding 20% (Figure 2).

It is important to note that although Model 2 incorporated a distribution to represent variability in residues on insects, based on more relevant data than were used in Model 1, there are several important sources of uncertainty regarding this distribution. First, the distribution of RUDs for chlorpyrifos may not be the same as that for the unnamed pesticides to which the data refer. Second, the insects collected in the field studies are unlikely to be representative of those taken by blue tits, because they were collected by pitfall trapping at ground level whereas blue tits feed in the tree canopy. Third, the application method and environmental conditions may differ significantly between the UK and the USA. Fourth, the use of RUDs assumes a linear relationship between application rate and residue, which seems reasonable but is untested. The true residues could therefore be either higher or lower than those used in Model 2. Ideally, these sources of uncertainty should be included in the model; this could be depicted as a broad band of uncertainty around the exceedance curve for Model 2 in Figure 2.
Model 3 – incorporating variability in the foraging behaviour of birds

Model 3 explored the effect of using more realistic estimates of the proportion of their food that blue tits obtain in treated areas ($PT$ in equation (1)). The assumed value of 1 used in Model 1 is an extreme worst-case. This was replaced in Model 3 by a distribution based on measurements in the field. Ideally, these would be measurements of the proportion of food actually taken by blue tits from sprayed orchards, but such measurements are practically impossible to obtain. Instead, it was assumed that the proportion of food taken in orchards is equal to the proportion of time birds spend in orchards, as measured by radio-tracking of 23 blue tits caught in and around apple orchards in the UK (Figure 3). These data were incorporated in the model as a non-parametric ‘General’ distribution fitted by RiskView (© Palisade Corporation). Model 3 therefore sampled 1000 values at random from this General distribution and
simply substituted them for $PT$ in the calculation of daily dose using equation (1). All other aspects of the model remained as in Model 2, including the General distribution for residues in insects.

![Histogram showing distribution of time spent in orchards.](image-url)

**Figure 3.** Distribution of time spent in orchards and therefore potentially exposed to pesticide applications by 23 blue tits, as measured by radio-tracking.

**Results**

The results for Model 3 (Figure 2) show lower mortalities than Model 2, because most blue tits spent only a small proportion of their time in orchards and therefore received lower exposures. For example, the exceedance curve for Model 3 shows a 19% chance of mortality exceeding 1%. The exceedance curve is a flexible way of presenting results as it shows the probability of any given level of mortality. Alternatively, results can be tabulated for specific levels of mortality, or the average mortality over all 1000 simulations can be given (Table II).
Table II. Summary of results for Models 1-3

<table>
<thead>
<tr>
<th>Model</th>
<th>Average mortality</th>
<th>Probability that mortality exceeds 1%</th>
<th>Probability that mortality exceeds 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>53%</td>
<td>0.84</td>
<td>0.73</td>
</tr>
<tr>
<td>Model 2</td>
<td>9%</td>
<td>0.51</td>
<td>0.27</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.4%</td>
<td>0.19</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Care is needed to interpret the results correctly. Model 3 refers to the part of the blue tit population that lives around orchards, because the radio-tracking data used in the model relate to blue tits that were caught in and around orchards. The result in the third column of Table II can therefore be interpreted as showing a 19% chance that mortality of blue tits living around orchards exceeds 1%. Note that the estimated mortalities are estimates of the overall mortality for a large number of birds chosen at random from different orchards. Individual orchards would show higher or lower mortalities, as the distribution of residues used in Models 2 and 3 mainly derives from variation between orchards (see introduction to Model 2, above). If desired, Model 3 could be restructured to produce distributions showing the variation in mortality between orchards. Alternatively, Model 3 could be expanded with additional data to include blue tits living away from orchards, in which case all the estimated mortalities would be lower.

Model 3 can also be used to make statements about other small insectivorous species of birds, provided their foraging behaviour is similar to blue tits (i.e. similar to Figure 3). In this case, the interpretation of the results in Figure 2 and Table II is slightly different: they estimate the proportions of these species that will exceed given levels of mortality. For example, the results indicate that 19% (about 1 in 5) of these species will suffer mortality greater than 1%, and the average mortality taking these species together would be 1.4% (Table II). As before, this refers to those individuals living around orchards.

It is important to note that although Model 3 incorporated a distribution to represent variability in bird behaviour (Figure 3), based on field measurements, there are several important sources of uncertainty regarding this distribution. First, it was derived from a limited sample of blue tits and orchards. Second, radio-tracking was conducted at various times throughout the season when pesticides are applied; it is conceivable that birds use orchards more intensively immediately after applications, either because spraying is a response to high insect levels, or if insects become more easily available to birds after spraying. Third, the proportion of time spent in orchards may not be representative of the proportion of food obtained there. Ideally, all significant sources of uncertainty should be included in the model; this could have been depicted as broad bands of uncertainty around each of the exceedance curves in Figure 2, and wide confidence limits on the results in Table II.
Appendix 1

Model 3 was still very simple. It took account of uncertainty regarding the toxicity of chlorpyrifos to blue tits and variability in residues and bird behaviour but, as noted above, it did not include all the uncertainties affecting these parameters. The parameters TFIR, FDR and W in equation (1) were treated as fixed; in reality they are affected by both variability and uncertainty. Furthermore, the model omitted other potentially important processes including the potential for birds to actively avoid contaminated food (which could reduce exposure) and the potential for exposure by non-dietary routes (which could increase exposure). Finally, it was assumed that all the parameters were independent, which may not be true. All these things could, and ideally should, be included in the model unless there was evidence that they were unimportant.

Discussion

As would be expected, the level of risk decreases markedly from Model 1 to Model 3, because worst-case assumptions regarding C and PT are being replaced by distributions based on more realistic data (Figure 2 and Table II). A similar decrease could be shown by using point estimates of C and PT (e.g. means or specified percentiles) in a refined deterministic assessment, but this would not reflect the full range of variability in these parameters. The ability to incorporate variability is one of the key advantages of the probabilistic approach, and provides a more complete description of risk for the decision-maker.

Probabilistic approaches also provide a means of incorporating uncertainty concerning the model parameters, as illustrated by the treatment of toxicity in the models. Similar methods could be applied to other forms of uncertainty, to provide the decision-maker with a clearer understanding of how they affect the assessment outcome; this could be represented either by putting confidence bands around exceedance curves or confidence limits on numerical estimates of risk. This is another substantial advantage over deterministic methods, which account for uncertainty using simple assessment factors that generally have not been derived from an objective analysis of uncertainty. However, more work is needed to apply uncertainty analysis more comprehensively in probabilistic assessments for pesticides, as is illustrated by the many uncertainties not quantified in this paper.

Another important advantage of quantifying uncertainty, not explored in this paper, is that it reveals which sources of uncertainty have most impact on the risk estimate. This provides an objective basis for specifying which additional data are needed to refine the assessment, and should increase the cost-effectiveness of the assessment process.

It is frequently suggested that probabilistic assessment requires more data than current approaches. In fact, as the example in this paper shows, a limited
probabilistic assessment can be conducted with the same minimum dataset that is currently used for the initial deterministic assessment. Many more uncertainties need to be incorporated, as already mentioned, but some methods of uncertainty analysis are specifically designed for working with limited and/or subjective information, without requiring every parameter to be quantified precisely (4). More work is needed to explore the applicability of these methods to pesticide risk assessment.

Conclusions

Probabilistic risk assessment methods offer substantial advantages for assessing the impacts of pesticides on birds, and may provide a practical solution to many of the difficulties which are encountered using current procedures. However, the present example was exploratory and was not intended as a model for regulatory purposes. Substantial further work is required to implement probabilistic methods and achieve a consensus on how they should be used. Amongst other priorities (4), there is a need to be more comprehensive in incorporating uncertainty, and to explore the applicability of methods designed for use with limited and/or subjective information.

Acknowledgements

The author is grateful to ECOFRAM project members, Mark Clook, Dwayne Moore, Scott Ferson and many other individuals for ideas and discussion; to Monty Mayes for encouragement; to Pierre Mineau for providing toxicity data; to Alain Baril for help with the analysis of uncertainties in Model 1; to Dave Fischer for providing residue data; to colleagues at CSL for the data in Figure 3; and to the UK Department for Environment, Food and Rural Affairs for funding.

References

1. *ECOFRAM Terrestrial Draft Report*; US Environmental Protection Agency Ecological Committee on FIFRA Risk Assessment Methods (ECOFRAM); Available at www.epa.gov/ecotox/, 1999.
2. *ECOFRAM Aquatic Draft Report*; US Environmental Protection Agency Ecological Committee on FIFRA Risk Assessment Methods (ECOFRAM); Available at www.epa.gov/ecotox/, 1999.


Comparison of a deterministic approach to ecotoxicological risk assessment with a probabilistic approach

A paper submitted to
The Environmental Panel of the UK Advisory Committee on Pesticides
(Paper SC 10992 for the meeting of 23 May 2001)¹.

by

Dr Andy Hart
Central Science Laboratory
Sand Hutton
York, YO41 1LZ
UK

Summary
Ecotoxicological risk assessments of pesticides currently conducted under European Directive 91/414/EC are deterministic, i.e. they use single point estimates of toxicity and exposure. Recently, there has been increasing interest in probabilistic approaches. Probabilistic approaches use distributions for toxicity and exposure, instead of point values, to take account of variability and uncertainty.

As part of ongoing discussions on probabilistic approaches, the UK Pesticides Safety Directorate requested preparation of this paper, which was subsequently submitted to the Environmental Panel of the Scientific Committee on Pesticides¹. It presents both probabilistic and deterministic approaches to a single worked example, to enable a direct comparison.

Introduction
Currently under European Directive 91/414/EC, ecotoxicological risk assessments are based on single point estimates of toxicity and exposure. This process produces either a ‘toxicity:exposure ratio’ or ‘hazard quotient’ which is then compared to a regulatory trigger value. If this trigger value is breached, no authorization can be granted ‘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 ‘appropriate risk assessment’ usually involves higher tier toxicity studies or, more occasionally, refined exposure estimates. Both of these refinement steps provide single refined estimates of the potential exposure or effect of a plant protection product.

¹ Note: this paper is slightly different from the version submitted to the Environmental Panel. The summary has been added, and some formal introductory text for the Panel was deleted. The technical material is unaltered.
In recent months the avian and mammalian deterministic risk assessment approach has been developed at the EU level by a small expert working group. This group has produced guidance on how to refine the risk by refining each part of the exposure component (European Commission, 2001). This procedure has therefore moved the risk assessment on slightly as it provides several ways to address the risk, however this approach still does not take full quantitative account of uncertainty and variability associated with key parameters.

At the same time there has been increasing interest in probabilistic approaches to pesticide risk assessment. Probabilistic approaches use distributions for toxicity and exposure, instead of point values, and produces distributions to measure the risk. A number of case studies have been produced by various authors, to explore the application of these methods to pesticides. In many cases, as well as being probabilistic, these case studies have taken account of more factors than is currently normal (e.g. incorporating additional routes of exposure) and expressed risk in new ways (e.g. % mortality rather than a toxicity-exposure ratio).

In order to provide a more direct comparison of the two approaches, i.e. deterministic and probabilistic, the following worked example has been developed. Both approaches are applied in parallel to the same basic toxicity dataset and the same exposure assumptions. Both produce toxicity-exposure ratios as output, calculated using the same equations. It is hoped that, by comparing the two approaches in this direct way, their respective advantages and disadvantages can be assessed.

**Deterministic approach**

The deterministic assessment presented in this paper is taken from the draft Guidance Document on Risk Assessment for Birds and Mammals, where it is presented as a worked example (European Commission, 2001).

The example involves a fictitious insecticide, to be used at the rate of 150 g a.s./ha on cereals between growth stage 11 and 32 (early growth, in the spring).

For the purposes of this paper, only the acute risk to birds is considered. There is a single avian acute toxicity study with an LD50 of 38 mg a.s./kg bw.

The deterministic assessment starts with a tier 1 assessment and then proceeds with the exploration of several refinement options, applying the approaches recommended in the Guidance Document. All of the information used in refining the deterministic assessment, and the way in which it is used, are taken without alteration from the worked example in the Guidance Document.

**Probabilistic approach**

The probabilistic risk assessment uses the same starting point and refinement steps as the deterministic one. It uses precisely the same data and assumptions as the deterministic assessment, except that the assumptions are expressed in more quantitative forms which are explained below. The probabilistic assessment produces toxicity-exposure ratios (TER’s), calculated using exactly the same equations as the deterministic TER’s.

The probabilistic assessment uses distributions instead of fixed values for one or more of the inputs. In each step of the assessment, the TER is calculated many times using different values selected at random from the input distributions. Hence the initial output for each probabilistic assessment is a distribution of TER’s. Summary statistics are taken from these distributions and tabulated below for comparison with the deterministic TER’s.
The calculations were carried out with the spreadsheet program ‘Excel’, which is installed on most office computers, together with a readily-available add-on program called ‘Crystal Ball’. The calculations are reproducible: other people would obtain the same results using the information provided below and a few details of the program settings (see Appendix).

**First tier risk assessment**

**Deterministic assessment**

According to the draft Guidance Document, the appropriate indicator species for this pesticide is a 500 g herbivorous bird with a food intake rate of 328 g/day wet weight. The standard deterministic calculations for acute exposure of this indicator species and the resulting TER are shown in Table 1.

**Interpretation**

The deterministic TER is less than the threshold value of 10 specified 91/414/EC Annex VI, so a refined assessment is required.

**Table 1. First tier risk assessment, comparing deterministic and probabilistic results.**

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Toxicity</th>
<th>Exposure</th>
<th>Deterministic TER</th>
<th>Probabilistic TER’s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rel. Food intake</td>
<td>RUD</td>
<td>Appl. rate</td>
<td>ETE</td>
</tr>
<tr>
<td>Acute</td>
<td>38</td>
<td>0.66</td>
<td>142</td>
<td>0.15</td>
</tr>
</tbody>
</table>

**Key to Table 1:**

Toxicity – acute avian LD50, mg a.s./kg body weight
Rel. Food intake – food intake per unit body weight (kg/kg)
RUD – Residue per Unit Dose on short grass, mg a.s./kg wet weight grass
Appl. Rate – application rate in kg a.s./ha
ETE – exposure estimate, calculated as = Rel. Food intake x RUD x Appl. Rate
Deterministic TER – ratio of Toxicity/ETE
Probabilistic TER’s: 10th and 90th percentiles of distribution of TER, and the percentage of the TER distribution which falls below the threshold of 10.

**Probabilistic assessment**

The probabilistic assessment uses exactly the same calculation and data as the deterministic one, except for the pesticide concentration factor (RUD). In practice, pesticide residues are variable. The value used for RUD in the deterministic assessment (142) is actually the 90th percentile of a distribution estimated from historical data on other pesticides (as explained in Appendix II of the draft Guidance Document). It is therefore logical to use the same distribution in the probabilistic assessment, i.e. a lognormal distribution with a natural mean of 76 and a 90th percentile of 142. However, the probabilistic assessment uses the whole distribution, not just the 90th percentile. It therefore gives a distribution of TER’s. Three statistics from the probabilistic TER distribution are given in Table 1: the 10th percentile, the 90th percentile, and the percentage of TER’s which are below the threshold of 10.

The 10th percentile of the probabilistic TER is exactly equal to deterministic TER (2.7, see Table 1). This is because both results correspond to the 10th percentile of the RUD distribution.
Interpretation

The draft Guidance Document implies that the distribution of RUD values represents spatial variability. The spatial units of this variability are not specified, but can be inferred as relating to variation between fields (i.e. pesticide residues vary from field to field). The 10th percentile of the probabilistic TER (2.7, Table 1) can therefore be interpreted as the TER which will occur on a field with 10th percentile residues. In other words, 10% of fields have a TER of 2.7 or less.

It is important to bear in mind that these results refer to specifically to birds which feed entirely on the treated crop, as this is assumed by both the deterministic and probabilistic calculations (this assumption is relaxed in refined assessments, see later).

The 90th percentile TER is 14.1 (Table 1). Taking this together with the 10th percentile gives a measure of the range over which the TER varies. In this example, if birds feed entirely on the treated crop, then on 80% of fields they will experience a TER of between 2.7 and 14.1.

The probabilistic assessment can also estimate the percentage of the TER distribution which falls below the Annex VI threshold value of 10. In this example, if birds feed entirely on the treated crop, the TER will be less than 10 on 77% of fields (Table 1). Conversely, the TER will be above threshold on 23% of fields. This way of expressing the risk would presumably lead one to the same conclusion as the deterministic assessment – that further refinement of the assessment is required.

In summary, the deterministic assessment is simpler to calculate but only gives a single TER. In the real world, the TER varies from field to field. Probabilistic assessment describes the distribution of risk, and estimates the percentage of TER’s which fall below the critical regulatory trigger of 10.

Refined risk assessment - concentration

Deterministic assessment

In the above first tier risk assessment, the concentration on treated food, in this case short cereal shoots, was estimated using historical data for other pesticides.

In trying to refine the concentration estimates, it is possible to examine the residues data submitted by the Notifier. This is usually presented in Section 4 of the dossier and section 6 of the monograph. For the acute risk assessment, data are required on the initial residues immediately after application. Unfortunately, these type of data are rarely available for new products with the above use pattern. In this example it is assumed the Notifier has generated residue data by sampling the crop immediately after application. These samples dates produced the following results: average residue is 8 mg/kg, 90th percentile residue is 15 mg/kg.

When the 90th percentile of the measured residues is used to re-calculate the deterministic assessment, the resulting TER is 3.8 (Table 2).

Interpretation

The deterministic TER is still below the Annex VI threshold of 10. Therefore, the acute risk has not been adequately addressed by the refined residue data, and further refinement is required.
Table 2. Refined risk assessment, using measured values for pesticide residues.

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Toxicity</th>
<th>Exposure</th>
<th>Deterministic</th>
<th>Probabilistic TER's</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Rel. Food intake</td>
<td>Init. Conc.</td>
<td>ETE</td>
</tr>
<tr>
<td>Acute</td>
<td>38</td>
<td>0.66</td>
<td>15</td>
<td>9.9</td>
</tr>
</tbody>
</table>

Key to Table 2:
Init. Conc. – initial concentration of pesticide (90th percentile of measured values)
ETE – exposure estimate, calculated as = Rel. Food intake x Init. Conc.
Other terms as in Table 1.

Probabilistic assessment

As before, the probabilistic assessment uses the same data as the deterministic assessment but uses a whole distribution for residues instead of just the 90th percentile. The example in the draft Guidance Document did not specify the shape of the distribution, so we chose one which fitted the basic information provided. This resulted in residues ranging between 0 and about 27 mg/kg.

The results of the probabilistic assessment are shown in Table 2.

Interpretation

The distribution of TER’s is wider than in the first tier assessment, as judged by the range between the 10th and 90th percentiles. The proportion of TER's falling below the Annex VI threshold of 10 is 60% (compared to 77% previously).

Following the same reasoning as before, the result can be interpreted as meaning that the TER will fall below the Annex VI threshold on 60% of fields.

It is important to note that the assessment still refers to birds feeding entirely on the treated crop. While this may happen sometimes, it represents a worst case. It would therefore be helpful to refine the assessment using more realistic assumptions about feeding behaviour.

Refined risk assessment – feeding behaviour

Deterministic assessment

Proportion of diet obtained in treated area (PT)

In the first tier risk assessment it is assumed that individuals obtain all their dietary requirements from the treated area, i.e. the proportion of diet obtained in treated area (PT) = 1. In reality it would extremely rare that this was always the case. Using information outlined in the draft Guidance Document, it may be possible to reduce the default PT value of 1 to a more realistic figure. In order to do this data from radiotracking studies may, if they are available, help, however it is appreciated that these will be rarely available. Therefore, an alternative option would be to carry out an appropriate literature search to try and determine the proportion of the diet that may be obtained from the treated area. However, before doing this,

---

2 For the purposes of this paper we assumed the measured residue data followed a Weibull distribution. The parameters of the Weibull distribution (location 0, 50th percentile 7, 90th percentile 15) were chosen such that the mean (8) and 90th percentile matched the values quoted in the Guidance Document.
key species that may be exposed should be identified. With the above scenario the major
types of birds of concern are geese.

There is much evidence of geese grazing short grass and cereal shoots, for example Greylag
goose, Brent goose and Canada goose are identified as an appropriate indicator species. There
is currently a lack of information to indicate the proportion of diet obtained from the treated
area, i.e. there is a lack good quality quantitative data to indicate whether geese spend all day
and hence obtain all there food from the treated area. However, it is known that the areas
where the product will be used are predominantly coastal and where geese feed on arable field
they only do so for the part of the day the tide is in. There is no quantitative data to indicate
the exact time that this is, however it is estimated that out of a maximum 8 hour feeding
period, geese are only on arable crops for 4 hours. This means that the proportion of diet
obtained in treated area (PT) can be reduced from 1 to 0.5.

The example in the Guidance Document emphasises that this refinement step is only possible
due to knowledge regarding the likely use pattern of the product and the behaviour of geese in
that area.

**Proportion of different food types in the diet (PD)**

In the first tier it is assumed that all the food consumed by the bird is young cereal shoots. In
order to refine this factor, it would be ideal to have data on the likely composition of birds
diets in the field when the pesticide was being applied. Unfortunately specific data of this type
is rarely available, however, data are available on what birds eat at different times of the year.

Data from the public domain on the Brent goose indicated that it will eat grass and maintain
its body weight, however this information came from a study where Brent geese had been
kept on pasture so therefore it is of limited use for refining the risk assessment. Data on
Canada goose indicated that between the months of October and March grasses made up 33%
of the composition of the gizzards of Canada goose. It should be noted that this was 33% by
volume and that the study was conducted in the USA. Data on the Greylag goose indicated
that grass occurred in 73% of the stomachs of greylag goose sampled between November and
February. Further data indicated that between the months of March and May, grass made up
96% of the stomach contents of sampled geese. Between the months of September and
November grass and cereal seedlings made up 17% of the stomach contents, whilst between
December and February grass made up 60%. The remainder of the diet consists of grain,
potato spoil and turnips. These studies were carried out in the UK.

From the above, it can be seen that accurate refinement of PD is difficult and the published
data available is only likely to help on a qualitative basis. For example, the above data
indicate that grass (and it is assumed cereal shoots) will be consumed by geese and it will
make up a significant proportion of the diet. From the data presented on the greylag goose, it
can tentatively be concluded that for the time period of interest (i.e. February to April) grass
made up between 60 and 96% of the diet. The remainder of the diet is made up of harvest
spoil and grain left on the surface post harvest. These food items are assumed to have no
pesticide residues.

**Interpretation**

The estimates of PD and PT derived above were used to re-calculate the deterministic TER
(Table 3). Note that the deterministic assessment now produces a range for the TER,
corresponding to the range estimated for PD. The lower estimate (8.0) is below the Annex VI
threshold of 10, but the upper estimate (12.8) is above it.
On the basis of this result, the example in the draft Guidance Document concludes that on balance the acute risk to birds has been adequately addressed. No further explanation is given.

Table 3. Refined risk assessment, using information on bird feeding behaviour.

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Toxicity</th>
<th>Rel. Food intake</th>
<th>Exposure</th>
<th>Deterministic TER</th>
<th>Probabilistic TER's</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Init. Conc.</td>
<td>PD</td>
<td>PT</td>
<td>ETE</td>
</tr>
<tr>
<td>Acute</td>
<td>38</td>
<td>0.66</td>
<td>15</td>
<td>0.6 – 0.96</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Key to Table 3:

PD – proportion of diet which is grass
PT – proportion of diet obtained from treated area
ETE – exposure estimate, calculated as = Rel. Food intake x Init. Conc. x PD x PT
Other terms as in Tables 1 and 2.

Probabilistic assessment

The probabilistic assessment in Table 3 uses essentially the same assumptions as the deterministic one.

Based on a discussion of literature data, the example in the draft Guidance Document concludes ‘tentatively’ that grass comprises between 60 and 96% of the diet. The probabilistic assessment therefore assumes that PD lies between 0.6 and 0.96. As the Guidance Document did not indicate a preference for any particular value within this range, the probabilistic assessment assumes that all values between 0.6 and 0.96 are equally likely.

Although the example in the draft Guidance Document makes it clear that PT is uncertain, it ultimately uses a fixed value of 0.5. Therefore PT is also fixed at 0.5 in the probabilistic assessment at this stage.

Interpretation

Results are shown in Table 3. The probabilistic assessment gives a much wider range of TER’s (80% lie between 9.7 and 71.5) than the deterministic one (range 8.0 – 12.8). This is because the deterministic TER’s refer only to those fields with 90th percentile residues whereas the probabilistic assessment refers to all fields.

The 10th percentile of the probabilistic TER’s falls between the two deterministic TER’s (as would be expected from the input data) and close to the Annex VI threshold of 10. Eleven percent of the TER distribution falls below the threshold. The deterministic assessment shows that TER may fall below 10 but not how often this will happen; the probabilistic assessment does both.

Following the same reasoning as before, the result can be interpreted as meaning that the TER will fall below the Annex VI threshold on 11% of fields.

However, it is clear that there is significant uncertainty about the assumptions concerning PD and PT – the former is described in the draft Guidance Document as ‘tentative’, the latter as an estimate without a quantitative basis. Therefore, it may be useful to examine the assumptions a little more closely, before moving to a judgement of whether the risk is acceptable.
Examination of uncertainty affecting PD and PT

Uncertainty affecting the estimates of PD (proportion of grass in the diet) and PT (proportion of diet obtained in treated area) was examined by conducting a further refinement of the probabilistic assessment.

As argued earlier, values between 0.6 - 0.96 seem most likely for PD. However, values outside this range are certainly possible: a value of 0.33 was reported for Canada geese. In principle, the average value for PD could vary between 0 and 1, but seems unlikely to be very close to zero (no grass in diet). A customised distribution is used to represent this in the probabilistic assessment, with the probability rising linearly from zero for PD = 0 to a maximum at PD = 0.6, remaining level until PD = 0.96 and then decreasing linearly to zero at PD = 1.

The draft Guidance Document bases the estimate of PT on qualitative information that ‘where geese feed on arable fields they only do so for the part of the day the tide is in’. It was estimated that this occurs for half of the maximum feeding period of 8 hours. Even if the basic premise is accepted, the argument leading to the estimate of 0.5 is affected by many sources of variability and uncertainty. Examples include: the location of intertidal feeding sites is not specified; the period for which intertidal feeding sites are available will vary widely, following the tide cycle (which itself varies geographically and temporally); and the feeding rate of geese may differ between arable fields and intertidal feeding areas.

It is conceivable that, for individual geese on a particular day, PT may vary over the entire range from 0 to 1. If we consider a local population of geese, it seems less likely that the average value for PT is close to 0 or 1. However, one would hesitate to rule out an average of, say, 0.75 (e.g. if intertidal feeding sites were available for only 2 of 8 hours feeding) or, conversely, 0.25. Therefore we assume that the average value for PT lies between 0.25 and 0.75, with the most likely value around 0.5. A triangular distribution was used to represent this in the probabilistic assessment.

We must bear in mind that these estimates of the uncertainty in PD and PT are subjective. Nevertheless, if we examine how much they affect the result of the risk assessment, it should help us judge whether the uncertainty is likely to matter. The results are shown in Table 4.

Table 4. Refined risk assessment. The deterministic calculations are unchanged from Table 3, but the probabilistic assessment now takes account of uncertainty in the estimates of PD and PT. This changes the probabilistic TER’s and also provides confidence limits for them. See Tables 1-3 for key.

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Toxicity</th>
<th>Rel. Food Intake</th>
<th>Init. Conc.</th>
<th>PD</th>
<th>PT</th>
<th>ETE</th>
<th>Deterministic TER</th>
<th>Probabilistic TER's</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute</td>
<td>38</td>
<td>0.66</td>
<td>15</td>
<td>0.6 – 0.96</td>
<td>0.5</td>
<td>3.0 – 4.8</td>
<td>8.0 – 12.8</td>
<td>11.1* (6.6 – 40)</td>
</tr>
</tbody>
</table>

* values in brackets are approximate 95% confidence limits for the probabilistic TER results

Interpretation

At first sight the probabilistic results in Tables 4 suggest a slightly lower risk than in Table 3. This is because, in accounting for uncertainty in PD, we reduced its average value compared to the previous assessment.
Appendix 2

However, it is important to note that the assessment is, in effect, averaging the risk for different values of PD and PT. The distributions for these variables represented (at least partly) our uncertainty or lack of knowledge about PD and PT. In the real world, PD and PT may both be towards the lower end of our distributions, in which case the real risk would be lower. Or they may both be towards the upper end, in which case the risk would be higher.

We can explore this by making the probabilistic assessment hierarchical. This enables us to put approximate confidence limits on the results, showing how much lower or higher they could be depending on the uncertainty in PD and PT. 95% confidence limits are shown in brackets in Table 4.

The 95% confidence limits for the percentage of TER’s falling below the Annex VI threshold are 0% and 28%. This range represents the effect of our uncertainty about PD and PT on our ability to say how often the TER breaches the Annex VI threshold.

Following the same reasoning as before, the result can be interpreted as meaning that the TER is likely to fall below the Annex VI threshold on about 10% of fields, but there is a small chance that the percentage is as low as zero or as high as 28%.

The fact that the 95% confidence intervals are quite wide, shows that the uncertainty about PD and PT has a substantial influence on the risk. If it were decided that the uncertainty was too high for a regulatory decision to be made with confidence, then an obvious option would be to reduce the uncertainty by obtaining better information about PD and PT. This might take several forms: a more detailed evaluation of existing data, modelling, or field studies to measure the feeding behaviour of geese in relevant habitats.

Limitations

There are a number of limitations which affect the two approaches (deterministic and probabilistic). These have some effect on the comparison between the two approaches, and should definitely be borne in mind if they were being used in a real assessment.

When calculating the TER’s, no allowance is made in either approach for differences in toxicity between the test species and geese. The draft Guidance Document assumes that these factors are addressed by the use of 10 as the threshold value in Annex VI. In fact, as the draft Guidance Document points out, a factor of 10 is likely to cover 75% of species. Uncertainty due to between-species variation could be included in the probabilistic assessment (and would broaden the distribution of TER’s obtained), but then it would be inappropriate to compare the result to the Annex VI threshold of 10.

In this paper, both the deterministic and probabilistic assessments used population averages for input variables. The value used for relative food intake is an estimated population average; the estimates used for PD and PT are implicitly, and the LD50 is the dose required to kill half the population. Consequently, both assessments estimate the risk to an hypothetical ‘average’ individual. Taking account of individual variation in these variables would produce a lower TER in the deterministic assessment and broaden the distribution of TER’s obtained in the probabilistic assessment.

There are many more limitations which affect both approaches, most of which could only be addressed by substantially complicating the underlying model or expanding its scope (e.g. to incorporate additional routes of exposure). Part of the reason these factors are not considered in current deterministic assessments is because they are not sufficiently understood, i.e. they are uncertain. It might be easier to incorporate these factors using probabilistic approaches as they are specifically designed to handle uncertainty.

One limitation specific to the probabilistic assessment in this paper is that it assumes the initial concentration, PT and PD are independent of one another. In practice PD and PT are
unlikely to be independent, because birds probably obtain different foods from treated and untreated habitats. If there were a positive correlation between PD and PT (which seems plausible), then the probabilistic assessment will underestimate the frequency of high exposures. To protect against this, the probabilistic assessment could be repeated with another method which does not assume independence.

Conclusion
This paper provides a direct comparison of probabilistic and deterministic approaches to the same worked example, taken from the draft report of the EU expert group. It was prepared to assist the Environmental Panel of the Scientific Committee on Pesticides in discussing the relative advantages and disadvantages of deterministic and probabilistic approaches. In order to avoid prejudicing those discussions, no opinions are expressed here.

Acknowledgement
The author is grateful to Mark Clook of PSD for valuable discussions during the preparation of this paper, and for commenting on previous drafts.

Reference

Appendix
The following details, together with the information provided in the paper, should allow reproduction of the probabilistic results.

The probabilistic assessments were conducted with Crystal Ball 2000 (version 5.1) and Microsoft Excel 97 (version SR-2). The random number seed in Crystal Ball was set to 999.

Non-hierarchical assessments used Latin Hypercube sampling and were run for 50,000 iterations.

The hierarchical assessment (to obtain the confidence limits in Table 4) used Monte Carlo sampling. The inner loop varied concentration and was run for 1000 iterations, the outer loop varied PT and PD and was run for 250 iterations.
Objective
To take account of uncertainties affecting acute risks to birds using probabilistic methods, while staying as close as possible to current EU procedures and using the same data.

Deterministic approach
The deterministic example was taken from the Draft Guidance Document on Risk Assessment for Birds and Mammals (European Commission, 2001):

- fictitious insecticide, avian LD50 38 mg/kg
- applied to cereals in the spring at 150 g/ha
- indicator species for risk assessment - 500g herbivorous bird (e.g. goose) eating 328g wet weight per day

Acute exposure (ETE) was calculated as:

\[ ETE = \text{FIR} \times C \times PD \times PT \]

Where:

- \( ETE \) = acute exposure mg a.s./kg bw/day
- \( \text{FIR} \) = Food intake rate in g/g bw/day (\( = 328/500 \))
- \( C \) = Concentration of pesticide in contaminated food
- \( PD \) = Proportion of diet made up by this food type
- \( PT \) = Proportion of this food which is obtained in treated area

The output of the deterministic approach was a TER (toxicity/exposure ratio):

\[ \text{TER} = \frac{\text{LD50}}{ETE} \]

Probabilistic approach
The probabilistic approach was used to take account of variability in C, and uncertainty in PD and PT.

C - residues on treated cereals
The deterministic approach estimated C as:

\[ C = \text{RUD} \times \text{AR} \]

Where:

- \( \text{RUD} \) = Residue per Unit Dose (142)
- \( \text{AR} \) = application rate of pesticide (0.15 kg/ha).

The RUD of 142 is the 90th percentile of a lognormal distribution based on the residue data of Fletcher et al. (1994).

To be consistent with this, the probabilistic approach used the same lognormal distribution.

PD - proportion of food which is treated
The Guidance Document example referred to published studies which showed that:

- captive geese can survive on 100% grass
- the mean proportions of grass in the gut contents of wild geese in several studies were 33%, 73%, 60% and 96%, with the last two figures considered most relevant

On this basis, the deterministic approach calculated two TERs, one based on PD = 0.6 and the other on PD = 0.96.

For the probabilistic approach, the data were interpreted as showing mean PD could lie anywhere between 0 and 1 but was most likely to lie between 0.6 and 0.96 (Figure 1).

Results
The deterministic approach produced two values for the TER: 5.6 (for PD = 0.96) and 9 (for PD = 0.6). These values are both below the EU regulatory threshold of 10.

The probabilistic approach used 2D Monte Carlo (in the program Crystal Ball) to estimate distributions for the TER. From these distributions I estimated the proportion of fields on which the average goose has a TER under 10.

For the probabilistic approach it was considered that the mean for PT could lie anywhere between 0.25 and 0.75, but was most likely to be around 0.5 (Figure 2).

A simple generic probabilistic approach?
For other pesticides with the same use pattern, the exposure scenario is the same. Only the LD50 and application rate are required to read the probabilistic estimates from Figure 3. Obviously it would be essential to define the inputs very carefully before using generic scenarios in a routine way.