NDAWG

National Dose Assessment Working Group

NDAWG/1/2005

An Overview of Uncertainty in Radiological Assessments

National Dose Assessment Working Group: Subgroup on Uncertainty and Variability

Contents

1 Introduction

This document has been written by the Uncertainty and Variability Sub-Group of the National Dose Assessment Working Group (NDAWG) as part of the NDAWG's aim to facilitate the development of coherent and transparent methods for the assessment of radiation doses to members of the public.

The document is intended to provide those considering carrying out assessments that take account of uncertainty and variability with an overview of the issues involved and warning of the potential pitfalls. It is not intended to give detailed guidance to those carrying out such assessments. References are given to more detailed information where appropriate. Although the document is concerned with the assessment of public doses from routine aerial and aquatic discharges of radioactive material, examples and lessons from other fields (such as the disposal of solid radioactive waste) have been included, where relevant.

Most assessments of the doses received by members of the public from routine discharges of radioactivity have been carried out deterministically, ie, single values of the dose received have been calculated using point values for the input parameters required for the calculation. No attempts have been made to calculate the ranges of the doses that could be received by members of the public in such assessments. This is in line with the guidance on principles for the assessment of prospective public dose issued by the Environment Agencies together with the FSA and NRPB (EA, 2002). Principle 12 of the guidance states: 'Where the assessed critical group dose exceeds 0.02 mSv/y, the uncertainty and variability in the key assumptions for the dose assessment should be reviewed.' The guidance does not call for full probabilistic uncertainty analyses to be carried out. In most cases, the guidance suggests that a qualitative or semi-quantitative review should be undertaken. A workshop held to discuss the implications of variability on setting limits for authorised discharges (Walsh et al, 2000) concluded 'Studies of variability and uncertainty in critical group dose assessments have an important role in improving understanding of the dose assessment process. However such studies should not be carried out routinely as part of the authorisation process.'

The management of uncertainty is not just a technical exercise. It is difficult to quantify uncertainty and in most cases the quantification of uncertainty is itself uncertain. Uncertainty is, in part, socially constructed and its assessment includes subjective judgements.

Those carrying out such assessments should consider a number of issues before commencing:

- Who is the assessment being carried out for?
- What decisions will be made based on the assessment? Will inclusion of uncertainty and variability improve those decisions?
- Will incorporation of uncertainty and variability improve the assessment?
- What are the major sources of uncertainty and variability? How will these be kept separate in the analysis?
- What are the time and resource implications of including uncertainty and variability? Is this effort justified?
- Are the necessary skills and experience available?
- What methods of incorporating uncertainty and variability are to be used? Have the strengths and weaknesses of those methods and other methods that could potentially be used been evaluated and compared?
- How will the results be communicated to the public and decision-makers?

Consideration of these issues is not straightforward and expert advice and guidance should be sought before commencing the assessment.

2 Terms

There is no commonly accepted system for characterisation of uncertainty. Uncertainty is difficult to define and there are many classification systems in use. However, for the purposes of this report, the following defines the term's sensitivity, uncertainty and variability and includes a brief discussion of the different types of uncertainty. More information on these topics can be found in IAEA, 1989. In the remainder of this report the term 'uncertainty' is used in a very loose sense.

Sensitivity analysis: Sensitivity analysis is the study of the effect of changes in input values on the output from a model. This can be done by either varying a single parameter at a time to see the effect on the output or by varying a suite of input parameters simultaneously. It enables the user to identify the parameter or groups of parameters to which the model is most sensitive and, as such, can be used to direct research programmes.

Uncertainty: Uncertainty measures the lack of knowledge of the system under investigation, which in radiation dose assessment terms will relate to how well doses of interest can be estimated. For example, how well are the parameter values in a calculation of dose known? If further investigations can reduce the uncertainty in these parameter values by increasing the accuracy and precision with which they are known then this is epistemic or so called Type B uncertainty (IAEA, 1989). This is applicable to a parameter that is thought to have a well-defined value, but due to inevitable experimental difficulties there is some uncertainty about that value. In many dose assessment applications, a detailed knowledge of the processes involved is not required and a simpler parametric representation can be employed that captures the essential details. This adds modelling uncertainty by simplifying relationships but allows the average parameter value to represent the process adequately. For example, the transfer coefficient for a radionuclide between cow's intake and milk is uncertain and is determined by a multiplicity of physiological processes, but a single average value could be determined by a suitable experiment. In principle, carrying out further investigations to improve knowledge can reduce uncertainties. However, uncertainty is not simply the absence of knowledge. Uncertainty can still prevail in situations where further information becomes available. Also, new information can either decrease or increase perceived uncertainty by revealing the presence of complexities previously unknown or poorly understood. In other words, more knowledge does not necessarily imply less uncertainty. Though it may reveal uncertainties that were previously hidden, it may not help to resolve them.

The sources of uncertainty in the predictions from models can be grouped into broad categories as illustrated in figure 1 and outlined here:

- Measurement uncertainty. This is the uncertainty in the field or laboratory data on which models are based, eg, lack of precision, inaccuracy, sampling and analysis errors.

- Parameter value uncertainty. This is caused by not knowing the most appropriate values to select for the various parameters of a model. Lack of data that could have been collected but have not, or data that may be practically immeasurable (too expensive and resource intensive). There may be conflicting evidence or different data sets available. Parameter value uncertainty can also arise when the parameters of a model are not closely related to measurable quantities, as this can result in ambiguities of interpretation of available data.

- Conceptual modelling uncertainty. This is the uncertainty associated with forming a coherent representation of the processes involved in the system being modelled based on the available data. General considerations of simplicity, adequacy and underlying physical principles will govern the selection of an appropriate model where a choice might exist. Model structural error is often overlooked when performing uncertainty analysis with the implicit assumption that the model is a good 'fit' to the environment that it purports to represent. Many environmental models perform badly against observations (Beven, 2002).

- Computational uncertainty. This arises from the representation of the selected model in computational terms. It includes the use of simplifying assumptions, discretisation and numerical methods of solution.

- Scenario uncertainty. This concerns uncertainties which cannot be adequately depicted in terms of chances or probabilities, but which can only be specified in terms of (a range of) possible outcomes. In dose assessment this source of uncertainty includes the need to make assumptions about the habits of animals in the food chain and human behaviour.

- Ignorance. Although not a manageable category of uncertainty, the recognition of ignorance allows for the fact that "we don't know what we don't know" and that there are inherent limitations to the reduction of uncertainty.

Variability: Variability refers to the actual differences that occur both in transfer in different environments and between individuals within a group; for example, differences in how much of a particular food is eaten or where individuals spend their time. Variability is also known as objective uncertainty, aleatory uncertainty or Type A uncertainty (IAEA, 1989) and can refer to those parameters for which there is no single correct value, and only a probability distribution of values can be specified. For example, the atmospheric conditions at the time of a future accident cannot be specified in advance, and this uncertainty is type A. It is not possible to reduce variability by improving knowledge.

The difference between uncertainty and variability: It is possible that some parameters have both an extrinsic uncertainty due to the limitations of measurements and models and an intrinsic variability. Thus, in many cases the differences between uncertainty and variability are not clear.

3 The role of expert judgement in analyses of uncertainty and considerations in the elicitation of expert judgements

The identification and quantification of uncertainties in many fields require the application of expert judgements. These judgements may be made in an *ad hoc* fashion or may be elicited using structured procedures. The elicitation of expert judgements by informal or formal procedures is potentially associated with many pitfalls, as discussed in the Annex to this report. A summary of key considerations is given below.

- In the context of assessments of routine discharges of radioactive wastes, the environments of interest have generally been well characterised and mathematical modelling of radionuclide transport through those environments is a mature discipline. Therefore, the main potential role for formal elicitation is likely to be in the definition of parameter value distributions for use in either generic or site-specific assessments.

- The process of expert elicitation does not add to the information available, rather it translates what is known into a usable form. Specifically, formal elicitation procedures cannot legitimately be used to overcome ignorance. However, they can be used to interpret available qualitative and quantitative information in the context of a particular radiological assessment.

- Elicitations can either be from individuals or groups. Group techniques are often preferred, because they are more effective at drawing out information by interactions between the experts. However, in a group context, there are possibilities of bias additional to those arising from the elicitation of individual experts. To a considerable degree, the possibilities of bias inherent in individual or group elicitations can be limited by the use of well-structured procedures. The individual components of such a structured procedure are summarised in the Annex.

- Whether elicitations are undertaken in a group context or from several individual experts, there is no guarantee that a consensus view will, or should be, reached. Furthermore, it is legitimate for the experts to conclude that the elicitation as originally formulated addresses an inappropriate parameter and for them to redefine the parameter during the elicitation, so that they can give legitimate guidance on its distribution of values. Thus, expert elicitation of parameter value distributions may have implications for the conceptual or mathematical structure of the underlying models.

- The use of formal procedures of expert elicitation is highly resource intensive. For most routine assessments, it would not be commensurate with the limited radiological impacts of the waste discharges to utilise this approach. Such procedures are likely to be most appropriate to sites where discharges could result in doses approaching the relevant constraints, or in the derivation of generic parameter value distributions for application at a wide variety of sites.

4 Potential pitfalls involved in uncertainty analysis in dose assessments

The examples given in section 6 and the vast majority of other studies involving uncertainty are limited to considering parameter uncertainty and do not consider the other types of uncertainty outlined in section 2. The ranges in parameter values can be chosen to take some account of conceptual or computational uncertainties, although this is not good practice, but these are not directly taken into account in uncertainty studies. This is not thought to be a major issue for radiation dose assessments relating to routine waste discharges.

- In many cases, the ranges in parameter values are chosen by the people carrying out the study and so are susceptible to bias. An alternative is to use some type of expert elicitation to obtain the ranges but, as outlined in section 3, there are potential bias and resource issues associated with this approach. In particular, the experts may not be able to consider the parameters used in the models but have to consider quantities that can be directly measured. The experts may be asked for the range on the 'best estimate' or mean value of a parameter, but will tend to give a full range for the parameter value based on their knowledge. If the experts disagree on the range for a particular parameter and a simple arithmetic aggregation procedure is used, there will be a tendency for the range used in the final analysis to be unduly large. Also, expert elicitation is a very time consuming and hence expensive undertaking.

- For many parameters there is insufficient information to fully characterise a distribution of possible values and judgement has to be used. It may not be possible to distinguish between uncertainty, eg, in the average value for a parameter, and variability, the range in the parameter that will occur naturally. Often the experts or analysts carrying out the uncertainty study will only feel happy to estimate a maximum and minimum value for a parameter together with the best estimate or most likely value. In many cases, there will be a tendency to concentrate on the area of the distribution that will give the highest doses and to err on the side of caution, which will skew the resulting analysis. In general, it is better practice to apply caution to the results of a realistic assessment than to apply caution in the selection of input parameter values, as the propagation of cautious assumptions through a model can have unintended consequences.

- In many cases a number of parameters are considered which together represent the transfer of radionuclides through the environment and hence the radiation dose. For example, the deposition of a radionuclide onto pasture and the subsequent transfer to milk is represented by a series of parameters such as the interception of the activity by the surface of the grass, removal from the grass, the amount of intake by a cow and the transfer of activity from the intake to milk. The ranges for each of the parameter values may seem reasonable, but when they are combined the resulting range in transfer from deposition to milk could be far greater than observed or than is considered feasible. This can be due to a number of causes, including correlations between the parameters or constraints on combinations of values that can occur.

- In carrying out a probabilistic uncertainty analysis, it is necessary to specify the shape of the distribution and this may affect the results of the analysis. As noted, often the only values available to characterise the distribution are the maximum, minimum and some best estimate or reference value. One way to represent this is through using a triangular distribution. Such triangular distributions are often skewed such that the best estimate is less than the mean of the distribution and there is a higher probability of getting a value higher than the best estimate rather than of getting a lower value. The desire not to underestimate doses or risks can further contribute to the skewed nature of the distribution. In some cases, for example intake rates for food, there is a cut off value, say zero, and this can further skew the distribution. The use of triangular distributions can have a significant effect on the final results of an uncertainty study (Cabianca et al, 1998 and GRNC, 2002). One effect is that the estimate of dose using all best estimate values is a low percentile on the distribution of doses and significantly below the mean or median value of the distribution. Other distributions, such as normal or lognormal are often more plausible representations of environmental variables.

- One of the most difficult aspects to incorporate into an uncertainty analysis is correlation between parameter values. Although statistical techniques are available to take account of correlations, it is not easy to decide the extent to which different parameters are correlated. Two parameters are correlated if when one changes the other also changes in some related manner. For example, the interception by plants of activity deposited from atmosphere is closely related to the standing biomass density of the plant. Also, the behaviour of a particular element in soil is likely to affect both the extent to which it will migrate down the soil column and its uptake through a plant's root system. However, the degree of correlation is hard to quantify. Environmental processes include non-linear dynamic relationships between variables and 'chaotic' systems which give rise to the observed complexity at the real world scale of interest.

- In studies considering uncertainty in dose assessments, the range in the estimated doses can often be found to be fairly narrow. This occurs where there are multiple radionuclides and exposure pathways leading to the estimated dose. The narrow range can cause people to question the results of the uncertainty study and this was the case for a study in France (GRNC, 2002). However, the nature of dose assessments means that there is a tendency for different factors to cancel each other out. For example, a high intake of one food is balanced by a low intake of another food. The same is seen when dose assessments are carried out by different organisations; there may be many differences in parameter values but the final results can sometimes be surprisingly similar.

- Closely related to Uncertainty is Quality. Key to the ultimate acceptability of an assessment is rigorous quality control of the models and the data, effective checking and peer review with good communication maintained between the modellers and the suppliers of the data and also with the users of the outputs.

- Unless regulatory targets and limits are expressed in terms of defined confidence limits (or other specified uncertainty indicators, bands or margins), there is no easy way to use the distributions of results arising from assessments incorporating uncertainty and variability to judge compliance with the target or limit. Use of the mean of the distribution, the expectation value or a best estimate to compare with the target or limit will result in a failure to use available information about the shape of the distribution.

- Concentrating on the uncertainty due to lack of knowledge of parameter values can overlook other aspects such as methodological or scenario uncertainty.

- Effort can unnecessarily be expended on eliciting distributions for parameters that do not significantly affect the result. A probabilistic assessment is not "a priori" better than a deterministic assessment incorporating consideration of the qualitative and quantitative factors affecting the result obtained.

- In many cases there is a lack of empirical data to quantify uncertainty and there are severe constraints on the time, effort and financial resources that can realistically be expended in obtaining better, more representative data.

- Care should be used when interpreting the extremes of the distributions of results. Values from the extremes of the distributions can represent the combination of several possible, but unlikely, circumstances giving rise to an unrealistic combination.

- Uncertainty distributions can mask compounded pessimisms and cause unnecessary public alarm at prospective doses.

5 The presentation of the results of a probabilistic dose assessment

If an uncertainty analysis is carried out, a large number of results are generated and there is then the question of the most meaningful way to present them. Various studies have been undertaken on the presentation of uncertainty and probabilistic results, but there is no consensus on how this is best achieved. The challenge is to find uncertainty expressions that both match scientific practice and can be understood by lay people. Results from one study by Walsh et al (2000) are given here. Possibilities are to present the results graphically using some form of probability distribution which can take different forms. Box plots can also be used to give a graphical representation of value, spread and symmetry of a range of numbers and are particularly useful for comparative purposes. The properties of a distribution can be quantified using a table of results, although this does not give much information about the shape of the distribution.

The presentation of results from a dose assessment depends on the audience and information required, but it is important that sufficient information is presented to characterise fully the distribution. For an assessment of doses the following characteristics could be presented: median, geometric mean (in the case of a log normal distribution), percentage above a chosen dose limit/constraint and values at chosen percentiles (eg, 5^{th} and 95^{th}). When discussing/comparing distributions, many authors simply include the values of the 5^{th}, 50^{th} and 95^{th} percentile. The minimum and maximum values of dose distributions should be used with caution or not at all. If the distribution is made up from results of a limited number of runs the extremes of the distribution will not be well characterised.

The presentation of probabilistic dose estimates was considered in a study for FSA (Collier and Wright, 2002). Greenstreet Berman carried out an initial review of best practice in presenting probabilistic dose estimates and carried out a pilot trial using a worked example to explore presentations of variability. This showed that it is possible to convey an understanding of the approach and the sources and significance of variability in dose. A follow up study considered communication of uncertainty in dose assessments (Collier and Shaw, to be published).

Doses are often presented and reported as 'crisp numbers', significant to two (or more) digits. If a dose is presented as 0.053 mSv, that means that it is different from 0.052 and 0.054 mSv or in other words is known to an accuracy of about 2%! There are typically no 'error bars' or other explicit indicators of uncertainty for the doses presented in published reports. To quote doses to a single digit would give an odd appearance to the reader, whereas three digits is certainly excessive and so two digits may be a reasonable compromise. However, an unintended consequence of this stylistic convention is to convey an impression of accuracy in the results that would be difficult to justify given the underlying uncertainties. Excessive precision is confusing and misleading and it might be worth adding a comment to say that two significant figures are used for presentational purposes and do not imply that degree of accuracy.

6 Examples of relevant uncertainty studies

The following are some examples of studies that have included an investigation of the uncertainties in the results of a dose assessment. Primarily these are for routine releases of radionuclides to the environment (both retrospective and prospective), but also include other types of assessment. Many of the UK studies were carried out on behalf of the Food Standards Agency or the Environment Agency.

- *Distribution of risk in exposed groups from routine releases of radionuclides to the environment* (Jones et al, 2003). This study examined the distribution of risk to members of exposed groups from routine discharges by considering the influence on the estimated risk of variability and uncertainty in the parameters of the calculation. The ratio of the 5^{th} to 95^{th} percentiles of dose received by the exposed groups for the two sites considered was found to be typically 3 or 4. The spread of risk across the exposed groups considered was estimated to be typically an order of magnitude or more. The greatest single contributor to the spread of risk in an exposed group was the risk per unit dose.

- *Uncertainty in dose predictions via aerial pathways* (Charles, 2004). Work has been carried out by Westlakes Scientific Consulting to assess the uncertainty in calculating radiation doses. This particularly focussed on the uncertainty in dose predictions via aerial pathways and the uncertainty in the iodine-129 transfer to milk exposure pathway. The study found that further refinements of the cow and pasture models are of little benefit in reducing the uncertainty in dose from iodine-129 in milk unless more data are available.

- *Uncertainties in assessing doses to members of the public due to radioactive discharges to the marine environment* (Brownless et al, 2002). The various uncertainties associated with the calculation of water concentrations (WAT model) and doses in marine dose assessments (ADO) were considered. Parameter value ranges appropriate to the whole of the UK were determined via informal expert elicitation. It was found that the most significant uncertainties were associated with sediment parameters, *viz*, distribution coefficient (K_D) and sediment load. For specific radionuclides that vary from a conservative to particle-reactive mode of behaviour the maximum uncertainty range for concentrations and doses can be of the order of 500 times.

- *A Probabilistic Modelling Suite for the Marine Environment* (Grzechnik et al, 2002). The WAT & ADO deterministic suite of dose assessment models was modified using a Monte-Carlo simulation driven by a Latin Hypercube random sampling routine. Uncertain parameters (as defined by Brownless et al, 2002) were represented as distributions and propagated through the models. Case studies were conducted for the entire UK and the Sizewell specific situation, with various suggestions for the inclusion of variability via input distributions and the presentation of outputs. Site-specific modelling was found to reduce uncertainties by approximately one order of magnitude.

- *Representation of Habits Survey Data for Probabilistic Dose Assessments* (Grzechnik, 2003). Various scenarios for the inclusion of uncertainty and variability in habits survey data were investigated for aquatic probabilistic dose assessment models. The ratio of 95th to 5th percentiles (uncertainty) was found to be as high as 30 when the entire habits survey group was included. However, scenarios which used critical group habits (upper range habits rates) were found to give uncertainty ratios approximately 10 times lower.

- *Parameter Correlations in Marine Dose Assessments* (Grzechnik, 2003). An investigation of the correlations of parameters in the probabilistic versions of the WAT & ADO models was undertaken. It was determined that proportionality exists between suspended sediment load and sedimentation rate, and this must be taken into account in all future marine probabilistic dose assessments.

- *Variability in critical group doses from routine releases of radionuclides to the environment* (Cabianca et al, 1998). The aims of this study were to identify those factors in a critical group dose assessment whose variability will lead to variation in radiation doses and to assess the distribution in individual doses within the overall critical group due to the likely ranges in the input parameters. Illustrative results for two sites showed that the dose received by a critical group may cover a significant range.

- *Variability in critical group doses: the implications for setting authorised limits for discharges* (Walsh et al, 2000). This study considered the implications of the variability in critical group doses in relation to the system of radiological protection. A workshop was held to discuss the issues and the findings are reported.

- *Analysis of sensitivity and uncertainty in the risk of leukaemia due to the nuclear installations in the Nord-Cotentin region* (GRNC, 2002). The Groupe Radioécologie Nord Cotentin in France carried out this study as part of a large programme of work to investigate whether the incidence of childhood leukaemia in the region could be due to operations at the nuclear sites situated there. They considered possible ranges in the input parameters to the assessment of risks, except that they did not consider the uncertainty associated with internal dosimetric data, nor that associated with risk coefficients. They found that the resulting distribution in risk was narrow with less than a factor of three between the 5th and 95th percentile.

- *Radiological assessment of sources of environmental contamination in the Southern Urals* (Aarkrog et al, 2000). A dose assessment was carried out for population groups in the Southern Urals as part of a larger study to consider the radiological impact of past, present and potential sources of environmental contamination in the area and strategies for remedial measures. Current doses to the population of Brodokalmak on the Techa River were estimated probabilistically taking into account possible ranges in the input parameters used for the dose assessment. The distributions in dose were found to be narrow with 90% of the doses calculated being within a factor of about 2 of the median value.

- *Radiological Assessment under Uncertainty in a Solid Waste Management Context*. Performance assessments undertaken by the US DOE in respect of the proposed spent fuel and high-level waste disposal facility at Yucca Mountain, Nevada incorporate parametric uncertainty in all components of modelling of radionuclide transport from the emplaced wastes to the biosphere. In particular, the biosphere model used includes parametric uncertainty to compute probability density functions for biosphere dose conversion factors, which are used to relate radionuclide fluxes from the geosphere to radiation doses to humans (Rautenstrauch et al, 2003). This biosphere modelling is underpinned by a detailed report setting out the basis for the input parameter value distributions used in the model (Wasiolek, 2003).

- *Accident consequence uncertainty analysis* (Goossens and Kelly, 2000). This was a major study to estimate the uncertainty associated with the prediction of the various endpoints in probabilistic accident consequence assessment codes as a result of potential nuclear accidents. It was carried out as collaboration between the European Commission and the United States Nuclear Regulatory Commission. A central feature of the study was formal elicitation of expert judgement (see Section 3) to define the parameter ranges used in the uncertainty analysis. A number of reports and papers have been published on this study and its findings. These include papers on the use of expert judgement.

7 Summary and conclusions

As this overview of the issues has shown the incorporation of uncertainty and variability in the assessment of doses received by members of the public from routine discharges is not straightforward. There are problems in ensuring that all sources of uncertainty and variability have been taken into account. Identification of appropriate ranges of the values of the parameters used is not easy and the presentation of the results is not straightforward.

Those who wish to carry out such assessments need to consider these issues and the associated pitfalls. Expert advice and guidance should be sought before commencing the assessment.

8 References

Aarkrog A, Simmonds J, Strand P, Christensen G and Salbu B (2000). Radiological assessment of past, present and potential sources to environmental contamination in the Southern Urals and strategies for remedial measures (SUCON). RISØ National Laboratory, Roskilde Risø-R-1243.

Arkes HR and Hammond KR (Eds) (1986). Judgement and Decision Making: An Interdisciplinary Reader, Cambridge University Press.

Beven K (2002). Towards a coherent philosophy for modelling the environment. Proceedings: *Mathematical, Physical & Engineering Sciences*, **458** (2026), 2465-2484(20). The Royal Society.

BIOMASS (2003). "Reference Biospheres" for solid radioactive waste disposal. Report of BIOMASS Theme 1 of the BIOsphere Modelling and ASSessment (BIOMASS) Programme, IAEA-BIOMASS-6, International Atomic Energy Agency, Vienna.

Bonano EJ, Hora SC, Keeney RL and von Winterfeldt D (1990). Elicitation and Use of Expert Judgement in Performance Assessment for High-level Radioactive Waste Repositories. NUREG/CR-5411. Prepared for the US Nuclear Regulatory Commission by Sandia National Laboratories.

Brownless GP, Grzechnik MP, Round GD and Pidcock AJ (2002). Uncertainties in assessing doses to members of the public due to radioactive discharges to the marine environment. CEFAS Environment Report RL 12/01.

Cabianca TRA et al (1998). *The variability in critical group doses from routine releases of radionuclides to the environment*. Chilton. NRPB-M952.

Charles K (2004). Problems associated with probabilistic dose assessments – the practical aspects and experience gained from actually trying to do it. Westlakes Scientific Consulting. Paper prepared for the NDAWG sub-group on uncertainty and variability.

Collier GD and Wright MW (2002). Presentation of probabilistic dose estimates – Pilot project for FSA. Greenstreet Berman, Reading.

Collier D and Shaw J (to be published). Communication of uncertainty. Greenstreet Berman, Reading.

Cooke RM (1991). Experts in Uncertainty: Opinion and Subjective Probability in Science. Oxford University Press.

Environment Agency, Scottish Environment Protection Agency, Department of the Environment, Northern Ireland, National Radiological Protection Board, Food Standards Agency (2002). Authorisation of discharges of radioactive waste to the environment. Principles for the assessment of prospective public doses. Interim guidance.

Goossens LHJ and Kelly GN (2000). Expert judgement and accident consequence analysis. Special issue of *Rad Prot Dosim*, **90** (3).

Grzechnik MP (2003). Representation of Habits Survey Data for Probabilistic Dose Assessments. Extended Abstracts, Radiation Protection Symposium of the North West European RP Societies, 2-5 June 2003, Utrecht, Netherlands, 155.

Grzechnik MP (2003). Parameter Correlations in Marine Dose Assessments. CEFAS Environment Report RL 17/03.

Grzechnik MP, Round GD, Brownless GP and Camplin WC (2002). A Probabilistic Modelling Suite for the Marine Environment. CEFAS Environment Report RL 3/02.

Group Radioécologie Nord Cotentin (2002). Analyse de sensibilité et d'incertitude sur le risque de leucèmie attribuable aux installations nuclèaires du Nord-Contentin.

IAEA (1989). Evaluating the reliability of predictions made using environmental transfer models. Safety Series 100, IAEA, Vienna, Austria.

Janis IL and Mann L (1977). Decision Making. The Free Press, New York.

Jones KA, Bexon AP, Walsh C, Haywood SM, Jones AL, Fayers CA, Haylock RGE and Phipps AW (2003). Chilton, NRPB-W45.

Kahneman D, Slovic P and Tversky A (1982). Judgement Under Uncertainty: Heuristics and Biases. Cambridge University Press.

Lock A (1987). Integrating Group Judgements in Subjective Forecasts, In: Wright G and Ayton P (Eds). Judgmental Forecasting. John Wiley and Sons Ltd, 109-127.

Phillip LD (1987). On the adequacy of judgmental forecasts, In: Wright G and Ayton P (Eds). Judgmental Forecasting. John Wiley and Sons Ltd, 11-30.

Rautenstrauch KR, Smith AJ and Andrews R (2003). Technical Basis Document No 12: Biosphere Transport, Revision 1. Bechtel SAIC Company, LLC, 1180 Town Center Drive, Las Vegas, Nevada 89144, USA. September 2003.

Sutherland SS (1992). Irrationality: The Enemy Within. Constable.

Thorne MC and Williams MMR (1992). A review of expert judgement techniques with reference to nuclear safety. *Progress in Nuclear Energy*, **27**, 83-254.

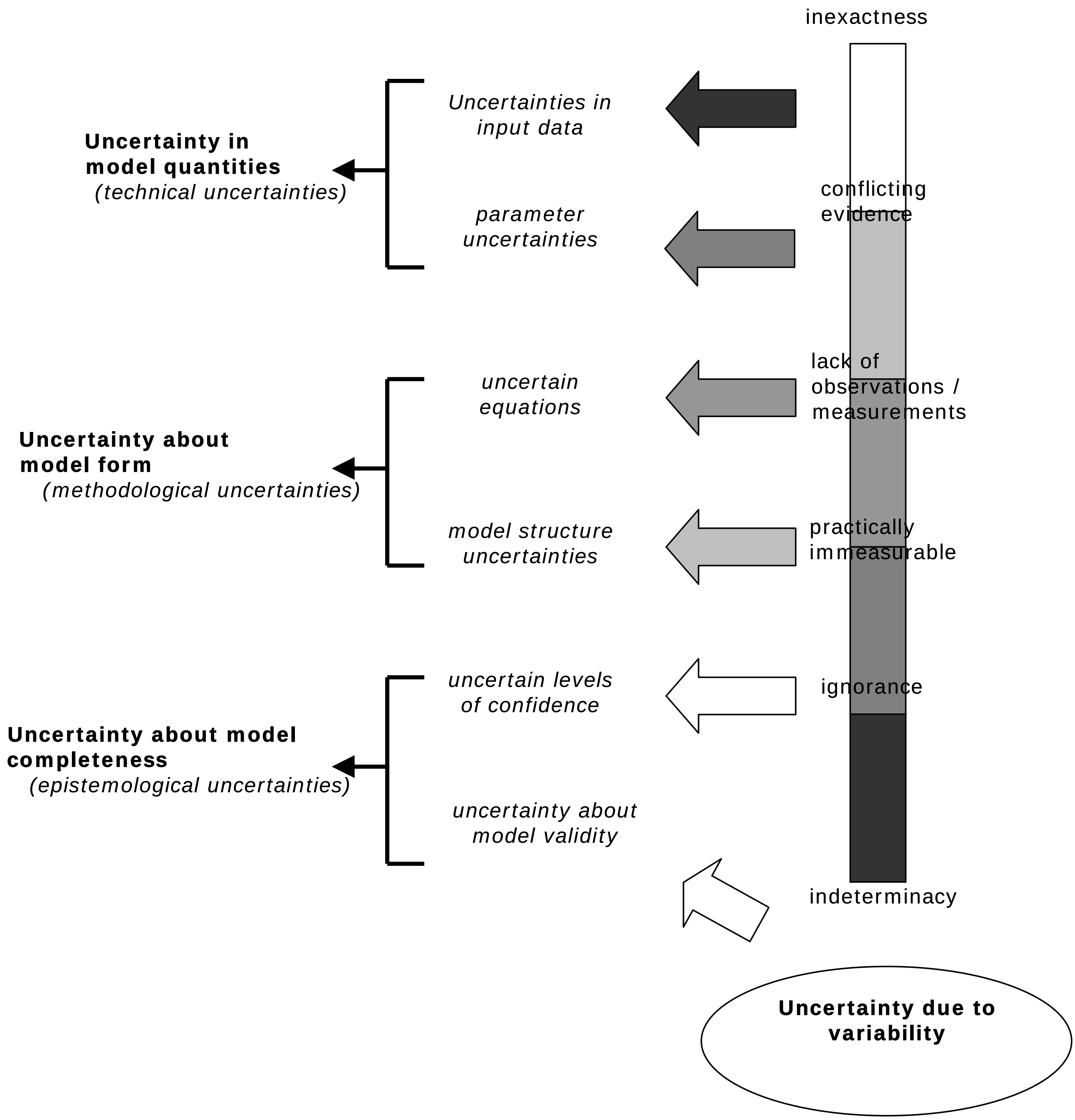
Tversky A and Kahneman D (1974). Judgement under uncertainty: Heuristics and biases. *Science*, **185**, 1124-1131.

Walsh C, Jones KA and Simmonds JR (2000). *Variability in critical group doses: the implications for setting authorised limits for discharges*. Chilton, NRPB-M1221.

Van Asselt MBA (1999). Uncertainty in Decision Support: From Problem to Challenge. ICIS working paper I99-E006. International Centre for Integrative Studies (ICIS), University of Maastricht, The Netherlands.

Wasiolek MA (2003). Environmental Transport Input Parameters for the Biosphere Model. DOE Report ANL-MGR-MD-000007 REV 01, June 2003.

Figure 1: A modellers view on uncertainty (van Asselt, 1999)



Annex: The Role of Expert Judgement in Analyses of Uncertainty and Considerations in the Elicitation of Expert Judgements

The identification and quantification of uncertainties in many fields require the application of expert judgements. These judgements may be made in an *ad hoc* fashion or may be elicited using structured procedures. The elicitation of expert judgements by informal or formal procedures is potentially associated with many pitfalls. There is an extensive literature on techniques of elicitation and many case studies have been reported (eg, Dalrymple and Willows, 1992; Goossens and Kelly, 2000). This information should be consulted and reference should be made to appropriate experts in the field when a requirement for expert elicitation has been defined. In this Annex, a brief overview of the issues that arise is provided, as background to the extensive and detailed discussions that are available elsewhere.

In both safety analyses and environmental impact assessments, expert judgements are invoked at various stages. These have been identified (Thorne and Williams, 1992) as:

- Definition of the phenomena and factors that have to be taken into account;
- Selection of particular mathematical and computational tools to represent those phenomena and factors;
- Selection of data values, ranges or distributions for use in the various mathematical and computational tools;
- Interpretation of the results of the modelling studies.

In some contexts, considerable emphasis has been placed on the first two of the above steps. For example, in the context of geological disposal of solid radioactive wastes, the need to demonstrate comprehensiveness of assessments has led to the use of formal techniques, such as FEP (Features, Events and Processes) analysis and the use of interaction matrices, in the process of developing conceptual models of the system of interest. These techniques have been described in detail, with illustrative examples, in BIOMASS (2003), as has the translation of such conceptual models into mathematical representations.

However, in the context of assessments of routine discharges of radioactive wastes, the environments of interest have generally been well characterised and mathematical modelling of radionuclide transport through those environments is a mature discipline. Therefore, the main potential role for formal elicitation is likely to be in the definition of parameter value distributions for use in either generic or site-specific assessments.

The process of expert elicitation, does not add to the information available, rather it translates what is known into a usable form. Specifically, formal elicitation procedures cannot legitimately be used to overcome ignorance. However, they can be used to interpret available qualitative and quantitative information in the context of a particular radiological assessment. In achieving this context-specific interpretation, care has to be taken to avoid the various subjective biases that can arise (Cooke, 1991; Sutherland, 1992). These biases include (Tversky and Kahneman, 1974 and; Kahneman et al, 1982):

Anchoring: In which undue weight is given to the conventional value, first value given, or information from a previous assessment;

Availability: The tendency to give too much weight to readily available data or recent experience, which may not be representative;

Coherence: The tendency to assign probabilities to events on the basis of the ease with which coherent accounts can be constructed of how those events could arise;

Overconfidence: Over-estimation by an individual of their ability to make quantitative judgements;

Representativeness: The tendency to place more confidence in a single piece of information that is considered to relate directly to the issue under consideration than to a larger body of more generalised information;

Satisficing: Comprehensiveness can be sacrificed to expediency by considering only a limited number of options and picking from among them;

Motivational: Arising from various sources including a desire to influence a decision, a perception that the expert will be evaluated based on the outcome of the study, a desire to appear knowledgeable and authoritative, or a desire not to contradict a stand taken previously;

Unstated Assumptions: These can constrain the degree of uncertainty expressed, or bias the result, if the question answered is not identical to the question asked, or assumed to have been asked;

Value Laden Context: The value orientations and biases of an analyst, institute, discipline or culture can shape the ways that questions are framed, data are selected, interpreted or rejected, methodologies are devised, and explanations or conclusions are formulated.

Elicitations can either be from individuals or groups. Group techniques are often preferred, because they are more effective at drawing out information by interactions between the experts. However, in a group context, there are additional possibilities of bias due to a tendency to conformity ('Groupthink'), inhibition of contributions from less secure individuals, eg, because of status considerations, and premature closure of discussion (Janis and Mann, 1977; Lock, 1987). Where group techniques are used, a structured elicitation procedure can help overcome such biases. Formal techniques for group elicitation have been summarised by Lock (1987), Bonano et al (1990) and Dalrymple and Willows (1992), amongst others. Several types of participant are required. These comprise generalists, specialists and normative experts.

Generalists oversee the undertaking of performance assessments, whereas specialists provide the judgements required. Normative experts (or facilitators) require a sound theoretical and conceptual knowledge of probability and techniques for eliciting judgements and they need to be knowledgeable about the psychological processes occurring in the specialists' minds as they are processing the information to produce requested results.

Dalrymple and Willows (1992) have outlined the steps to be included in such an elicitation. Following selection of the experts, which itself requires careful consideration (Bonano et al, 1990), the steps comprise:

Exploration of bias: In which any interest in the outcome of the exercise is explored;

Definition: An unambiguous definition of the parameter to be elicited is agreed;

Assessment: Methods for assessing or measuring the parameter are explored; this provides insight into the participants' depth of expertise, reminds them of the different types of technique available, and ensures that they consider the measurement uncertainty and limits of applicability associated with each technique;

Factors and assumptions: This covers factors that a parameter depends on and assumptions that participants would make when assessing their uncertainty about a parameter;

Sources of uncertainty: Participants are asked to list the sources of uncertainty about the parameter;

Conditioning: Participants are encouraged to stretch the range of values that the parameter could adopt by being asked to describe scenarios in which high and low values could arise;

Encoding: Encoding of the probability distribution function for the parameter is undertaken by locating points on its range and asking for assessments of the probability that the value of the parameter is less than each selected value;

Verification: The cumulative probability distribution function and probability distribution function implied by the encoding are displayed to participants, in case there are any anomalies that they wish to correct;

Confirmation: Notes taken at the meeting are written up and distributed to participants for comment, together with the elicited probability distribution function.

Where judgements are elicited from individual experts, the above steps are generally applicable. However, additionally, consideration has to be given as to how those judgements should be aggregated. Although a broad range of techniques has been proposed, all require additional judgements by the analyst on the worth of each expert (Thorne and Williams, 1992).

Whether elicitations are undertaken in a group context or from several individual experts, there is no guarantee that a consensus view will, or should be, reached (Thorne and Williams, 1992). Furthermore, as Dalrymple and Willows (1992) have emphasised, it is legitimate for the experts to conclude that the elicitation as originally formulated addresses an inappropriate parameter and for them to redefine the parameter during the elicitation, so that they can give legitimate guidance on its distribution of values. Thus, expert elicitation of parameter value distributions may have implications for the conceptual or mathematical structure of the underlying models.

References

BIOMASS (2003). "Reference Biospheres" for solid radioactive waste disposal. Report of BIOMASS Theme 1 of the BIOsphere Modelling and ASSessment (BIOMASS) Programme, IAEA-BIOMASS-6, International Atomic Energy Agency, Vienna.

Bonano EJ, Hora SC, Keeney RL and von Winterfeldt D (1990). Elicitation and Use of Expert Judgement in Performance Assessment for High-level Radioactive Waste Repositories. NUREG/CR-5411, Prepared for the US Nuclear Regulatory Commission by Sandia National Laboratories.

Cooke RM (1991). Experts in Uncertainty: Opinion and Subjective Probability in Science. Oxford University Press.

Dalrymple GJ and Willows M (1992). Dry Run 3: A Trial Assessment of Underground Disposal of Radioactive Wastes Based on Probabilistic Risk Analysis – Volume 4: Elicitation of Subjective Data. DoE Report No DoE/HMIP/RR/92.043.

Goossens LHJ and Kelly GN (2000). Expert judgement and accident consequence analysis. Special issue of *Rad Prot Dosim*, **90** (3).

Janis IL and Mann L (1977). Decision Making. The Free Press, New York.

Kahneman D, Slovic P and Tversky A (1982). Judgement Under Uncertainty: Heuristics and Biases. Cambridge University Press.

Lock A (1987). Integrating Group Judgements in Subjective Forecasts. In: Wright G and Ayton P (Eds). Judgmental Forecasting. John Wiley and Sons Ltd, 109-127.

Sutherland SS (1992). Irrationality: The Enemy Within. Constable.

Thorne MC and Williams MMR (1992). A review of expert judgement techniques with reference to nuclear safety. *Progress in Nuclear Energy*, **27**, 83-254.

Tversky A and Kahneman D (1974). Judgement under uncertainty: Heuristics and biases. *Science*, **185**, 1124-1131.