Workshop on the Use of Human Data in Risk Assessment 23-24 February 2004, Cardiff

Workshop Report No. 3

Brussels, November 2004

ECETOC WORKSHOP REPORT No. 3

European Centre for Ecotoxicology and Toxicology of Chemicals
4 Avenue E. Van Nieuwenhuyse (Bte 6), B-1160 Brussels, Belgium.

Use of Human Data in Risk Assessment

CONTENTS

EXECUTIVE SUMMARY

Available human data are not always consistently addressed in chemical risk assessment, while questions of data quality often inhibit their interpretation and use. ECETOC, with the aim of initiating a consensus on this topic, has organised an International Workshop on the Use of Human Data in Risk Assessment, in collaboration with the International Programme on Chemical Safety (IPCS), OECD Environment Directorate and European Commission's Institute of Health and Consumer Protection (IHCP). This report presents the proceedings and outcome of the Workshop, held in Cardiff, UK (February 2004).

Good quality human data are not readily accessible and their usefulness is frequently constrained by the lack of information on exposures. A clear framework should be developed that enables human data from different sources to be collected and assessed in terms of quality and application in risk assessment. The framework must be built on guidelines for (i) good human data collection practice, (ii) good exposure assessment practice and (iii) good risk assessment practice, and better networking of existing information sources.

An action plan is required to improve the recording of existing data, while incentives should encourage the provision of human data of improved quality for risk assessment purposes.

1. SUMMARY AND CONCLUSIONS

There is a consensus emerging that data from human studies are not consistently addressed in chemical risk assessment, while questions of data quality often inhibit their interpretation and use.

ECETOC, in collaboration with the International Programme on Chemical Safety (IPCS), OECD Environment Directorate and European Commission's Institute of Health and Consumer Protection (IHCP), has held an International Workshop on the Use of Human Data in Risk Assessment in Cardiff, UK (23 - 24 February 2004). The purpose of the Workshop was to discuss the principles that govern the use of human data in chemical risk assessment, with the aim of initiating a consensus view on this topic. Over 30 experts from government, academia and chemical industry participated in the Workshop.

The Workshop reviewed the interpretation of human data for risk assessment in the occupational and public health, and outdoor environment settings, respectively. Workshop participants discussed a proposed framework to guide the use of human data in the risk assessment process, including criteria for characterising the quality of study design, exposure and effect data. The role of, and the process for consideration of null epidemiologic data in risk assessment were considered. Three break-out groups examined the suitability of the proposed quality criteria for a series of selected case studies focusing on the use of human data on (i) acute (local) effects, (ii) chronic effects data and (iii) the relevance of different data sources.

The Workshop recognised that useful and good quality human data are found in many forms such as case reports, poisons centre records, health surveillance results, self-administered surveys and epidemiological studies. Many data are not readily accessible and their usefulness is frequently constrained by the absence of adequate contextual information on exposures and other quality considerations.

The Workshop recommended that a clear framework should be developed that enables human data from different sources to be collected and assessed in terms of quality and application. The framework should be easily understandable and applicable in a consistent and transparent manner. The framework must be built on solid foundations which should cover the identification and/or development of guidelines for (i) good human data collection practice, (ii) good exposure assessment practice and (iii) good risk assessment practice, and better networking of existing information sources.

A sustained and collaborative effort is required to develop and implement an action plan to improve the recording of existing data for risk assessment purposes. This should involve a range of organisations, and includes raising awareness and training of key groups (risk assessors and data collectors).

2. WORKSHOP OVERVIEW

2.1 Introduction

Risk-based decision-making and standard setting are increasingly being used across Europe and elsewhere, but existing chemical risk assessment processes do not consistently address available human data. Although positive human findings often affect the outcome and interpretation of risk assessment, contradictory (and particularly "negative" or "null") human data are often regarded as being less reliable than animal findings. Indeed, even when human data are available, questions of data quality often inhibit their interpretation and use. Consequently, no clear consensus exists among the scientific and regulatory communities on how to use human data in the risk assessment process. There is therefore a need to develop a framework by which human data can be consistently recorded, collected and optimally used in the risk assessment process.

ECETOC, in collaboration with the International Programme on Chemical Safety (IPCS) [a], OECD Environment Directorate and European Commission's Institute of Health and Consumer Protection (IHCP), has held an International Workshop on the Use of Human Data in Risk Assessment in Cardiff, UK (23 - 24 February 2004). The purpose was to discuss the principles that govern the use of human data in chemical risk assessment, with the aim of initiating a consensus view on this topic. Thirty-one experts from national (7) and international (6) governmental agencies, academia (10) and the chemical industry (8, from ECETOC member companies) participated in the Workshop providing a balance in participation among government, academia and industry (Appendix A).

2.2 Workshop structure

As preparation for the Workshop, ECETOC, with support from the Long-range Research Initiative of the European Chemical Industry Council (Cefic), commissioned Dr G. Swaen of the University of Maastricht (NL), to undertake a critical review of the approaches available for assessing the quality of human occupational health data. The review included a proposed framework for how human data could be used in risk assessment. It was circulated to participants ahead of the Workshop together with abstracts from each of the invited speakers. The critical review will be published in a peer-reviewed scientific journal.

[a] A collaborative programme of the World Health Organization (WHO), the International Labour Organization (ILO) and the United Nations Environment Programme (UNEP)

On the first day of the Workshop, invited speakers described the use of human data in the occupational, outdoor and public health settings and proposed criteria for characterising the quality of study design, exposure and health effects data. The role of null epidemiological data was reviewed. On the second day three break-out sessions allowed participants to discuss issues in more detail focusing on specific case studies, to illustrate the use of acute and chronic human data and relevant sources of human data for chemical risk assessment.

2.3 Workshop aims and objectives

The primary aim of the Workshop was to initiate a consensus view on the principles that govern the use of human data in the risk assessment process. Specific objectives were to:

1. Propose criteria that would enable the quality of data to be assessed and weighed, in order that it can be consistently used with confidence;
2. suggest how human data of different (but known) quality ought to be used in the risk assessment process (and what constraints may also be appropriate);
3. identify any constraints and/or limitations that might be placed upon the use of such data, including areas where its use may be currently inappropriate.

In addition, it was hoped that the Workshop would begin to identify any guidance that might be needed to assist risk assessors to more consistently use human data and to enable future epidemiological research to be designed in such a way that they would have the most value for the risk assessment process.

2.4 Workshop programme

Following an introduction by Prof. J. Bridges (University of Surrey, UK), Chairman of the Workshop, there were presentations by Dr S. Fairhurst (UK HSE), Dr T. Meredith (WHO-IPCS) and Prof. T. Eikmann (Giessen University, D) on the interpretations of human data for risk assessment in the occupational and public health, and outdoor environment settings, respectively. Dr G. Swaen (Maastricht University, NL), proposed a framework for how the criteria for the quality of study designs and the quality of data on human exposure and effects can guide the use of human data in the risk assessment process. Dr R. Schnatter (ExxonMobil, USA) explained the role of, and the process for consideration of, null epidemiologic data in risk assessment (Appendix B).

Participants then split up into three break-out sessions to discuss the suitability of the proposed quality criteria when applied to a series of selected case studies for the use of human data on (i) acute (local) effects and (ii) chronic effects data and (iii) the relevance of different human data sources (Appendix C).

3. THE WORKSHOP

Prof. J. Bridges, Chairman of the Workshop, gave an overview of the chemical risk assessment process and emphasised that supporting work and documentation must be credible, convincing, clear, precise, current and cost-effective. He highlighted some of the dilemmas in achieving these outcomes at a time of diminishing resources i.e. increasing numbers of assessments, shorter turnaround times, the need for new and improved methodologies to replace out-of-date animal tests, and the need for improved transparency and quality. Prof. Bridges concluded that under such circumstances it was timely to address how better use of available human data might assist risk assessment processes and at what particular points in the process improvements in the use of human data would be of highest priority.

3.1 Interpretation of human data for risk assessment in the occupational setting

S. Fairhurst
Industrial Chemicals Unit, UK Health and Safety Executive

Dr Fairhurst drew on his experience with European classification and labelling activities and with the establishment of occupational risk management measures to describe the interpretation of human data for risk assessment in the occupational setting. He highlighted that human data is a key information source for toxicologists in understanding the toxicological hazards of substances, the risk to health posed by specified exposure conditions and in appreciating the effectiveness of existing control strategies. By having data of direct relevance, the occupational setting ought to be one of the prime arenas in which human data are used for risk assessment and standard-setting purposes. However, while human data are highly desirable, the information available about human experience is often disappointing, in terms of extent and quality. There are several reasons as to why this might be so. There also seems to be a difference in expectations between industry and regulators and between different regulatory systems as to the impact that human data should have on decision-making. Dr Fairhurst illustrated this situation by referring to "negative" human data; this could be seen as either a contribution to the overall appreciation of toxicity (where other strands of "positive" evidence raise a regulatory concern) or as vindication that current exposure conditions pose no problem (and hence no regulatory action is required).

Dr Fairhurst offered an assessment of the role of human data in relation to the full range of toxicological endpoints. In relation to skin and respiratory sensitisation (asthmagenicity), he felt that these toxicological endpoints should be of critical

importance in terms of using human data, particularly due to the absence of robust animal models for respiratory sensitisation. He pointed out the different interpretations that could be put on the absence of reliable human data for these endpoints. It could be that one could conclude (i) that the chemical lacked these hazardous properties because there had been extensive human experience and no cases, or that the quality of the data precluded any judgement being made, or simply (ii) that there were no data available leading to the conclusion that there was simply an absence of knowledge. Where a chemical has been shown to have sensitising properties, there were other deficiencies, such as the lack of dose-response information in many cases. The extent and usefulness of human data on different toxicological endpoints varied, being relatively rich for some endpoints, e.g. single exposure acute disturbances of the central nervous system for certain solvents. There were quite a few examples of specific chemicals and toxicological endpoints where extensive and reliable human data had been used in the occupational setting, e.g. hydrogen sulphide (cardiovascular effects), kaolin (respiratory tract effects), cadmium (kidney effects) and lead (various effects), although there may still be debate about all the details of the picture revealed. With respect to carcinogenicity, substantial and very valuable human data exist for some substances such as benzene, 1,3-butadiene and hexavalent chromium, but for other chemicals such as perchloroethylene the available human data was much more problematic, with different bodies reaching different conclusions from the evidence available. In general, human data for reproductive toxicity were often inconclusive and difficult to interpret. Again the example of perchloroethylene was used to illustrate this point.

Dr Fairhurst also drew attention to the use and potential value of other types of human data. For example, he advocated that use of toxicokinetic data from human subjects and/or tissue samples and physiologically-based pharmacokinetic modelling could help in the interpretation of the significance of experimental animal data for human health, and in route-to-route extrapolation. Useful data on skin penetration could also be gained from studies using human tissue.

In summary, Dr Fairhurst promoted a view that the weight of all available data should be used for risk assessment including the full range of human data. He encouraged toxicologists to embrace the use of human data in their work and for toxicologists, clinicians, occupational hygienists and epidemiologists to work together more closely in setting up new studies and also in interpreting existing data. Collaborative projects, and improved awareness and training in each other's disciplines about how human data could be applied to risk assessment, were suggested as practical ways ahead.

The discussion among Workshop participants made it clear that for many substances there were many data gaps and that acquisition and interpretation of existing human data should be prioritised. However, in all cases the collection of data should be designed with the potential use of the data in mind. It was proposed that priority for data collection could be given to those types of data that are already used to good effect, or have the potential to be used better. The importance of new sources of data including industry sources of unpublished monitoring and surveillance data was also highlighted. It was recognised that further efforts were needed in this area to improve uniformity and harmonisation. The view emerged that if better data were needed, as they clearly were, then those providing the data should not be disadvantaged by providing data over and above standard requirements. Nor should the additional data or the prospect of searching for additional data become opportunities for "fishing expeditions". Guidance about data quality and use of human data was essential.

3.2 Interpretation of human data for risk assessment in the outdoor environment

T. Eikmann
Institute of Hygiene and Environmental Medicine, Justus-Liebig University, Giessen, Germany

Dr Eikmann presented the use and interpretation of human data for the assessment of risk in the outdoor environment based on ambient air monitoring, biomonitoring and the monitoring of human effects in North-Rhine Westphalia, Germany. He described how the recording and interpretation of human data for the assessment of influences of the outdoor environment had changed in the last few decades. In the 1960s and 1970s environmental influences were commonly correlated with the frequency of particular diseases, specific health-related effects or other personal data rather than human biomonitoring and human effect monitoring, which only became possible with technological developments in instrumentation. He explained that in break-out with the availability of this new data epidemiological studies had begun to consider concurrent effects and broader environment impacts such as health-related quality of life and registration of environmental annoyance.

Dr Eikmann concluded that human biomonitoring was most useful as the basis for derivation of internal tolerable substance concentrations and for formulating hypotheses about correlations between the concentration of substances and their effects. Human biomonitoring data was considered to be the golden standard in combination with ambient air monitoring. However, it was not possible to carry out biomonitoring studies for all substances, e.g. formaldehyde, particles and fibres. Epidemiological studies provided good data for general morbidity and mortality,

but were not adequate for individual environmental problems and required considerable expenditure.

Dr Eikmann discussed the example of particulate matter in the outdoor air and the evidence on human health effects. From long-term studies there is sufficient evidence that long-term exposure to particulate matter implies serious health effects. Short-term studies demonstrate associations between particulate matter concentrations and human down to the lowest concentrations measured. A no-effect concentration or a threshold cannot be derived from these data. The existing data do not allow drawing conclusions about the effects of ultra-fine particles. Dr Eikmann concluded this part of his presentation by stating that data from epidemiological studies are not a sufficient basis for deriving no-effect threshold concentrations; and special risk groups may influence the results of epidemiological studies.

Dr Eikmann used the example of assessing the impact of traffic on the health of children to identify some of the challenges in environmental health. He concluded that there were insufficient studies on environmental exposure of children as most epidemiological studies have addressed adults and diseases and/or conditions associated with higher age such as cardiovascular or cardiopulmonary diseases.

3.3 Interpretation of human health data for risk assessment in the public health setting

T. Meredith
International Programme on Chemical Safety (IPCS), World Health Organization (WHO), Switzerland

Dr Meredith described the relevant activities of the International Programme on Chemical Safety (WHO/ILO/UNEP). He reviewed what is meant by risk assessment in the context of public health broadly and described the recent IPCS initiatives to improve the use of existing observational human data and how these initiatives are being advanced in concert with the chemical-specific risk assessments, the development and harmonisation of risk assessment methodology, and activities relating to poisons information, prevention and management, chemical incident alert, surveillance and response, and capacity building.

The risk assessment of chemicals for the protection of public health follows the same process as in other settings, namely hazard identification, hazard characterisation (or dose-response assessment), exposure assessment and risk characterisation. It is an accepted IPCS principle that well-documented observational and clinical

epidemiological studies have a clear advantage over animals in providing the most relevant information to the species of interest. However, Dr Meredith stated that further development and harmonisation of risk assessment methodologies were essential in order to put this principle into practice, together with the development of mechanisms for collecting and disseminating clinical and exposure data from human observations. The recommended further priorities for action agreed by over 120 governments worldwide in the context of the Intergovernmental Forum on Chemical Safety (Bangkok, 2003) have recently underscored these needs. IPCS has been invited by the Forum to take the international lead for developing guidance and mechanisms for collecting human data.

Dr Meredith suggested a number of opportunities for using human data to better inform the risk assessment process, including validation of new and existing methods of hazard identification, developing more realistic exposure scenarios (including aggregate and cumulative exposures), and improving cross-species extrapolation and development of weight-of-evidence approaches for risk characterisation. He drew attention to the IPCS draft guidance document for chemical-specific adjustment factors (CSAF) and how this might provide a framework for justifying the use of specific adjustment or scaling factors rather than default values in risk assessment. He welcomed input from clinical toxicologists and pharmacologists in developing case studies on how this guidance could be applied.

Dr Meredith reviewed progress with recent IPCS initiatives relevant to human data, including work underway to demonstrate the feasibility of collecting harmonised human data from poisons centres on a multi-national basis, a survey of the sources of human data used, and the policies and practices for use of human data in risk assessment. Plans for a prospective study to collect human observational data to meet particular risk assessment needs were introduced.

Dr Meredith reflected that it was important for the risk assessment of chemicals to continue to have a strong international focus to ensure benefits from harmonisation of approaches. Significant progress has been made in the adoption of key terminology, concepts and methodologies for hazard and exposure assessment, allowing increased collaboration in the systematic investigation of the many thousands of chemicals in trade. Harmonisation also helps to ensure common approaches to human health protection in the outdoor environment and in occupational and public health settings as well as providing the possibility for consistent approaches across different chemical sectors. Human observational data also provided opportunities for enhancing the evidence for successful risk management practices.

The deliberations of the World Summit on Sustainable Development (2002) have strengthened the concept of public health settings by reconfirming that human health is at the centre of sustainability and development, and by establishing new goals to ensure that chemicals are not used in ways that harm human health and the environment. Supporting this, the World Health Assembly resolved in 2003 that all countries should take full account of the health aspects of chemical safety in order to develop a more strategic approach to international chemicals management.

Current public health issues that support future improvements in the interpretation and use of human health data for risk assessment include the development of improved tools and sources of surveillance data, the need for better understanding of the magnitude and extent of chemicals exposure resulting in injury, ill-health and disease, and the need to understand differences in susceptibility among vulnerable populations and individuals. Dr Meredith explained that work on these issues requires the development of more strategic alliances between risk assessors and clinical toxicologists, other public health specialists such as epidemiologists, forensic toxicologists, and occupational health and safety specialists, and those involved in poisons centres and chemical accident preparedness, response and follow-up. IPCS would continue to develop its work programme to include such collaborations.

In discussion, Workshop participants highlighted the role that poisons centre data have in identifying exposures of concern. Experience to date had highlighted the many challenges in documenting accidental exposures and the difficulty of sometimes obtaining specific information on chemical identities. The use of harmonised reporting formats and terminology, developed originally by the IPCS (INTOX Data Management System) for aggregating poisons centre data, was suggested as a basis for developing a more harmonised terminology and data reporting formats in other areas, e.g. occupational health and safety reporting.

3.4 A proposed framework for development of quality considerations on design characteristics, exposure and effect data, and evaluation

G. Swaen
Maastricht University, NL

Dr Swaen gave an overview of a framework to enable the quality of study design, exposure and effect data to be taken into account for the evaluation of chemical substances. In introducing his background paper, Dr Swaen discussed some of the reasons why human data are not always used in risk assessment and some of the specific weaknesses of human data. In common with several other presenters, Dr Swaen drew attention to the limited discussion about how human data can

optimally be used for risk assessment despite the general consensus that human data forms the most direct evidence for human health risks from chemicals. Although, in principle, human data should always be given priority over experimental animal data, they should not be considered mutually exclusive.

Dr Swaen proposed that there were two important aspects of health effect data to determine the weight of any findings in the proposed framework for using human data, namely the specificity of the effect and the timeframe in which the effect appears (the latency), both acute, sub-acute and long-term. On this basis human data could be put in six categories, (i) acute, specific; (ii) sub-acute, specific; (iii) long-term, specific; (iv) acute, non-specific; (v) sub-acute, non-specific and (vi) long-term, non-specific. For each category he proposed a number of criteria for design, exposure and effects that could enable high-quality human data to be confidently identified for use in risk assessment.

Dr Swaen identified a number of types of observational human data, including case reports and time-series, cross-sectional, case-control and cohort studies, and health surveillance programmes. He proposed that all had a role to play in the risk assessment process from hazard identification through providing supportive information to the determination of the dose-response assessment, e.g. confirming the predicted range, or, in some cases, providing a basis for quantitative risk assessment. He added that the type of health effect under consideration could justify the application of different standards of quality.

The discussion focused on a number of issues, including:

- The use of "null" human data (no association between exposure and effect parameters) and the variety of approaches for defining such data and how their quality could be ascertained;
- how human data should be weighed against animal data and the importance of using both types of data to help overcome shortcomings in either data set;
- the need to take into account other information such as information on the mode-of-action of a chemical as well as animal and human toxicology information;
- the need for good archiving and recording of data in epidemiological research and health surveillance to enable reassessment and further analysis, particularly given the challenges involved in the long-term collection of exposure data;
- whether or not new systems of monitoring such as post-marketing surveillance of chemicals could be useful.

3.5 The role of null epidemiologic data in risk assessment

R. Schnatter
ExxonMobil, USA

Dr Schnatter defined null epidemiological data as a study or a body of data consistent with the hypothesis of no relationship between exposure and disease. He explained that such a definition could also include studies in which no-observed effects have occurred in human populations.

To understand the preferred role of null data in risk assessment, Dr Schnatter explained the need to (i) define the strengths and weaknesses of these data, (ii) understand the requirements and stages of the risk assessment process, and (iii) understand the strengths and weaknesses of competing data on which the risk assessment may be based. Evaluating the strengths and weaknesses of null data should not differ from the evaluation of human and/or epidemiologic data in general. The adequacy of the study design, the lack of bias, the control of confounders and the adequacy of exposure and disease definitions are paramount. For null data, the precision of the risk estimate is often of heightened importance. Dr Schnatter showed that statistical power is often not the most relevant issue and stated that more attention should be given to the related concept of study size and precision.

Dr Schnatter explained that the use of null epidemiologic data has been restricted principally to hazard identification, or first stage of the risk assessment process and that more limited experience had been acquired applying null studies to the dose-response assessment stage. The use of null data for this purpose seemed limited in general by weaknesses in the reporting of exposure data in epidemiological studies. He proposed that, even if exposure is not quantified within a given study, strong epidemiologic data could be used more in risk assessment validations and/or sensitivity analyses. He explained that when there is a mixture of positive and null epidemiologic data, it is essential to use the strongest quality data in risk assessment, whether they are positive or null.

Dr Schnatter emphasised that, if study outcomes are not explained by exposure levels or study strengths, it is important to identify sources of heterogeneity of the results through meta-analysis or pooled analysis techniques. Finally, when both animal and human data exist, the strengths and weaknesses of both bodies of data should be used to assess the optimal use of the epidemiological and animal data.

Dr Schnatter introduced a proposed scaling scheme for the use of null ("negative") human data in risk assessment. This scheme included the categorisation of the

quality of human data, categorisation of other complimentary data, such as *in-vitro* data and positive human data from different sources. In discussion, Workshop participants supported the use of all available information in risk assessment and the proposed scaling system as a logical way to provide guidance on data of different qualities.

3.6 Break-out session A: Use of human data on acute (local) effects due to (long-term) exposures

C. Money (Leader) and D. Schwela (Rapporteur)

The aim of this session was to:

1. Investigate the utility of the proposed assessment scheme when applied to commonly encountered human data;
2. identify areas where such data may be commonly compromised and hence which provide opportunities for improving the integrity of information collection and reporting;
3. identify areas for which data are missing and additional data will be valuable and contribute to the risk assessment of these chemicals.

Appendix D provides details of the case studies used and the specific learnings associated with each example. Based upon the case studies, participants of the break-out group (Appendix C) concluded that:

- Acute human data vary in their inherent quality and are available in different forms and from different sources, e.g. case reports, poisons centre records, workplace health surveillance activities, self-administered surveys, and formal scientific or epidemiological investigations;
- there is a need for acute effects data to be accompanied by information on exposure so that they can be considered more fully. A number of relevant issues were identified including the nature of concomitant exposures and the background incidence of different toxicological effects;
- there is a need for better characterisation and understanding of the quality and robustness of reporting systems in order to aggregate data from multiple sources or from different reports. Guidance on the establishment of recording systems and on the minimum data to collect for specific purposes would be useful in order to have more confidence in using human data.

3.7 Break-out session B: Use of human data on chronic effects due to long-term exposures

G. Swaen (Leader) and R. Schnatter (Rapporteur)

In this session a number of examples of chemicals causing chronic effects due to long-term exposures were presented, e.g. vinyl chloride monomer, welding fume and acrylonitrile (Appendix E). These were followed by a number of statements to initiate discussion. The following key points emerged from the discussion:

- The availability of dose-response information over a range of exposure concentrations is critical for risk assessment, including the determination of the no-effect level.

In the evaluation of the quality of human studies, the Hertz-Piccioto criteria of 1995 [a] should be used. In order to be able to perform risk characterisation, dose-response must be adequately assessed. For example, in a cross-sectional study on welding fumes, the only available data indicated that exposures varied between 5 and 200 mg/m^3, leading to an average annual reduction in FEV_1 of 0.5 to 2% for the total cohort. Participants did not regard these data to be of sufficient quality and specificity to adequately assess dose-response. For a dose-response relationship to be established, individual exposure and effect data are needed over a range of exposure concentrations. Some participants doubted if these data would be of sufficient quality even for hazard identification. It was stressed that sufficient confidence in the hazard identification was required in order that any related risk characterisation could proceed. Another example was presented in which only an exposure range was available and no exposure concentration at which no effects were noted, specified. It was concluded that these data were inadequate for risk characterisation, since it was not possible to estimate a no-effect level.

Workers are usually exposed during their job at different levels and times of exposure (in contrast to animal experiments where the dosing regime is kept constant). The variation of the exposure leads to a variety of problems regarding the selection of the "optimal" variables for characterising the exposure. There are several ways proposed in the literature how to describe the exposure, such as cumulative and average exposure, and duration of exposure. Also some proposals about lagging are available, e.g. ignoring exposure in recent years. Time intervals of 10 to 15 years are often used. But these proposals are of limited value for analysing human data. Therefore, research should focus on (i) development of methods which take into account the individual pattern (time and level) of the exposure, and (ii) development of adequate statistical methods for analysing these data.

[a] Hertz-Piccioto I. 1995. Epidemiology and quantitative risk assessment: A bridge from science to policy. 1995. *American Journal of Public Health* 85:484-491.

- The difficulties of using cancer registries and health surveillance systems for showing the absence of effect.

The example of vinyl chloride monomer (VCM) was used to discuss what data were needed to prove that a given low exposure dose does not induce angiosarcoma. It was recognised that the current case register can never prove the null hypothesis (i.e. no relation between exposure and effects), but can be used to make clear that there is no or perhaps a very small health effect for rare outcomes. One should always be aware that there might be some under-diagnosis, i.e. a case of angiosarcoma possibly not reported or misdiagnosed. Thus health surveillance can help to determine the priority of assessing the risk of a particular chemical. The group recommended continuing health surveillance, certainly in exposure situations with specific and known health effects.

- The importance of using the full range of toxicological information, including information on mode of action in addition to rodent carcinogenicity studies and human epidemiological studies, e.g. in the assessment of acrylonitrile which was associated with null epidemiological data.

Where several human studies might be available, the best quality studies should form the basis for risk assessment. Moreover, pooled analysis or meta-analysis on the good quality studies may provide additional insight.

The example of 1,3-butadiene was presented as a basis to consider in case of conflicting results: in production there appears to be an excess of leukaemia, while in polymerisation, there seems to be an excess of lymphoma. Participants agreed that there are limits to epidemiology and that risk assessors must recognise that in some instances where knowledge is limited, no definite conclusion can be drawn. Butadiene appeared to be one such case and hence it would be prudent for risk management purposes to treat it as a carcinogen.

Another example was the possible association between cellular phone use and the risk of brain cancer. It was clear that the one positive result should be seen as a hypothesis generating and not hypothesis testing. In that particular study, only specific sub-analyses showed some association, and participants agreed that, by doing more analyses than originally planned for, the study would become a "fishing expedition". Caution must be taken to avoid this and authors should report the original hypothesis. If more analyses were performed, it should be mentioned as such in the report. In discussion, this example was used to explain the importance of a fixed protocol for the statistical analysis. The positive result was thought to be a chance finding, resulting from lack of adherence to the original protocol for the statistical analysis and reporting the findings as original hypothesis.

Finally, statistical modelling techniques for human data risk assessment were discussed. It was agreed that the results depend to a great extent on the statistical model used. The choice of the model should be based on mechanistic considerations; it was not possible to distinguish the appropriate model on the basis of empirical analyses (e.g. threshold or not).

3.8 Break-out session C: Relevance of different data sources for risk assessment

L. Onyon (Leader) and P. Amcoff (Rapporteur)

The aim of this session was to identify the different sources of human data available for risk assessment, to discuss their relevance and what could be done to improve the confidence in using human data in risk assessment processes. The following key points emerged:

- The main obstacle to use of human data is its limited accessibility.

This obstacle stood in the way of being able to address whether existing data was being used to the maximum extent and to demonstrating and exploring its relevance and use. Improving the dialogue between generators of human data and risk assessment practitioners, together with an increased mutual understanding of why and how data may be generated and their possible use, was seen as a key to helping to overcome the obstacle. Ways to encourage routine reporting of human data, including among industry, and ways to encourage greater review and use of data, e.g. by allowing access to published epidemiological data in a way that encouraged independent analysis, were seen as complementary and supportive measures.

- Priority areas for improving the use of human data include better information about the circumstances and levels of exposure, toxicokinetic data and information that enables validation of toxicity test methods for which animal models are not available.

Understanding the availability of human data for different toxicological endpoints and at different points in the risk assessment process was discussed in order to identify priority areas to target. Better information about the circumstances and levels of exposure was identified as critical, as human data are often, even when they are available, poorly reported and characterised in terms of exposure. Further use of human data to refine traditional models used in toxicology was considered important. This included making the most of toxicokinetic information to improve the understanding of inter-species differences and extrapolation of animal toxicology

results to humans. Related to this, the use of human data to help gain a better understanding of modes of action of chemicals was seen as key point for development. Existing human data that enabled the potential validation of toxicity test methods for which animal models are not used or are not available, e.g. eye irritation and respiratory sensitisation, should also rank highly. It was concluded that risk assessors should be encouraged to systematically identify information that would be useful to improve certainty in their assessments. This would allow those collecting and reporting information to consider ways to increase the availability of these data.

- A framework for enabling improved use of human data will be useful for risk assessors. However, this work needs to be considered in stages, including checklists for data quality and refinement and extension of existing principles for assessing weight of evidence.

Participants discussed the proposal to develop a framework for considering quality in study design and evaluation of human exposure and effects data. The initiative taken by ECETOC, and Dr Swaen's proposal in particular, were welcomed. Further work in this area was seen as critical to increase confidence in the use of human data, by improving transparency and consistency in the presentation and weighting of different types of data as a basis for conclusions concerning hazard identification and dose-response assessment. A process of wide consultation and consensus building in its additional development was also recommended. Focus on development of a checklist of considerations for weighting of different types of human data and consideration of its relative contribution in relation to other types of data (such as those on toxicity in animals and mode of action) and illustration through case studies was suggested. Initial work could build upon the criteria classically used to establish causality, namely those developed by Bradford-Hill [a], although other criteria, such as those postulated by Henle-Koch in the field of microbiology [b], might also be relevant. The use of epidemiological data in risk assessment, and in particular improving confidence in presenting and using null epidemiological data was another area where guidance would be useful. The need for training risk assessors to enable their assessment of complex human data (e.g. prospective epidemiological studies) was also recognised.

[a] Bradford-Hill A. 1965. The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine* 58:295-300.

[b] Evans A.S. 1976. Causation and disease: The Henle-Koch postulates revisited. *Yale Journal of Biology* 49:175-195.

- There is need for fora which bring together experts from different fields of expertise to discuss the use of human toxicological data.

The initiative of ECETOC for holding the Workshop was appreciated, particularly as a forum for bringing together different experts from the fields of clinical and regulatory toxicology, occupational health and safety and experimental toxicology together with risk assessors. An open discussion of issues between experts from government, academia, industry and non-governmental organisations should be pursued. ECETOC, IPCS and other organisations could consider promoting targeted fora or workshops to further the understanding and resolution of obstacles in the use of existing human data.

4. RECOMMENDATIONS

4.1 Recommendations

The Workshop made the following recommendations:

1. Useful and good quality human data can be found in many forms such as case reports, poisons centre records, health surveillance results, self-administered surveys and epidemiological studies. Many data are, however, not readily accessible and their usefulness is frequently constrained by the absence of adequate contextual information on exposures and other quality considerations.
2. A clear framework should be developed that enables human data from different sources to be collected and assessed in terms of quality and application in risk assessment. Any framework should be easily understandable and applicable in a consistent and transparent manner.
3. The evaluation framework must be built on solid foundations, which should cover the identification and/or development of guidelines for (i) good human data collection practice, (ii) good exposure assessment practice and (iii) good risk assessment practice, and better networking of existing information sources.
4. A sustained and collaborative effort is required to map out an action plan to improve the recording of existing data for risk assessment purposes. This should involve a range of different interests and organisations, and include awareness raising and training of key groups (risk assessors and data collectors).
5. A discussion needs to be initiated on how to provide incentives for mechanisms that encourage the contribution of human data of improved quality for risk assessment purposes.

4.2 Way forward

The conclusions and recommendations of the ECETOC Workshop were presented at a subsequent IPCS meeting (held in Cardiff from 25 to 27 February 2004) on the collection, reporting and use of human data, designed to identify and develop proposals for strategic alliances and work between those involved in risk assessment, poisons centre development and chemical incident alert, surveillance and response.

The review paper commissioned for the ECETOC Workshop on the use of human data in risk assessment has been submitted for publication in a scientific journal.

Both ECETOC and IPCS are committed to progressing the recommendations of the workshop. IPCS, as part of its work programme, is currently developing a matrix of relevant international initiatives. This work will enable information gaps to be identified and addressed, with the aim of describing a suitable framework by 2006.

ECETOC's work programme for the period will also include activities which address considerations affecting the scientific basis for such a framework and the development of supporting practices and procedures.

APPENDIX A: LIST OF PARTICIPANTS

Name		*E-mail*	*Affiliation*
P	Amcoff	patric.amcoff@oecd.org	OECD, Paris, France
J	Bridges	j.bridges@surrey.ac.uk	University of Surrey, UK
JM	Calheiros	j.m.calheiros@mail.telepac.pt	Portuguese ATSDR / University of Porto, Portugal
J	Descotes	jacques.descotes@chu-lyon.fr	Poison Center and Pharmacovigilance unit, Edouard Herriot Hospital, Lyon, France
N	Edwards	nick.edwards@gstt.nhs.uk	London Poisons Centre (IPCS)
T	Eikmann	thomas.eikmann@hygiene.med.uni-giessen.de	Giessen University, Germany
S	Fairhurst	steve.fairhurst@hse.gsi.gov.uk	HSE, Bootle, UK
W	Fischer	werner.fischer@cibasc.com	Ciba Specialty Chemicals, Basel, Switzerland
U	Gundert-Remy	u.gundert-remy@bfr.bund.de	Bundesinstitut für Risikobewertung, Berlin, Germany
D	Heederik	d.heederik@iras.uu.nl	DECOS, Utrecht, Netherlands
T	Jovaisa	jovaisa@tox.lt	Lithuanian Poison Centre, Vilnius, Lithuania
R	Kroes	rmkroes@worldonline.nl	IRAS University of Utrecht, Netherlands
O	Ladefoged	ol@vfd.dk	Danish Institute of Food Safety and Nutrition, Søborg, Denmark
L	Levy	lsl1@le.ac.uk	MRC Institute for Environment & Health, University of Leicester, UK
B	Meek	bette_meek@hc-sc.gc.ca	Health Canada, Ottawa, Canada
T	Meredith	mereditht@who.int	IPCS, World Health Organisation, Geneva, Switzerland
L	Mølhave	lm@mil.au.dk	University of Aarhus, Environmental and Occupational Medicine, Denmark
C	Money	chris.money@exxonmobil.com	ExxonMobil Petroleum & Chemical, Brussels, Belgium
L	Onyon	onyonl@who.int	IPCS, World Health Organisation, Geneva, Switzerland
S	Munn	sharon.munn@jrc.it	European Commission, JRC, Ispra, Italy
D	Pallapies	dirk.pallapies@basf-ag.de	BASF, Ludwigshafen, Germany
R	Schnatter	a.r.schnatter@exxonmobil.com	ExxonMobil Biomedical Sciences, Annandale, USA
D	Schwela	dieter.schwela@jrc.it	European Commission, JRC, Ispra, Italy
K	Straif	straif@iarc.fr	International Agency for Research on Cancer, WHO, Lyon, France
G	Swaen	g.swaen@epid.unimaas.nl	Maastricht University, Netherlands
J	Tomenson	john_tomenson@ici.com	ICI , Runcorn, UK
K	Ulm	kurt.ulm@imse.med.tu-muenchen.de	Institut für Medizinische Statistik und Epidemiologie, München, Germany
E	van Vliet	pw.van.vliet@gr.nl	Health Council of the Netherlands, The Hague, Netherlands
M	Wilks	martin.wilks@syngenta.com	Syngenta Crop Protection, Basel, Switzerland
G	Würtzen	gwurtzen@eur.ko.com	Coca-Cola, Hellerup, Denmark
K	Ziegler-Skylakakis	kyriakoula.Ziegler-skylakakis@cec.eu.int	European Commission, Luxembourg
G	Aldridge	gloria.aldridge@cardiffandvale.wales.nhs.uk	Cardiff Centre of the UK National Poison Information Service, Cardiff, UK
H	Vrijhof	henk.vrijhof@ecetoc.org	ECETOC, Brussels, Belgium

From left to right:

1st row:

G. Swaen, Maastricht University, Netherlands
C. Money (Co-chair), ExxonMobil Petroleum & Chemical, Brussels, Belgium
D. Schwela, European Commission, JRC, Ispra, Italy
L. Mølhave, University of Aarhus, Environmental and Occupational Medicine, Denmark
M. Wilks, Syngenta Crop Protection, Basel, Switzerland

2nd row:

B. Meek, Health Canada, Ottawa, Canada
K. Ulm, Institut für Medizinische Statistik und Epidemiologie, München, Germany
T. Eikmann, Giessen University, Germany
G. Würtzen, Coca-Cola, Hellerup, Denmark
J. Calheiros, Portuguese ATSDR / University of Porto, Portugal
J. Bridges (Chair), University of Surrey, UK
E. van Vliet, Health Council of the Netherlands, The Hague, Netherlands
K. Straif, International Agency for Research on Cancer, WHO, Lyon, France
L. Levy, MRC Institute for Environment & Health, University of Leicester, UK
T. Meredith, IPCS, World Health Organisation, Geneva, Switzerland
J. Descotes, Poison Center and Pharmacovigilance unit, Edouard Herriot Hospital, Lyon, France
N. Edwards, London Poisons Centre, UK

Back row:

W. Fischer, Ciba Specialty Chemicals, Basel, Switzerland
H. Vrijhof, ECETOC, Brussels, Belgium
D. Pallapies, BASF, Ludwigshafen, Germany
G. Aldridge, Cardiff Centre of the UK National Poison Information Service, UK
T. Jovaisa, Lithuanian Poison Centre, Vilnius, Lithuania
O. Ladefoged, Danish Institute of Food Safety and Nutrition, Søborg, Denmark
K. Ziegler-Skylakakis, European Commission, Luxembourg
U. Gundert-Remy, Bundesinstitut für Risikobewertung, Berlin, Germany
R. Schnatter, ExxonMobil Biomedical Sciences, Annandale, USA
S. Fairhurst, HSE, Bootle, UK
S. Munn, European Commission, JRC, Ispra, Italy
R. Kroes, IRAS University of Utrecht, Netherlands
J. Tomenson, ICI, Runcorn, UK
D. Heederik, DECOS, Utrecht, Netherlands

APPENDIX B: WORKSHOP PROGRAMME

Programme - Day 1

Chairs:		**Jim Bridges** **University of Surrey, UK** **Chris Money** **ExxonMobil, B**
12.00 - 13.00	Registration desk open; refreshments	
13.30 - 14.00	**Introduction**	**Jim Bridges,** **University of Surrey, UK**
14.00 - 14.45	**Interpretation of human data for risk assessment in the occupational setting**	**Steve Fairhurst,** **HSE, UK**
14.45 - 15.30	**Interpretation of human data for risk assessment in the outdoor environment**	**Thomas Eikmann,** **Giessen University, D**
15.30	Coffee/tea break	
15.45 - 16.30	**Interpretation of human data for risk assessment in the public health setting**	**Tim Meredith,** **IPCS, CH**
16.30 - 17.15	**A proposed framework for development of quality considerations on design characteristics, exposure and effect data, and evaluation**	**Gerard Swaen,** **Maastricht University, NL**
17.15 - 18.00	**Role of null epidemiology in risk assessment**	**Robert Schnatter,** **ExxonMobil, USA**
20.00	Workshop dinner	

Programme - Day 2

09.00 - 11.00	**Break-out sessions to discuss the suitability of the proposed quality considerations when applied to a series of selected case studies: 3 discussion leaders / 10 - 12 participants / group** **A. Use of human data on acute (local) effects due to (long-term) exposures** **B. Use of human data on chronic effects due to long-term exposures** **C. Relevance of different data sources and impact on risk assessment**	
11.00	Coffee or tea	
11.30 - 12.30	**Panel discussion** Reports from 3 working groups (Rapporteurs) and plenary discussion	**All speakers**
12.30 - 13.00	**Conclusions**	**Jim Bridges**
13.00	End	

APPENDIX C: BREAK-OUT SESSIONS - PARTICIPANTS AND QUESTIONS FOR DISCUSSION

C.1 Session A: Use of human data on acute (local) effects due to (long-term) exposures

C. Money (Leader), N. Edwards, T. Eikmann, U. Gundert-Remy, T. Jovaisa, R. Kroes, T. Meredith, D. Schwela (Rapporteur), M. Wilks

There are a variety of sources of human data that can describe acute (or local) health effects. Such sources range from isolated case reports and poisons centre incidents to the findings of structured studies and surveillance programmes. The extent to which these sources might be reliably used within the risk assessment process is dependent not only on the inherent quality of the available data, but the purposes to which it is intended to be used. The aim of this break-out session is to discuss the range of data sources that are potentially available to risk assessors and, through the use of suitable case studies, to address the following questions:

1. Can minimum "quality criteria" be defined for any type of information source and to what extent might these affect the areas of the risk assessment process (effect identification; identification of lead effect; identification of LOAEL/NOAEL; shape of dose-response curve; *etc.*) to which the data could be applied?
2. When seen in the context of animal findings, how reliable are human data for the range of acute (and local) effects that may be present?
3. Are any sources of acute human data so inherently unreliable that they should be ignored for routine consideration within risk assessments?
4. Are there areas where such data may be commonly compromised and hence which provide opportunities for improving the integrity of information collection and reporting?
5. What role has null studies in the interpretation of acute effects? Can criteria be established which might assist in evaluating the importance of null findings associated with any particular data source? In what situations can the absence of reported effects be relied upon to conclude that there are no effects?

In order to help discussions, a number of specific examples of human data on acute health effects will be presented and discussed. The discussions during the break-out session will be summarised for the concluding panel discussion.

C.2 Session B: Use of human data on chronic effects due to long-term exposures

G. Swaen (Leader), J.-M. Calheiros, W. Fischer, O. Ladefoged, L. Levy, D. Pallapies, R. Schnatter (Rapporteur), K. Straif, K. Ulm, G. Würtzen

If good quality human data and good quality animal data are available, should we give priority to the human data, even if the animal data would result in more stringent exposure limits than based on the risk assessment on the human data only?

1. Even if standard operational procedures for quantitative risk assessment can be developed and consensus about these can be reached, risk assessment will always contain a substantial degree of expert judgement and therefore must be done on a case-by-case basis.
2. In epidemiologic studies it is not always possible to rule out the confounding effect of other risk factors and sometimes alternative explanations for an increased risk can be put forward. Should we take these alternative hypotheses seriously only if there is empirical support for these alternative explanations, and what are the minimum requirements for this empirical support? In other words: should we turn around the burden of proof and only put aside the human data if there is a substantial indication that a confounder is at work?
3. Should observational human data of good quality in which no adverse effects are observed be treated in a similar fashion as human studies in which effects have been observed? Is the application of statistical confidence limits around "negative" (null) studies to statistically exclude a certain risk over conservative, leading to unnecessary stringent exposure limits?
4. Industry should design and put into practice mechanisms to monitor the occurrence of adverse health effects of their products. A parallel with post marketing surveillance can be drawn.
5. Within the EU there is legislation that requires testing before new chemicals can be used and introduced to the market. Should these requirements be expanded to include human data on possible exposure and some form of monitoring exposed individuals? An example of a new product could be genetically modified crop.
6. A joint effort should be undertaken to preserve the original data of both animal and human studies on which formal risk assessments have been made for the future.

C.3 Session C: Relevance of different data sources for risk assessment

L. Onyon (Leader), P. Amcoff (Rapporteur), J. Bridges, J. Descotes, S. Fairhurst, D. Heederik, B. Meek, L. Mølhave, S. Munn, J. Tomenson, E. van Vliet, K. Ziegler-Skylakakis

The syndicate group was organised around a small number of questions and based discussions on the types of human data available, the availability of data for different toxicological endpoints and exposure and the relevance of human data at different stages of the risk assessment process, i.e. priority setting, hazard identification, hazard characterisation, exposure assessment and risk characterisation (risk management).

C.3.1 Types of human data

- Epidemiological data, looking for associations between hazards and dose-response
- Routine data collections, poisons data, adverse event notification schemes, occupational health surveillance (OHS) reporting, coroners' reports
- Biological monitoring/personal sampling
- Human kinetic studies - observational clinical data
- Published and unpublished industry studies
- Anecdotal reports

C.3.2 Availability of data for different toxicological endpoints

- Acute toxicity - local, systemic
- Sensitisation
- Short-term toxicity
- Genotoxicity
- Carcinogenicity
- Reproductive toxicity (fertility and development)
- Toxicokinetics

1. Given the assumption that we would all like to see more human data and have greater confidence in using it for risk assessment, what can be done to achieve greater relevance (and availability) of data?
2. What are the obstacles that may discourage risk assessors from using human data and issues that could be addressed as a priority?
3. Human data seems to be commonly used as confirmatory to studies in animals. Under what circumstances do we envisage human data playing a greater role than now?

4. What are some of the reasons for collecting human data and would there be any value in pooling information on certain endpoints?
5. What could be the role of ECETOC in promoting maximum use of human data?

APPENDIX D: CASE STUDIES ON THE USE OF HUMAN DATA ON ACUTE (LOCAL) EFFECTS DUE TO (LONG-TERM) EXPOSURES

D.1 Purpose

1. To investigate the utility of the proposed assessment scheme when applied to commonly encountered human data;
2. to identify areas where such data may be commonly compromised and hence which provide opportunities for improving the integrity of information collection and reporting;
3. to identify areas for which data are missing and additional data will be valuable and contribute to risk assessment of these chemicals.

D.2 Background

Human data that describe local (acute) effects are available from a number of different sources. These data invariably differ in their inherent quality. Accordingly, the extent to which they can be reliably used within the risk assessment process varies too. The following case studies describe typically encountered instances where human data may be cited in support of a substance and where a decision will need to be taken concerning the admissibility and value of the human experience.

D.2.1 Case study 1

Testing initiated under the HPV programme has identified substance X as being a dermal sensitiser (positive guinea pig maximisation test, GPMK). Substance X has been manufactured for over 50 years at a single site in Europe. The current plant physician reports that no cases (n = 250) of dermal sensitisation have ever been recorded at the plant (or detected during annual medical examination). No specific skin surveillance programme is followed, as the substance has not previously been considered to present a significant dermal risk. The substance is a solid and hygiene practices at the site are not high (apart from the GPMK result, the substance is otherwise classified as non-hazardous) and thus exposures are likely to have been elevated in the past.

- How reliable are the local human data when describing the lack of human dermal sensitisation potential of the material? (It can be assumed that no case reports have been recorded/reported elsewhere.)
- If the data are useful, then in what areas of the RA process might the information be applied (viz. dose-response evaluation, *etc.*)?
- What other information would make the data more useable/reliable?

- Are there lessons arising from how the data are described that can be applied in general, either to the systems of workplace health monitoring, which generate the data, or to the way these data are reported.

Case study 1 learnings

The human data are potentially useful but do not have a role in the interpretation of the animal findings as currently presented.

- There is a need for the data to be supported with relevant exposure information, particularly concerning the magnitude of dermal exposures (and would be relevant for helping to contextualise the relevance of the human findings in the context of sensitisation potency). Quantitative exposure data would be ideal, but failing this, good qualitative information would also be valid. Information on the nature of the processes, including variations in likely particle size, would also benefit the risk assessment of the substance.
- Key information is also missing concerning the quality of the human experience (labour turnover rates; potential impact of healthy worker effect; nature of the health surveillance programmes in operation; integrity of local record keeping; *etc.*) and hence the weight that might be placed on it.
- The nature of the animal findings also requires further explanation in order to better evaluate their relevance in the context of actual human experiences, e.g. are the dosing levels yielding effects heroic when compared to those experienced by occupational groups.

D.2.2 Case study 2

Several cases (n = 20) of severe eye irritation (acute pain persisting for 5 - 10 minutes post exposure with no long-term consequences or sequelae) have been reported by various European poisons centres following the use of a branded spray consumer oven cleaner. All of the ingredients of the cleaner have been tested *in vivo* and none is classified as an irritant. The reported cases are all amongst elderly women (> 65 years). The cleaner has been marketed for 18 months and around 10,000 cans are manufactured daily.

- How reliable are these data likely to be when seen in the context of the animal data?
- What relevance should be placed on the reported human experience?
- How should the quality of the human information be assessed?
- Why, within the context of the proposed scheme, might such information find use within the process of risk assessment?

- Might the nature of the symptoms suggest a need for any prudent immediate risk management measures, e.g. product recall or media communication?

Case study 2 learnings

The human data are interesting and potentially relevant for informing policy making but would not be sufficient to initiate any action without further information being available.

- There is a need for the data to be supported with relevant exposure information. Questions such as with what circumstances (of exposure) the incidents were associated and whether these were normal conditions or they represent some form of unusual use pattern are all critical when interpreting accurately the relevance of human experience. The importance of recording relevant exposure information at poisons centres was noted, together with the implications of this for training.
- Key information is also missing concerning the quality of human experience (consistency of the classification) and hence the weight that might be placed on it. These aspects reinforce the need for consistency of classification of human experiences within poisons centres (and elsewhere). The potential need for prospective follow-up of cases was noted.
- It would be beneficial if the relevance of the findings were capable of wider investigation via the ability to interrogate poisons centre records across the area of sale of the product. Are these isolated instances or do they reflect a wider trend that has simply not been identified or reported elsewhere? The need for poisons centres records to be flagged in order to monitor potentially unusual trends and the difficulties of searching free text in such records were noted.

D.2.3 Case study 3

Little is known about the specific health status of populations in the metal-working industry, although both irritant and contact dermatitis are commonly encountered conditions. An isolated report suggests that the prevalence of keratoses (a pre-malignant skin lesion) appears high (approximately 30%) in one exposed working groups (n = 20). Recent animal data (18-month skin painting study with no other reported adverse findings) indicate that (repeated) contact with substance Y can result in skin cancer. No knowledge of the mechanism of action is available, but the substance is a known dermal irritant and is used as an additive in metal working fluids (for machining the specialty alloys). The substance has only been manufactured for 10 years.

- What reliability and/or weight can be placed on the reported human information in the light of the animal findings?
- Would the human data suggest that the substance is a likely human carcinogen?
- Is the information of use in any other part of the risk assessment process?

Case study 3 learnings

The human data are potentially relevant but much necessary information is missing before they would be sufficient to be considered to be substantive. As currently presented, the human data would not have much bearing on the interpretation of the animal findings.

- There is a need for the data to be supported with relevant exposure information. In this case, there is very little such information provided and there is a need not only to include information that describes occupational exposures to the substance but also exposures to other (occupational and non-occupational) potential causes of the condition, e.g. sunlight, use of skin care products, etc.

Key information is missing concerning the quality of the human experience, particularly data on the background incidence of the condition in normal populations, the distribution of the keratoses on the body (are they in locations associated with dermal contact such as the hands and arms, for example), and the diagnostic criteria employed. These factors materially affect the weight that might be placed on the report. The potential need for prospective follow-up of cases was noted.

D.2.4 Case study 4

Two confirmed cases of respiratory allergy caused by a reactive dye (reactive pink 107) have been reported by a regional occupational health clinic. Both cases have occurred amongst craft dyers when using the dye for mouth spraying via a nebuliser (no quantitative exposure information is known). The dye has been marketed for 2 years. The company selling the product operates an adverse health effects register, based upon reports received from its customers, in support of its TSCA obligations. The company argues that based upon the fact that no other cases have been reported in either the literature or its own register and that the reported cases most likely occurred due to exposures associated with clear conditions of misuse, the dye does not result in allergy under normal conditions of use and hence should not be classified as a respiratory sensitiser.

Assuming the data from the occupational health clinic are reliable,

- What relevance has the company's adverse effects register in terms of its ability to provide any relevant risk assessment information? What basic considerations might be expected for such surveillance systems if their findings are to be considered useful for RA purposes?
- How reliable is the scientific literature as a source of human data (and particularly so for newer chemicals)?
- Can national adverse effects registers such as TSCA be effectively mined to release useful human data?
- If the data are useful, then in what areas of the RA process might the information be applied (viz. dose response evaluation, etc.)?

Case study 4 learnings

Despite there being areas where the quality of the human data, as presented, might be further improved, it was felt that the information on the reported cases is sufficient to be of use in classifying the substance as a respiratory allergen. However, it was also noted:

- That case data such as these often are not supported by adequate relevant exposure information. This makes their interpretation difficult and limits their use when informing risk management strategies.
- That data that better inform about quality of the human experiences is often lacking. In this instance, the cases have been confirmed by a specialist centre. But, if this were not the case, then the data would be insufficient to classify and would warrant more detailed follow-up. It was considered that integrated reporting systems for key occupational (and non-occupational) diseases should be developed at the national and regional levels in order to improve the overall quality and continuity of data that can help inform risk assessment in these areas.
- The fact that an internal monitoring system for adverse effects was in operation was applauded. This is seldom routine. However, it was felt that guidance needs to be developed that describes the minimum considerations for such systems if they are to provide data useful for risk assessments. Systems that are simply based on ad hoc reporting are not inherently reliable. If such a system were operational, then this would provide valid data on the likely potency of the allergen.

D.2.5 Case study 5

Following complaints received from its members, a trade union undertakes a self-administered questionnaire (90% response rate) amongst workers at a chemical plant that identifies that 60% of the respondents report that they feel nauseous during the formulation of a product. The product has been manufactured for several years without complaint, but recently was re-formulated with the addition of a new additive (substance Z). The levels of substance Z in the workplace air are about 50 ppm. The supplier of substance Z reports that no such complaints have been reported during its manufacture (although exposures might be expected to be < 5 ppm). The manufacturer also reports that no adverse findings have been identified during animal testing (including 90-day repeat-dose toxicity studies). The trade union seeks the advice of the regulatory authority in trying to establish an exposure level that will not result in further complaints. The local management of the firm argues that self-reported studies are fundamentally biased and must therefore be discounted.

- Can self-administered questionnaires (SAQ) ever be used to reliably determine the presence or absence of health effects? Can any criteria be defined which would tend to suggest such findings might be given more or less weight?
- In the case described above, can any credence be given to the findings of the SAQ study undertaken by the trade union? Do the findings have any relevance in the context of the request to establish a safe working level?
- If the data are useful, then in what areas of the RA process might the information be applied (viz. dose-response evaluation, *etc.*)?

Case study 5 learnings

Insufficient time was available to discuss the case study in detail. However, it was noted that human data derived from SAQs are potentially useful and can constitute a basis for action in themselves but do need to meet basic epidemiological criteria, particularly any influence of participation or recall bias, and the need for the findings to be supported with adequate exposure information. The findings, as currently presented, would not be sufficient to establish any regulatory exposure limit but are such as to warrant a more thorough investigation of the situation.

APPENDIX E: CASE STUDIES ON THE USE OF HUMAN DATA ON CHRONIC EFFECTS DUE TO LONG-TERM EXPOSURES

E.1 Purpose

1. To exchange views on the validity, reliability and interpretation of human data on chronic effects due to long-term exposures.
2. To further develop consensus statements on these issues.

E.2 Background

Long-term health effects of exposure to chemicals are a matter of great concern. Providing evidence for such effects requires complex data sets and labour intensive research, usually following long latency periods. Research in this area is further complicated by limited availability of exposure data and possibly confounded by differences in distribution of other risk factors. Given this complex research, in many cases there is room for discrepancies with respect to the weight of evidence that should be given to study results and the subsequent interpretation of research findings. In this parallel session a number of case studies were presented and discussed to attempt to derive consensus on research needs and the interpretation of study results on chronic effects of long-term exposures.

E.2.1 Case study 1

In a prospective study of 47 welders with an average follow-up of ten years, an average annual reduction of pulmonary vital capacity between 0.5% and 2% is observed. Ambient air concentrations have been measured occasionally and were between 5 and 200 mg/m^3 of total dust. Can we derive an occupational exposure limit based on this study?

Case study 1 learnings

This study lacks specificity both with respect to the exposure situation and to the health effect. The session participants regarded the study to be insufficiently informative and not suitable to serve as the basis for the construction of a dose-response curve.

In addition the data cannot be used to estimate a no-effect level. Furthermore, we are dealing with a non-specific health effect and adjustment for other known risk factors such as cigarette smoking and concomitant exposure in the workplace have not been

made. Had the study been designed in an appropriate manner and had the exposure and effect data been available on an individual basis, the study could have been a valuable data source for risk assessment.

In conclusion, the study described above was thought not to be sufficiently informative to serve as a basis for risk assessment.

E.2.2 Case study 2

In a cross-sectional study of 100 workers exposed to formaldehyde concentrations averaging 0.22 ppm (8-hour average) in workplace air, 50% of the workers report to experience discomfort and irritation of the upper respiratory tract. Industry argues that in normal work situations with similar exposure concentrations these symptoms are not reported.

What type of data with what sample size should be collected to refute the findings above?

Case study 2 learnings

It is possible that the observed effects are due to peak concentrations that may regularly occur. A new study on this association should attempt to improve the exposure data in such a way that the information is more specific and available on an individual basis. A new study should also include a non-exposed comparison group since the reported symptoms are not specific. Such an approach would enable to investigate the shape of the dose-response curve. A dose-response analysis could provide evidence for a no-effect level. A new study should take into account excursions or peak concentrations.

E.2.3 Case study 3

In the late nineteen seventies a number of cases with angiosarcoma of the liver have been reported in workers exposed to vinyl chloride monomer (VCM). These observations have later been confirmed in extensive retrospective cohort studies in other factories. After the first reports the exposures to VCM have been greatly reduced and now are a fraction of the exposures that have occurred in the past. An industry-wide health surveillance system has been set up to monitor the incidence of angiosarcoma in the VCM industry. The data collected by means of this surveillance system indicate that the incidence of these tumours has strongly declined and that no cases have so far been reported in workers who have not experienced the high-exposure situations in the past. Under what conditions can it be concluded that the current VCM levels do not increase the risk for angiosarcoma of the liver?

Case study 3 learnings

Angiosarcoma of the liver is generally recognised as a specific effect, strongly related to VCM exposure. The health surveillance system maintained by the industry can provide very important evidence on the current risk for this disease. However it is known that cases of angiosarcoma are misdiagnosed as hepatocellular cancer cases and thus can be missed by the health surveillance system. Secondly, the latency period of angiosarcoma is not exactly known and it can be expected to be longer in low-exposure situations than in high exposure situations. In summary it was concluded that the health surveillance system cannot serve as complete evidence for a lack of risk for this specific effect, because of unknown completeness of the health surveillance system and because of the possibly longer latency required in low-exposure situations.

E.2.4 Case study 4

Chronic inhalation studies in rats show that acrylonitrile can induce several types of tumours, including brain tumours. These effects already occur at exposures to 20 ppm.

A meta-analysis on 25 retrospective cohort studies of acrylonitrile workers reveals no evidence for any excess cancer risk, but most studies evaluated relatively low-exposure situations. Current risk assessments are based on the positive animal data; the good quality, and extensive, human data play only a marginal role.

When can we use negative human data for risk assessment?

Case study 4 learnings

The workshop participants regarded this question as a complex matter. There is an essential difference between animal data and human data. Risk assessment should use all available good quality evidence. Human observational data can only rarely be used to exclude risks, and it is doubtful if they can ever be used to prove the non-existence of risks, because there is always the chance of a false negative outcome. Calculated relative risks must be regarded as point estimates with a certain confidence interval around them. Human data of good quality can be compared with good quality animal data in order to establish the validity of the animal model. If health effects require metabolic pathways, it is always questionable if an animal model such as the rat is representative for humans. If there is toxicological evidence for differences in metabolism between the rat and humans that can explain the differences in effect, such a comparison can be quite valuable. However, in the context of positive animal data, null human data should remain to be treated with great caution and necessary conservatism. The group felt that using 95% confidence limits to exclude certain excess risk is too conservative, since it involves two-sided testing. One-sided testing with a 95% degree of confidence was regarded to be more appropriate.

E.2.5 Case study 5

The possible carcinogenic effects of butadiene have been studied in butadiene production workers as well as polymerisation workers. In a cohort of production workers an excess of leukaemia mortality has been observed, whereas a study in the polymerisation industry reported an excess of non-Hodgkin's lymphoma. Under the assumption that these effects should not be attributed to exposure to other chemicals, the results are not in support of each other. How should these results be interpreted and what new data could provide further insight in the matter?

Case study 5 learnings

Particularly since chronic animal tests also provide evidence for a carcinogenic effect of butadiene, the results of the human studies should be interpreted with caution. The general agreement was that there are limits to what epidemiological studies can investigate and that there is always a potential for false positive or false negative results. In cohort studies many endpoints are investigated and, given all these uncertainties, a certain degree of inconsistency between epidemiological studies should be expected. Thus the workshop participants agreed that these human data should be treated with great caution and that the data indicate that butadiene may be a human carcinogen.

E.2.6 Case study 6

In a case-control study of brain tumours and cellular phone use an Odds Ratio of 1.2 (95% confidence interval: 0.80 - 1.60) was found. If the side of the head where the cellular phone was held was taken into account an Odds Ratio of 2.5 (95% confidence interval: 1.3 - 4.9) for acoustic neurinomas emerged. A second comparable study was negative. How should these data be interpreted?

Case study 6 learnings

The interpretation of human data, when results are more or less conflictive, is always difficult and in these instances particular interest should be given to the original research objective of the individual studies. The overall finding of the first study was essentially a negative one. Only the further secondary analysis revealed a statistically significant association with cellular phone use. It is questionable if this secondary analysis was already a part of the main research objective and it is possible that the researchers conducted several secondary analyses and then only reported the ones with positive findings.

E.2.7 Conclusions

It was concluded that in scientific research a clear distinction must be made between the findings with respect to the original study hypothesis and that the results of secondary analyses should always remain acknowledged as such in the reports. There is a clear distinction between hypothesis generating and hypothesis testing research and the stronger weight of evidence clearly lies with hypothesis testing projects.

Apart from these case studies a number of discussion points were put forward to further develop consensus on these matters. A first question discussed was if good quality human data should always be given priority to good quality animal data. In practice, if the animal data are negative and the human data are positive the priority is given to the human data and if the animal data are positive and the human data are negative the priority is given to the animal data. This is a built in mechanism of conservatism and must be recognized as such. Only if sound mechanistic data are combined with good quality negative human data can positive animal data be laid aside as not relevant for humans.

A second question concerned the role of expert judgement in risk assessment. The participants agreed that there is a need for developing standard operational

procedures for quantitative risk assessment and that a strong effort should be made to achieve this. However, there will always be an important role for expert judgement and risk assessment should be carried out on a case-by-case basis. No data set is exactly comparable with another and there still can be other important parameters that should be taken into account.

Thirdly, it was discussed to what extent and in which situations possible confounding should or should not be taken into account. The problem of confounding only applies to non-specific effects and in situations in which confounding is likely to be an issue. For instance if the effect is lung cancer, smoking is a likely confounder and should always be taken into account. However if the effect is brain cancer or lymphoma, for which no strong and clear confounders are known, confounding is not such an issue and study results in which no adjustment has been made for the effect of other factors should still be considered as reliable and informative.

A fourth interesting topic for discussion was the weight of evidence given to null human data. Basically the same quality criteria should be applied to positive or negative studies. However, the weight of evidence of negative studies decreases if the sample size is limited and if the exposures have been low or not well documented. Again, it was noted that one-sided testing on a 5% level is more appropriate than two-sided testing, since the *a priori* hypothesis is focused on an adverse effect on health and not on a positive effect on health.

Next, the implications of the need for human data for newly introduced chemicals were discussed. The current testing requirements do not include human data in the form of monitoring or health surveillance data. It was felt that health surveillance could only contribute relevant information if there is a clear indication for the potential target organ. Then the surveillance system could focus on the incidence of diseases in that target organ. It is questionable if clinical tests are sensitive enough to pick up early adverse health effects in exposed populations.

Finally, the importance of good documentation and archiving of human studies was stressed. Since exposures are more and more reduced it is expected that many risk assessments in the future will still rely on studies of humans that have already been conducted. Therefore it is of the utmost importance to have available a good documentation and archiving of these studies for future reference and evaluation. Perhaps an effort should be made to preserve the original datasets of human studies so that sophisticated statistical analyses can be performed on these datasets in the future.

APPENDIX F: ORGANISING COMMITTEE

G. Swaen (Chairman)	University of Maastricht NL - Maastricht
P. Amcoff [a]	OECD F - Paris
C. Money (Co-chairman)	ExxonMobil B - Machelen
L. Onyon	IPCS CH - Geneva
S. Pakalin [a]	European Commission (JRC-IHCP) I - Ispra
D. Schwela	European Commission (JRC-IHCP) I - Ispra
R. Schnatter [b]	ExxonMobil USA - Annandale
H. Vrijhof (Secretary)	ECETOC B - Brussels

[a] Part-time
[b] Corresponding

ECETOC PUBLISHED REPORTS

Workshop Reports

No.	Title
No. 1	Workshop on Availability, Interpretation and Use of Environmental Monitoring Data 20-21 March 2003, Brussels
No. 2	Strategy Report on Challenges, Opportunities and Research needs arising from the Definition, Assessment and Management of Ecological Quality Status as required by the EU Water Framework Directive based on the workshop EQS and WFD versus PNEC and REACH - are they doing the job ? 27-28 November 2003, Budapest