Evaluation of Standards and Models for Probabilistic Exposure Assessment

Final Report

- Executive Summary -

Dessau-Roßlau (Germany), November 2007
Evaluation of standards and models for probabilistic exposure assessment - Xprob

The project was carried out by a consortium consisting of: University of Bielefeld, Fac. Public Health; University of Hamburg / Institute of Biometry and Epidemiology; Epidemiology Working Group of the Hamburg Science and Health Department; University of Bremen, Institute of Risk Science Environment Health; Institute of Public Health North Rhine-Westphalia; Lower-Saxony Health Department.

Exposure assessment is an essential element of scientific risk assessment. For the successful application of (probabilistic) exposure modelling, both science based models and sufficient data of good quality together with guidance on their adequate usage are necessary. To distinguish between scientific risk assessment and administrative risk management, precise requirements on the subject, objective and level of protection have to be defined for transparent reflection of scenarios by exposure models. Variations of exposure characteristics in the population have to be described in the same way as uncertainties with regard to scientific derivation since both will effect the evaluation result. A tiered approach is recommended, beginning with the "traditional" form of exposure assessment based on average and reasonable most exposed case assumptions ("point estimates") and complementing it with distribution-based methods. The latter, however, is only applied in "unclear" situations – and if sufficient data are available. The data base for exposure factors was evaluated with a special focus on data from Germany. A comprehensive literature search and analysis of data sources revealed that there are pronounced data gaps. A methodology for deriving exposure factor distributions was developed and applied to the data sources. Recommendations on reference values for various exposure factors were derived and documented in a standardised format. However, in order to become standards, these reference values should be authorised by legitimate bodies. A database (RefXP) was developed and implemented for structured documentation and retrieval of the information gathered. Distributions as well as point estimates are provided for exposure factors. Methods and sample applications of distribution-based exposure modelling are described. The methodology developed for deriving distribution-based exposure factors can provide the basis for setting new standards in Germany. "Good practice" principles were derived which are intended to facilitate the application of the distribution-based methodology in exposure assessment. With its contribution to quality improvement of case-related risk assessments, the project result might also support areas such as health impact assessments as well as national and European chemicals legislation.
Executive Summary

The quantitative characterisation of health-related risks is an important element of defining transparent, rational priorities in regulations and policies. In Germany, quantitative exposure and risk assessment is gaining in importance. In addition to dose response assessment (i.e. the quantitative description of the toxicological hazard potential and the dose effect relationship), human exposure assessment (i.e. the description of the distribution of exposure conditions in the population) constitutes an important pillar of risk assessment.

Exposure modelling describes a person’s intake of harmful substances – its exposure – via various routes (inhalative, oral or dermal) and via various media (air, drinking water, food, dust etc.) for combined assessment. Many variables influencing the intake, such as inhalation rates, food consumption quantities, or anthropometric data (e.g. skin surface, body weight), are to be assessed independently from the pollutant and constitute the basis for so-called exposure factors. The provision of handbooks quantifying such exposure factors serves the harmonisation of risk assessment in various regulatory fields, ensures a permanently high scientific level in selection and application of the exposure factors and reveals information deficits and research needs.

Under the Environment and Health Action Programme (Aktionsprogramm Umwelt und Gesundheit - APUG), the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, through the Federal Environment Agency funded the project “Evaluation of standards and models for probabilistic exposure assessment” (Xprob). The project was carried out between September 2002 and November 2005 and aimed to update the “Standards for Human Exposure Assessment” published in 1995 by the Committee for Environmental Hygiene (AUH\(^1\) – now LAUG\(^2\)). A second goal was to update the methodological recommendations according to the present state of research. Following the intentions of APUG, the project should make a contribution to harmonising risk assessment in Germany. Under the management of the University of Bielefeld, a project group was established comprising the universities of Bremen and Hamburg, the Lower-Saxony State Health Institute as well as the Institute of Public Health NRW (lögd). This final report documents the most important results of the Xprob project. A CD containing the RefXP database, developed within the project, which includes an interactive data access for about 700 exposure factors in a standardised format accompanies this report.

Human exposure assessment is a fundamental element of scientific risk assessment and can be used both retrospectively for assessing human exposures which have already occurred and prospectively for assessing future exposures. To distinguish between (scientific) risk assessment and (administrative) risk management, precise requirements have to be defined concerning the subject, objective and level of protection which are then reflected in the exposure model in a transparent way.

---

1. AUH – Ausschuss für Umwelthygiene
2. LAUG – Länder-Arbeitsgruppe Umweltbezogener Gesundheitsschutz
Executive Summary

In this context, variations in exposure characteristics in the populations have to be described in the same way as uncertainties with regard to scientific derivations and their effects on the final result of the assessment.

As a result of the project, a tiered approach is recommended starting with the traditional approach of human exposure assessment based on (high-end) point estimates and complementing it in situations where relevant exposure cannot be ruled out. As a second step – if sufficient data is available – the exposure model should be recalculated with distribution-based methods. In this way the capacities required for distribution-based assessments only need to be provided for exposure situations that justify their input.

Exposure modelling is an important complement to the human biomonitoring (HBM) approach because it traces the influences of real environmental exposure and is able to estimate future exposure levels.

For evaluating the existing data base for the exposure factors in Germany, the project group used the AUH report as the starting point. The data of this AUH report on exposure factors were compared with the data and the data evaluation approaches in the handbooks for other populations. For selected topics comprehensive data and literature searches were carried out. The search was primarily concentrated on data and literature sources for Germany issued after the publication of the AUH report. These sources were analysed based on a standardised procedure and used for deriving recommendations for exposure reference values.

Under the project, a methodology was developed which aims at describing the distribution of exposure factors, instead of normatively defining the distribution. This methodology was developed based on an US Environmental Protection Agency (US EPA) approach, which is consistent with respect to statistical criteria. Moreover, an age stratification procedure was developed, optimally reflecting the differences between the various age categories described by empirical data. In doing so, very different data sets can be analysed using a uniform methodology and a consistent procedure for deriving reference values. This derivation procedure is transparent; in particular, the achieved correspondence between empirical and proposed distribution is described.

A description of this methodology is given in the first part of the final report. The developed evaluation routine implemented as a SAS programme has been made available to the expert public and can be found in a separate directory on the database CD.

The derived distributions of human exposure factors are to be seen as an extension and update of the standard exposure factors published in the AUH report. Thus, it is still possible to use these standards in their present (updated) form, e. g. as a basis for point estimates.

In addition to secondary analysis of existing data sets, literature reviews were also carried out to assess the quality of data sources and to fill existing data gaps. Moreover, for some exposure factors only data from literature are available. This applies in particular to inhalation, dermal uptake as well as soil and dust ingestion. A
detailed discussion of these topics is given in the second part of the final report. Here, recommendations for reference values in Germany are derived which, however, need to be agreed by authorized bodies such as for example the German States’ Working Group on Environment and Health (LAUG) so that they can be established as standards.

The following evaluations were made:

Anthropometry

With regard to human anthropometric data, several data sets of different quality are available. For body weight and size as well as body mass index (BMI) recommendations for reference values could be taken from survey data and from literature. Distribution-based reference values could only be derived for adults. The existing data gap with regard to updated and representative data particularly for children and young people could in the near future be filled by evaluating data from the National Children and Youth Health Survey (KiGGS\(^3\)) with the Xprob methodology. In this way, updated survey data which can be recommended without restrictions would also be available for children and adolescents and could be used for distribution-based human exposure assessments.

Time budget

Various data sets from the 1990s were considered and evaluated with regard to the “time budget and activity distribution” aspect. This includes variables describing times and places of a person’s stay as well as his/her patterns of activity in quantitative terms. The data situation has considerably improved since the publication of the AUH report. Data on a person’s activity patterns and whereabouts are, however, not always adequately retrieved in German surveys for exposure modelling purposes.

Representative studies on the time budget are available for the age groups 5 to 17 years as well as 23 to 79 years. Information on single variables such as for example sleeping times for children is however not available for all age groups. For children under 5 years of age, no such data are available for Germany.

All time budget data sets could be stratified in accordance with the Xprob methodology, i.e. gender-specific evaluations were carried out and the age classifications in the form of 25 – 34; 35 – 44; 45 – 54; 55 – 64; 65+ could in most cases be maintained. For children, only one data source was available on this topic. It could also be evaluated with regard to age and sex.

The data situation for activity patterns continues to be unsatisfactory. The 1998 Federal Health Survey provides information on different activity degrees, but these do not allow to distinguish between indoor and outdoor activities. For want of adequately updated data, the data set compiled by AUH and ICRP\(^4\) is, with some restrictions, still recommended for human exposure assessments.

Another possibility of closing this gap may be provided by the time budget survey of the Federal Statistical Office. For the purpose of human exposure assessments the survey could be recoded and reanalysed. With its methodology, the collection of data through diary entries, this survey provides a good approach for future surveys such as for example a survey on a person’s activities and whereabouts. The recoding and reanalysis if this survey was not carried out in the Xprob project because it would have gone beyond the scope of the

\(^3\) Kinder- und Jugendgesundheitsurvey

\(^4\) International Commission on Radiological Protection
project. KiGGS should moreover soon provide time budget information for the age groups of children and teenagers.

Further methodological research is needed with regard to correlations between variables and mixed distributions. One approach to finding a solution consists in the stratification of the particular data sets by further variables such as for example employment situation or real working time. This variable was collected under the 1998 Federal Health Survey, but due to lacking capacities in the Xprob project, a stratified evaluation of these activity patterns was not carried out.

Nutrition

Data sets on food consumption are collected for different reference periods (consumption per day, week and month) which can all be relevant for specific problems. Since major studies have presently not been completed yet or are not available as Public Use Files for secondary analysis, the derived distribution-based reference values are still based on the data material of the National Food Consumption Survey of 1985-89. For a future extension of the reference values it is recommended to document all studies mentioned in the present report in the form of distributions and as Public Use Files. For the age categories 14 years and older, the second National Food Consumption Survey will replace its predecessor. A corresponding analysis of the upcoming data should be done for different reference intervals and food categories in order to provide as many different scenarios as possible with specific and current data. If an exposure model considers several food consumption variables and if considerable correlations have to be expected among them, the distribution-based exposure model should be calculated as an empirical simulation based on individual data of the survey. Therefore, the survey data should be made available as Public Use Files providing its complete empirical data set.

The data collected in Germany on the consumption of self-grown fruits and vegetables are restricted to garden plot holders. For a secondary analysis under the Xprob project, only regional data from a North Rhine-Westphalian study could be used. This data exclusively provide information for adults living in the Ruhr area.

Although domestic tap water constitutes an important potential source of exposure, most food consumption surveys fail to distinguish between tap water and bottled water. For the consumption of tap water, data are only available from the DONALD survey for infants (0 – 36 months) and for adults (18 years and older) from the 1998 German Environmental Survey. For the latter, statistical distributions could be derived for different age categories. However, the response categories only provide for a very rough quantification of intake volumes which particularly for the upper consumption levels for coffee/tea are not represented by measured data. Representative data for children and adolescents are also being collected by the present KiGGS study.

Oral soil and dust intake

The discussion and derivation of suitable reference values for oral soil and dust intake are exclusively based on sources from literature.

Since publication of the AUH report, two new empirical studies on soil ingestion were carried out. The US-American study carried out in 1996, however, reveals considerable problems concerning the conduct and the evaluation of the analyses and is therefore not suited for deriving recommendations as exposure standards. In a recent study from Germany, sample sizes for the individual age groups are so small that they also cannot be used for deriving recommendations as reference values. By looking at the total number of all usable studies
and reanalyses, reliable ranges of mean values, medians and upper percentiles can be established for soil ingestion in children. A detailed stratification in the children’s age group is not possible due to the small sample size in the individual studies. Recommendations for reference values on oral intake levels of house dust cannot be derived because of lacking data.

For usable estimates of soil ingestion, additional empirical studies involving larger sample sizes are required in Germany.

Inhalation

New representative data sources which could be used for deriving reference values for inhalation rates could not be identified for Germany. Therefore, data from literature have to be used. For inhalation rates only preliminary recommendations can be made since the presently available data pool has to be regarded as unsatisfactory.

For the distribution-based human exposure assessment US-American data can be recommended with restrictions due to the non-European origin of these data. For point estimates in the first tier of the tiered approach, data from the AUH or ICRP can be used in the future. Due to a clear dependency of the breathing rates to the basic metabolic rates at least a body weight related approach for exposure assessment is recommended. From the probabilistic point of view an empirical study with a large sample size is recommended. Such a study would allow gender and age-specific analysis incl. derivation of statistical distributions. The activity dependencies should be taken into account. Different activity levels and covariation with age, sex and body weight should also be considered in this context.

Dermal uptake

New recommendations for dermal exposure factors have not been derived: a limited data and literature search didn't result in locating new information.

Use of drinking water for non-nutritional purposes

Data on the use of drinking water for body hygiene (showering, washing or bathing) or other activities such as cleaning, washing of clothes or preparing vegetables can only be found to a very limited extent. The available data and literature, respectively, provide only very rough information on the amount of water used per person and per day or activity. Therefore, no recommendations on the use of drinking water for non-nutritional purposes can presently be formulated.

Dwelling characteristics

For Germany, no data could be identified for dwelling space volumes and air exchange rates. Only with regard to the size of the living area surveyed under the socio-economic panel (SOEP) distributions could be derived. The SOEP contains further data on dwelling characteristics which will in the future still have to be analysed further (e.g. the construction year of the building). The data situation on the length of occupancy, however, is better. Here, age- and gender-specific reference values can be recommended (distribution-based).

A detailed discussion of the data quality can be found in the second part of the report. In the RefXP database developed within the project, the distributions and characteristics of the individual exposure factors are provided, structured by different topics regarding exposure factors. The selected procedure and the provision of results in the form of a uniformly structured database offer clear advantages compared to the isolated and non-uniform presentation used hitherto. In this
database detailed information is given on the data source of the exposure factor, its empirical distribution, on the distribution functions and its parameters with the best fit to the empirical data, in each case stratified by age and gender. Access to the exposure factors is supported by an interactive surface. Alternative data sources are also displayed. The provided data can be imported into modelling software. A user manual for usage of the database is given as part of this report. The database is structured in such a way that other institutions may contribute to the data sets in the future.

The methods of the distribution-based exposure modelling are described in this report and applied in example case studies, together with literature reviews, data searches and analyses as well as the development of a methodology for secondary analysis of existing data sets on exposure factors.

Core items of the report are “Procedures of exposure modelling in a real population (incl. its variations)” and “Description and documentation of uncertainties in modelling and of its results”. In this connection, the good practice concept of distribution-based exposure modelling is intended to serve the quality assurance and transparency in modelling, in handling the empirical input data and calculations as well as the documentation and interpretation of the results of the modelling process. For modelling, a tiered approach is recommended which will lead to more complex models, depending on the specific problem of concern. For evaluating the data basis, criteria have been derived and, if necessary, adjusted under the Xprob project.

Two scenarios serving as example case studies and one additional case study were used to review and further develop the methodology under the Xprob project. In this context, possible limitations of the methodology and further need for research were revealed. All examples have been taken from important human exposure assessment application areas.

The scenario "Living on a soil contaminated site" was focused on pointing out the possibilities of overcoming data gaps and adjusting derived human exposure standards for lifelong exposure (0 – 75 years) respectively. From this scenario, both substantial improvements could be developed for the data retrieval in the database as well as detailed instructions for dealing with age stratifications and data gaps. These instructions are also included in the report.

The scenario "Prognosis of tetrachloroethylene exposure of the German population" was used to find out whether the population’s exposure to tetrachloroethylene (PER) as an air-borne pollutant can be estimated using the newly derived data. In a tiered approach, partial models were developed for the input variables "indoor and outdoor air pollution" as well as for "uptake and distribution in the body" to estimate the expected exposure and body burden (PER concentration in the blood), respectively. In this context, it was explored if historically available empirical data on body burden of PER can be estimated by exposure modelling sufficiently precise with regard to its amount and distribution in the population. When addressing this problem, considerable data gaps were revealed particularly concerning the distribution of human activities, the intensity of these activities in relation to breath rates as well as the characterisation of the corresponding places where persons spent their time. This
application test clearly showed that factual dependencies between model variables and existing correlations have to be included in the modelling procedure. In case of high variability of the input data, the prognostic quality of the exposure model with respect to the central tendency, variation and upper percentile of the body burden is regarded to be good. The modelling estimates correspond to results from population studies (cross-sectional comparison) as well as to measured HBM values (comparison of individual data) in a specific study for evaluating remediation measures. Existing uncertainties with respect to selected variables, available data, models, calculations and outcome variables are described in the report.

In an additional case study, an existing data set was re-analysed to evaluate the distribution-based methodology. The data were taken from a human biomonitoring study on the mercury and arsenic burden of people living in a former miners region in the northern Palatinate area and an unaffected control group. Within the case study, a quantitative exposure model was defined both for mercury and arsenic, with due consideration of the relevant exposure routes. The required exposure factors were taken from the RefXP database or from literature. The mercury model allows making comparatively good predictions of mercury concentrations excreted via the urine. However, the exposure factors recognised as being most relevant in the mercury exposure model are "amalgam fillings" and "profession". No information is given in the RefXP database on these factors. In contrast, exposure factors from the RefXP database (e.g. drinking water and consumption of fish) were used in the arsenic exposure model. Here, however, only a relatively low proportion of the variance of arsenic concentration measured in the urine can be explained by the model.

Even in future it will be difficult to concretely distinguish between variation and uncertainty in exposure modelling. The transparent description of uncertainties when using (high-end) point estimate exposure modelling techniques helps to justify the necessity of further modelling and to explain differences in the quality of various assessments.

The interim and final results of the project each were presented to the expert public and discussed at two workshops with international participation. Both workshops are documented in specific volumes already published. The ideas and suggestions of the workshop discussions are included in this final report. International expert opinion was included in the project. On this occasion it became clear that the results of the German Xprob project are also relevant to EU-wide efforts on the harmonisation of exposure assessment and exposure-related databases.

The exposure reference values recommended by Xprob and the methods of probabilistic exposure modelling can be applied to many areas in which exposure assessment is carried out. Apart from application in the field of event-related risk assessment, applications are in particular possible in the field of health impact assessment (HIA) as well as in national and EU chemical policies.

For health impact assessments distribution-based models can be used for the prognosis of future exposures as well as for describing variabilities of exposure in the population and uncertainties of the results obtained. Additional benefits are seen in
the systematisation of the exposure assessment procedures, for the typology of information, and for analyses of subpopulations.

Within the European chemicals legislation and its modifications by the REACH process (Registration, Evaluation and Authorisation of Chemicals), present considerations concerning the details of the requirements for exposure assessment call for a tiered approach for developing exposure scenarios and for estimating exposure levels. Distribution-based exposure modelling is presently mainly recommended for the highest tier. Here, the Xprob results might contribute to the detailed technical implementation of the exposure assessment. Parts of the Xprob report may be adapted as REACH-specific requirements.

For several areas of the exposure factors, various data gaps were identified for Germany. For anthropometric exposure factors recommendations can be given on a person’s body weight, body height and BMI, whereas no or only limited data are available on a person’s body fat percentage, body surface or blood volume. On food consumption, various nationwide surveys carried out with different methods are available or currently being conducted which can all be used for future recommendations. However, not all results are presently available as Public Use Files so that the Xprob methodology would have to be used by the data holders themselves. Additional data are missing for specific population groups (e.g. vegetarians, self-supporters, pregnant women), for specific oral intake rates (e.g. soil and dust, drinking water consumption) or are missing in the form of longitudinal studies on lifelong nutrition habits. For the time budget, only recommendations on single variables can presently be made. Here, a survey is missing in which data on a person’s activities and the places where he/she is staying are jointly collected, because of the existing dependencies between time budget and activity patterns. The same applies to dwelling characteristics for which in particular data on room air volumes and airing habits are missing, providing information about typical air exchange rates.

The newly derived exposure factors underline the high relevance of national surveys and how important it is to repeat surveys on a regular basis for follow-up, i.e. changed food consumption habits or modifications of other exposure factors.

Due to capacity shortages in the project not all ways of analysing the existing data sources could be carried out. But the way of how to carry out these analyses, i.e. the methodology and the software program, has been shown. These might be used for generation of new and more up-to-date proposals for exposure factors. The development of an extension of the methodology of categorial characteristics, however, for adjusting frequency distributions is recommended.

An extension of the Xprob methodology to include further areas such as occupational health and consumer protection is also possible.

Methodological research issues concern the optimal apportionment of variation and uncertainty in practical exposure modelling. Deriving the distribution of correlated exposure factors at present requires simulation with the help of empirical data. Here, straightforward concepts for considering correlations in simulation calculations are
missing. These concepts are particularly necessary when the modelling of correlations is not supported by empirical data but by mere assumptions. The same problem applies to the distinction between subpopulations. Up to now, stratifications have only been made with regard to gender and age. Further stratifications, however, are conceivable. Consistent procedures are missing for cases where the stratification variable is not known and the stratification can only be carried out with the help of empirical data (so-called mixed distributions).

With regard to modelling, the development of a tool kit containing standardised elements for the respective scenario seems to be reasonable. In this way, the concept of exposure standards would also be transferred to the modelling level. Finally, also in the field of model evaluation further standards of comparison and quality factors are required which can help to assess the quality of modelling.

The alleged unambiguousness of point estimates is at first sight suspended under the distribution-based approach. While for the point estimate specific values are agreed on beforehand and thus implicitly a specific protection level is aimed at, the distribution-based analysis requires to explicitly determine a desired protection level.

Additional application scenarios should contribute to further demonstrate the methodology of probabilistic modelling. Moreover, one strand of development should address the comparison of measured and modelled exposures to show the possibilities and limitations of these two complementary approaches.

The report contains fundamental guidance on how to deal with probabilistic modelling. The provision of a limited number of regularly occurring application situations as "standard scenarios" could in general contribute to an agreement on exposure modelling and in particular on distribution-based exposure modelling.

The distribution-based exposure assessment contributes to transparency and is state-of-the-art in science today. Dealing with empirical data as well as the characterisation of heterogeneity in the population and uncertainty about the results obtained are challenging issues for risk communication. The repertoire of regulative exposure assessment, even if not used on a daily basis, includes the following: the methodology of using existing data and assessing new data respectively, quality assurance of the results obtained, and the use of modern techniques. The increasing transparency of distribution-based risk assessment, however, compensates for the disadvantage of higher complexity. Simple answers can be provided much quicker and at lower costs but often they require costly adjustments afterwards.

In addition to the general exposure factors which are provided in the database, more specific problems might require exposure information which is available in data sources but calls for more specific analyses. Instructions on how to deal with these cases are still to be developed. In view of the data gaps existing for more specific problems, instruments should be developed with which information on the distribution of individual exposure factors in Germany could be collected with limited resources and on a timely basis. In particular, it has to be verified if existing, regular surveys (e. g. marketing panels) can also be used for such problems.
Executive Summary

Under the Xprob project, proposals for exposure reference values have been developed which should be seen as a contribution to updating the AUH report. These proposals are intended to be legitimised by the proper organisations. By this, they will be a basis on which standardised and harmonised exposure assessment might be conducted. The AUH values will maintain their value and validity until then. For this reason, nearly all standards of the AUH report have been included in the RefXP database.

All in all, distribution-based exposure standards can contribute to a harmonisation of the assessment of environmental health risks in Germany. Their implementation into administrative policy, however, calls for additional support and practice-oriented developments. In this context, the LAUG intends to continue to support the procedure of defining reference values for exposure factors throughout Germany.

The methodology developed under the Xprob project for deriving distributions for exposure factors can serve as a basis for new reference values for exposure factors in Germany. The "good practice" principles described in the report will facilitate the use of the methodology in concrete exposure assessments. Thanks to the tiered approach, the new methodology is consistent with the present procedure; it provides additional support in those cases where a more detailed analysis of the exposure is required.

For filling existing data gaps, the methodology developed should be used for evaluating more recent studies which are up to now not yet freely available, and in particular for future studies. In general the attempt should be made to provide all national surveys to the expert public as Public Use Files.

Distribution-based exposure modelling is suited for identifying important influence factors and for prognoses made in the context of prospective health impact assessments. It is thus a reasonable complementation to the human biomonitoring approach.

Population-related exposure models which are based on distribution-based exposure factors can be used for the purposes of health monitoring, of impact assessments and for approval procedures in chemicals legislation (REACH).

The Xprob project has served to establish competence in Germany which will assert itself at the European level. The JRC of the EU is presently interested in cooperation with the Xprob project to support in particular the field of exposure factors with updated and valid data. Further developments should be aimed at using this competence in a profitable way at the national level (e.g. "reactive" exposure assessments and prospective impact assessments) but also at the international level such as for example under the REACH procedure.