Methodological Issues in Using Epidemiological Studies for Health Risk Analysis

M. Nurminen\textsuperscript{a} & C. Corvalán\textsuperscript{b}

1. HEALTH RISK ANALYSIS AS A PART OF THE RISK ASSESSMENT PROCESS

Today, governing bodies at the local, national and international level must increasingly make difficult decisions (for example, in the fields of energy, industrial manufacturing, traffic, food and waste management) which would ideally be based on weighing the health and environmental cost of a technology against its economical and social benefits (McMichael, 1989). This requires that health effects can be quantified, yet data for this quantification are often limited. Quantitative risk assessment (QRA) must nevertheless be carried out for regulatory purposes. Because the result is often presented as a single number (for example, excess number of exposed disease cases), it may appear to be a scientific certainty, and therefore has an obvious appeal among regulators and decision makers. Despite its apparent objectivity, QRA assessment is dependent on a series of assumptions and subjective choices which can have critical effects on the resulting risk estimates. Therefore, it is of prime importance that QRA be built on solid scientific bases.

\textsuperscript{a} Finnish Institute of Occupational Health; Helsinki, Finland

\textsuperscript{b} Office of Global & Integrated Environmental Health, World Health Organization, Geneva
Health risks in human populations are increasingly being assessed by the use of empirical data from epidemiological studies. The basic reason for this development is the fact that the evidence from animal experiments is of uncertain relevance for human populations. When animal data are used to estimate risks of human exposure, two major extrapolations are required: first, interspecies dose extrapolations to adjust for differences between humans and laboratory animals, possibly affecting the response to the toxic agent; second, the statistical bioassay models used in risk estimation to extrapolate from the high doses used in animal experiments to the much lower doses to which humans are likely to be environmentally exposed rely on assumptions that can often be criticised as being untenable (Hartley and Sielken, 1977). Epidemiological studies, coupled with retrospective exposure assessments, can yield more defensible estimates of likely human health risks than those obtained from statistical models based on animal studies. Recent work evaluating benzene furnishes an example (Paustenbach et al., 1992). A vexing problem that often renders human volunteer laboratory experiments infeasible is that the complex mixtures of exposures observed in the environment are not amenable to realistic dose-response assessment. On the other hand, with non-experimental epidemiological methods it is fairly easy to demonstrate that occupational groups sustain excess risks at high exposure levels. However, detecting health risks in relation to environmental exposures is far more difficult. The research situation is also problematic in that the model assumptions cannot be tested because the disease occurrence is often rare in the low-exposure range. Thus it is not surprising that different models may result in considerably disparate risk estimates when applied to the same data.

When the actual health effects are subtle, extremely large populations need to be studied to establish an exposure-disease relation in a follow-up investigation. However, the cost of cohort studies can be prohibitive. Using the ecological approach permits the study of very large populations at a decidedly reduced cost, and the advances of epidemiological study designs have brought about ways of improving the reliability of such research. It may be feasible to supplement aggregated data with sampled data at the individual level in a hybrid epidemiological analysis (Prentice and Sheppard, 1995). For example, aggregate data studies combining specific disease rates are often routinely available, providing comprehensive exposure and covariate data on random subsamples of modest size from populations in a multigroup ecological study, and can thus yield more cost-effective and reliable results than the data from studies of entire cohorts.

Health risk assessment is based on the premise that exposure causes disease. The process of accurate exposure/dose assessment is not straightforward; yet it is particularly important when animal data are used to predict human risks because of the need to extrapolate between species, and because of the relatively high doses administered to animals compared to the relatively low levels in
human exposures. Once a relation — often nonlinear — between exposure to an agent (e.g., pollution, toxic chemicals, hazardous wastes) and human disease risk has been found, an appropriate reaction is needed from society. QRA examines all relevant information: toxicology data collected from animal experiments, health data gathered in epidemiological studies and disease registers, knowledge about the mode and mechanism of exposure action, and the estimates of the sizes of exposed human populations at each anticipated dose level, and the level of exposure.

Focusing on epidemiological studies, the different phases of a risk assessment process vis-à-vis the types of epidemiological strategy may be arranged as in Figure 4.1 (cf. Partanen and Rantanen, 1991).

![Figure 4.1](image)

*Figure 4.1*. The phases of a risk assessment process and the associated epidemiological strategies.
Assessment of health risks is the establishment of a statistical relation between exposure to an identified hazard and the disease risk it induces. In cancer risk assessment, for example, techniques of ‘molecular’ epidemiology (Schulte and Perera, 1993), such as biological markers (Hulka and Margolin, 1992) and biologically based pharmacokinetic models (Edler and Rausch, 1994), have been used to estimate doses in target organs and cells as well as to elucidate mechanisms of exposure-disease relations observed in classical epidemiological studies. These studies that bridge the gap between laboratory experimentation and population-based epidemiology have also been called “transitional” epidemiological studies (Hulka, 1991). Two main types of quantitative estimates can be presented, together with ranges reflecting their uncertainty:

- first, the risk corresponding to a given level of exposure/dose, for example, expressed in terms of excess risk or the number of extra disease cases
- second, the exposure/dose corresponding to a given level of risk, such as the exposures estimated to cause adverse health outcomes in a certain percentage of exposed subjects.

The first type of estimate is obtained in a risk analysis, whereas the second type is obtained in what has sometimes been termed hazard analysis. The risk characterization part of QRA then summarizes and interprets the information collected from the estimation of exposure and health effects as well as discusses the limitations of the methodology and identifies the uncertainties in the risk estimates.

Miettinen (1985) has remarked that ‘descriptive relations bear on such passive matters as prognosis setting and risk assessment, whereas knowledge of causal relations is the basis for interventions, that is, for wilful alterations of the outcome through perturbations of the determinant.’ Aggregated measurements of disease occurrence in population groups are usually associated with descriptive epidemiology, whereas measurements made at the level of the individual subject are associated with analytical epidemiology. However, to make a stark distinction between the two branches of the discipline would not be sensible as it would neglect the role of the aggregate-level (‘ecological’) evidence in quantifying the relationship between environment and disease. QRA aims to provide an answer to the question: ‘What would the health gains of preventive action be?’ Clearly, informed decision-making can benefit from analyses that estimate the level of exposure and the magnitude of the health effect. However, the upshot of a completed risk assessment process is not a risk prediction in a hypothetical situation, but an evaluation of the efficacy of an actual risk prevention or reduction programme by using methods of intervention epidemiology. The final phase is risk management by a regulatory control of the hazard or by resorting to other preventive measures.
This paper reviews methodological issues and problems pertinent to the application of epidemiology in health risk analysis, as well as discusses concerns in the presentation and interpretation of results from such an activity. The estimates of health impact obtained from risk analysis are illustrated by studies in the field of air pollution epidemiology.

2. METHODOLOGIES FOR ANALYSIS OF ENVIRONMENTAL AND HEALTH RISKS

The methods developed for the Health and Environment Analysis for Decision-making (HEADLAMP) project (WHO, 1995) are generally based on *routinely* collected data. The importance of taking an integrated view of both environmental exposure and health effects in risk management was recognised as a key characteristic of the HEADLAMP process. In these circumstances, the use of either environmental data alone or health data alone may be the only means of identifying environmental hazards to the health of a population. Ideally, corroboration of either the environmental data or the health data is required to establish a relationship. This involves cross-checking both environmental and health data: checking health evidence of potential problems associated with environmental exposure data available, or conversely establishing whether the potential health risk implied by the existence of a hazard is substantiated by excess morbidity or excess mortality.

*Example 1.* A number of epidemiological *cancer surveillance* systems have been established in many parts of the world (mainly in developed countries) with the avowed aim of detecting associations between cancers and occupational exposures (e.g., Dubrow et al., 1987; Siemiatycki et al., 1981) which may lead to the generation of hypotheses about previously unknown carcinogens; for example, one such programme played a role in suggesting a linkage between electromagnetic radiation and leukaemia (Milham, 1985). However, since the extent or distribution of occupation-related cancer in many countries is largely unknown, efforts to identify previously undiscovered carcinogens is a lesser priority than the quantification of the effects of substances already known or suspected to be carcinogenic. This is particularly the case of developing countries but also in developed countries (Australia for example).

*Example 2.* If the aim of epidemiological *hazard surveillance* is to identify and control exposures to carcinogenic substances in the workplace, then it would appear intuitively that the most expeditious approach would be to go directly to industry and identify the exposures. In practice, however, it is not easy to identify all exposures. This view appears to be supported by experience in Finland, where a register of employees occupationally exposed to carcinogenic substances and processes was established (in 1979) with the aim of identifying all workplaces where there is potential exposure to known human carcinogens. It is extremely unlikely, that all uses of carcinogens in the Finnish industry are being identified with this hazard surveillance system. The Finnish Institute of
Occupational Health (FIOH) has shown that: (i) small industries in particular are underreporting the use of registrable substances; (ii) fewer exposures and fewer exposed workers are registered than expected from prior estimates; and (iii) no exposure data exist on 50 of 138 registrable substances, (J. Rantanen and P. Heikkinen, Carcinogens in the work environment. Address to Symposium on Experimental and Epidemiologic Applications to Risk Assessments of Complex Mixtures. Hanasaari, Espoo, Finland, May 1989).

However, the register does not include workers who are exposed to substances that are probable human carcinogens according to the classification of the International Agency for Research on Cancer; in 1993, for example, 80,000 Finnish workers who were exposed to quartz dust, which then was a probable human carcinogen, were not registered (FIOH. 1995). Even if it were possible to identify every use of every known or suspected carcinogen, this in itself would present an impossible logistical problem of trying to ensure that the risks were adequately controlled in all cases. Furthermore, much of the effort would be unproductive, since many of the uses of carcinogenic substances would involve insignificant risk, because of minimal exposure or infrequent use.

Technically, it is possible to link the data from the FIOH ‘register of employees occupationally exposed to carcinogens’ on an individual basis with the Finnish Cancer Registry data. However, cancer risk assessments that would exploit this linkage would be handicapped by the lack of data on relevant covariates; for example, the registers contain no information on personal smoking habits, a major risk factor in many cancers, because of confidentiality or other restrictions on the availability of individual data. Therefore, it is unlikely that the FIOH register data can be used, for example, for the estimation of the etiologic fraction of particular occupational exposures in the causation of cancers. Nevertheless, there are indications that the mere establishment of this hazard surveillance system may have reduced the use of carcinogenic materials by way of alerting industries to use substitute products: in 1993, 15,000 new notifications were made to the FIOH register; this figure is 5% less than in the previous year. However, cautious interpretation of these statistics is in order, considering that the economic recession deepened sharply in Finland in 1993, resulting in a reduced use of raw materials (incl. carcinogenic substances), in increased unemployment rates, and in fewer vacant jobs, especially in the construction industry.

Essentially, routine data on environmental conditions are most often available at the aggregate level. These tell one about the risks potentially faced by a group, but tacitly assume that all members of the group experience the risk equally (and thus potentially experience the health consequence equally). Conversely, routine data on morbidity rates or mortality rates alone give only average risks expected by individuals, without data on the materialised risks (disease events experienced by individuals). Both types of data should be used with caution on their own, since ideally a linkage should be established in order to make accurate judgements about potential policy solutions.
Assessment of the health risks associated with environmental exposures involves three methodological steps:

1. literature review for finding or building a risk model for exposure-effect relations
2. quantification of exposures, estimation of the sizes of the exposed populations, and attachment of exposure levels to subpopulations
3. prediction of the impact of a change in exposure levels on health effects.

Example 3 Ostro (1996) has recently prepared a report in which he presented a methodology for and estimates of the health effects of particulate matter in Santiago, Chile. using detailed air quality information and considering a broad range of potential health outcomes. Basically, he followed the above three-step strategy in the risk analysis process. The development of dose-response functions first involves the selection of scientific studies that are thought to provide the best estimates of an effect. Ostro gave several applied criteria for an epidemiological study to be included as a basis for such functions: a qualified study should:

- have a proper design and methodology
- recognize and attempt to minimize confounding, and omitted covariates
- control for the effects of seasonality and weather
- analyze reasonably completely the air pollution and health data
- provide for an airborne particulate measure that could be reasonably converted into particulate matter less than 10 microns in diameter (PM10)
- involve relevant levels of air pollution
- address clinical outcomes or identifiable changes in behaviour.

The next step in the quantification involved determining the relevant population. Based on air pollution dispersion models, metropolitan Santiago was first divided into a grid of 289 areas. A set of isopleths — connected areas having nearly equal annual average concentrations — were then graphed. Finally, the populations were assigned exposure values (pollution concentrations) based mainly on their area of residence. The third step in the risk analysis process involved the choice of a 'target' concentration, guided by some current annual average standard for PM10.

3. ISSUES AND PROBLEMS OF HEALTH RISK ANALYSIS

Risk analysis is not a true linkage method in the sense that local health data are not utilised, except perhaps in the form of the established exposure-health relation as a basis for risk predictions. Thus, the usefulness of risk analyses lies in the fact that these can be performed in areas where insufficient health outcome data are available. Risk analysis methods are the least resource-intensive, quick to execute, and allow rapid risk predictions. The success of the
risk assessment process, however, depends on a number of issues such as the choice of the risk models, the development of prediction formulas, and the adequacy of exposure assessment. There is generally a great deal of uncertainty and a number of shortcomings associated with all of these issues. A discussion of these limitations is necessarily dependent on the objective of the research and the practical setting in which the study is conducted. While it would be next to impossible to discuss in generic terms the problems that practising risk analysts encounter in such diverse areas as cancer epidemiology or in pollution epidemiology, some cautionary remarks are given here.

3.1 Limitations of Exposure and Health Data

A crucial constituent in any attempt to estimate the quantitative health effect of a given level of exposure is the reliability of the method assessing the level of exposure in the studied populations. However, direct measurements are rarely available, and often some semiquantitative method for estimating exposures from limited information must be employed. Besides magnitude, valid representation of the actual exposure has to account for the composition and duration of exposure as well as for the time since the start of first exposure. The important practical question then is: What preventive measures should be taken to ensure valid estimates?

Example 4. Exposure to crystalline silica (silicon dioxide) dust can cause occupational disease unless it is appropriately controlled. In the past, Australian control strategies have been designed to prevent the occurrence of silicosis. At workplaces in which existing standards have been enforced by the inspectorate, and modern control measures have been rigorously applied, there appears to be little evidence of adverse health outcomes, although the data may be incomplete. The various Australian standards, ranging from 0.15 to 0.2 mg/m³ depending on the state, are historically based on the occurrence of silicosis in sandstone workers in Sydney. But, Worksafe Australia (1993a) has recently promulgated a report on crystalline silica in which lung cancer is considered as one of the health effects that should be taken into account in determining exposure standards.

Australian occupational environmental data on the extent of silica exposure and health effects are limited. While some industries, such as mining, have good exposure monitoring records and compensation registers on silicosis, little information is available on industries such as manufacturing and construction. Therefore an approach to national risk assessment was needed to supplement existing records of exposure monitoring and data on health effects.

Health outcome data was derived from compensation systems and it was indicative of past exposure to silica. The largest numbers of workers compensated by the New South Wales Dust Diseases Board came from the manufacturing, construction, and mining industries. Other health outcomes such as lung cancer or chronic obstructive airway disease have not been assessed for
work involving exposure to silica. However, specific studies on lung cancer and
airway disease in Western Australia revealed that these diseases are more
prevalent among silica-exposed workers than among the general population, but
it was unclear to what extent their increased frequency was due to silica
exposure, cigarette smoking, or the joint operation of the two. Berry (1996) has
recently reviewed the evidence that there may be an increase in the incidence of
lung cancer in those exposed to silica. He concluded that "Although it has not
been demonstrated that silicosis is a necessary precursor of any lung cancer that
may be due to silica exposure, there is evidence that any increase in lung cancer
risk will be greater in those with silicosis than in those without. Thus
control measures effective in reducing silicosis will also be effective in reducing
any excess in lung cancer."

For the purposes of a QRA the members of the National Institute of
Occupational Health and Safety (NIOHS), Australia, Expert Working Group on
Crystalline Silica (Worksafe Australia, 1993a) used their expertise on the
hygienic conditions and processes at the work sites and supplementary
information from companies, where available, to evaluate the median level of
exposure in the occupational subpopulations and the number of workers actually
exposed to silica. Yet, owing to their crude nature, these data have to be
examined with reservation. Also, because the use of expert judgement is often
essentially subjective, it is important that the decisions reached and criteria used
are well documented, as was the case in this QRA (Worksafe Australia, 1993a).

### 3.2 Applicability of Exposure and Dose Estimates for
Risk Modelling

A big concern in QRA is that the results from epidemiological studies derived in
exposed populations cannot always be applied directly to another population
subjected to different exposure conditions. In air pollution epidemiology, for
example, it is likely that in applying results from developed countries to cities in
developing countries, one underestimates the true health effect of pollution,
because the baseline health status is likely to be poorer in these regions (Ostro,
1994). Moreover, the different methods used to estimate the exposure/dose-
related risk may need to be adapted to address the question at hand. Apart from
the statistical and other methodological uncertainties in the conversion, the
extent to which the exposure measurements were representative of the working
environment is essential.

*Example 4 (cont'd).* Using empirical models for risk, based on published
Canadian epidemiological experience on hard-rock miners (Muir et al., 1989),
Nurminen et al. (1992) predicted the occurrence of silicosis in the Australian
labour force currently exposed to crystalline silica dust. In the exponential model
for risk of silicosis (see Section 3.5), the investigators used a cumulative index of
exposure, which was a product of the median level and average duration of
individual worker experiences in the particular industry or occupation or age
categories. This approach, which was also used in the Canadian (Muir et al.,

- 107 -
1989) and Chinese (Pang et al., 1997) studies, was followed because, in general, it is the implicitly adopted strategy for dust control measures. However, the use of such an index implies, for example, that the increase in the relative risk of silicosis caused by exposure to, say, a silica dust level of 0.1 mg/m³ for 40 years is the same as that caused by exposure to a level of 0.4 mg/m³ for 10 years. Nurminen et al. (1992) also used South African risk estimates for lung cancer (Hnizdo and Sluis-Cremer, 1991) to estimate the excess risk of lung cancer possibly due to silica exposure in the Australian workforce.

Using revised assessments of South African silica exposure data and a new South African epidemiological study of dose-related silica risk (Hnizdo and Sluis-Cremer, 1993), Legh et al (1997) revised their original estimates of quantitative absolute dose-related risk and lung cancer. The original study required a conversion factor between historical South African exposure measures (shift-weighted respirable silica surface area (RSA) particle years) and the measure used in Australia (full shift respirable silica gravimetric concentration): 1 mg/m³ respirable quartz-years was taken to be equivalent to 3,385 RSA area particle-years. It would now appear that the factor was in error by about 2-to-3-fold, in such a direction that 1 mg/m³ = 8,000 RSA, due to the effect of new data (on shift length weighting, silica content of dust, and calculation of mass from particle geometry and density). With this correction, the risk estimates increased. At 0.1 mg/m³ work-life excess risk for silica-related lung cancer increases from 0.3% to about 0.8% resulting in 26 extra cases per year for Australia rather than 10. The recent South African cohort study on dosage assessed the silicosis risk in relation to dose to be up to 30 times greater than did the Canadian study (Muir et al., 1989) originally used by Nurminen et al. (1992). When the South African study (Hnizdo and Sluis-Cremer, 1993) was used, the prediction was that at 1 mg/m³ 106 new silicosis cases, instead of 11, would be diagnosed annually. Although the South African study could also have underestimated the gravimetric dose conversion factor by up to 3 times, this would not explain the 10-fold risk difference between the two studies.

### 3.3 Confounding and Multicollinearity Posed by Multiple and Mixed Exposures

In risk analysis, one should be careful when assessing different exposures separately. In situations where two different exposures have the same health endpoints, one should ensure that the estimates are adjusted for the effect of both exposures. In the case of increased mortality resulting from high levels of PM10, for example, increased mortality has also been associated with increasing levels of sulphur dioxide (SO₂). If these two measures are highly correlated, only one variable would remain statistically significant when modelling the contribution of both. However, when modelled separately, they may both show significant associations with mortality, and it would be erroneous to make separate estimates for each. This would result in an exaggerated number of deaths. In this case, only one measure should be used, perhaps the one with the more complete data, even if the results can be expected to be conservative.
However, the selection of variates and the specification of the model form is difficult, and it is the investigator who is responsible for the choices based on his/her research aims and interests.

*Example 4 (cont’d).* In the Australian study the investigators used the epidemiological relation between silicosis in Ontario hard-rock miners and cumulative exposure to silica, which may not be directly applicable to the conditions of Australian industrial work sites. The Canadian study cautioned about possible coexposures when risks due to silica dust in one industry or occupation are compared with those reported in other contexts. For example, some forms of crystalline silica, such as cristobalite and tridymite, may have a greater fibrogenic potential than silica itself. On the other hand, silica mixed with high concentrations of inert coal dust and silicates may diminish the apparent toxicity of the silica fraction. Similarly, high concentrations of clay minerals in silica dust may reduce the risk of silicosis among workers in the brick industry. Another problem is the assumption that the South African data can be used as an estimate of the silica-induced risk of lung cancer when there is possible confounding with the effects of radon daughters.

*Example 5.* Schwartz and Dockery (1992) considered pollution and mortality in Philadelphia during the years 1973-1980. They used the Poisson regression model to study the effect of total suspended particles (TSP) and SO$_2$ on daily mortality counts from all nonexternal causes of death and from cardiovascular and respiratory causes. Both TSP and SO$_2$ were associated with total mortality when entered separately into the model that also included the variates ‘weather’ and ‘season’ as potential confounders. However, in the multivariate model that included both pollution measures, the effect of TSP changed little from the univariate model, but the effect of SO$_2$ was halved.

In a reanalysis of the Philadelphia data, Moolgavkar et al. (1995) found that the pollution measures TSP, SO$_2$, and ozone (O$_3$), were positively associated with total mortality, but they concluded that the correlations between these measures preclude the designation of a single component of the pollution mixture as the cause of mortality. A third analysis of the Philadelphia data for the period 1973-1990 was carried out by Li and Roth (1995). After comparing an array of models, the investigators concluded that the association of pollution variates was highly dependent on model specification, and that no single model could be selected as final.

In a subsequent debate the first two teams of investigators (Dockery and Schwartz, 1995; Moolgavkar et al., 1996; Dockery and Schwartz, 1996; Moolgavkar and Luebeck 1996) exchanged views on how the issue of confounding by copollutants and time trends should be approached. Two alternative methods include restriction of the study base and multivariate modelling. In either case, caution should be exercised in the interpretation of small relative risks, particularly when residual confounding is a real concern.
3.4 Population Heterogeneity as a Source of Uncertainty

The QRA process involves in practice many sources of uncertainty, including the presence of population heterogeneity. In environmental health linkage analysis, risk statistics are usually collected and presented at very highly aggregated levels. If they are averaged over many distinct individual types, the results may have little relevance for any specific individual.

Example 4 (cont’d). Each of the approximately 136,400 workers in the Australian labour force, who were assessed as exposed to silica dust in their work, belonged to exactly one of the 1,000 possible categories in a 50 × 20 industry-by-occupation cross-classification. (Actually, there were 665 subpopulations with a nonzero exposure intensity.) This classification was considered to be specific enough for monitoring the population at risk. In all industries, only an estimated 10% of the workers were exposed to silica at levels above 2.0 mg/m³. Such exposure matrices are generally assumed to have an error structure, in which the average of the true doses for all subjects in an exposure assignment group is equal to the assigned value. Consequently, if the true dose-response is linear, the estimated slope of a linear relationship will not be biased toward the null (Armstrong, 1990; Hatch and Thomas, 1993). Close study of the exposure matrix led to the incontrovertible conclusion that urgent attention should be paid to the working conditions of certain occupations across all industries. For instance, in each industry with drilling plant operators, this occupation was the one with the highest exposures. However, the immediate need for improved hygiene would not be so clear if the population were heterogeneous with respect to exposure.

Strictly speaking, the homogeneity assumption probably does not hold in most practical situations since unrecognised risk factors presumably subject different people to different background disease risk. As pointed out by Robins and Greenland (1991), the case that the heterogeneity of background disease risk of exposed populations is almost always quite severe. This is because there are likely to be unmeasured genetic and environmental factors that vary across individuals and thus strongly affect individual risk. Also some persons are more ‘susceptible’ to an underlying risk factor (or set of factors) than others, and sufficiently so to make them contract a disease after being exposed. Under heterogeneity, one can still estimate the cumulative incidence rate by the number of cases over the population size, but must recognise that these estimates represent average risks and rates in the population. However, usual variance estimators may be biased upward (away from the null) when risks are heterogeneous and not low. The implication is that the practitioner of risk analysis should always check for heterogeneity before presenting aggregate population statistics. If heterogeneity is discovered, then aggregate population risk estimates may be misleading. The populations should either be split into more homogeneous subpopulations, or the statistics be presented with due caution for interpretation.
3.5 Risk Modelling and Building of Prediction Formulas

Two quantities of importance to public health can be estimated in the course of risk analysis, namely individual risk (i.e. the probability that an individual will develop a disease as a result of exposure in a specified time period) and the population risk (i.e. the expected number of cases of disease attributable to exposure in the population under study in a specified time period). These two measures may have different regulatory implications: the regulatory authorities may wish to evaluate either the risk to individuals who are exceptionally highly exposed, or the risk to a large population whose average exposure could be much lower.

The risk of developing a disease due to environmental or occupational exposure to agents at various levels can be assessed by means of a statistical model for exposure-effect relations. However, the concept of risk being a probability measure pertains to an individual. In epidemiology, risk can be estimated as a cumulative incidence rate in a population. Risk functions describe the change in risk as a function of a change in the exposure index. A cumulative exposure index is formed as the product of the intensity and duration of exposure. These entities can then be translated to the predicted number of disease events caused by the exposure in question.

Example 4 (cont'd). For silicosis, the risk was quantified by a model relating cumulative respirable silica exposure (particle-years) to the cumulative incidence rate. The investigators assumed an exposure-effect relation in which the effect was proportional to the power of the exposure:

$$ R_\alpha = \alpha (L \times D_\alpha)^\beta $$

where $R_\alpha$ stood for the risk in the $\alpha$th age category, $L$ was the dust level or intensity of exposure (mg/m³), $D_\alpha$ was the average attained duration of exposure (in years) for the subjects in the $\alpha$th age category, and $\alpha = 0.00109$ and $\beta = 1.72$ were the model form parameters that were estimated from the experience of hard-rock miners in Ontario (Muir et al., 1989). As an indication of the variability of the risk estimates, the investigators chose the range of values for the parameters $\alpha$ and $\beta$ indicated by the results of fitting separate models for different diagnostic definitions of silicosis (based on reader consensus of the radiological category).

The preceding risk model assumes, first, that silica dust is a necessary cause of silicosis. In other words, there is no risk ($R = 0$) if there is no cumulative exposure ($L = 0$ or $D = 0$). Second, it is the accumulation of silica over the years, that is, the product of level and duration ($L \times D$) that determines the risk and not the intensity of exposure in itself. However, several additional demographic assumptions were necessary to predict realisations of risks in terms of the numbers of people sustained.
Omitting details, the expected number of silicosis cases (E) was computed as

\[ E = S \times T \times I \]

where \( S \) stood for the size of the industrial subpopulation, \( T \) was the follow-up time, and \( I \) was the incidence rate of silicosis (in units of cases per year). The rate \( I \) was considered a weighted average of the age-specific incidence densities, \( I_a \). The latter densities were solved from the relation between risk (estimated by cumulative incidence rate) and incidence density, specifically:

\[ R_a = 1 - \exp(-I_a \times D_a) \Leftrightarrow I_a = -(\log(1-R_a))/D_a \]

But, since the silicosis risks were small, an accurate approximation was:

\[ I_a \approx R_a / D_a \]

This above example illustrates the many stages in risk model building. Although Nurminen et al. (1992), in their risk analysis, provided an extensive explanation of the assumptions (mathematical, statistical, demographic) underlying the modelling, there were also occupational hygienic and toxicological variabilities that could make the error range even more uncertain. To allow for the uncertainties associated with many of the model variates, different scenarios should be considered by using different inputs for the prediction formula. However, by setting the error limits high enough to swamp the uncertainty for each of the many variates separately — but not necessarily for all the variates simultaneously — risk assessments may consider scenarios that will rarely, if ever, come true.

4. RISK CHARACTERISATION

The very important part of QRA, risk characterization, summarizes and interprets the information collected from previous activities, and identifies the limitations and the uncertainties in risk estimates. When the results from exposure and effect estimation are at hand, the next task for the analyst is to present the information to the decision makers in such a form that they can readily act on it. Because the results produced by different methods of health risk analysis take on different tabular and graphical form, it is important to have guidance on how to convey information in a meaningful and useful manner. The interpretation of the results, both by the risk assessor and the risk manager (who is often a public decision maker) and later by the governmental and non-governmental organizations as well as the general public, may be critically dependent on the methods used to present the results. This is especially crucial in the linking of environment and health data, since the decision makers may not be well versed in the specialized statistical methods. Moreover, there is a need to present the results of health risk analysis in such terms that they can be easily transformed into inputs for a societal or individual cost-benefit analysis.
The quantitative estimate of risk is the result of main interest to the health agency or risk manager in arriving at decisions. It provides the link between the analysis and later policy work. Two basic and most commonly used quantitative measures of risk are increased individual risk and the number of excess cases of disease or disease burden ('body count').

Example 4 (cont'd) Using the models for risk presented in Section 3.5 above, Nurminen et al. (1992) predicted the occurrence of silicosis in the Australian labour force currently exposed to crystalline silica dust. As a result of an 0.9 (diagnostic range 0.4–1.9) average work-life risk, approximately 1010 (range 380–2410) silicosis cases were predicted for the next 40 years among the estimated 136,400 men exposed at current silica dust levels [0.61–0.8 (average 0.694) mg/m$^3$]. Currently 77% of the labour force at risk is exposed to silica dust levels of ≥0.1 mg/m$^3$. With this level as the limit about 440 (range 140–120) silicosis cases would appear in 40 years Adopting this level as the national exposure standard would reduce the risk of silicosis cases to 0.4 (range 0.1–1.0)%.

The average work-life risk of silicosis for an individual exposed to silica dust and the excess numbers of silicosis in the exposed population were estimated above separately for the currently prevailing exposure levels and for the lower control limits. To provide some perspective, the results of risk assessment are often expressed as small decremental risks. Thus, a risk analyst might interpret the results as follows. If an exposure level of 0.2 mg/m$^3$ were the standard adhered to, then there would be a 12% reduction in the excess number of silicosis cases; alternatively, if the exposure standard were set at 0.1 mg/m$^3$, a 36% reduction would be predicted. This would mean that the corresponding numbers of cases of silicosis prevented would be 230 and 570.

This type of characterization gives the kind of information risk managers need to make informed decisions regarding the necessary magnitude of reaction and whether a range of risk-reduction measures should be considered.

As noted previously, uncertainty is an important issue in health risk analysis. Adequate documentation of uncertainty is imperative in the characterization of risk. The discussion of assumptions and uncertainties should highlight the major limitations of the analysis, that is, remark on the relative importance of the various sources of variation (both sampling and nonsampling error).

Risk analysts frequently present information in terms of probability measures. Probability distributions can be difficult for a nonspecialist to interpret. One problem is that while cumulative incidence rates (estimates of risk) allow one to read a median (and percentiles of the distribution), the mean value cannot be determined from the plot in the case of a non-symmetric (e.g. log-normal or truncated) distribution. To avoid misinterpretation of a cumulative distribution function, it is recommended that it be plotted in alignment with a graph of an incidence density curve using the same horizontal scale, and that the mean be marked on both curves (Ibrekke and Morgan, 1987). Another problem is that the
accuracy of statistical estimation decreases away from the centre of a probability distribution (usually the median or mean) towards its tails. Thus, it is easier to characterize accurately the risk of exposure for about 95% of the population, than the risk of various special groups (e.g. occupationally exposed workers), who can be exposed to particularly high doses (the 95-99% group). This is unfortunate because it may be at the extremes of the distribution where prevention is most needed — for example, the estimation of the probability of rare but serious incidents or accidents is less reliable than that of more frequent events.

A new methodology for sensitivity analysis in public health risk assessments is possible using the Monte Carlo techniques (Thompson et al., 1992). These extended methods begin with the conventional estimation of exposure and risk, and continue by modelling key inputs as random variates described by probability density functions. For example, the problems associated with the use of overly conservative assumptions (which are likely to overestimate risk) and the need to properly account for small but highly exposed populations can be dealt with Monte Carlo techniques. This approach imparts much more information to the risk manager concerning the distribution of the likely values of each parameter of the risk reduction model than do single point estimates based on fixed parameter values. To allow for the uncertainties with the specification of the model variates, different scenarios regarding the assumptions can be proposed. Overall, the simulation technique provides a simplified quantitative way to estimate the probability distributions for exposure and health risks within the validity of the model used.

The final estimates derived from a quantitative risk assessment need to be put into words to explain the information. Qualitative discussion of the work must state the assumptions and limitations of the analysis and note key factors modifying the risk and its variation that cannot be quantified. The presentation of the results of a risk analysis should not preclude its direct application to policy making. While the line between risk analysis and decision-making is often not clear cut, the risk analyst must present the results in such a way that policy questions can be answered. Risk managers may even assume that the primary justification for performing risk assessments is that they contain cost-benefit analyses. For these applications, the Monte Carlo techniques have been invaluable tools (e.g. Worksafe Australia, 1993b).

The linking of environmental exposures to health outcomes is frequently done via a regression model for risk, for example, a multiple logistic regression. Whatever method is used, presentation of the results after allowance for covariates should be in a form similar to that which would be given if no covariates were included in the risk function. The quoting of coefficients in the logistic model does not achieve this, and is artificial since the logit transformation would not be necessary had there been only the one risk factor of
interest, and no covariates. This does not mean that the risk-odds ratio would not be useful as an auxiliary parameter in risk modelling. A risk analyst should, however, proceed to estimate more informative measures of exposure effect, such as the absolute excess risk (risk difference) or the relative excess risk (risk ratio minus one). (See Nurminen, 1995.)

While the line between risk analysis and decision-making is often not clear cut, the risk analyst must present the results in such a way that policy questions can be answered.

Epidemiological data often need to be interpreted in order to be of use for health policy making. Traditional epidemiology is mostly concerned with the increased incidence associated with exposure to a specific risk factor, whereas policy makers are more interested in the reduction of risk after the cessation of exposure. The importance of a risk factor for the incidence of a disease in a population is usually expressed as the etiological fraction, i.e. the proportion of the total incidence of the disease that can be attributed to that risk factor in the population. This post hoc measure indicates the proportion of incidence that could have been prevented by the total elimination of that risk factor in the population.

However, since prevention will usually not eliminate but merely reduce the prevalence of an environmental risk factor, a measure was developed to estimate the expected impact of a change in prevalence of a risk factor on the incidence of a disease, the potential impact fraction (Morgenstern and Bursic, 1982). It indicates the incidence that may be avoided by a planned intervention programme as a proportion of the incidence that would be expected to occur in that population without preventive intervention. The potential impact fraction can be calculated when the prevalences of exposure to a risk factor in the population and the corresponding incidence density ratios or risk ratios are known.

The potential impact fraction in the traditional epidemiological literature assumes an immediate reduction of excess risk after termination of exposure. However, this risk reduction may take many years to achieve, so that the estimates of effect will have to incorporate a time dimension. To achieve this objective, an applied epidemiological methodology based on the preventive impact fraction has been developed to help apply existing epidemiological knowledge to decision-making in health policy (Gunning-Schepers, 1989). The computer simulation model, PREVENT (Gunning-Schepers et al., 1993), can estimate the health benefits for a population of changes in risk factor prevalence. The PREVENT model is a useful tool for policy makers because it will present the results in graphic or tabular form for the intermediate output variates of
etiolologic fraction, trend impact fraction, and potential impact fraction, and for
the following final output variates: disease-specific mortality, total mortality,
disease-specific mortality difference, potential years of life gained, actual years
of life gained, survival curves, and life expectancy at birth.

A preventive intervention programme is often difficult to sell politically since its
effects take so long to become apparent. The situation can be even worse:
effects will seldom become apparent as real reductions in risk because of the
demographic changes in the target population over time. This does not mean
that prevention will not have a beneficial effect despite competing death risks,
say, in an aging population. It does mean, however, that in order to see the real
effects, it is important to show what would happen without the preventive
intervention, and not merely compare the effects to the current level of
mortality. The potential utility of simulation models such as PREVENT for
policy making lies in their ability to provide more precise quantification of
effect estimates by appreciating a time trend, multiple risk factors, and the
interaction between the effects of intervention and the demographic evolution of
the population (Gunning-Schepers et al., 1993).

This section reviewed some of the issues which need to be addressed in
translating the results of a quantitative risk assessment into implications for
policy or action. While the derivation of risk estimates is (in principle at least) a
straightforward input-output process, the presentation of results in interpretive
terms is a data-analytic challenge for the practitioner of applied epidemiology.
Whatever results are presented, they should reflect the impact of a change in
exposure conditions on the disease incidence.

5. ESTIMATES OF HEALTH IMPACT OBTAINED FROM
HEALTH RISK ANALYSIS

The following study examples will give some indication of the risk estimations
that have been presented in recent health risk analyses. All three examples are
concerned with adverse health effects related to air pollution in developing
countries. The examples are meant to mainly illustrate the kind of outputs that
are produced by the estimations; for more detailed results (e.g. interval
estimates) as well as a discussion of the limitations and characteristics of these
studies, we refer to the original sources.

Example 1 (cont'd)  Ostro (1996) presented estimates of health risk for the
population of Santiago, Chile. The results suggested that reductions in
particulates would result in significant health benefits. For example, for the
Santiago population of 4.4 million, lowering the annual PM10 to 30 µg/m³
would prevent 5,200 (central estimate, low estimate 2,400, high estimate 8,200)
premature deaths and 13.5 (central estimate; low estimate 9.6, high estimate
16.9) million days of restricted activity in adults due to respiratory illness each
year.
Example 6. Romieu et al. (1990) studied the health effects of urban air pollution in Latin America and the Caribbean region. The authors derived an age-specific estimate of the population exposed to different levels of particulate matter in order to make risk predictions. They examined total mortality based on a dose-response function derived by Evans et al. (1984) who, through meta-analysis, estimated the regression (age-adjusted) equation as:

\[ \text{Number of Excess Deaths} = 0.45 \times \Delta \text{TSP} \]

where the number of excess deaths is the change in the total mortality (annual deaths per 100,000) reflecting the change in the TSP concentration, \( \Delta \text{TSP} \). Assuming a total of 81 million persons, with 43 million exposed to an annual geometric mean of TSP of 100 \( \mu \text{g/m}^3 \); 23.5 million to 150 \( \mu \text{g/m}^3 \); and 14.5 million to 250 \( \mu \text{g/m}^3 \), and taking 75 \( \mu \text{g/m}^3 \) as a guideline value, they estimated the corresponding excess number of deaths per year as follows:

\[
0.45 \times (100-75) = 11.3 \text{ per 100,000, or 4,859 in 43 million persons} \\
0.45 \times (150-75) = 33.8 \text{ per 100,000, or 7,943 in 23.5 million persons} \\
0.45 \times (250-75) = 79.0 \text{ per 100,000, or 11,455 in 14.5 million persons}.
\]

This comes to a total of 24,257 excess deaths per year. Assuming a mortality rate of 5 per 1,000, then 24,257/(0.005 \times 81,000,000) = 0.06, or 6% of the annual mortality.

Using other relevant relations, the authors were able to estimate the increment in pulmonary symptoms and disease among children. For the three pollution levels, increases of 2.8%, 8.3% and 19.3% in frequency of chronic cough were estimated, with the excess number of children affected estimated at 2.3 million. For adults, they estimated 65 million days per year of respiratory-related restriction in work activity. For the elderly, the excess number of cases of chronic bronchitis was estimated at 105,800.

Example 7 In 1990, the U.S. Agency for International Development (USAID) and the U.S. Environment Protection Agency (USEPA) financed a study to evaluate the health risks related to environmental problems in Bangkok, Thailand. The study also aimed to test and adapt methods developed by the USEPA (USAID, 1990). Even though the application of known dose-response relations to the estimated human exposures is in theory quite straightforward, in practice it becomes difficult due to the many assumptions that need to be made. The concerns arising in this study included the lack of measurement for many air pollutants. Several hundred different contaminants are expected to be found in a polluted city like Bangkok, but available ambient measurements included only six. These were monitored in only a few places in the large city, and only mean annual measurements were available. The crude assessment must therefore tacitly assume that the places where the monitors are located will represent the average levels of pollution in the city. That the unmeasured pollutants pose negligible effects on health; that all individuals are equally exposed. Breathing average levels of pollution based on the annual mean, every day, and that there are no differences across individuals. In addition, indoor air pollution, or the time people spend indoors was not taken into account. The authors took the
view that most risk assessments represent a plausible worst case. This viewpoint may be true when assuming, for example 24-hour outdoor exposure levels for all subjects, but may not hold for low annual (average) exposures resulting from short periods of very high exposure levels.

To derive the dose-response estimates, the authors used the calculations of Ostro (1983). Based on the data obtained from the Health Interview Survey (consisting of 50,000 interviews) conducted by the National Center for Health Statistics, Ostro related the TSP levels to restricted activity days. The respondents were questioned about illnesses that resulted in work loss or reduced activity days over the previous two weeks. Questions regarding socioeconomical and demographic factors, smoking habits, and occupation were also included. To control for possibly confounding factors, the regression equations included variates for age, sex, presence of a chronic condition, race, marital status, annual income, annual mean temperature and precipitation, population density, occupational status, and the number of cigarettes smoked per day. The results of the ordinary regression were used in the risk analysis. The relation between restricted activity days in the receptor population and ambient TSP was obtained as:

\[ \text{Number of Restricted Activity Days} = 0.00282 \times 26 \times \Delta \text{TSP} \times S \]

where the factor 26 is the adjustment from a 2-week recall period to a full year, \( \Delta \text{TSP} \) is the change in TSP concentration, and \( S \) is size of the population.

In the Bangkok study it was assumed that there was no restricted activity below 75μg/m³. Then, for each monitoring station, the estimated population in the area around the monitor was obtained, and the above equation was calculated separately. The total number of restricted activity days is then given by the summing for all areas. For example, in 1986 the mean TSP level for Sukhumvit (an urban residential area) was 120 μg/m³. The estimated population in the area of the monitor was 781,000. Therefore, according to the equation above, there were about 2.6 million days of excess restricted activity days in the population (which is over 3 days per person per year, on average).

6. CONCLUDING REMARKS AND SUGGESTIONS

The methods used for risk estimation unavoidably give very approximate projections because they usually involve a myriad of assumptions whose tenability cannot be verified with certainty. Therefore the presentation of mere point estimates of the expected risks and excess numbers would undoubtedly give a misleading impression of precision. Confidence intervals or similar range estimates should be computed for the risk parameters (and thereby for the number of disease cases attributable to exposure) to provide some idea of the variability of the assessment process. But possible model misspecifications of the exposure-response relation can override sampling errors.
Providing decision makers estimates of environmental risks without corresponding estimates of the associated uncertainty is not in keeping with good epidemiological practice.

The outcome of the issues and problems raised above is the overall conclusion that the results obtained from environmental health risk analysis often cannot be quantified precisely. Thus, to allow for the uncertainty associated with the risk assessment process, different scenarios should be considered.

Given the unavoidable and often very large differences in quantitative risk estimates between studies, the statisticians presenting their risk analysis results to risk managers or decision makers should consider the following suggestions (Cox and Ricci, 1989):

- Risks should be presented in a sufficiently disaggregated form (showing risks for different subgroups) so that key uncertainties and heterogeneities are not lost in the aggregation.
- Confidence bands around the predictions of statistical models are useful, but uncertainties about the assumptions of the model itself should also be presented.
- Both individual risks and population risks should be presented, so that the equity of the distribution of individual risks in the population can be taken into account.
- Any uncertainties, heterogeneities, or correlations across individual risks should be identified.
- Sensitivity analyses for key assumptions should be used extensively.

Finally, a practical suggestion for the communication of the results of health risk analysis: Applicable display methods for the visualisation of aggregated environment and health data are readily available in the literature (e.g. Cleveland, 1993). More importantly, the graphical display of the distribution of risk predictions as a function of a change in exposure level may be a vital link between risk perception and informed judgement.
REFERENCES


