chapter twenty

Epidemiology in health risk assessment

A.E.M. de Hollander

20.1 Introduction: Why epidemiological data in health risk assessment?

The identification and quantification of human health risk associated with exposure to chemicals is a complex process in which a variety of disciplines are involved, such as toxicology, epidemiology, clinical medicine, chemical subdisciplines (analytical chemistry, organic chemistry, biochemistry), and biostatistics. All contribute, but none provide a complete picture. Among these disciplines epidemiology is becoming increasingly important. As will be pointed out in the next sections of this chapter, this is largely due to the growing scientific awareness that the relevance of results obtained in experimental animals to human health is limited. In Section 20.2, some important methodological limitations of epidemiology in studying environmental (including nutritional) risk factors are discussed. Section 20.3 indicates the prospective role of epidemiology in risk assessment and the way in which methodological limitations may be overcome.

20.1.1 A safer world for rats?

The information on toxicological risks from food contaminants and additives (both natural and man-made) is mainly derived from toxicological studies in animals. Compared with epidemiological studies, these studies have the advantage of an experimental design. All conditions are maintained constant except for the factor of interest: exposure to a certain substance. In toxicological studies, exposure and exposure conditions (such as housing, diet, and climate) can be carefully monitored and controlled. Histopathological and biochemical methods offer possibilities to study adverse responses with high sensitivity. However, toxicological research is not meant to make this world safer for rats and mice. It is also improper to deal with humans as if they were rats weighing 70 kg.

Translation of results obtained in experimental animals to human populations as a step in quantitative risk assessment requires three important assumptions:

- animals under laboratory conditions and human populations respond alike;
- the response to (high) experimental exposure is relevant to human health and may be properly translated to environmental exposure (including food intake) levels which are often orders of magnitude lower;
- (standard) experiments in animals reveal all responses to a substance which are potentially adverse to humans.

All three assumptions may be challenged. Consequently, they may give cause to substantial uncertainty in quantitative risk assessment.

The sensitivity of species (or strains or even individuals) to a toxic substance can differ dramatically. For instance the LD_{50} (see Section 8.9.1) of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) differs several orders of magnitude from one species to another. In hamsters, it is almost 10,000 times higher than in guinea pigs. There are not only differences in intensity. TCDD-induced tumorigenicity may also involve different organs in different species.

Although TCDD is believed to be a non-genotoxic carcinogen in rodents, results of human studies do not give rise to great concern. Epidemiological research on Vietnam veterans exposed to the TCDD-tainted defoliant Agent Orange did not reveal any effect. Even among a population of highly exposed people living near the chemical plant that exploded in Seveso in Italy in 1976, the only irrefutable effect was chloracne (rash). Only recently (and not undisputed in the scientific press) an increase of 50% in the cancer risk was found in a large cohort of workers occupationally exposed to high levels of TCDD for long periods. It mainly involved soft tissue sarcomas, while in experimental animals liver tumors were predominant. The exposure levels were about 500 times higher than the exposure levels human populations are likely to experience. In a low-exposure group, no increased cancer risk was shown, although the exposure was still 90 times the average environmental level. These results suggest that humans are by far less sensitive to TCDD than laboratory rodents. However, there is no guarantee humans will be the least sensitive species if other toxic substances are involved.

Quantitative risk assessment based on animal experiments means the translation of results obtained in genetically homogeneous experimental animals under well-controlled laboratory conditions to a free-living, heterogeneous human population exposed to a wide variety of risk factors affecting the state of health. In addition to the fact that experimental animals are roughly equally sensitive to toxic effects of the substance they are exposed to, they do not smoke, and do not drink, do not have dangerous occupations or hobbies, and do not have unhealthy dietary or sexual habits which may obscure the effects of the substance under investigation.

In most experiments, the test animals are only exposed to one toxic substance at the same time, while humans are generally exposed to a variety of chemicals. Exposure to mixtures or combinations of different substances may have unpredicted (and unpredictable) health effects as a result of all sorts of interactions between the components. This has been an issue of concern among toxicologists for more than a decade now. However, there is still no satisfactory answer to the question how to deal with combined actions in health risk assessment.

Since the (statistical) sensitivity of animal experiments is limited by the number of animals in the exposure groups, toxicologists are often forced to use relatively high exposure levels to ensure the detection of potentially adverse effects. As Theophrastus Bombastus von Hohenheim, better known as Paracelsus, already stated five centuries ago: "Alle Dinge sind Gift und nichts ist ohne Gift. Dasz ein Ding kein Gift ist, macht allein die Dosis" (only the dose determines toxicity). Substances humans cannot live without, such as oxygen and water are toxic at doses lower than ten times the normal intake. Thus, one may query the significance of effects which are observed at "unphysiologically" high exposure levels. Illustrative for this dilemma is the recent discussion in the American scientific press in which the relevance of animal experiments for carcinogenicity is called into question by several prominent toxicologists. In order to avoid false negative results chemicals are tested in chronic animal studies at maximum levels which are tolerated without clear signs of toxicity. Some toxicologists argue that these chronic exposures almost by definition lead to severe disturbances of homeostases. Responses associated with tumor promotion, such as excessive cell proliferation to compensate for cytotoxicity and disturbance of hormonal balances, have been observed at these levels. The authors pointed out that it would only be surprising if those lifelong disturbances would not be expressed in altered tumor incidence rates. In view of the fact that almost half of the tested chemicals appeared to be carcinogenic, some authors suggest that carcinogenicity revealed by these studies may often be an artifact of the experimental design, which is of no relevance to most environmental exposures.

Another subject of scientific dispute is the way in which results of animal experiments should be translated to real-life exposure levels. In Chapter 18, several models for extrapolation to low dose levels have been described. All lack information on what happens at the low doses to which humans are actually exposed. For instance, estimates of bladder cancer risk from saccharin, made on the basis of data obtained in rats, varied by as much as six orders of magnitude, depending on the assumptions used to translate from high to low dose levels. Estimates ranged from 0.001 cancers per million exposed, using the multi-hit method, to 5200 cancers per million exposed, using the single-hit method. In contrast, a review of 13 case-control studies in humans led to the conclusion that there was no consistent association between saccharin intake and the incidence of bladder cancer.

Only in human studies it may be verified whether environmental (including dietary) exposure to so-called rodent carcinogens indeed increases the risk of cancer. When assessing human health risk based on animal studies, the question ought to be asked whether these studies will reveal all relevant responses. One has to consider the possibility that the toxicological methods are "blind" to more subtle responses, which may have great impact on public health in the long term.

The case of oral contraceptives is very illustrative. Before their introduction, toxicity experiments in rodents revealed that female sex hormones could induce breast tumors. Considering the underlying mechanisms and low-dose extrapolation, it was concluded that no such effects were to be expected in women using oral contraceptives. However, the most important side effect of "the pill," the disturbance of blood coagulation, has never been found in experimental animals. Other responses which may not be easily revealed in animal experiments are minor neuropsychological disorders, such as chronic headaches,

concentration disturbances, forgetfulness, and depressiveness. These symptoms have been reported in painters exposed to organic solvents during long periods of their lives.

20.1.2 Risk communication

Epidemiology may contribute to a rational public and political awareness of the risks of daily life. Descriptive epidemiological studies may help health authorities to see the state of public health and its relationship with environmental problems in true perspective. With relatively simple statistical parameters of the health impact of serious diseases, one can inform the public and policy makers on the importance of certain risk factors. This may be useful in setting priorities for the funding of research, prevention, and control programs. For instance, one may rank diseases in terms of potential years of life lost and then conclude that cancer is by far the most serious threat to public health, followed by coronary heart disease, and traffic accidents (see Figure 20.1, based on Canadian health statistics). The diagram in Figure 20.2 shows the lost life expectancy for an individual, caused by several risk factors.

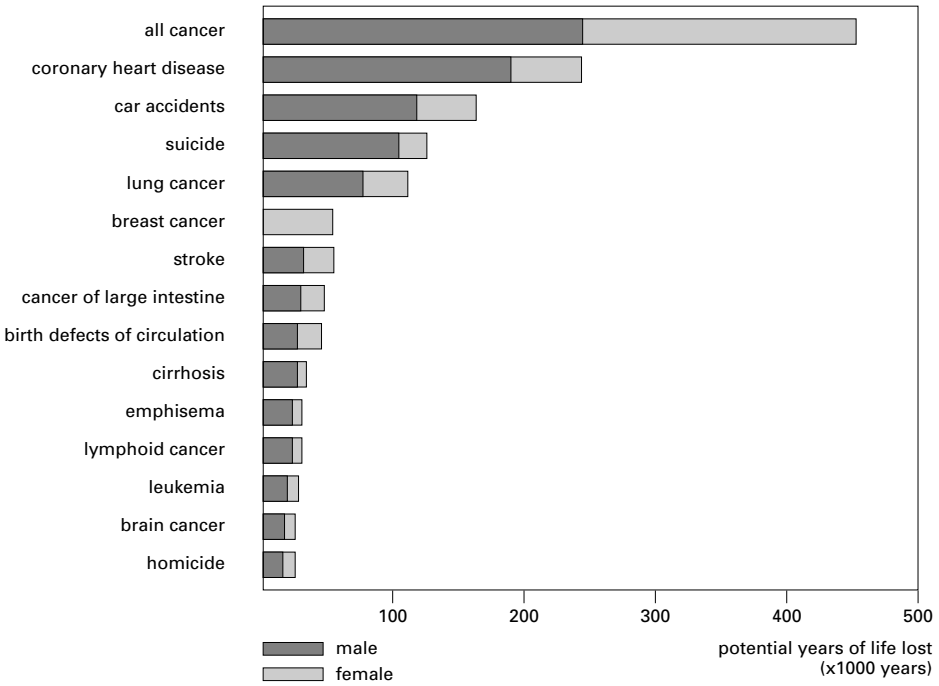

Figure 20.1 Health problems ranked by potential number of years of life lost.

A recent review of worldwide trends in age-related cancer mortality showed rapid increases in the prevalance of various types of cancer (e.g., of the brain, the central nervous system, breast, kidney). These could not be explained by increased accessibility of health care records, changes in disease registration, improvements of diagnostic technology or by life-style trends. The investigators suggested that these trends reflect an increase in environmental or occupational exposures to carcinogenic factors. Such descriptive studies may stress the need for more research and preventive measures to reduce exposure to carcinogenic agents. However, public health is a complicated subject that may be looked upon from many different angles. To find the right method to measure health impact is not easy. During the last decades public health science and management have focused on the

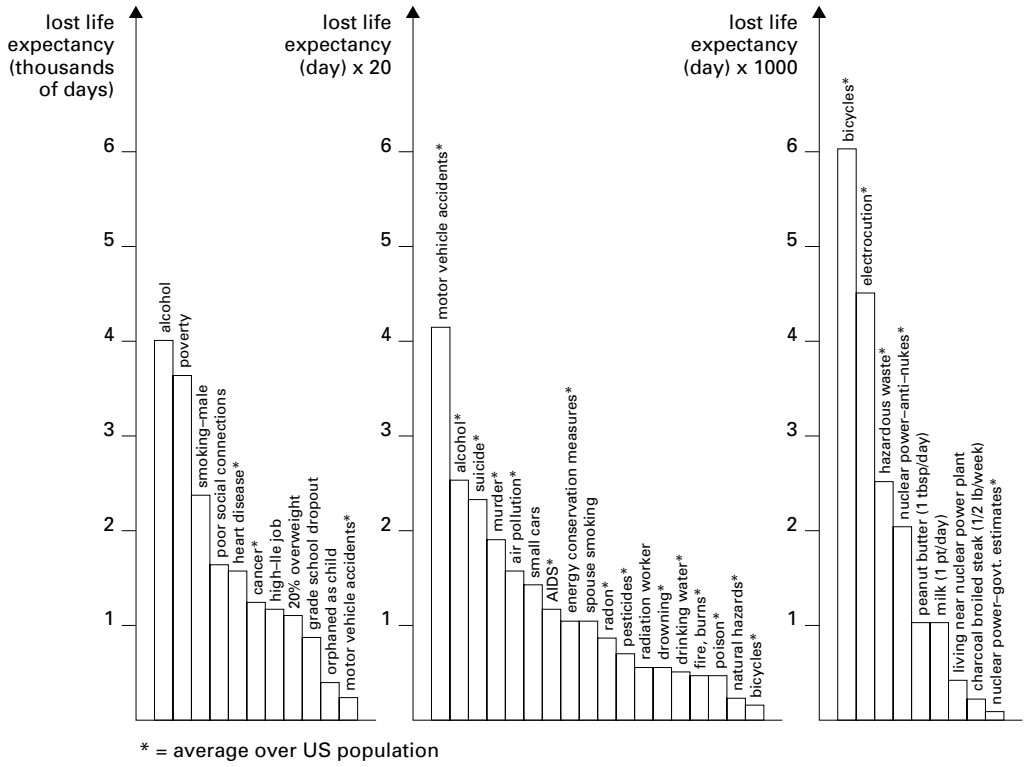

Figure 20.2 Comparison of risks. Asterisk designates average risk spread over the total US population: others refer to risks of those exposed or participating. The ordinate scale is shown at the left. The heights of the bars are multiplied by 20 in the center section and by 1000 in the right section. The first bar in each of these reproduces the last bar in the previous section, showing the effect of scale change. (Source: Cohen, 1991.)

prevention of (early) death. Nowadays many have recognized the fact that there is more to life than dying, as the saying goes. The area of special attention for public health policy is shifting from prevention of fatal disease to improvement of quality of life by reducing the period of dependence and disability in elderly life. Prevention should aim at chronic morbidity as in the cases of rheumatoid arthritis, chronic obstructive pulmonary disease, diseases of the eyes and ears, diabetes, dementia, and other multi-factorial diseases of old age which cause severe disability. The diagram in Figure 20.3 shows the principles of reduction of morbidity.

The great challenge for public health care is to make the disease-curve move faster to the right than the death-curve in order to reduce the black area in the diagram that represents severe disability. It is the concept of the ideal car that disintegrates completely after exactly ten years of loyal service without any prior mechanical problems. Some authors argue that for the purpose of health risk assessment, attention should be focused on risk factors associated with chronic morbidity rather than on pursuing every carcinogen that shows itself in animal experiments or occupational epidemiology.

It should be noted that there is no scientific consensus yet on how one should measure public health. The above examples were given to show that public health is more than just counting bodies, or the sum of years of life lost.

The public's perception of the threats of daily life may sometimes be seriously biased. Often the public opinion is fixed on the effects of popular risk factors such as

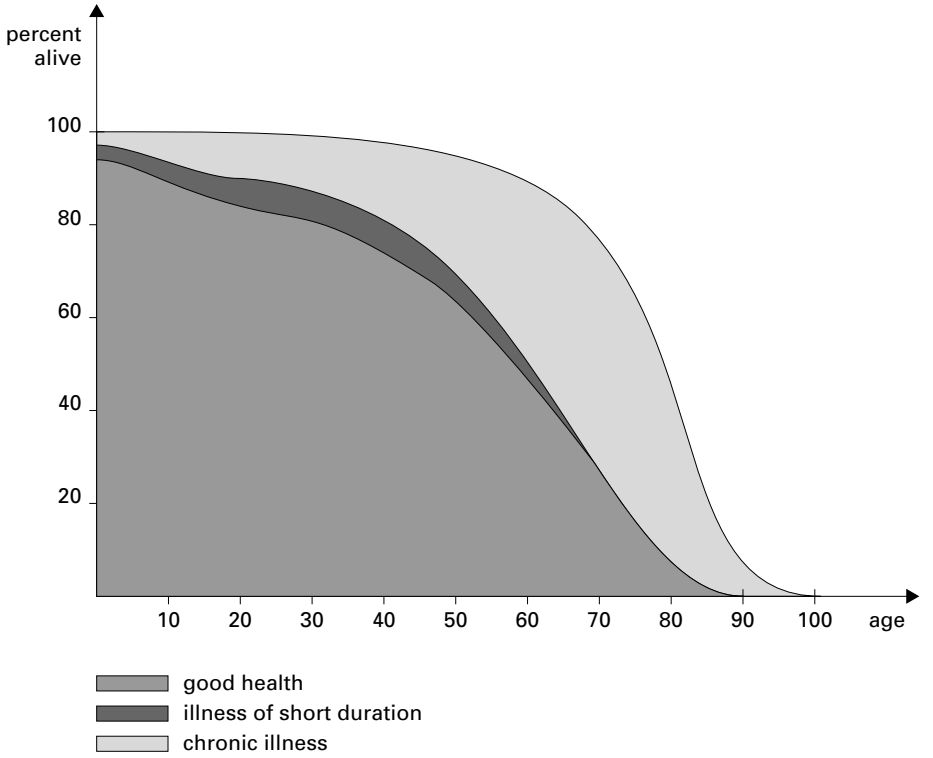

Figure 20.3 The burden of disease.

asbestos, TCDD, benzene, and radiation. However, the risks from these factors are, if verifiable, in fact often trivial compared to the dramatic health effects of factors such as traffic, smoking, dietary habits, occupational exposures, and poverty (wealthier is healthier).

Epidemiological data may give a good notion of the real risks of daily life by transfering information on health risk from the expert sphere to the public sphere in a way that appeals to lay-people. This can be achieved, for instance, by comparing the effects of different risk factors on public health. Although an oversimplification, the risk yardstick in Figure 20.4 might be a first step.

20.2 Limitations of epidemiology in risk assessment

In contrast to toxicology, epidemiology is an observational science. This is both its strength and its weakness. No extrapolation from one species to another is required. However, the study of free-living populations has important methodological limitations which have to be taken into account if the results are to be applied in health risk assessment.

20.2.1 Time interval between exposure and response

Epidemiological studies in which food components are tested for their adverse health effects are considered unethical in Western society. Furthermore, some health effects of major concern need a considerable amount of time to manifest themselves. In the case of chronic diseases like cancer, the induction period may span almost one generation. As in the case of exposure to industrial asbestos, it takes several decades until carcinogenicity

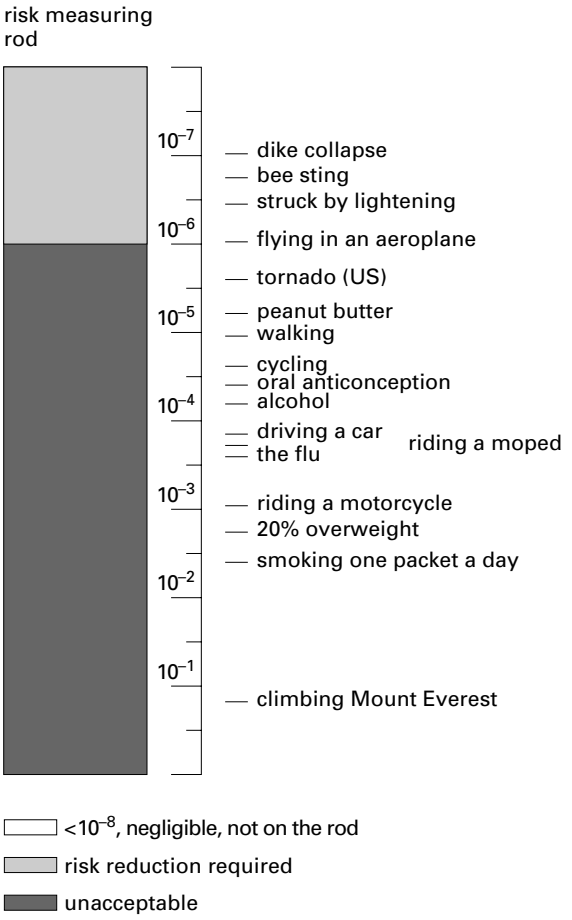

Figure 20.4 Measuring risks.

can be observed. Meanwhile, many people may already have suffered from hazardous exposures. Thus, for the screening of new chemicals or new applications, epidemiology is a useless tool and one has to rely on toxicological research.

On the other hand, epidemiology may provide possibilities to detect health risks after a certain food additive or contaminated food has found its way to the market. This so-called post-marketing surveillance is sometimes applied when new drugs are introduced.

20.2.2 Low sensitivity

Since World War II numerous associations between potential causal factors and major diseases have been found and confirmed. These were strong associations with large relative risks (ratio risk in persons exposed to the factor of interest/risk in the unexposed). Examples are smoking and lung cancer, excercise or dietary habits and coronary heart disease, urban pollution and lung disease, industrial asbestos exposure and mesothelioma. The results hardly needed statistical analysis to be convincing. Most of the evident risk factors have been revealed and confirmed. The associations epidemiologists are looking for nowadays are likely to be more subtle. Owing to computer programs, the design and statistical analysis of studies have become more sophisticated. Nevertheless, the results often fail to be convincing.

One of the problems epidemiologists have to deal with is the low sensitivity of their methods. Slight increases in the risk of exposed compared to unexposed will remain unnoticed, because of chance, lack of heterogeneity of exposure, inadequate parameters of exposure and/or response, insufficient time elapsed since onset of exposure, or bias and confounding.

If the prevalence of a disease is high, a slight increase in relative risk has great public-health impact. If approximately 40% of the male population dies of cardiovascular disease, even a slight increase in the relative risk caused by a certain risk factor leads to a significant increase in the number of deaths per year.

In the next subsections, the most important aspects of the sensitivity of epidemiological methods are discussed. For a more detailed description of epidemiological studies, the reader is referred to Chapter 15 of Part 2.

Intermezzo

Example

In the summer of 1972, a chemical factory accidentally distributed ten 25 kg bags of polybrominated biphenyls (PBBs, a flame retardant with carcinogenic potency) to dairy farms throughout Michigan for use as cattle feed additive. The chemicals caused widespread deaths among cattle, calf loss, and decreased milk production, but the mistake was not discovered until 9 month later. Since PBBs persist in tissue, it was possible to measure levels of the contaminants in cattle and dairy food products. This monitoring led to widespread quarantining and destruction of animal food products in which PBB levels exceeded certain limits.

Reports on various diseases in families on contaminated farms led to epidemiological studies designed to detect possible PBB-related health effects. Cases of PBB-syndrome consisted largely of subjectively reported symptoms, such as headaches, joint pain, loss of appetite, and skin rash. Epidemiological studies failed to show associations between PBB serum levels and reported symptoms in PBB exposed family members. Low levels of PBB were demonstrated throughout the whole population of Michigan, but not in other neighboring states or in Canada. The chemicals were found to concentrate in mother's milk. Continuous surveillance of health effects was carried out in exposed families. No evidence of increased disease incidence, including cancer, has yet been revealed. However, it may take decades before the carcinogenic potency of a substance can be observed in epidemiological research.

20.2.3 Chance

In epidemiological studies, the effects of differences in exposure to a risk factor are compared. The sensitivity of this type of study is limited, because groups may differ in the outcome of response variables by chance alone. The role of chance can be illustrated by calculating the confidence intervals for an observed relative risk equal to 1 as estimated with the standardized mortality ratio of a cohort study or the odds ratio from a case-control study (Tables 20.1 and 20.2). Table 20.1 shows how the relative risk estimate improves with an increasing number of cases. Table 20.2 shows that the reliability of the relative risk estimate depends on the number of cases as well as on the proportion of controls exposed. Given an estimated relative risk of 1, with a certain confidence interval (95% is normally used), these intervals give the range of results which may be observed by chance alone.

Even if there is no effect of exposure, an industrial medium-sized cohort study may show an increased risk of 20 to 30% by chance alone. If one compares cases and controls

Table 20.1 95%-Confidence intervals for an estimated standardized mortality ratio of 1, based on the results of a cohort study

Number of responders in the exposed group	Comparison with an unexposed group of the same size
100	0.750–1.334
200	0.816–1.225
1000	0.915–1.094
5000	0.959–1.042

Source: Krewski et al., 1989.

Table 20.2 95%-Confidence intervals for an estimated odds ratio of 1, based on the results of a case-control study (unmatched analysis, equal number of cases and controls)

Percentage of controls exposed	Number of cases			
	100	200	1000	5000
5%	0.222–4.498	0.354–2.822	0.656–1.524	0.832–1.202
25%	0.502–1.992	0.620–1.612	0.812–1.231	0.913–1.096
45%	0.551–1.815	0.661–1.512	0.835–1.197	0.923–1.083

Source: Krewski et al., 1989.

Table 20.3 Total sample size required to detect an increased risk in epidemiological cohort studies with same-size exposure groups

Relative risk (%)	Risk in unexposed cohort (%)		
	1	10	20
1.05	1,032,609	114,096	64,497
1.10	266,433	29,569	16,796
1.25	46,546	5,233	3,014
1.50	13,231	1,519	894
2.0	4,083	487	298
4.0	789	107	73
5.0	538	77	55

Source: Krewski et al., 1989.

for exposure to a risk factor (in so-called case-control studies), the role of chance is even more prominent. In a case-control study with as many as 200 cases, one might find an increased risk ranging from 50 to 200% even if there is no real association between risk factor and response. This variation due to chance decreases with the size of the population. An important implication of this phenomenon is that negative results of epidemiological studies offer no guarantee that effects are actually absent.

The effect of chance implies that very large cohorts are necessary for the detection of low risks. Table 20.3 shows that cohort studies involving several thousands of subjects are required to detect relative risks lower than 2 if the disease concerned is rare in the unexposed cohort.

20.2.4 Exposure

To identify and quantify risk associated with environmental exposure, one needs to design a natural experiment in which identifiable populations differ distinctly in exposure level. This is often difficult. The case of smoking and lung cancer was clear-cut, owing to the fact

that there was an extreme heterogeneity in exposure levels between populations of smokers and non-smokers. However, if everyone in the Western world would have smoked a pack a day no epidemiological study would have been able to identify smoking as the predominant cause of lung cancer. Lung cancer would probably have been believed to be genotypical. If everyone is exposed to the predominant risk factor, the distribution of lung cancer is largely determined by genetic predisposition for the disease. In general, one might argue that the wider a particular environmental risk factor is spread, the less it explains the distribution of a disease. If a cause is ubiquitous, no epidemiological method will be able to detect it, because all methods rely on differential exposure to a risk factor.

This implies that epidemiological studies on health risk from food components will become increasingly difficult since there is a tendency towards buying mass-produced foods from supermarkets. In addition, increasing mobility eliminates the effects of differences in local environment. Health effects of certain food contaminants can only be studied if populations distinctly differ in exposure. Examples of such natural experiments concern populations living in areas where the water or soil is contaminated with nitrate, arsenic, or cadmium, resulting in a relatively high intake of these contaminants with food or drinking water. In epidemiological research one can compare the incidence of a particular disease in a certain area with national statistics or with the incidence in uncontaminated regions. Comparing people who use artificial sweeteners or drink coffee with people who do not, is another example of studying distinct differences in exposure. If the exposure conditions differ little from each other, the sensitivity of epidemiological methods is often too low.

It is generally recognized that exposure of defined populations is often hard to quantify. Exposure assessments based on questionnaires, as in the case of dietary patterns, are inherently inaccurate. Exposure levels in ambient and indoor air, food, or drinking water are often poorly documented, if monitored at all. If the disease under investigation follows chronic cumulative exposure, as in the case of cancer or cardiovascular disease, exposure patterns may be determined by factors, such as time-activity patterns, housing, occupation, and dietary habits. These may change drastically in the period between the start of exposure and the onset of the disease. In general, continuous monitoring of exposure is not feasible.

Often, the exposure variable is a poor surrogate, because it does not properly reflect the actual exposure. In studies on the association between passive smoking and lung cancer, non-smokers were classified as belonging to the group of passive smokers, depending on the smoking habits of their spouses. This classification may reflect the exposure to environmental tobacco smoke inaccurately, because exposure takes place in a variety of environments, at work, during public transport, or in all sorts of public places such as pubs and theaters. The relevant latency period for lung cancer ranges over several decades. During such a period, changes may occur in smoking habits, working environment or even in the spouses involved.

Intermezzo

Example

In the city of Antofagasta, Chile, the water supply contained high concentrations of arsenic during the period between 1958 and 1970. In this desert area, the water has to be brought from the Andes, where arsenic is found in naturally high concentrations, over a distance of 300 km. Copper mining was another probable source of arsenic contamination. In the early 1960s, a large number of citizens (children in particular) were found to have an abnormal skin pigmentation. Clinical investigations revealed that this pigmentation was

accompanied by other symptoms such as chronic cough, chronic diarrhea, abdominal pain, and blood vessel disorders.

Using arsenic concentrations in hair and nails as exposure markers, a clear association was shown between arsenic exposure and intoxication symptoms. Further, a significant higher prevalence of abnormal pigmentation was found in comparison with a control population.

In retrospective studies, epidemiologists have to rely on the memory of participants to make an assessment of exposures. Unwittingly, this information can be seriously flawed. In case-control studies, a well-known type of information bias is caused by the fact that cases (people with the disease) or their relatives tend to overreport their exposure to a risk factor compared to controls, especially if an association between the risk factor and the disease concerned has been suggested in the media. This may then become a self-fulfilling prophecy.

In epidemiology, measures or "surrogates" of exposure are often of a qualitative nature. At best, they discern the less exposed from those who are more exposed. Sometimes, several groups differing in exposure can be defined depending on the quality of the exposure data. Misclassification of exposed as unexposed or the other way round tends to dilute the association between exposure and disease, as it diminishes the differences in exposure between the groups to be compared.

This can be illustrated by the results of three large cohort studies (including about 90,000 workers in rubber or petrochemical industries) on the relationship between benzene exposure and leukemia. With relatively high statistical power due to the large sizes of the cohorts, all three studies failed to show any effect of long-term benzene exposure. However, further analysis of the job characteristics revealed that many workers classified as exposed, had been exposed to negligible concentrations in the past. If subpopulations of workers with the highest exposures were defined and compared with unexposed, an effect became apparent.

Quantitative health risk assessments are based on quantitative relationships between exposure and response. This requires quantitative exposure data. Often such data are unreliable or lacking, and consequently have to be estimated by "educated guessing." This is very important, because apart from the extrapolation model, a quantitative risk assessment is as reliable as the exposure data. The latest episode of the continuing debate in the US on the occupational exposure standard for benzene is illustrative. This standard is based on a quantitative evaluation of an increased leukemia risk that was found for cohorts of workers who where followed since the 1950s and '60s and for which historical exposure data were provided by chemical industries. Recently, it was revealed that the actual exposure levels in the past had been much higher. As it appeared, the information on exposure levels provided by the industry had been carefully brought in accordance with occupational exposure standards which applied in the '60s. Because the exposure levels used in the quantitative risk extrapolation were too low, the true risk may have been overestimated. Consequently, the discussion on adjustment of the current occupational exposure standard has started again.

Intermezzo

Example

In April of 1956 a 6-year old girl was taken to a hospital in Minamata, Japan, with gait and speech defects as well as delirium. Many cases followed in all age groups. After several years, it was found that methylmercury was the cause of the epidemic of neurological

disorders. Mercury derivatives from various small industries had accumulated in sediment and fish in Minamata Bay. Several hundreds of families of fishermen incurred injury. Among the exposed families there was at least a tenfold increase in birth defects, in particular severe cerebral palsy.

20.2.5 Response

Most diseases can be looked upon as multi-causal. A variety of factors contribute to the incidence. When comparing the outcomes of responses in different exposure groups, epidemiologists try to make adjustments for risk factors such as age, sex, socioeconomic status, smoking, and dietary habits. In practice, this implies that exposure groups are selected in such a way that these factors are equally distributed. If known risk factors are equally distributed over the exposure groups, comparison may reveal an association between exposure to the risk factor and the response concerned. The more a risk factor contributes to the development of a disease, the more epidemiological studies are likely to identify it and be able to quantify its influence.

Intermezzo

Example

In Iraq an outbreak of intoxications due to the consumption of fungicide-treated grain occurred in 1971–72. Of a total of 6530 hospital admissions, 459 patients died. The symptoms were typical for organic mercury intoxication. Analytical data showed a definite association between the consumption of bread baked from fungicide-treated grain and the mercury concentrations in hair and blood.

Examples of predominant risk factors are smoking (historical obduction data suggest that lung cancers were relatively rare before the "smoking era" started), asbestos (almost 80% of mesothelioma incidence can be traced back to occupational asbestos exposure in the past), and vinyl chloride (only two cases of angiosarcoma, a blood vessel tumor of the liver, in a cohort of workers alerted a medical company doctor to suspect the PVC-monomer). Toxic effects following the intake of food are often associated with accidental contamination. In the history of epidemiology there are many examples of this kind. In 1956 the accumulation of methylmercury in fish from Minamata Bay in Japan caused severe neurological disorders in many people as well as neurological birth defects (see also Section 20.2.4). In 1981 contaminated cooking oil caused an epidemic of acute lung disease in Madrid, Spain, with more than 20,000 cases over a period of three months. The contaminant caused severe immunotoxic effects. Toxic effects are often caused by components of vegetable origin. In India many people suffer from paralysis of the lower limbs, caused by the consumption of chick peas. In Africa a syndrome of severe neuropathological symptoms is caused by chronic hydrocyanic acid intoxication as a result of the poor man's diet consisting only of cassava. However, this kind of one-to-one associations between risk factors and diseases (disease specificity) is rare. In most cases, the risk factor under investigation is one of several, which reduces the sensitivity of epidemiological studies substantially.

If an association between the risk factor under investigation and another risk factor is found, one has to take into account the possibility of confounding. For example, comparison of an urbanized population with a rural population will show that the lung cancer incidence rate is higher in the first group. Thus, one might conclude that exposure to the

urban ambient air induces lung cancer. However, the difference in lung cancer incidence may well be explained by differences in smoking habits and occupation, two major risk factors for lung cancer associated with urban life.

Inaccurate response parameters may also dilute the association between a risk factor and a disease and thus decrease the sensitivity of epidemiological methods. This can be best illustrated by studies on the association between radiation exposure and leukemia. A substantial risk increase could have escaped observation in epidemiological studies, if all types of leukemia had been combined. Only by studying the various types of leukemia individually, the association between exposure and acute nonlymphatic leukemia was revealed.

Intermezzo

Example

In 1955 the first cases of an "epidemic" disease causing severe bone pain and multiple bone fractures were reported by a local general practitioner in a rural area in Japan. The disease, referred to by the Japanese as Itai-Itai disease ("ouch-ouch"), was clinically diagnosed as osteomalacia in most cases. Epidemiological studies showed that the disease occurred throughout the region. Only when it was found that the disease was associated with tubular kidney damage, a link with cadmium poisoning could be demonstrated. Local food and drinking water appeared to contain high levels of cadmium.

20.3 Prospectives

The earlier mentioned limitations imply that the role of epidemiology in quantitative assessment of toxicological risks from food intake is often relatively modest:

- epidemiological studies are only feasible after human exposure has taken place and sufficient time has elapsed for disorders to develop;
- it may not always be possible to define exposure groups that differ substantially and consistently in their exposure to the risk factor of interest;
- the sensitivity of the available methods is often low; small health risks may not be noticed due to chance, lack of heterogeneity of exposure, inaccurate parameters of exposure and/or response, bias or confounding;
- negative results, even from impeccable studies, cannot prove the absence of an effect.

In spite of these limitations many investigators claim that the advantages of epidemiology are not yet fully exploited in toxicological risk assessment. In this section, several aspects of optimizing the application of epidemiological data in risk assessment will be discussed.

20.3.1 Toxicology and epidemiology as complementary disciplines

In identifying, quantifying, and verifying human health risk, toxicology and epidemiology are complementary. Both disciplines have their limitations. To put it bluntly: animal studies may only predict what will happen to other individuals of the same species under the same circumstances, while epidemiology often tries to compare incomparable free-living human populations. On the other hand, using the biblical metaphore: 'the lame

leading the blind,' together they may bring us where we want to be. If toxicological and epidemiological data are evaluated in coherence, most of the individual shortcomings may be overcome. Animal studies can provide a first screening of hazardous chemicals used in food production, while the existence of toxicological risks at realistic exposure levels is most adequately studied in humans. Further, explorative epidemiology may reveal unknown associations between environmental factors and diseases, but only toxicological research can elucidate the underlying mechanisms.

20.3.2 Analysis of all available data, meta-analysis, and publication bias

Considering the limitations of the disciplines involved, human health risk assessment should comprise all available data, human and animal, positive and negative results. As far as epidemiological data are concerned, especially the validity of the study design should be carefully scrutinized. A number of important questions have to be addressed. What is the validity of measurements of exposure and response variables? How can the pitfalls of epidemiology, i.e., confounding, selection, and information bias be circumvented? Did the study generate a hypothesis by computerized analysis (some investigators use the metaphore "torturing") of existing data sets or did it test an a priori hypothesis by an appropriate study design?

Lack of statistical power can sometimes be overcome by combining (pooling) data from different studies through so-called meta-analysis. While in primary epidemiological research, populations of exposed and unexposed are the subjects, in meta-analysis the results of these primary studies function as such. In meta-analysis, weight can be assigned to the results of the individual studies, according to the reliability or the sample size of the primary studies. Meta-analysis offers the opportunity to study how variables affect study results. Such variables may be date of publication, exposure levels, study design, definition of exposure, definition of disease, methods of statistical analysis, and geographic location of studies.

In analyzing the results of a number of studies, one methodological pitfall of great concern remains: publication bias or bias to positive results. Editors of scientific journals are more likely to publish results indicating positive associations, even if they are only weak and from a methodological point of view inadequate, than results of impeccable studies indicating no association at all. (The book on train accidents which never happened on railroads which were never built, was never written.) Authors may (subconsciously) be inclined to refrain from completing work that does not offer the promise of positive findings, publication, and peer recognition. The consequence may be that a literature study comprising all available animal and human data still gives a distorted picture of human health risk.

20.3.3 Causality

An old truth is that by itself, epidemiology will never be able to prove the causality of an association between a risk factor and a disease. The likelihood of causality is often evaluated using the following criteria:

- an appropriate *time sequence* of exposure and onset of response, which includes the time interval necessary to induce the response;
- *consistency* of an association in various populations, under various circumstances, observed in studies with different designs;
- *strength* of an association, expressed in terms of relative risk estimates. The higher the relative risk, the more unlikely an association based on bias or confounding;

– *dose–response relationship*. Higher exposures should lead to increases in response intensity or frequency (though less simple dose-response relationships may be found);
– *biological plausibility*. Agreement with knowledge on underlying mechanisms.

The latter criterion takes us back to toxicology. Epidemiology may validate the relevance of toxicological data for human health. Conversely, toxicological data, combined with clinical findings, may be necessary to validate associations suggested by explorative epidemiology.

20.3.4 Future

In the coming years, the role of epidemiology in health risk assessment is likely to become more prominent. Several developments will contribute to this.

– The reliability of exposure assessment will increase as a result of the development of appropriate biomarkers for exposure as well as the improvement of exposure monitoring and modeling. An example of exposure assessment using biomarkers is the monitoring of DNA-adducts in white blood cells of humans exposed to genotoxic substances. The amount of adducts found in human DNA samples not only reflects exposure to certain genotoxic substances, but also overloading of certain defense mechanisms, such as DNA-repair.
– Biomedical research provides a growing number of biological variables (biomarkers) that are early (accurate) indicators of the development of diseases before the symptoms appear. Examples of early indicators of diseases are changes in immunoglobulin composition, (liver) enzyme activity and lung function parameters. Spectacular developments in molecular biology provide tools to relate specific mutagenicity of substances to activation of oncogenes in human tumor cells. The use of such biological response indicators in epidemiology may improve the specificity of associations between exposure and response' and may enable early interventions.
– Improvement of disease monitoring will offer possibilities to study the distribution of diseases in time and space more adequately.
– Cooperation between epidemiologists and toxicologists may lead to more sensitive methods to analyze and extrapolate available data on human health risks.

Reference and reading list

Bertazzi, P.A., C. Zocchetti, A.C. Pesatori, S. Guercilena, M. Sanarico, L. Radice, Ten-Year Mortality Study of the Population Involved in the Seveso Incident in 1976, in: *Am. J. Epidemiol.* 129, 6, 1187–1200, 1989.

Buffler, P.A., The evaluation of negative epidemiologic studies: the importance of all available evidence in risk characterization, in: *Regul. Toxicol. Pharmacol.* 9, 34–43, 1989.

Checkoway, H., N.E. Pearce, D.J. Crawford-Brown, *Research methods in occupational epidemiology*. New York/Oxford, Oxford University Press, 1989.

Cohen, B.L., Catalog of risk extended and updated, in: *Health Physics* 61, 317–335, 1991.

Doll, R., Health and the environment in the 1990s. *Am. J. Publ. Health* 82, 933–941, 1992.

Feinstein, A.R., Scientific standards in epidemiologic studies of the menace of daily life, in: *Science* 242, 1257–1263, 1988.

Fries, J.F., L.W. Green, S. Levine, Health promotion and the compression of morbidity, in: *Lancet*, 1989, 481–483.

Hattis, D. and K. Silver, Human interindividual variability: a major source of uncertainty in assessing risk for noncancer health effects. *Risk Analysis* 14, 421–432, 1994.
Hertz, Picciototto, Epidemiology and quantitative risk assessment: a bridge from science to policy. *Am. J. Public Health* 85, 484–491, 1995.
Kaldor, J., N. Day, The use of epidemiological data for the assessment of human cancer risk, in: Hoel, D.G., R.A. Merrill, F.P. Perera, (Ed.), *Risk Quantitation and regulatory policy*. Banbury Report 19, Cold Spring Harbor Laboratorium, 1985.
Krewski, D., M.J. Goddard, D. Murdoch, Statistical considerations in the interpretation of negative carcinogenicity data, in: *Regul. Toxicol. Pharmacol*. 9, 5–22, 1989.
Krewski, D., D. Wigle, D.B. Calyson, G.R. Howe, Role of epidemiology in health risk assessment, in: *Cancer Res.* 120, 1–24, 1990.
Lilienfeld, A.M., D.E. Lilienfeld, *Foundations of Epidemiology*. New York/Oxford, Oxford University Press, 1980.
Ozonoff, D., Conceptions and misconceptions about human health impact analysis. *Environ. Impact Assess. Rev.* 14, 499-415, 1994.
Rinsky, R.A., RE: Benzene and leukemia: a review of the literature and a risk assessment, in: *Am. J. Epidemiol*. 129, 5, 1084–1086, 1989.
Rose, G., Environmental factors and disease: the man made environment, in: *Brit. Med. J.* 294, 963–965, 1987.
Rothman, K.J., *Modern Epidemiology*. Boston/Toronto, Little, Brown and Company, 1986.
Rothman, K.J., Methodologic frontieres in environmental epidemiology. *Environ. Health Perspect.* 101, 19–21, 1993.
Skene, S.A., I.C. Dewhurst, M. Greenberg, Polychlorinated Dibenzo-*p*-dioxins and Polychlorinated Dibenzofurans: The Risks to Human Health. A Review, in: *Hum. Toxicol.* 8, 3, 173–203, 1989.
Taubes, G., Epidemiology faces its limits. *Science* 269, 164–169, 1995.
VeFlorey, C. du, Weak associations in epidemiological research: Some examples and their interpretation, in: *Int. J. Epidemiol.* 17, 950–954, 1988.
World Health Organization (WHO), *Guidelines on studies in environmental epidemiology*. Geneva, Environmental Health Criteria 27, 1983.