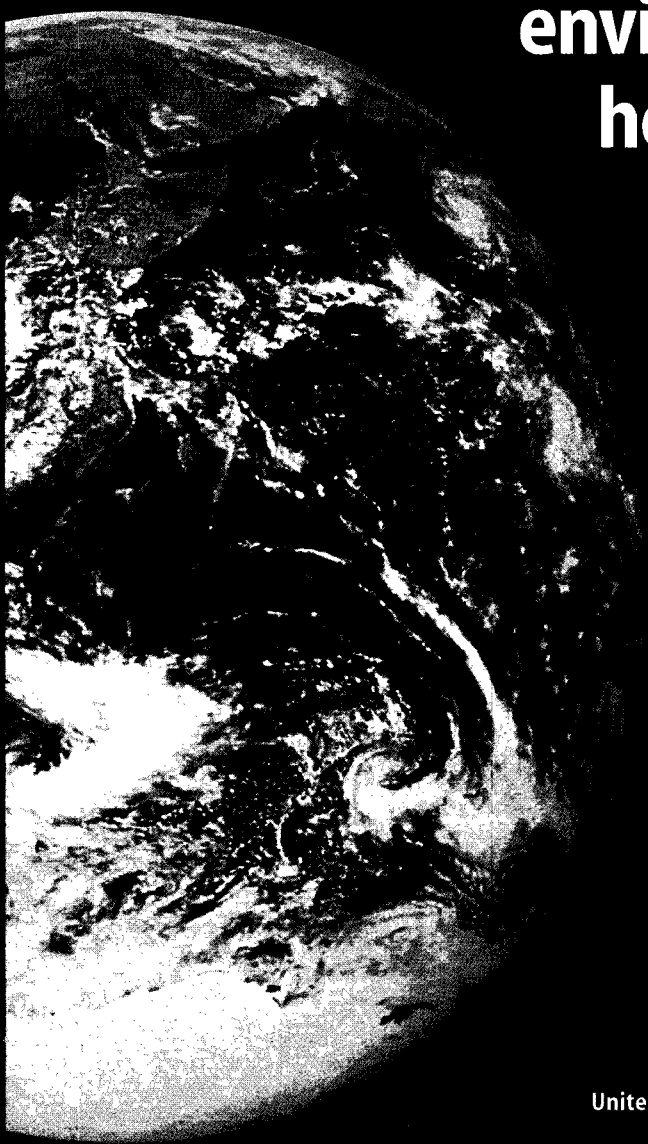
WHO/EHG/95.26
English only
Distr.: Limited

Linkage methods for environment and health analysis

General guidelines

Edited by
D. Briggs
C. Corvalán
M. Nurminen



UNEP
United Nations Environment Programme

United States Environmental Protection Agency

Office of Global and Integrated Environmental Health
World Health Organization
Geneva

Linkage Methods for Environment and Health Analysis

General guidelines

A report of the Health and Environment Analysis for Decision-making (HEADLAMP) project

Linkage Methods for Environment and Health Analysis

General guidelines

A report of the Health and Environment Analysis for Decision-making (HEADLAMP) project

United Nations Environment Programme

United States Environmental Protection Agency

World Health Organization

Office of Global and Integrated Environmental Health
World Health Organization
Geneva, 1996

Cover photo: WHO/NASA

Contents

Preface

The Health and Environment Analysis for Decision-making (HEADLAMP) project is a joint collaborative project between the United Nations Environment Programme (UNEP), the United States Environmental Protection Agency (USEPA) and the World Health Organization (WHO).

HEADLAMP aims to make valid and useful information on the local and national health impacts of environmental hazards available to decision-makers, environmental health professionals and local communities. It uses methodologies in environmental epidemiology, human exposure assessment and other health and environment sciences to collect and analyse data to produce information that can be understood easily and used as a basis for action.

The project was initiated in late 1993 with a feasibility study to identify and adapt methods that could be used at the local level in combination with routinely collected health and environment data, to estimate the health impacts of environmental contamination. Field studies were later carried out in Accra (Ghana) and São Paulo (Brazil) to examine data availability and quality as well as the potential for linking health and environment data. One of the project's important milestones was a consultation with international experts, held in Geneva in August 1994. Several of the papers presented during that consultation were published in a special issue of the *World Health Statistics Quarterly* (48;1995), under the title, "Health and environment analysis and indicators for decision-making". Additionally, a report entitled *Epidemiologic Methods for Linking Health and Environment Data for Decision-making* was reviewed, and the future direction of the project determined. This book is a revised version of part one of that report, and deals with general issues of data linkage. Part two deals in more detail with the analytic methods.

During 1995, seven field studies were conducted in Calcutta (India), Cape Town (South Africa), Cotonou (Benin), Dar es Salaam (Tanzania), Managua (Nicaragua), Manila (the Philippines), and Talcahuano (Chile). The studies focused on the development and testing of environmental health indicators, and their links to local decision-making. Preliminary findings were reviewed at a field studies meeting held in Nairobi, Kenya in October 1995. A book on the findings of the field studies is being planned, following the recommendations of the meeting. Instructional materials and workshops on HEADLAMP methods, and the identification, collection and use of environmental health indicators, are also being planned as part of future capacity-building activities.

We gratefully acknowledge the valuable comments of the many reviewers of this book who included members of a review team established at the first HEADLAMP consultation and composed of M. Akerman, U. Blumenthal, H. Pastides, P. Quénel, M. Salinas and J. Songsore. Valuable comments were also provided by M. Conomos, A. Dahl, C. Dora, T. Fletcher, G. Goldstein, H.N.B. Gopalan, K. Katsouyanni, M. Krzyzanowski, J. Leigh, A.J. McMichael, S. Markowitz, B. Nussbaum, C. Stephens, and J. Stober. Thanks are also due to J. Johnson for preparing the layout of this document.

Carlos Corvalán,
HEADLAMP coordinator,

Office of Global and Integrated Environmental Health
World Health Organization, Geneva, March 1996

Health and Environment Analysis for Decision-making

C. Corvalán[a] *& T. Kjellström*[a]

1. INTRODUCTION

The *Health and Environment Analysis for Decision-making (HEADLAMP)* project is aimed at improving information support for environmental health policies. HEADLAMP makes valid and useful information on the local and national health impacts of environmental hazards available to decision-makers, environmental health professionals and the community. It combines methodologies in environmental epidemiology, human exposure assessment and other health and environment sciences to produce and analyse data, to convert these data into information, and to present this information so that it can be understood, interpreted and acted upon by those responsible for environmental health protection. The information created via local and national HEADLAMP applications will help monitor progress towards sustainable development as recommended in Agenda 21 (UN, 1993).

Important elements of HEADLAMP are methods for linkage of health and environment data, the use of environmental health indicators used to quantify and monitor the local situation, and the interpretation and translation of resulting information into the decision-making process. This chapter gives an overview of the HEADLAMP project and its background. A series of papers, published in the *World Health Statistics Quarterly (Volume 48(2), 1995)*, present further details on HEADLAMP methods and examples of field studies, based on materials discussed at the first expert meeting on HEADLAMP held in Geneva in August-September 1994.

2. THE HEALTH AND ENVIRONMENT SITUATION

Human exposure to pollutants in the air, water, soil and food — whether in the form of short-term, high level episodes, or longer-term low level exposures — is a major contributor to increased morbidity and mortality. The disease burden attributable to these exposures is not known with any degree of certainty, however, because levels of general environmental pollution fluctuate greatly,

[a]Office of Global and Integrated Environmental Health, World Health Organization, Geneva

methods for analysing the relationships are incompletely developed, and the quality of available data is generally poor. Precise measures of the association between pollution levels and health outcomes are therefore rare. Exposure to environmental pollution is also usually involuntary and people may be ignorant of their exposure or its possible effects; as a result they may exert little control over exposure risks. Biological and chemical agents in the environment are nevertheless responsible for the premature death of millions of people and to the disablement of hundreds of millions more every year (WHO, 1992a). Certain environmental hazards affect very large populations. In large cities, for example, millions of people may be exposed to severe outdoor air pollution. In addition, indoor air pollution is a widespread and serious problem particularly in developing countries. World-wide, at least a hundred million people are potentially affected by respiratory diseases associated with air pollution (WHO, 1992a). The need to control harmful exposures is therefore evident.

The ability to link health and environmental data, and thereby to understand relationships between levels of exposure and health outcome, is clearly vital in attempts to control exposures and protect health. This capability is particularly important for countries in which issues of environmental pollution have traditionally taken second place to demands for economic development. In many of these countries, environmental pollution is rising while populations are undergoing rapid expansion, particularly in urban centres. Levels of exposure and their potential health effects are therefore liable to increase. At the same time, awareness is growing of the links between economic growth and environmental protection, and of the need to develop strategies for sustainable development which both preserve the environment and enhance quality of life. In this context, decision-makers urgently need information on the health impacts attributable to environmental pollution in order to assess the implications of their decisions, compare the potential effects of different decisions and choices, and prevent irreversible and costly health and environmental damage.

Standards and guidelines against which to assess levels of environmental pollution are now widely available. WHO has developed environmental quality guidelines for different pollutants in the air (WHO, 1987), water (WHO, 1993), food (FAO/WHO, 1989) and the workplace (WHO, 1980, 1981, 1982, 1983, 1984, 1986). These guidelines are based on epidemiologic and toxicologic studies and indicate the maximum environmental levels, or the levels of human exposure, considered acceptable in order to protect human health. Nevertheless, individual susceptibility to pollution varies, so some persons may still experience adverse health effects at levels below the maximum recommended levels. Moreover, in many areas of the world these levels are frequently exceeded, in some places by as much as several times the guideline levels, and actions to reduce human exposure may be difficult or very costly. Adverse impacts on human health may thus be expected to continue in these areas. In such cases, analysis of health and environment data provides a valuable tool for obtaining

estimates of the health impact of pollution, which can be used to set priorities for action.

Many epidemiologic studies have been undertaken to analyse the relationships between specific forms of environmental pollution and health outcome. Most of these have been in developed countries, and the methods used are not always easily applicable to other settings, especially if high quality data are unavailable. In particular, major problems often exist in obtaining either health or, more particularly, environmental exposure data at the individual level. As a consequence, it is normally necessary to rely on so-called "ecological" methods, in which the statistical unit of observation is a population rather than an individual (Beaglehole et al., 1993).

A serious limitation in conducting these studies concerns the measurement of exposure in individuals. Routinely collected environmental data are widely available in most countries. In addition to the data commonly collected by national and local authorities, extensive monitoring is also carried out via the *Global Environment Monitoring System* (WHO, 1990, 1991; UNEP/WHO, 1993). These networks provide data on pollution levels at specific sites, which can then be used to characterise average exposures for geographical regions. Environmental data are also often compared with guideline values or standards for maximum recommended levels in order to determine levels of compliance with prevailing policies. Seldom, however, are the data used to quantify the potential health impacts. Equally, although many countries routinely collect health outcome data in the form of morbidity and mortality statistics, attempts are rarely made to link them to environmental or other factors in order to attribute outcomes to their cause.

The analysis of data on health and environment, as a basis for estimating the health impact of pollution and setting priorities for action, thus remains an urgent need in many parts of the world. One of the main aims of the HEADLAMP project is to provide this capability. Its objective is to contribute to an ongoing process of monitoring and policy review in which repeated assessments of health and environmental status are used, first to develop and then to revise and update appropriate actions to reduce exposures. To this end, HEADLAMP is based on the linkage of environmental and health data, using relevant epidemiological methods as well as tools of environmental analysis. The linkage methods are designed both to control for extraneous determinants and to account for changes in the underlying population structure. Given that linkages are often conducted using data spanning relatively long time periods, they also account for artifactual changes resulting from changes in disease classification or changes in exposure monitoring methods. By providing this capability, HEADLAMP methods may also play a significant role in *Environmental Health Impact Assessment (EHIA)*, the aims of which are to

predict the health impacts of development projects likely to introduce new forms of pollution or increase existing pollution levels.

> The analysis of data on health and environment, as a basis for estimating the health impact of pollution and setting priorities for action, remains an urgent need in many parts of the world.

3. SPECIFIC ENVIRONMENTAL HEALTH PROBLEMS

Human exposure to pollution may occur in many different situations and via a range of different pathways. Pollution may be encountered, for example, in the general environment, in the occupational environment, or in an individual's personal or domestic environment. Human exposure in any of these environments may occur via the air, water, food or soil.

3.1 Air Pollution

Air pollution is a general term that describes the admixture of potentially harmful substances within the air we breathe. The most well-documented of these substances (and those usually monitored on a routine basis) include sulphur dioxide (SO_2), nitrogen oxides (NO_X, including NO and NO_2), carbon monoxide (CO), ozone (O_3), lead (Pb), and total suspended particles (TSP, also known as suspended particulate matter, or SPM, of which the respirable particles are of most concern, e.g. particulates of up to 10 μm in size, or PM_{10}). The major sources of these pollutants is the combustion of fossil fuels for energy generation, industrial processes and transportation, and of solid fuels, such as coal and wood, for domestic purposes. Evidently, the combination and concentration of outdoor air pollutants varies from city to city, according to the quantity and composition of fossil fuels used. But they also depend on other environmental factors such as the geographical and meteorological characteristics of the area concerned. Air pollution is different from other forms of pollution in that, once the pollutants are in the air, exposure cannot be easily avoided. If high levels of air pollution are occurring in a city, therefore, it may be expected that a large proportion of the population will be exposed. Nevertheless, levels of air pollution may vary markedly even at the local scale, especially in the case of low-level emissions (e.g. from road transport). Short-term variations in pollution levels will also occur due to variations in emission activity. In addition, levels of exposure will vary depending on the proportion of time people spend outdoors, and to the ability of the individual pollutants to enter the indoor environment.

Indoor air pollution is in some situations considered to be more serious than outdoor air pollution. This is due to the tendency for the entrapment of pollutants indoors, resulting in higher concentrations. In addition, most people spend a much larger proportion of their life indoors, than outdoors, often in close proximity to indoor emission sources. Indoor air pollution is an especially serious problem in some developing countries. It is also often unrelated to outdoor pollution levels. In many rural areas, for example, where ambient air pollution is low, the use of biomass fuel in unventilated houses causes pollution concentrations much higher than those in even the worst polluted cities. This form of pollution is likely to affect women and children more severely because they spend the longest time indoors. Chen et al. (1990) identify solid-fuel-fired cooking and heating stoves as a major source of indoor pollution and conclude that the evidence argues strongly that this source of indoor air pollution is a risk factor for chronic lung disease in adults, especially among women. Coal smoke may also be a risk factor for cancer in women. The health effects in children, however, are of special concern. Combustion-related pollutants are a risk factor for acute respiratory disease in young children, which is one of the main causes of infant and childhood morbidity and mortality in developing countries (WHO, 1992b).

Levels of suspended particulates in the most polluted indoor environments in developing countries may reach levels that are several times higher than the daily averages measured in cities with severe air pollution problems. Nevertheless, one of the main differences between outdoor and indoor air pollution is that the latter typically has a larger temporal range, periods of no exposure alternating with periods of high concentrations (e.g. associated with use of cookers or fires). Although outdoor air pollution also shows marked diurnal, weekly and seasonal variations, the amplitude of variation is generally less and the pollutants involved are hardly ever absent (i.e. they provide a source of ubiquitous exposure).

3.2 Water Pollution

Water pollution is a pressing problem in many areas of the world, irrespective of the level of development. The main pathway of exposure is through use of contaminated drinking water. Most drinking water is obtained from ground water or surface water and can be contaminated by the presence of physical, chemical and biological agents. Exposure may also occur as a result of bathing or washing in contaminated water.

Biological pollution is often of greatest concern, particularly in less developed countries and in rural areas. Diarrhoeal disease due to faecal water pollution is a widespread problem and a major cause of infant deaths (Martinez et al., 1993). Chemical pollutants in water include nitrates and nitrites, pesticides, volatile

organic compounds, and heavy metals such as arsenic and lead, and to a lesser extent, mercury, cadmium and other metals.

Both chemical and biological contamination may be difficult to remove from water supplies. In the case of ground waters, for example, rates of turnover of the water may be extremely slow — often in the order of hundreds or thousands of years — so that, once contaminated, the waters are not easily cleaned. In the case of surface waters, pollutants are often stored in the sediments on the bed and banks, and slowly released over many years. As a result, contamination often continues long after the original source of pollution has been removed.

3.3 Food and Soil Contamination

Environmental pollution can also be transmitted through food and via the soil. The soil may be chemically contaminated with a wide range of pollutants, including pesticides and heavy metals such as lead or cadmium. Agricultural activities, industry, landfill and emissions from road transport are all important sources of soil pollution. Once in the soil, many pollutants may be held for considerable periods of time, bound to the soil particles. Over time, however, many of the pollutants are likely to be leached from the soil into the ground- or surface-waters, while others are taken up by plants. Thereby they may enter the foodchain. Direct contamination of foodstuffs may also occur as a result of the deposition of pollutants from the atmosphere, through the use of contaminated irrigation water, by application of pesticides and other substances to growing crops or livestock, and through contamination during processing and distribution. Major concerns in these cases are contamination with chemicals (such as pesticides) or biological agents. The WHO Working Group on Infant Feeding estimated that up to 70% of the 1,400 million cases of diarrhoeal disease occurring worldwide each year in children under five years are due to pathogens transmitted through food (food, in this case, includes drinking water and water used in food preparation) (WHO, 1993b). In local "hot-spot" areas, metals such as mercury and cadmium have caused important outbreaks of poisonings; for example Minanata disease (WHO, 1976) and Itai-Itai disease (WHO, 1992c).

3.4 Radiation

Another environmental health hazard of concern is ionizing radiation. This is emitted from nuclear power stations, both as a result of accidents (e.g. Chernobyl) and routine operation. Natural emissions (e.g. from geological sources) are also important, for example in the form of radon gas which is a widespread source of indoor air pollution. The health effect of main concern that is due to general environmental exposure is cancer, but a number of other effects are being investigated following the Chernobyl accident (WHO, 1994).

Exposure to non-ionising radiation (e.g. from overhead power lines) is also a cause of growing concern, although the links to health remain more equivocal.

4. THE ENVIRONMENTAL HEALTH HAZARD PATHWAY

Environmental health hazards take many forms. They range from traditional hazards such as human faeces, in densely populated areas, to the wide mix of air pollutants emitted by road vehicles. The hazard pathway, however, is broadly similar, and is described in Figure 1.1. The starting point in most cases is some form of human activity, or, more rarely, a natural process which releases pollutants into the environment. The process of release is termed *emission*. Once in the environment, pollutants typically undergo a process of *dispersion*, during which they are transmitted through the environment via the air, water, food or soil. *Exposure* occurs when humans encounter the pollutants in the environment.

4.1 Emission Sources and Processes

The human activities responsible for emissions of pollutants are highly varied. Mining and quarrying, energy production, manufacturing, transport, agriculture, domestic activities and waste management are all major emission sources, but other sectors (such as tourism, forestry and commercial services) may also be important. In each case, emissions may occur throughout the complete product life cycle, from initial extraction of the raw materials through processing and distribution, to product use or consumption and final disposal. A wide range of emission processes are also involved. Energy combustion — for example in vehicles, manufacturing industry, electricity generation and home heating — is one of the most important emission processes, especially to the air. In addition, however, large quantities of pollutants are emitted through other processes, such as spillage of chemicals, the deliberate discharge of effluents, leakage and seepage from equipment or storage sites, leaching of agricultural chemicals, gaseous release of volatile materials (e.g. in petrol stations or from landfill sites), wear and tear of equipment (e.g. industrial machinery, vehicle tyres and brakes), and respiration and excretion by agricultural livestock. Because these activities and processes represent the starting point for environmental emissions they also represent the most effective point of prevention and control. Much environmental policy is therefore focused at trying to regulate these source activities, or to incorporate into them methods of emission control.

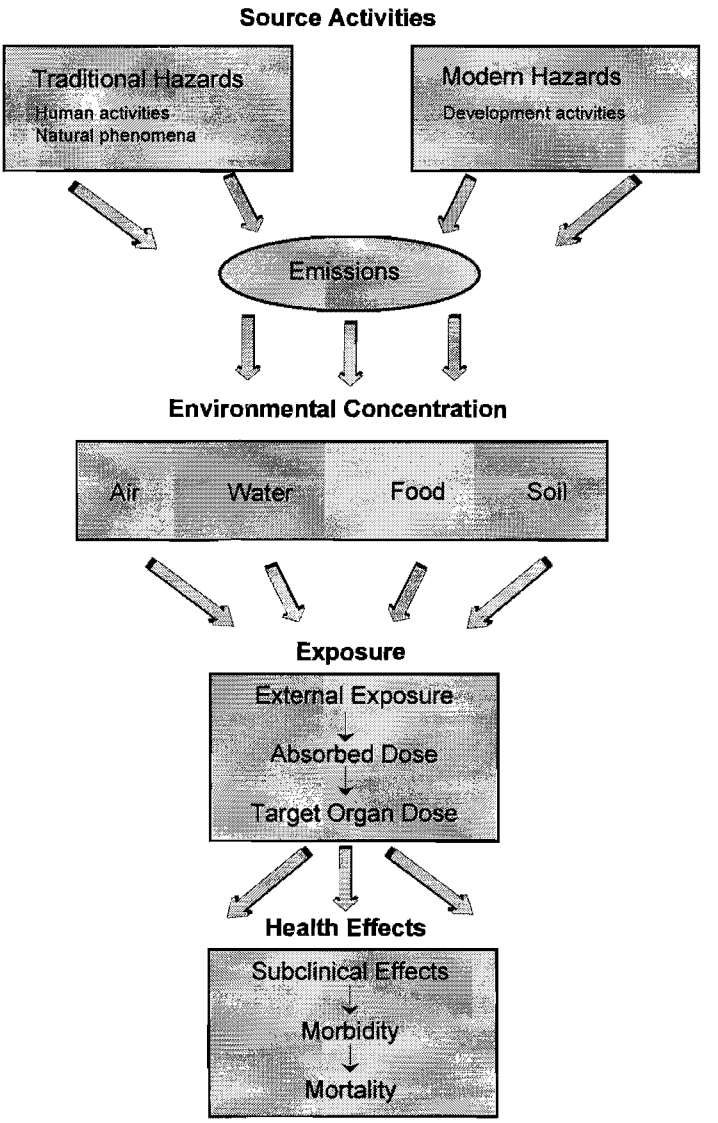


Fig. 1.1 The environmental health hazard pathway: conceptual framework at the individual level

4.2 Dispersion Processes

Once in the environment, pollutants may be dispersed via air, water, soil, living organisms and/or human products (e.g. food). The pathways of dispersion vary greatly, depending upon both the emission source and the pollutant concerned. Rates and patterns of dispersion also depend to a large extent upon the environmental conditions. Pollution dispersal in the air, for example, is affected by weather conditions (especially wind speed, wind direction and atmospheric stability), by the emission height (e.g. whether from ground level sources such as road traffic or from high level sources such as tall chimneys), and by the local and regional topography. Pollution dispersal in the soil is influenced by soil conditions, such as its texture, structure, degree of compaction and drainage characteristics. Dispersal by living organisms or human products depends upon the patterns of movement, contact and exchange which occur.

During dispersion pollutants undergo a wide array of changes and transfers. Dilution occurs due to admixture with the transporting medium (e.g. the air or water). Sorting and segregation of pollutants occurs on the basis of size, mass or density. Chemical reactions occur, breaking down the original pollutant or converting it into new compounds. Pollutants are also removed from the transporting medium through deposition; air pollutants, for example, are deposited due to settling out under the effects of gravity, by rainwash and by interception (scavenging) by plants and other obstructions.

The intensity of these different processes varies substantially over both time and space — and often over very short periods and distances. As a result, many pollutants show extremely complex patterns, especially in complex environments such as cities and towns where there are a large number of emission sources and major variations in environmental conditions. This complexity means that it is often very difficult to model or measure pollutant patterns and trends, and thus to predict levels of human exposure.

4.3 Exposure Processes

Pollutants enter the human body in a number of different ways — by inhalation, ingestion or dermal absorption. The amount of any given pollutant that is absorbed is often termed the *dose*, and may be dependent on the duration and intensity of the exposure. *Target organ dose* refers specifically to the amount that reaches the human organ where the relevant effects can occur. The first effects may be sub-clinical changes, which in turn may be followed by disease and in some cases even death. Brief examples are given below.

Air: contaminated air may enter the human body by inhalation of the air pollutants, but may also be absorbed through dermal contact. The most common health effects are associated with the respiratory system, particularly in more sensitive persons, such as children and the elderly. For example, particulate matter and sulphur dioxide, two very common air pollutants, may cause broncoconstriction, chronic bronchitis or chronic obstructive lung disease.

Water: contaminated water is usually absorbed by the human body by ingestion, but some contaminants may also be absorbed by inhalation or via dermal contact. Depending on the type of contamination, different vital organs may be targeted by different contaminants. For example, contamination with volatile organic compounds may affect the liver or the kidneys, causing hepatitis or kidney failure.

Food and soil: contamination by food and soil may take many forms. Lead contamination provides an example. Lead in food or soil is absorbed from the gastrointestinal tract (up to 50% may be absorbed in children compared to 10% in adults). Almost all organ systems can be potential targets for lead, including effects on hemobiosynthesis, the nervous system and on blood pressure.

Measuring exposures accurately and precisely is clearly of great importance when seeking to establish exact associations with health outcomes. Nevertheless, it is often impossible to measure exactly how much of the exposure of interest reaches the target human organ. Biological monitoring techniques can provide good estimates of dose but they are not always practical — or available — in assessing exposure to environmental pollutants. In some cases, individual exposure levels can be measured by using personal monitors. Almost all these approaches, however, are extremely costly and time-consuming, and consequently cannot easily be applied to a large number of individuals as part of a population study. More commonly, therefore, exposure is assessed indirectly, for example on the basis of measured pollution levels for a whole area. The measured levels are then used to give an exposure score to all individuals living or working within the area. Such an approach clearly ignores local or individual variations in exposure and results in mis-classification of exposure levels. These errors are likely to increase as the size of the individual areal units increases. Theoretically, therefore, the use of small areal units (e.g. at the scale of census districts or neighbourhoods rather than whole cities or departments) should help to improve exposure estimates. Unfortunately, the lack of pollution data often limits this approach. In addition there may be a significant time lag between exposure and health effects for many pollutants. This means that the health outcomes observed at present may be due to exposures which occurred many years or even decades earlier. Historic data on pollution levels are often especially sparse. Significant uncertainties in exposure classification consequently tend to occur, and the existence of a measurable concentration of a pollutant, even when higher than recommended levels, is not

always a sufficient basis to infer health effects. Moreover, exposure often occurs to a number of different pollutants, in combination, so environmental concentrations of one pollutant do not always give a good indication of potential health effects. Social and other factors may also act to distort or mask the association between exposure and health outcome. For all these reasons, health or environmental data *on their own* may give misleading impressions of the environmental health risks which exist. Instead, recognition, assessment and management of known environmental health problems, such as those described above, requires the use of both environmental and health data, and — where feasible — their combination through the use of appropriate methods for data linkage.

5. ANALYSIS AND INTERPRETATION TOOLS

The linkage of environmental and health data offers considerable benefits, but also poses many dangers if not carefully carried out. In linking the data it is all too easy to overlook the statistical problems and inconsistencies of the different data sets, or to misinterpret their apparent relationships. Valid linkage thus relies on the use of both valid data and appropriate linkage methods.

Many methods for data linkage have been developed in many different areas of application. Their suitability for linking environmental and health data, however, is often limited and always needs to be carefully assessed. Two important criteria must be considered in this context. On the one hand the methods must be simple, inexpensive to implement and operable with available data, thus allowing rapid assessment. If the methods are overly complex, requiring extensive resources and large amounts of additional data collection, few of the less-developed countries will be able to apply them, and even in more developed countries their use may be costly and result in delays in action. On the other hand, if their results are to be accepted as a basis for action, the methods must be scientifically credible and statistically valid. This means that they should be accurate, sensitive to the variations in the data of interest, and unbiased. They should also produce results that agree with those obtained from more detailed studies, for which the statistical precision can be quantified.

In practice, it must be admitted that these requirements cannot always be met — indeed, if they could, there would hardly be a need for individual-level studies. Even if they do not meet all these criteria, however, the methods may have considerable value. Results from ecological studies, for example, are useful if the potential biases can be identified and evaluated (though this is usually difficult and further individual-level results are needed for accurate risk assessment). At the very least, the results can show areas or issues requiring further, more detailed investigation. Researchers in countries where detailed, individual-level studies have not been performed also urgently need access to

methods which can help to shed light on the extent and health effects of specific forms of environmental pollution.

If detailed information on the exposure-response relationship of pollutants in different settings around the world was available, techniques of risk analysis could be used to estimate the impact of exposures on different populations without the need for new substantive research. This implies knowledge about exposures, estimates of the population exposed and of the health effects associated with the exposure in the form of a dose-response function. At present, this approach is possible to some extent, but the lack of information for many parts of the world (especially developing countries), and about many exposure-response relationships, acts as a major limitation. Thus, such quantitative risk assessments can often only be made by extrapolating the available study results from one country (often developed) to others (often less developed). The fact that the range of exposure levels, and the distribution of extraneous determinants, may differ substantially between populations inevitably limits the validity of this approach. In addition, assessments can only be reliably carried out for pollutants for which well researched exposure-response relationships have been established. Even then, uncertainty regarding the assumed association between environmental pollution levels and the actual exposures in individuals is a major constraint.

The limitations of risk analysis are therefore considerable. Nevertheless, it remains the only tool available for estimating the health outcomes of environmental pollution for areas where health monitoring is not undertaken, or for which data quality is poor. It is also the only feasible approach for obtaining crude estimates of health impacts in very large population groups. The development and application of well-tested methods of risk assessment is therefore an important priority.

At the same time, it is important to emphasise that — even when adequate control for extraneous determinants are incorporated — HEADLAMP methods should not be seen as substitutes for individual-level epidemiological studies. Instead they are seen as a useful alternative where opportunities for more detailed studies are limited, and as a means of extracting valuable information from routinely collected data. The long-term challenge remains the development of new forms of study design and data analytical techniques for environmental epidemiology.

> Valid health and environment data linkage relies on the use of both valid data and appropriate linkage methods.

6. DEVELOPING ENVIRONMENTAL HEALTH INDICATORS

Studies of environment and health, and of the linkages between them, can produce large volumes of data. If they are to support and improve decision-making, however, these data need to be translated into a clear set of messages, targeted at issues capable of management and control. One way of achieving this is through the development and application of *indicators*. The key characteristic of an indicator is that it converts "data" to useful "information". The development of appropriate environmental health indicators is clearly integral to the HEADLAMP approach. They are essential both to the design of HEADLAMP studies, and to the use of HEADLAMP results. Prior identification of indicators, for example, helps to set the agenda for the study by highlighting the issues which need to be investigated, and thus the data and methods needed. Equally, the indicators provide a means of converting the results into a language and form of direct relevance to decision-makers. In addition, once identified and established, the indicators provide a means of monitoring subsequent trends in environmental health, and hence of evaluating the effectiveness of any action taken.

Methods of data linkage and use of environmental health indicators are therefore invaluable tools for policy making and management. Reduction of exposures requires investment by people and authorities; given the shortage of resources for essential development activities in virtually all countries, sound and convincing information is essential to motivate such investment. The information required is likely to include clear specification of the problem, its importance, and the costs and benefits of possible response options. Providing this information requires the availability both of suitable methods of data analysis and linkage, and of indicators which can express the results of these analyses in terms which are understandable and relevant to the decision-maker.

7. THE HEADLAMP PROCESS

The HEADLAMP project has three defining characteristics which differentiate it from ad hoc epidemiologic studies. These are:

1. HEADLAMP is based on already known and scientifically established relationships between environmental exposures and health effects. On the basis of these relationships it is possible to define environmental health indicators which — within the context of HEADLAMP — are chosen for their potential value in the decision-making process. Research to establish new environment-health relationships is a related but separate activity.

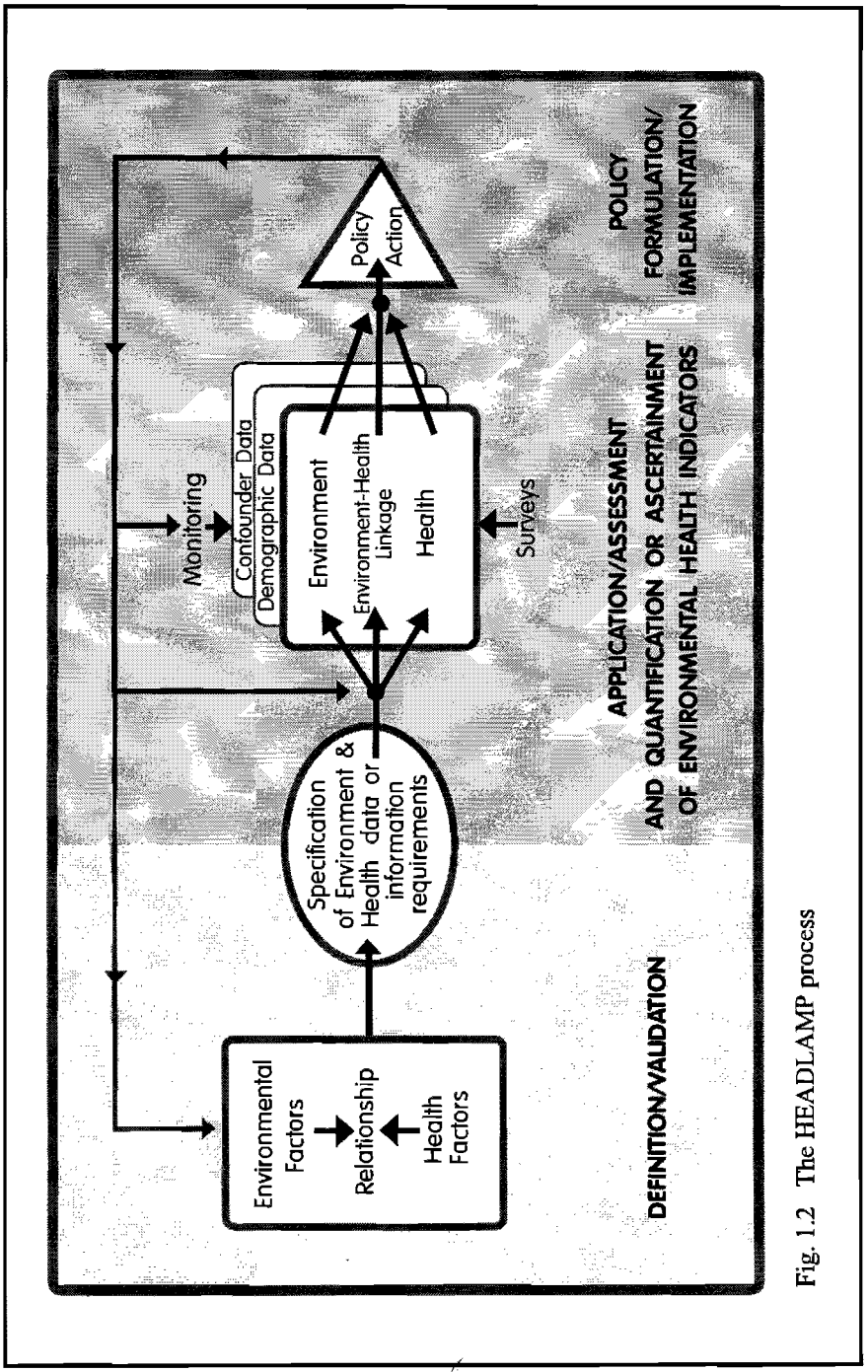


Fig. 1.2 The HEADLAMP process

2. The environmental health indicators used in HEADLAMP are usually based on the use of routinely-collected data. A major advantage of this approach is its cost-effectiveness. Data collection is expensive, and it is therefore important to obtain maximum possible value from the data through their repeated and most effective possible use. It also provides important feedback to the data collection process by helping to indicate needs for new or improved monitoring. To measure the relevant environmental health indicators, it may also be necessary to collect new data. In these situations HEADLAMP will encourage the use of appropriate, low-cost techniques.

3. The ultimate aim of HEADLAMP and its environmental health indicators is to obtain information on which to base preventive action aimed at environmental health problems. HEADLAMP is intended to be an on-going activity, focusing on information needs at local and national level. As such, it is designed to indicate environmental health trends, and to enable policy-makers and managers to assess the value and performance of their policies over time. National and local capacity-building is, therefore, also an integral part of the HEADLAMP approach.

Based on these characteristics, a framework for implementing HEADLAMP in the field has been developed and is summarized in Figure 1.2. Application of HEADLAMP methods is motivated by concern regarding specific environmental conditions and their potential adverse impact on human health. In practice, application of the HEADLAMP process follows three stages, reflecting the three characteristics described above.

The first stage of the process relates to the definition and validation of the problem. The known links between a defined environmental factor and its associated health outcomes provide the starting point. These links will already have been established in previous research and in the literature. Basic information requirements are identified at this stage.

The second stage involves the compilation, assessment and quantification of relevant environmental health indicators. During this stage, detailed data requirements are specified, taking account both of the specific setting in which the analysis is being conducted, and of the inevitable limitations of data availability. These data are obtained as far as possible from available routine data sources, but may be supplemented where necessary through the implementation of purposely-designed, rapid surveys. These data are then analysed to obtain information on environmental health effects or conditions. The variables produced through this process comprise the environmental health indicators. Depending on the problem and/or feasibility of obtaining all the relevant data, environmental health indicators may be derived from: health data (e.g. morbidity rates attributable to definable environmental factors),

environmental data (e.g. pollution levels with human health implications), or results of the linkage of environmental and health data (e.g. ecological studies).

The third stage comprises policy formulation/implementation. At this stage, relevant policy actions are defined and implemented on the basis of the trends and patterns shown by the environmental health indicators. In this context, the HEADLAMP process needs to be seen not as a one-off activity, but as part of a continuing cycle of assessment and action. Thus, repeated assessments may be undertaken at appropriate intervals in order to monitor changes in health and/or environmental status and to ascertain if any particular trend has been established. Thus, repeated assessment would contribute to the monitoring of the effects of policy implementation, provide support for changes in policy, and convey environmental health information to the public and other stakeholders. A decision to cease monitoring activities may be taken once pre-set targets have been met on a sustained basis.

Application of the HEADLAMP approach is aimed at improving protection against environmentally related disease and the promotion of a healthy environment. This is consistent both with the principles of sustainable development, as presented by the UNCED's Agenda 21, and the ideals in the Alma-Ata declaration of "Health for All". Agenda 21 recognizes that both insufficient and inappropriate development can result in severe environmental health problems. Thus, while development cannot occur without a healthy population, such development should in turn not create additional environmental health problems (UN, 1993). "Health for All" ideals of equity in health are also closely linked to environmentally related health problems. These explicitly recognize that some sectors of the population are adversely affected both by the characteristics of the environment in which they live, and through their access to health services. The implementation of HEADLAMP activities at the local level will therefore complement and support the endeavours already being taken within these initiatives. If effective environmental health decision-making and actions can be sustained and multiplied in many local situations, they will have a significant impact at the national and global levels.

REFERENCES

Beaglehole R, Bonita R, Kjellström T. *Basic Epidemiology*. Geneva: World Health Organization, 1993.

Chen BH, Hong CJ, Pandley MR, Smith KR. Indoor air pollution in developing countries. *World Health Statistics Quarterly* 1990;3:127-38.

FAO/WHO. Summary of acceptances: worldwide and regional Codex standards. *Codex Alimentarius* part 1, rev.4, 1989.

Martinez J, Phillips M, Feachem RGA. Diarrheal diseases. In: Jamison DT, Mosley WH, Measham AR, Bobadilla JL, editors. *Disease Control Priorities in Developing Countries*. Oxford: Oxford University Press, 1993.

UN. Agenda 21. *The United Nations Programme of Action from Rio*. New York: UN, 1993.

UNEP/WHO. GEMS/AIR — A global programme for urban air quality monitoring and assessment. Doc. No. WHO/PEP 93.7; UNEP/GEMS/93.A.1. UNEP, Nairobi, 1993.

WHO. *Environmental Health Criteria* No.1, Mercury. Geneva: WHO, 1976.

WHO. Recommended health-based limits in occupational exposure to heavy metals. *Technical Report Series* 647. Geneva: WHO, 1980.

WHO. Recommended health-based limits in occupational exposure to selected organic solvents. *Technical Report Series* 664. Geneva: WHO, 1981.

WHO. Recommended health-based limits in occupational exposure to pesticides. *Technical Report Series* 677. Geneva: WHO, 1982.

WHO. Recommended health-based occupational exposure limits for selected vegetable dusts. *Technical Report Series* 684. Geneva: WHO, 1983.

WHO. Recommended health-based occupational exposure limits for respiratory irritants. *Technical Report Series* 707. Geneva: WHO, 1984.

WHO. Recommended health-based limits in occupational exposure to selected mineral dusts. *Technical Report Series* 734. Geneva: WHO, 1986.

WHO. *Air Quality Guidelines for Europe.* WHO regional publications, European series No.23, Copenhagen: WHO, 1987.

WHO. Joint UNEP/FAO/WHO food contamination monitoring programme. Doc. No. WHO/EHE/FOS/90.2. Geneva: WHO, 1990.

WHO. GEMS/Water 1990-2000: The challenge ahead. Doc. No. WHO/PEP/91.2. Geneva: WHO, 1991.

WHO. *Our Planet Our Health.* Report of the WHO Commission on Health and Environment. Geneva: WHO, 1992a.

WHO. *Indoor Air Pollution from Biomass Fuel.* Report of a WHO consultation. Doc. No. WHO/PEP/92.3A. Geneva: WHO, 1992b.

WHO. *Environmental Health Criteria* No. 134, Cadmium. Geneva: WHO, 1992c.

WHO. *Guidelines for Drinking-Water Quality.* Vol. 1, Recommendations. Geneva: WHO, 1993(a).

WHO "Contaminated food: a major cause of diarrhoea and associated malnutrition among infants and young children". In: *Facts about Infant Feeding*, No. 3, April 1993(b).

WHO. *International Programme on the Health Effects of the Chernobyl Accident* (IPHECA). (Report of the management committee meeting, Geneva 16-17 March 1994.) Doc No. WHO/EOS/94.24. Geneva: WHO, 1994.

Development of Environmental Health Indicators

C. Corvalán[a], *D. Briggs*[b] *& T. Kjellström*[a]

1. INTRODUCTION

1.1 The Need for Information

The need for information to support environmental health is evident. In many parts of the developing world, the environmental health burden is increasing; even in the developed world, new pollutants are emerging which pose new threats to human health. Against this background, there is clearly an urgent need for action to reduce the environmental health burden, for example through:

- education and awareness raising to help individuals better appreciate the environmental risks to which they are exposed, and the personal opportunities which exist for risk avoidance and reduction
- environmental improvement to reduce the hazards involved, especially in those areas where human exposure may occur
- technological innovation, to develop new, cleaner and more sustainable methods of production
- demand control, to reduce the pressures from consumption and resource use

All of these actions are potentially costly. Many are also likely to be relatively long-term in their effect. All, therefore, require the availability of reliable information. Information is needed to help identify and prioritise the problems which exist; to help specify safe limits and environmental guidelines and standards; to define, evaluate and compare the actions which might be taken; to monitor the effects of these actions; to inform the numerous groups of stakeholders involved; to provide a rational framework for discussion and debate; and to guide the research and development needed for the future.

---

[a]Office of Global and Integrated Environmental Health, World Health Organization, Geneva

[b]Nene Centre for Research, Northampton, United Kingdom

1.2 The Problem of DRIPS

Over recent years the potential to gather data on both environment and health has greatly improved. In the area of environment, for example, earth-observing satellites produce an unprecedented volume of data on a regular basis (Simonett 1988). The establishment of global monitoring programmes such as the GEMS network (WHO, 1990; 1991; UNEP/WHO, 1993) and trans-national information systems such as the CORINE system (Briggs, 1995a) is helping to provide extensive data sets on the environment. Improvements in field monitoring techniques, modelling and computing have also greatly increased the supply of environmental data at both the local and regional level. Technical advances in the area of health data have perhaps been less dramatic, but the establishment of more sophisticated computer-based health reporting systems has helped to increase the supply of health data in many countries. For the future, use of geographic information systems (GIS) and other information technology — as in the HEGIS programme (WHO, 1993b) — offers the potential to bring together these health data at both the national and international level.

Contradictorily, the difficulties of using these growing volumes of data in decision-making have often increased. We are now faced with a problem which some observers have termed DRIPS — Data-Rich Information-Poor Syndrome. How can we keep track of this growing mass of data? How can we find and select the specific data we need — and how can we do that quickly, when we need them? How can we condense the huge quantities of data down to more manageable amounts? How can we extract meaning from the tangle of data before us? In many cases, therefore, the increased availability of data has been a quantitative rather than a qualitative improvement. It has given us a larger data resource, but often without the tools to exploit it. At the same time, the increased availability of data has had a self-amplifying effect, raising both the expectations of and demand for additional data. Having data for one period, for example, tends to beg the question of whether or not changes are occurring over time. Having data for one area creates pressures for other areas to collect similar data for comparative purposes. Above all, as we learn more about the world from the data we collect, we realise its complexity — and discover that we always need more data to describe it adequately.

Providing relevant information, in a form which all those involved can understand and accept, within the constraints of time and other resources, is thus a major challenge. It is not just a matter of collecting data. It requires the *selection* of information which is directly relevant to the task at hand. It requires the *translation* of this information into a consistent and coherent form. It requires the *presentation* of the information in a manner which is appropriate and acceptable to the different users.

Providing information in a form useful for decision makers requires the *selection* of relevant information, the *translation* of this information into a consistent and coherent form, and the *presentation* of the information in an accessible and acceptable manner.

1.3 The Role of Indicators

Environmental health indicators represent one means of meeting these needs. The term 'indicator' is derived from the Latin *indicare*, meaning to announce, point out or indicate. As such, indicators represent more than the raw data on which they are based. They provide a means of giving the data added value by converting them into information of direct use to the decision-maker. Indicators are thus a crucial link in the decision-making chain. Measurements produce raw data; data are aggregated and summarised to provide statistics; statistics are analyzed and re-expressed in the form of indicators; indicators are then fed into the decision-making process (Figure 2.1).

Within this context, an *environmental health indicator* can be seen as a measure which summarises in easily understandable and relevant terms some aspect of the relationship between the environment and health. It is a way, in other words, of expressing scientific knowledge about the linkage between environment and health in a form which can help decision-makers to make more informed and more appropriate choices.

Environmental health indicators can thus contribute to improved health management and policy everywhere. They are, however, of particular value in countries in which traditional problems of access to natural resources remain, and in which issues of environmental pollution have traditionally taken second place to demands for economic development. In many countries, indeed, problems of resource depletion, desertification and environmental pollution are rising while populations are undergoing rapid expansion. In recent years, awareness has been growing of the association between economic growth and environmental protection, and, in many countries, strategies for sustainable development which both preserve the environment and enhance quality of life are being implemented. If decision-makers are to take the actions needed to prevent irreversible and costly health and environmental damage, they urgently need reliable and relevant information on levels of environmental pollution and their links with human health.

In recent years much has been written about indicators, in many different fields of policy and management. Much effort has also gone into the construction of indicators for policy support, notably in the area of environmental policy. There is danger that excessive expectations have been raised in the process. Indicators

are not panaceas. They cannot of themselves solve problems. Nor can they avoid the need for difficult choices and decisions. But — if well-designed and sensibly used — they can play an important part in supporting decision-making. At the very least, they can give a common currency or language and a means of information exchange between the many stakeholders concerned. They can help to quantify the situation and emphasise its significance. They can help to simplify the information and present it in a form directly relevant to the question being addressed. They can highlight the trends or the questions involved and point to possible responses and solutions. Indicators thus help to lead the decision-maker towards the choices available, and to evaluate and compare the implications of these choices. Equally, they can provide a means of public information, and an opportunity for external scrutiny of decisions and policies.

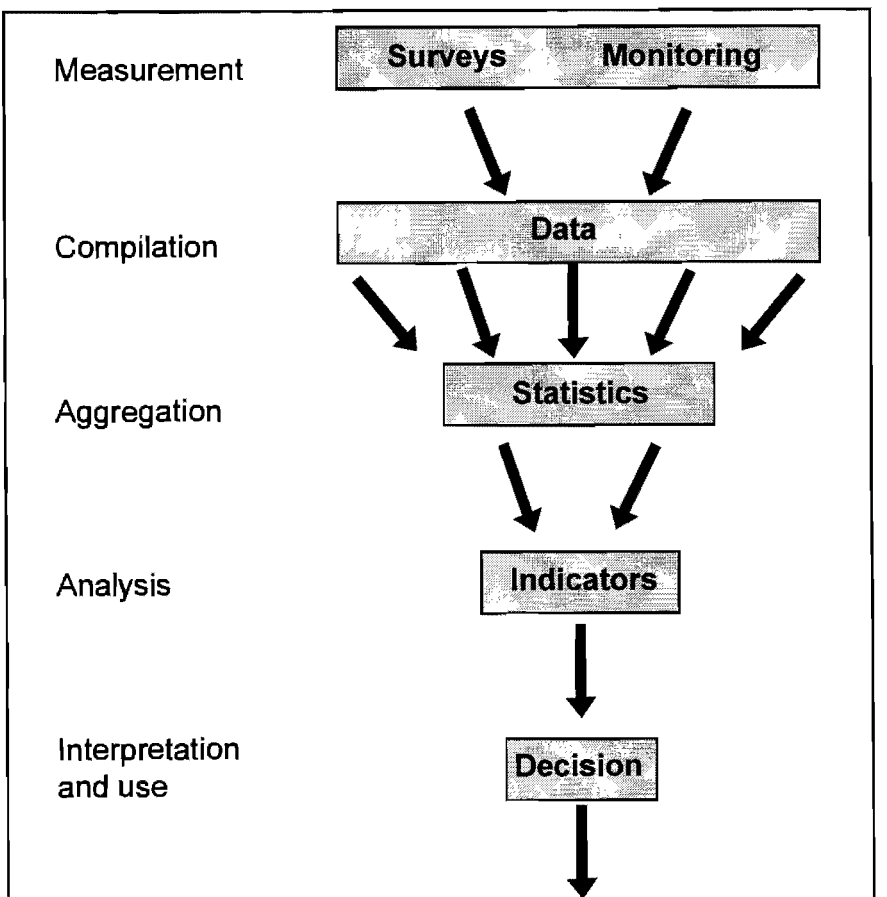


Fig. 2.1 The place of indicators in the decision-making chain

Indicators therefore have a major role to play in relation to the protection and management of environmental health, but if they are to be used both effectively and with validity, it is important that they are properly conceived and understood. This Chapter provides an introduction to the application of environmental health indicators for decision support. It analyses the concepts of environmental health indicators, describes some of the key issues and questions involved in their development, and illustrates their use.

2. BACKGROUND

The concept of indicators is far from new. The use of indicators has a long history, for example, in economics (e.g. indicators such as GDP and the unemployment rate), resource management (e.g. indicators of land suitability) and ecology (e.g. the use of "indicator species"). In recent years, however, there has been a marked growth in interest in the use of indicators in many other fields. The use of social indicators (e.g. of deprivation, poverty) is now widely accepted, while performance indicators are increasingly being used to monitor the activities of industry and the public services. Indicators have also become well-established in the fields of both environment and health.

2.1 Sustainable Development and Agenda 21

One of the most important stimuli for indicator development in the areas of environment and health has been the emergence of sustainable development as a guiding principle for policy, and the adoption in 1992 of Agenda 21 at the United Nations Conference on Environment and Development (UNCED).

Sustainable development has been defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission on Environment and Development, 1987). As such, sustainable development is both inherently concerned with the link between environment and human health and reliant on the provision of good information.

Developments which jeopardise human health — whether through pollution or resource depletion — are clearly not sustainable. Principle 1 of the Rio Declaration, for example, clearly stated the case:

> *"Human beings are at the centre of concerns for sustainable development. They are entitled to a healthy and productive life in harmony with nature."* (UN, 1993)

Chapter 6 of Agenda 21 takes this principle further by emphasising the fundamental commitment within sustainable development of "protecting and promoting human health".

Against this background, indicators are clearly needed which describe and monitor progress towards sustainability. Chapter 40 of Agenda 21 "Information for Decision-making" (UN, 1993), for example, stated that:

> *"Indicators of sustainable development need to be developed to provide solid bases for decision-making at all levels and to contribute to a self-regulating sustainability of integrated environment and development systems".*

Countries and international governmental and non-governmental organizations were called upon to develop the concept of indicators of sustainable development. The Statistical Office of the United Nations was given a special role to support this work and to promote the increasing use of such indicators. National programmes for indicator development have thus been set up in many countries to support environmental policy and state of environment reporting (e.g. Environment Canada, 1991; Adriaanse, 1993). The adoption of Local Agenda 21 has similarly encouraged the establishment of sustainability indicators by local governments and city authorities (e.g. Local Government Management Board, 1994; Sustainable Seattle, 1993). Internationally, organizations such as OECD (1993), UNEP/RIVM (1994), the World Resources Institute (1995), the World Bank (1994) and the Worldwide Fund for Nature and New Economics Foundation (1995) have attempted to construct core sets of indicators to monitor global environmental trends.

2.2 The Development of Environmental Health Indicators

Major developments in the construction and use of indicators have also occurred in the area of environmental health. The initiative for indicator development has been taken to some extent by national or local agencies (e.g. Alexander, 1994). The Scientific Committee on Problems of the Environment (SCOPE) also recommended that sustainability indicators should include measures of human impact and exposure (SCOPE, 1995). WHO has taken a leading role in developing both the concept and use of environmental health indicators. Ever since the WHO Programme for the Promotion of Environmental Health was established almost 50 years ago, the development of methods, and practical applications of the measurement of environmental health status, have been important concerns. The initial priority was to provide information on basic issues of drinking water, sanitation and shelter — and even today two of the most widely used indicators of environmental health status in a community are the percentages of a population that have access to drinking water and sanitation.

Since then, the emergence of "environmental epidemiology" (WHO, 1983) has helped to focus attention on the more complex linkages between environment and health, and has stimulated the need for more sophisticated and scientifically validated environmental health indicators.

In recent years, therefore, many WHO programmes have become involved in the establishment and use of environmental health indicators. Notable examples include the global and European Health for All programmes and the Healthy Cities programme (WHO, 1992b; 1995). Against this background, the first WHO meeting dealing specifically with environmental health indicators was held in Dusseldorf in 1992 (WHO, 1993a) and since then a number of consultations have been held to develop indicators (WHO, 1993b; 1994b). Based on these initiatives, several national and regional programmes have been established to construct environmental health indicators.

2.3 Definitions

An environmental health indicator may be defined as:

> *An expression of the link between environment and health, targetted at an issue of specific policy or management concern and presented in a form which facilitates interpretation for effective decision-making.*

Several aspects of this definition are worthy of emphasis. The first is that an environmental health indicator embodies a linkage between the environment and health. As such it is more than either an environmental indicator or a health indicator. *Environmental indicators* represent indicators which describe the environment without any explicit or direct implications for health. The vast majority of environmental indicators so far developed are of this type — for example, indicators of atmospheric emissions, surface water quality, designated areas or threatened wildlife species. *Health indicators* are indicators which describe the status of, or trends in health without any direct reference to the environment. Again, the majority of health indicators so far developed are of this type; examples include simple measures of life expectancy, or cause-specific mortality rates where no attempt has been made to estimate those health outcomes attributable to the environment.

Given knowledge of the relationship between environmental exposure and health effect, however, both environmental indicators and health indicators can be converted into *environmental health indicators*. An environmental health indicator is thus a measure which indicates the health outcome due to exposure to an environmental hazard. As such, it is based upon the application of a

known or postulated *environmental-exposure health-effect relationship*. In this context, two general types of can be distinguished:

- An *exposure-based indicator* projects forward from some knowledge about an environmental hazard to give an estimated measure of risk. Such indicators can be conceived as the combination of an environmental indicator with a known environment-health relationship. An example might be the imputed cases of cholera due to lack of access to safe water supplies, or the imputed cases of respiratory disease associated with exposure to traffic-related air pollution.

- An *effect-based indicator* projects backwards from the health outcome to give an indication of the environmental cause (i.e. the environmentally attributable health outcome). Examples in this case might be the proportion of the diarrhoea death rate due to water-borne infections, or the rate of childhood leukaemias attributable to exposure to environmental radiation.

The importance of this environment-health relationship within the concept of environmental health indicators cannot be over-emphasised. It is only through knowledge of this link that an environmental indicator or a health indicator can be translated into an environmental health indicator. An environmental health indicator is thus an environmental indicator *or* a health indicator *plus* a known environment-health relationship.

Another important characteristic of an environmental health indicator is its relationship with policy or management. Any environmental health indicator must have utility. To be useful, it must relate to aspects of environmental health which are both of relevance to the decision-maker and amenable to control. Equally, it must be expressed in a way which is pertinent to, and understandable by, the decision-makers concerned. In many circumstances, this requires that the indicator be expressed in terms of the health risk associated with a specific environmental hazard.

> An environmental health indicator is an environmental indicator *or* a health indicator *plus* a known *environmental-exposure health-effect relationship*.

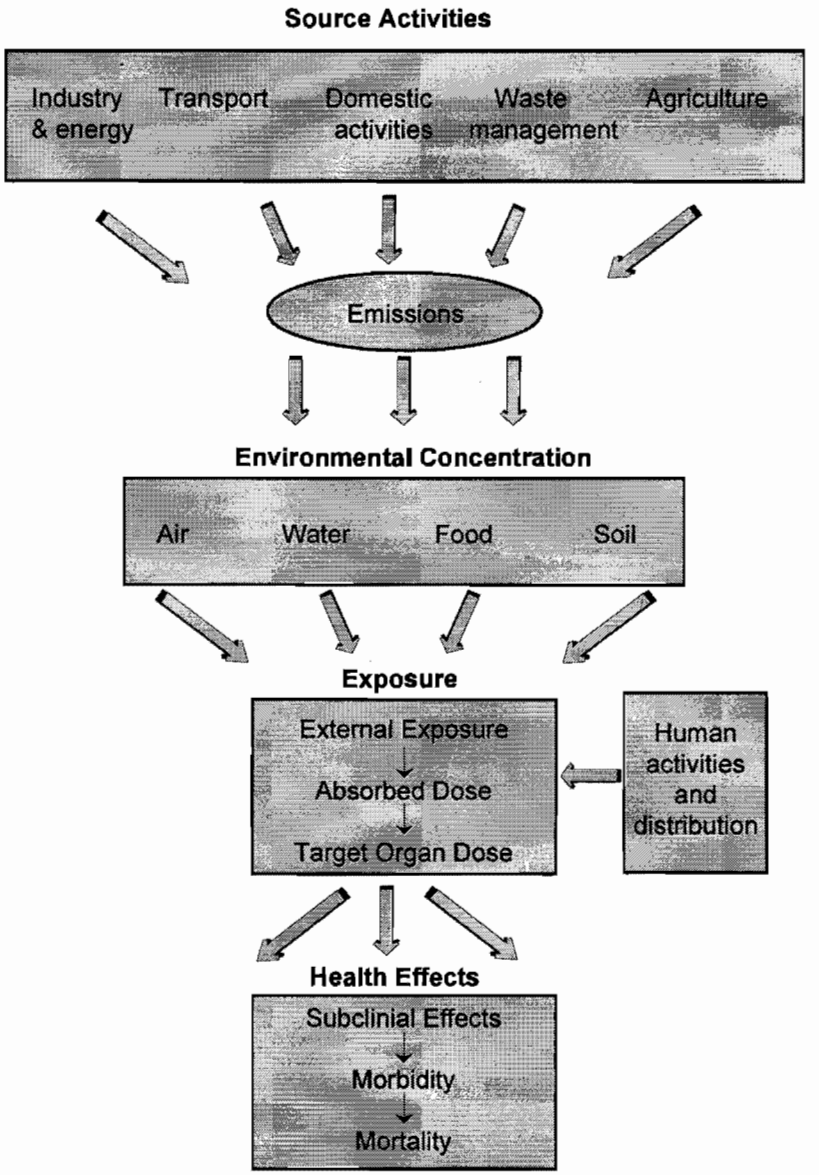


Fig. 2.2 The environment-health chain: the example of pollution

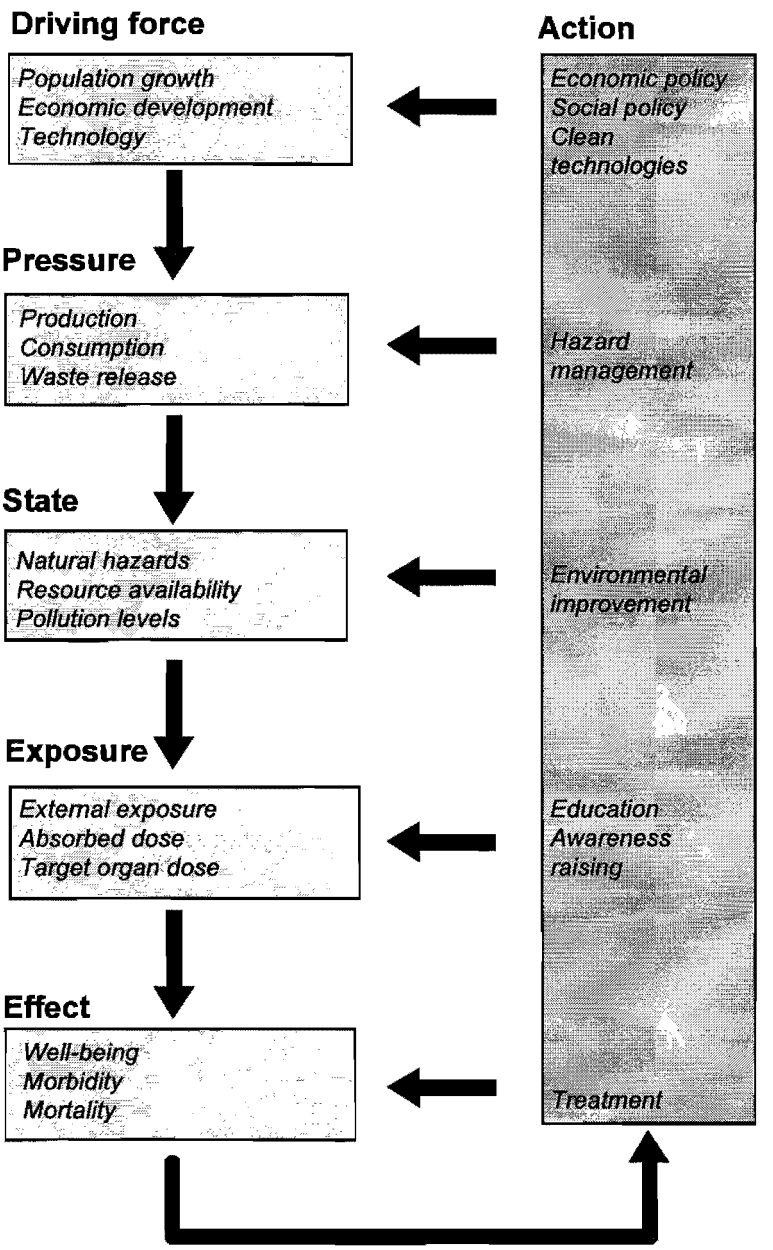


Fig 2.3 DPSEEA: a conceptual framework for the development of environmental health indicators

3. A CONCEPTUAL FRAMEWORK FOR ENVIRONMENTAL HEALTH INDICATORS

3.1 The Environment-Health Chain

As the preceding discussion has indicated, the link between environment and health operates through the exposure of humans to environmental hazards. These hazards may take many forms — some are wholly natural in origin; the majority derive from human activities and interventions. In all cases, however, health effects only arise if humans are exposed — often at a specific place and time — to the hazards which exist.

The environment-health chain is most clearly seen in the case of exposure to pollution (Figure 2.2). Most environmental pollutants are the product of human activities. These may be released into the environment in a variety of ways, and may then be dispersed and accumulate in different environmental media (e.g. the air, water, soil, food). Exposure occurs when humans encounter the contaminants within any one of these media. A range of health effects may then occur, from minor sub-clinical effects through illness to death, depending upon the intrinsic harmfulness of the pollutant, the severity of exposure and the susceptibility of the individuals concerned. The whole process is often driven by persistent forces which both motivate the creation of the hazard and increase the likelihood of exposure. Thus, population growth, economic development, technological change and — behind these — social organisation and policies may all lie at the root of the problem. Ultimately, it is often to these that action needs to be addressed.

3.2 The DPSEEA Framework

The environment-health chain illustrated by the example of pollution provides a useful organising framework for the development and use of environmental health indicators. To make it more generally applicable (e.g. to other forms of environmental hazards), and to set it more firmly within a decision-making context, it needs to be further conceptualised.

Over recent years, a number of attempts have been made to devise conceptual frameworks for indicator development. Of these, the one which has been most widely adopted has been the simple pressure-state-response (PSR) sequence, initially applied by OECD as a framework for state-of-the-environment reporting. In many ways, however, the PSR sequence has proved too limiting, and it has more recently been extended to include recognition of both the "driving forces" responsible for pressures on the environment, and of the effects which often precede the policy response (e.g. USEPA, 1994). Figure 2.3 further adapts these concepts to provide a specific framework — the DPSEEA framework — for the development of environmmental health indicators. Examples of indicators based on this structure are given in Table 2.1. A more elaborated example which includes "action" indicators at each step is given in Table 2.2.

Table 2.1 Examples of environmental health indicators within the DPSEEA framework

		Pollutant type		
Stage	***Process***	***Chemical*** (e.g. child lead exposure)	***Physical*** (e.g. ionising radiation)	***Microbiological*** (e.g. water contamination)
Driving Force	*Type of development or human activities*	Use of lead as a petrol additive	Shifts to use of nuclear energy generation	Population growth in areas of poor sewage treatment
Pressure	*Source activity*	Consumption of leaded petrol	Amount of radioactive material used	Amount of untreated waste produced
	Emissions	Tonnes of lead emitted from cars	Calculated emissions at nuclear facilities	Amount of untreated effluent
State	*Environmental levels*	Lead concentration in air	Radiation levels in air, water, food	Coliforms in water, food
Exposure	*Human exposure*	Calculated personal exposure to lead from all sources	Calculated exposures: workers; nearby residents	Estimated exposure to contaminated food/water
	Dose	Lead in blood	Personal dosimeters; Urine; Faeces	Serum analysis for hepatitis A and typhoid; faeces for cholera, shigella
Effects	*Early/ sub-clinical*	Behavioural disorders; Reduced IQ in children	Chromosomal abnormalities	Diarrhoea, fever, nausea
	Moderate /clinical	Anaemia; Increase in blood pressure	Genetic defects; Leukaemia; Cancer	Cholera, hepatitis A, typhoid, dysentery, gastroenteritis
	Advanced /permanent	Encephalopathy; Acute lead poisoning	Acute radiation sickness; Cancer	Death from dehydration

Table 2.2 Example of Occupational Lead Exposure

Stage	*Process*	*Descriptive Indicator*	*Action Indicator*
Driving Force	*Type of development or human activities*	Industrial/Occupational use of lead; Mining of lead	Technological innovation affecting use of lead; Education about hazards of lead
Pressure	*Source activities*	Specific uses of lead ; Tons of lead consumed (quantity produced and recycled)	Trends in lead use profile; Trends in quantity of lead used; Substitution for lead
	Emissions	Contamination of occupational and para-occupational environment	Availability and use of control technology
State	*Environmental levels*	Airborne lead concentrations; Lead dust concentrations (work and home)	Trends in ambient air and dust monitoring
Exposure	*Human exposure*	Blood lead; Blood zinc protoporphyrin (ZPP) Personal air sampling	Surveillance of blood lead and ZPP; Trends in personal air monitoring; Education about hazards of lead
	Dose	Blood lead; Bone lead (research tool)	Trends in blood lead (e.g. government registries)
Effects	*Early/ subclinical*	Deranged haem synthesis; Non-specific CNS symptoms; Abnormal nerve conduction velocity	Application of special surveys based in the workplace
	Moderate/ clinical	Abdominal and constitutional symptoms; Anaemia; Decreased renal function	Routine medical surveillance, employment-based
	Advanced/ permanent	Renal failure; Peripheral neuropathy; Encephalopathy	Periodic analyses of major morbidity and mortality; Clinical interventions

3.3 Driving Forces

Within this framework, the *driving forces* component (D) refers to the factors which motivate and push the environmental processes involved. One of the most important of these is population growth. Almost inevitably this results in more people being exposed to environmental hazards simply by virtue of the increased number of people living in the areas concerned. More indirectly, it tends to lead to the intensification of human activities within these areas, thereby contributing to environmental damage and resource depletion. In some cases, also, it results in expansion of human populations into more marginal zones. Here, the inherent instability of the environment may mean that the population is especially vulnerable to environmental hazards, while the environment in turn is especially sensitive to damage.

A wide range of other important driving forces also exist, including technological development, economic development and policy intervention. Historically, for example, advances in technology have greatly increased the capability of humans to modify their environment. While these advances have often provided for major improvements in human well-being and health in the short-term, in the longer term they have frequently resulted in accelerated rates of resource depletion and environmental pollution. Economic development — which itself both fuels and is fuelled by technological development and population growth — has similarly contributed to environmental contamination and instability by creating increased rates of consumption and attendant problems of waste and pollution.

3.4 Pressures

The driving forces within the DPSEEA model result in the generation of *pressures* (P) on the environment. These pressures are normally expressed through human occupation or exploitation of the environment. Pressures are thus generated by all sectors of economic activity, including mining and quarrying, energy production, manufacturing, service industries, transport, tourism, agriculture and forestry. In each case, pressures arise at all stages in the supply chain — from initial resource extraction, through processing and distribution, to final consumption and waste release.

One of the most important components of these pressures in the context of human health is clearly the release of pollutants into the environment. These releases may occur in a wide variety of ways, and into different environmental media. Energy combustion — for example in vehicles, manufacturing industry, electricity generation and home heating — is one of the most important emission processes, especially to the air. Large quantities of pollutants are also emitted through other processes, however, such as spillage of chemicals, the deliberate

discharge of effluents, dumping of wastes, leakage and seepage from equipment or storage sites, leaching of agricultural chemicals, gaseous release of volatile materials (e.g. in petrol stations or from landfill sites), wear and tear of equipment (e.g. industrial machinery, vehicle tyres and brakes), and respiration and excretion by agricultural livestock. Because these activities and processes represent the starting point for environmental emissions they also represent the most effective point of prevention and control. Once in the environment, pollutants may undergo a wide range of secondary transfers. Much environmental policy is therefore focused at trying to regulate source activities, or to incorporate into them methods of emission control.

3.5 State

In response to these pressures, the *state* of the environment (S) is often modified. The changes involved may be complex and far-reaching, affecting almost all aspects of the environment and all environmental media. Thus changes occur in the frequency or magnitude of natural hazards (e.g. in flood recurrence intervals or in rates of soil erosion); in the availability and quality of natural resources (e.g. soil fertility, biodiversity); and in levels of environmental pollution (e.g. air quality, water quality). These changes in the state of the environment also operate at markedly different geographic scales. Many changes are intense and localised, and often concentrated close to the source of pressure (e.g. habitat loss, urban air pollution, contamination of local water supplies). Many others are more widespread, contributing to regional and global environmental change (e.g. desertification, marine pollution, climate change). Because of the complex interactions which characterise the environment, almost all these changes have far-reaching secondary effects: modifications of one area or one component of the environment feeds through to affect others.

3.6 Exposure

Environmental hazards, however, only pose risks to human well-being when humans are involved. *Exposure* (E_1) thus refers to the intersection between people and the hazards inherent in the environment. Exposure is rarely an automatic consequence of the existence of a hazard: it requires that people are present both at the place and at the time that the hazard occurs.

The concept of exposure is best developed in relation to pollution. Exposure to pollution can occur in a number of different ways — by inhalation, ingestion or dermal absorption — and may involve a wide range of different organs. *External exposure* refers to the quantity of the pollutant at the interface between the recipient and the environment. It is often measured either using some form of personal monitor (e.g. passive sampling tubes for air pollution) or by

modelling techniques (e.g. based upon knowledge of concentrations in the ambient environment). The amount of any given pollutant that is absorbed is often termed the *absorbed dose*, and may be dependent on the duration and intensity of the exposure. *Target organ dose* refers specifically to the amount that reaches the human organ where the relevant effects can occur.

Recent advances in molecular biology and analytic chemistry facilitate development of *biomarkers* as measures of dose. The aim is to characterize absorbed toxins in their most biologically relevant form. After absorption, transport, and metabolism, exogenous toxins initiate biological changes. Identifying and quantifying the transformed toxin as close as possible to this point of initiation enhances accuracy in knowing the true biologically active dose. Biomarkers may also be used to characterize the early biological alterations, thereby increasing the fit between toxin and response. DNA- and protein-adducts have undergone recent intensive study as biomarkers in the study of cancer caused by environmental factors.

3.7 Effects

Exposure to environmental hazards, in turn, leads to a wide range of health *effects* (E_2). These may vary in type, intensity and magnitude depending upon the type of hazard to which people have been exposed, the level of exposure and the number of people involved. For convenience, a simple spectrum of effects can often be recognised. The earliest, and least intense, effects are sub-clinical, merely involving some reduction in function or some loss of well-being. More intense effects may take the form of illness or morbidity. Under the most extreme conditions, the result is death.

In this context the concept of *sentinel diseases* is of particular interest. Some diseases are very specific to environmental (or occupational) exposures; for example, asbestosis and mesotheliomas as indicators of asbestos exposure, silicosis as an indicator of exposure to silica dust, or leukaemia as an indicator of exposure to ionizing radiation. In practice, however, there are few diseases which can be used as sentinels of environmental exposures.

The concept of *risk analysis* is in many cases of greater utility. Given known exposures and knowledge of dose-response functions, it is possible to make reasonable estimates of the health burden of specific pollutants. The further elaboration of risk analysis methods will be an important contribution to the development of indicators, by providing inexpensive and rapid estimates of the health impact of specific environemtal exposures at the aggregate level.

3.8 Actions

In the face of these effects, society typically attempts to invoke a range of *actions* (A). These may take many forms and be targetted at different points within the environment-health chain. In the short term, actions are often primarily remedial (e.g. the treatment of affected individuals). In the longer term, actions may be protective (e.g. by trying to change individual behaviour and lifestyle to prevent exposure). Alternatively, actions may be taken to reduce or control the hazards concerned (e.g. by limiting emissions of pollutants or introducing measures of flood control). Perhaps the most effective long-term actions, however, are those that are preventive in approach — aimed at eliminating or reducing the forces which drive the system.

The most effective long-term actions are those which are preventive in approach — aimed at eliminating or reducing the forces which drive the system.

4. ISSUES IN INDICATOR CONSTRUCTION AND DESIGN

4.1 Criteria for Indicator Development

As the preceding discussion suggests, while indicators are intended to provide a simplification of reality, they are themselves far from simple. Unfortunately, this underlying complexity has not always been appreciated by those who have called for indicator development, and as a result indicators have often been considered in relatively fuzzy and ill-defined terms. Often, it seems, there is a tendency to confuse indicators with the general issue or theme to which they relate. The consequence of this is likely to be the development of poorly conceived or inadequate indicators. These pose a double jeopardy. They are likely to be a waste of time and effort, and they are likely to misinform, rather than inform, the users.

Good indicators, in contrast, require careful planning and design. They depend upon an understanding of the questions being addressed, of the way in which they will be used, and of the way in which the systems involved operate. In addition, they need to be very precisely formulated and often need to be tested before they can be used.

Fortunately, in recent years much has been learned about the development and use of indicators in a wide range of decision-making areas. On the basis of this

experience, a number of criteria have now been established for indicator selection and construction (e.g. Kreisel, 1984; UNEP/RIVM, 1994). These can be further adapted in relation to environmental health (Table 2.3). Not all these criteria can necessarily be achieved in all circumstances; problems of data availability, resources and the need for compatibility with previous indicator series may mean that some have to be sacrificed. In general, however, it may be argued that environmental health indicators should meet the majority of the conditions listed in Table 2.3 — and should certainly satisfy the first four criteria.

Table 2.3 Criteria for Environmental Health Indicators

1.	be based on a known linkage between environment and health
2.	be sensitive to changes in the conditions of interest
3.	be directly related to a specific question of environmental health concern
4.	be related to environmental and/or health conditions which are amenable to action
5.	be consistent and comparable over time and space
6.	be robust and unaffected by minor changes in methodology/scale used for their construction.
7.	be unbiased and representative of the conditions of concern
8.	be scientifically credible, so that they cannot be easily challenged in terms of their reliability or validity
9.	be easily understood and applicable by potential users
10.	be available soon after the event or period to which it relates (so that policy decisions are not delayed)
11.	be based on data which are available at an acceptable cost-benefit ratio
12.	be based on data of a known and acceptable quality
13.	be selective, so that they help to prioritise key issues in need of action
14.	be acceptable to the stakeholders

4.2 Position in the DPSEEA Chain

One of the most fundamental questions in developing any environmental health indicator is clearly at what position within the DPSEEA chain it should be targetted. In terms of environmental epidemiology, the links within this chain which have often generated most attention are those between exposure and effect. In terms of health policy and management, however, it is often the earlier steps in the chain which are of most interest. Many environmental health problems ultimately derive from relatively remote causal forces and events. Immediate sources of exposure thus represent little more than symptoms of the

problem. Desertification, for example, is often a consequence of deeper-seated social and economic causes. Pollution, equally, is often a symptom of inadequacies in industrial technology and policy controls. If the aim is to identify the underlying cause of the problem, and to take effective action at source, it is therefore essential to have indicators which allow the effects on health to be traced back to their underlying sources and causes.

Indicators from higher up the DPSEEA chain also tend to provide a better early warning both of impending environmental problems and of the effects of intervention. Detectable changes in the state of the environment and in human health, for example, typically lag some way behind changes in source activity or emissions — in the case of some effects, such as cancers, often by many years. Most preventive action, similarly, occurs at or close to the source of the hazard (e.g. by controlling emissions at source or through hazard management). The consequences of these interventions are likely to be shown by pressure indicators before they are seen in changes in either state or effect indicators.

A further reason for relying on indicators from higher up the DPSEEA chain is the practical one of data availability. Typically, data become more difficult to acquire as one passes down the chain. Thus, while there are normally abundant data on social and economic conditions and trends, much less is known about the actual pressures on the environment, less still about environmental conditions and almost nothing about actual exposures. As a consequence, *proxy indicators* of exposure commonly have to be used, derived from higher up the DPSEEA chain.

The use of indicators from higher up the exposure chain — whether in their own right or as proxies — is not without its dangers. As noted earlier, to be effective any environmental health indicator must be based on a clear and firm relationship between the environmental hazard and the health effect. Unfortunately, the further removed the indicator is from the health effect, the weaker this link is liable to be, for the DPSEEA cycle is an imperfect one. Each link in the chain is itself dynamic and uncertain; each step is subject to a wide range of influences and controls. The extent to which the driving forces are translated into active pressures on the environment, for example, depends upon the policy context, social attitudes and the pre-existing economic infrastructure of the area concerned. Whether these pressures cause detectable changes in the environment depends upon the ability of the environment to absorb and damp down the changes involved. Whether the environmental hazards, in turn, lead to health effects is determined by all the factors that control exposure and human susceptibility to its effects. It depends, therefore, on the form, duration, intensity and timing of exposure; on the social, economic and prior health status of the individuals concerned; on the quality and accessibility of the health system. Equally, of course, there is no certainty that action will be taken in response to the existence of environmental health

problems. This depends not only on adequate recognition of the problems concerned, but also on political will, economic and technological capability and public acceptance of the actions involved. As a consequence, indicators from higher up the chain must invariably be used and interpreted with care.

4.3 Specific Versus Composite Indicators

Indicators represent an attempt to simplify the complexity of reality into an easily interpretable measure. In order to describe reality, however, a large number of different indicators may potentially be needed — relating, for example, to the many different hazards and health outcomes of interest. In using specific indicators of this type, therefore, there is the possibility that decision-makers will be confronted with a bewildering range of information, much of it apparently contradictory in the message it gives.

Because of this problem, there have been many attempts in recent years to develop more synoptic or composite indicators that condense a wide range of information on different (but related) phenomena into a single measure or index. An oft-quoted example of this is the Retail Prices Index, which is used to show trends in inflation based upon a "basket" of goods. A composite indicator of human development has similarly been developed by UNDP (1990). Other examples are the various indicators of deprivation which have been widely used in social sciences and epidemiology (e.g. Carstairs and Morris, 1991; Jarman, 1984; Townsend, 1987), and the composite indicators of environmental quality developed by Inhaber (1976) and Hope et al. (1991). Composite indicators of this type are already used, also, in measuring land suitability (e.g. FAO, 1976). Similarly, many countries use composite indicators of stream water quality, as well as air quality indices.

A corresponding case can be made for the development and use of composite environmental health indicators. Because of the need for an explicit linkage with health, it is unlikely that the sorts of general indicator of overall environmental quality which have so far been proposed are of much utility. On the other hand, it may be useful in some cases to construct compound indicators either of total exposure to a specific hazard (i.e. covering all media and exposure pathways) or of groups of hazards. Thus, instead of producing separate indicators for exposure to each air pollutant, it might be possible to derive a composite indicator of exposure, including all air pollutants of interest. Ostensibly, indicators of this type have a number of benefits. By reducing the volume of information, for example, they facilitate the decision-maker's task. Equally, by taking account of the various pollutants to which people may be simultaneously exposed, they offer scope to allow for additive and synergistic effects.

Nevertheless, composite indicators also have many dangers and disadvantages. One problem is that such indicators require more data; the indicator is thus more than ever susceptible to gaps or weaknesses in data availability. More importantly, the results of the indicator depend to a great extent upon how it is constructed, what variables are used and how these are weighted and combined. Where the different components of the indicator are measured in the same units, it is theoretically feasible to combine them by simple addition or averaging. For example, the total pollutant concentration in the air can be calculated by summing the concentrations (in parts per million or microgrammes per cubic metre) of all the pollutant species of interest. Such a process does not necessarily make sense, however, since it assumes that all pollutants are of equal importance. In some cases, composite indicators also need to be constructed from individual components which are measured on different scales, so that simple arithmetic manipulation is not feasible. In these circumstances, the indicators need to be compiled on the basis of a suitable model.

Commonly, this involves some form of weighted aggregation. Where the indicator is intended to provide an index of health risk, these weights might be chosen to reflect the known harmfulness of each hazard (e.g. toxicity of each pollutant), though even this is problematic since the different pollutants may have different health effects. Complex interactions may also occur between the various pollutants, so that the overall effect on health cannot simply be conceived as the sum of the various parts.

Another widely used approach is based on the principle of limiting factors. This assumes that the condition of interest is defined by the state of the worst (or least optimal) factor. For example the indicator, "percentage of people with access to safe drinking water", assumes that all pollutants of potential health concern are below specified limits; exceedance of these limits by any one pollutant would render the water 'unsafe'. Equally, we might conceive an indicator such as "number of days of clean air" to give a general measure of levels of urban air pollution. Again, if any of the pollutants of concern exceeded recommended limits, the air would be classified as not clean.

The choice of model for compiling composite indicators of this type is clearly crucial. Unless an accepted model exists by which to convert the various components to a common measurement scale (e.g. to comparable measures of risk), the construction of such indicators is clearly likely to be somewhat arbitrary and open to challenge. It may also be difficult to test or verify composite indicators, since they do not relate to specific, measurable conditions. For the same reasons, it is difficult to establish clear standards and guidelines for composite indicators of this type, and interpretation of composite indicators can consequently be problematic.

4.4 Steps in Indicator Development

The development of reliable and effective environmental health indicators is not a trivial task. Indicators must be matched to their purpose: they must address the problem of concern, at the appropriate point in the environment-health chain, and at appropriate geographical and temporal scales and resolution. Both the data and the computational methods and models needed to construct them must be available. They must be expressed and presented in an easily understandable and usable form. Moreover, if the results of indicators are to be more widely applicable, if the indicators themselves are to be accepted by the many stakeholders concerned (e.g. scientists, politicians, the public), and if we are to learn from our collective experience in developing and using indicators, it is important that all these issues of design are carefully documented and open to scrutiny.

A large number of questions therefore have to be faced in designing and using indicators. The details of these questions vary depending upon the particular character of the indicator and its intended use. Figure 2.4 summarises the sequence of steps commonly involved. As this indicates, the main steps are as follows:

1. Specification of the problem to be addressed (i.e. the use of the indicator) and the user(s) concerned. The purpose might be defined in various ways, depending upon the interests of the user: for example, in terms of a specific environmental hazard (e.g. ionising radiation), a specific health outcome (e.g. childhood leukaemia), a specific policy or action (e.g. food hygiene legislation) or an underlying driving force (e.g. population growth).

2. Specification of the environment-health relationship on which the indicator will be based. This is essential if a valid environmental health indicator is to be identified. This relationship may be expressed in more or less quantitative terms (e.g. as an explicit exposure-response relationship) or as a general tendency (e.g. poor sanitation leading to higher rates of infection).

3. Specification of the point in the DPSEEA framework at which the indicator will be targetted. This will depend upon the particular interest and responsibilities of the user, but will also be influenced by the availability of relevant data and computational methods.

4. Specification of the parameter on which the indicator will be based — i.e. the particular measure of environment or health which will be used (e.g. atmospheric NO_2 concentration, cough and wheeze; or water quantity and infant diarrhoeal disease).

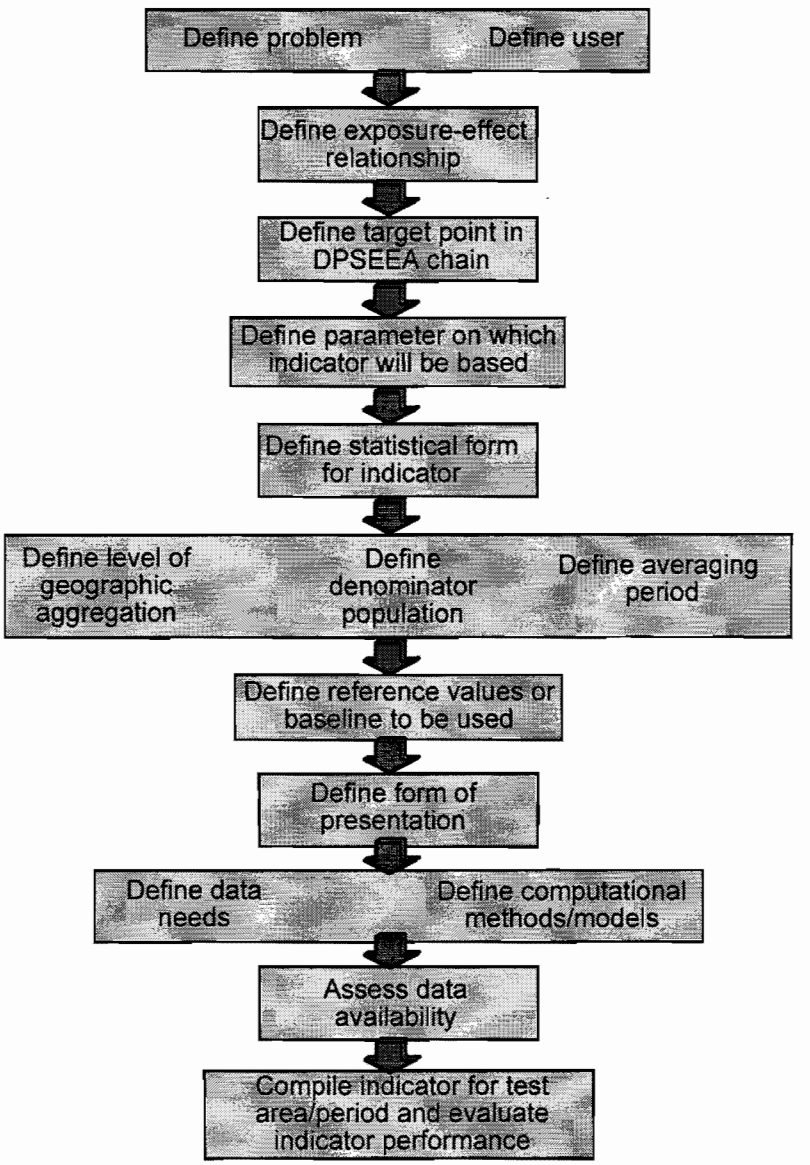


Fig. 2.4 Steps in the construction of environmental health indicators

5. Specification of the statistical form of the indicator. This step involves a number of considerations. Indicators can be presented in a variety of statistical forms: e.g. as simple frequencies or magnitudes (e.g. number of deaths), as rates (e.g. emission rates, mortality rates), as ratios (e.g. pollution level relative to the WHO guideline level, standardised mortality ratio), as measures of rate change (e.g. rate of population growth, rate of reduction in air pollution level), or in various more complex forms. The form chosen should reflect the purpose of the indicator.

6. Specification of the denominators and levels of aggregation required for the indicator (e.g. the level of geographic aggreggation, denominator population, averaging period).

7. Specification of the baseline or reference data against which the indicator will be standardised. This will need to reflect the statistical form of the indicator and the level of geographic aggregation, etc.

8. Specification of the form in which the indicator will be presented (e.g. graphically, as a map, as a simple statistic).

9. Specification of the data needs and models or methods required to compute the indicator.

10. Assessment of data availability and quality in the light of the foregoing specifications. At this stage, if relevant data are unavailable, it may be necessary to reconsider the indicator design (e.g. by choosing a proxy or by using a different level of aggregation).

11. Computation and testing of the indicator in a pilot area. This is a crucial step in order to determine whether the indicator is sensitive to the variations in the conditions of interest, whether the computational methods are sufficiently robust and the data adequate, and whether the results of the indicator are interpretable.

For the sake of clarity, these are presented here as a simple sequence. In reality, however, they are normally interactive and reiterative in form. Many of the questions of indicator design are inter-dependent, and need to be considered simultaneously. Many aspects of indicator design ultimately have to be amended in response to practical issues such as data availability. Until the indicator has been tested and used, it may not be certain that it operates effectively.

4.5 Quality Control

In that environmental health indicators contribute directly to decisions about human welfare and health, they inevitably carry a heavy burden of responsibility. Far-reaching and costly consequences can flow from their use. The validity of environmental health indicators is therefore of paramount concern. As the criteria in Table 2.3 emphasize, indicators must not only be scientifically credible and unbiased, but be based on accepted principles and open to scrutiny. The construction and use of environmental health indicators thus needs to include provisions for validation and quality control. These need to consider not only the way in which the indicator is designed but also the data, methods and models used in its application.

The issue of data quality is especially crucial in this respect, for many of the data used in constructing environmental health indicators are inherently unreliable. In the case of health data, for example, major problems typically include gaps and duplicates caused by errors in reporting, mistakes or inadequacies in georeferencing, differences in notification and referral procedures, discrepancies in diagnosis, misclassification of outcomes and sampling biases. Many of these problems are especially severe in relation to morbidity data, although where disease registers exist (e.g. cancer registers) or formal notification procedures apply (e.g. for many communicable diseases), data quality is often better.

In the case of environmental data, the problems are often even more serious. Most environmental data are derived from surveys or monitoring networks which provide data for only a small sample of points. Biases in sampling design therefore limit the representativeness of many of the data. Differences in sampling regime, measurement methods, classification procedures, reporting procedures and in the definition of the phenomena being measured all contribute further to uncertainties in the data. Many environmental data thus contain a major potential for error, and marked discrepancies may occur between data from different sources, making comparisons between different areas or studies difficult. Unfortunately, these errors are not always documented or immediately apparent, with the result that many users tend to treat environmental data as if they were error free (Briggs, 1995b).

Against this background, there is a clear need to treat indicators with care. Ideally, the data sources used need to be checked (e.g. by examining the genealogy of the data and by cross-validating the data against independent sources). As far as possible, the indicators themselves should also be tested for inconsistencies. Trends and geographic distributions should be carefully inspected to identify significant discontinuities, and these should be investigated to ensure that they are not artefacts of the data sources or methods used. Comparisons should be made between indicators to check for unexpected departures from established relationships. The definition of indicators should be checked to ensure comparability. Where feasible, margins of error should be assessed so that the true patterns or trends can be separated from "noise" due to uncertainty in the indicators. The definitions, methods and data sources used in constructing environmental health indicators should always be fully documented, in order to facilitate these quality checks.

4.6 Indicator Reporting

The way in which indicators are reported may have significant consequences for decision-making. An indicator can be measured as a point in time and reported as such or it can be measured at several points in time and reported as "change" over the time period. For example, the information provided by an indicator of health status differs if the health indicator is presented as the number of persons with the health outcome of interest, the proportion of the population with this health outcome or the annual change in that proportion. Similarly, for an environmental indicator such as desertification, a very different information is provided if desertification is described as the proportion of the area of a country that is classified as a desert, or the annual change in that proportion.

In the health field, available epidemiological tools provide a number of options for reporting on indicators. The occurrence of ill health (or good health) in a population is reported as either the number of existing cases of a disease (prevalence) or as the number of new cases of the disease occurring in a set time period (incidence). As the number of cases depends on the size of the population studied, prevalence and incidence are mostly reported as rates (e.g. number of cases per 1000 population). Moreover, these measurements can be reported as "change" when a comparison rate (expected "background" rate or the rate at some earlier time) is available.

5. TOWARDS A CORE SET OF ENVIRONMENTAL HEALTH INDICATORS

In recent years, considerable effort has been devoted to developing core sets of environmental and sustainability indicators for policy support. It might therefore be expected that similar core sets of environmental health indicators could usefully be constructed. Establishment of a core set of indicators would certainly offer a number of advantages. They could save time and resources, by avoiding duplication of effort in researching and developing new indicators. They could provide a basis for comparison between different areas and over time. They could help to establish standards for indicator development which would improve the general quality of information available to the decision-maker.

In practice, the construction of a core set of environmental health indicators is a much more difficult task than may be supposed. By definition, indicators need to be use-specific, so indicators developed for one application cannot readily be translated to another. Indicators therefore tend to be driven by prior concern about a problem. In some areas of application — such as the environment and economy — a broad consensus often exists about what these key problems are. Core sets of indicators can thus be developed on this basis. In the area of environmental health, however, this consensus is less well established, and many of the problems may be relatively local in their extent. The definition of core environmental health indicators is therefore more difficult. For these reasons, no attempt here is made to present formal lists of core indicators. Instead, as an example, a "matrix" of environmental health indicators is given in Table 2.4, which could serve as a model for specific indicator development.

Table 2.4 Environmental Health Indicators Matrix (Illustrative example) Page i of v

Driving Forces	Pressures	State	Exposures	Effects	Actions
Population changes and social conditions	Social, economic, demographic characteristics	Birth rate; age distribution; income distribution	Proportion of population living in poverty; of vulnerable age groups (in association with exposures of other driving forces)	Mortality, morbidity, disability (in association with other driving forces)	Education (particularly female), health care, birth control initiatives, income distribution
Human settlements and urbanization	Urbanization, urban migration; housing	Overcrowding; garbage disposal; noise levels. Indoor pollution - chemical (e.g. VOC[a], formaldehyde, isocyanyte, etc.); - physical (e.g. radon); - biological	Proportion of population living in disadvantaged areas; Proportion of time spent indoors; Proportion living in affected housing	Road accidents; crime rate; infectious diseases; mental health; neurobehavioural disorders; cancer; respiratory conditions	Service provision; health facilities; facilitate growth of smaller urban centres; improved housing
Water requirements	Quantity: - Inherent scarcity - Increased consumption Quality: - Natural - Pollution: (sewage, industrial effluent; urban run-off; agricultural run-off)	Water supply and sanitation - Formal access - Private system (e.g. wells) - informal market Industrial use Agricultural irrigation	Proportion of population without access to sanitation; with insufficient water; buying water from vendors	Morbidity & mortality resulting from: waterborne diseases (e.g. cholera); water-washed diseases (e.g. trachoma); water-based (e.g. schistosomiasis); water-related (e.g. malaria); water-dispersed (e.g. legionella)	Water conservation measures; use of treated urban wastewater for irrigation; increase public access to safe water/hygienic sanitation; pollution control legislation; community education

Table 2.4 Environmental Health Indicators Matrix (Illustrative example) Page ii of v

Driving Forces	Pressures	State	Exposures	Effects	Actions
Food and agriculture needs	Food production, diet; amount produced	Calories per person. Extent of land degradation; availability of water	Proportion of children with lower than acceptable calorie intake	Malnutrition; lower rate of growth in children; lowered immunity. (Risk mostly in developing countries, particularly for children)	Improved access, distribution; health education
	Microbiological contamination	Presence of microorganisms (measurements)	Consumption of contaminated food	Diarrhoea; typhoid fever; cholera, shigella, etc. (Risk to the general population)	Access to clean water; improved personal hygiene, hygienic sanitation, hygienic food production; pasteurisation, irradiation
	Toxic agents; type and amounts of chemicals used	Chemical additives; heavy metal releases in the environment; pesticides. Agricultural chemicals and organic wastes contaminating water supply	Population living in affected areas. Use (or lack of) of protective equipment for workers	Accidental poisoning; suicides. (Risk particularly to workers and population in developing countries)	Legislation, supervision; improved labeling; use of protective clothing

Table 2.4 Environmental Health Indicators Matrix (Illustrative example) Page iii of v

Driving Forces	Pressures	State	Exposures	Effects	Actions
Energy demand	Use of fossil fuels for transport, industry and home use (type and amount used)	Concentration of air pollutants; SO_2[b], TSP[c], CO[d], NO_x[e], ozone, lead, cadmium, mercury, arsenic, etc.	Proportion of urban dwellers, proportion of population living in areas where these pollutants exceed recommended levels	Respiratory conditions, carcinogenic effects, other pollutant-specific morbidity/mortality effects. (Risk to urban population)	Abatement expenditure, legislation re transport & industry, increased research in alternative power sources (solar, wind, etc.)
	Use of biomass fuel for cooking and heating (type and amount)	Indoor air concentration of pollutants, SO_2[b], TSP[c], CO[d], NO_x[e], hydrocarbons, aldehydes, acenaphthylene, benzene, phenol, cresol, toluene, polyaromatic hydrocarbons, etc.	Proportion of time spent indoors; in cooking areas	Respiratory conditions, CO[d] poisoning, risk of respiratory cancer. Accidental burns. (Risk to women and children in both urban and rural settings, in developing countries)	Improved access to improved stove designs; use of processed biomass fuels. Use of fossil fuels (gas)
	Use of nuclear energy (amount of radioactive material used)	Number and state of facilities; radiation levels	Personal monitoring (workers) Population living in surrounding areas	Leukaemia, cancer	Safety measures in place

Table 2.4 Environmental Health Indicators Matrix (Illustrative example) Page iv of v

Driving Forces	Pressures	State	Exposures	Effects	Actions
Industry development	Workplace (characteristics, type of industry, type and amount of hazardous materials used)	Workplace exposure levels (e.g. asbestos, silica dust, organic solvents, lead, mercury, cadmium, manganese, arsenic, nickel, aromatic amines, benzene, noise, etc.)	Monitoring exposures in workplace; in work-specific areas; in individual workers	Occupational diseases and accidents	Emission control measures, chemical safety legislation, epidemiologic studies, improved labeling, improved supervision
	Accidental releases (quantified emissions)	Short-term, high concentration of toxic substances (air, water)	Environmental measures in populated areas	Several including poisoning, cancer risk	Disaster prevention/ preparedness measures. Environmental Health Impact Assessment
	Toxic chemicals and hazardous waste disposal (quantified)	Nature & amounts of hazardous materials in the environment (measured)	Population living around hazardous waste disposal sites	Several potential health effects (pollutant-specific)	Legislation for safe disposal methods Supervision (e.g. against illegal dumping)

Table 2.4 Environmental Health Indicators Matrix (Illustrative example) Page v of v

Driving Forces	Pressures	State	Exposures	Effects	Actions
Global limits	Release of Chlorofluorocarbon (CFC) and other ozone-damaging chemicals	Stratospheric ozone depletion; Solar ultraviolet (UV) radiation at ground level	Proportion of time spent outdoors in specific locations. Use of (or lack of) protection	Skin cancers, ocular cataracts, immunosuppression	Legislation (Montreal Protocol)
	Release of "greenhouse gases"	Climate change: - temperature and precipitation change - increased climate variability - sea level rise	Population living in affected areas	Heat-related illness and mortality; redistribution and reemergence of vector- and water-borne diseases; new and reemerging infections; large-scale negative effects on nutrition	Research; monitoring; legislation (Framework Convention on Climate Change)

[a] Volatile Organic Compounds
[b] Sulphur dioxide
[c] Total Suspended Particles
[d] Carbon monoxide
[e] Nitrogen oxides

6. CONCLUSIONS

Population growth, technological and economic development, changing lifestyles and social attitudes, natural processes of change in the physical environment and the long-term impacts of past human interventions are all contributing to increasing problems of environmental health. To address these problems effectively, decision-makers require better information. This information needs to be reliable, consistent, targetted at the issues of real concern, available quickly, and available in an understandable and usable form.

Environmental health indicators provide one means of providing this information. To be effective, however, they need to be carefully designed and subject to rigorous quality control. In recent years, much progress has been made in developing indicators in a wide range of fields and for many different applications, much of it driven by the Earth Summit of 1992 and the adoption of Agenda 21. Progress on developing environmmental health indicators, however, has so far been slower, partly due to lack of consensus about the key issues which need to be addressed. Usable environmental health indicators also depend upon the existence of known and definable links between environment and health. Difficulties in establishing these relationships — due, for example, to the complexity of confounding effects and the problems of acquiring reliable exposure data — inhibit the practical utility of many potential indicators and make it difficult to establish core indicator sets. It is also important to appreciate that environmental health indicators are not panaceas. Indicators are only as good as the understanding which was used in developing them, and the models and data on which they are based.

Environmental health indicators consequently have limits, but if used within these limits, and with awareness of the limits, they can still make a major contribution to improved management and protection of public health. Well-designed and well-constructed indicators provide the capability to define more clearly the environmental health issues which need to be addressed, to prioritise these issues, to identify where action can best be taken, to compare the potential cost-effectiveness of different actions and strategies, to assess the effects of past or current action, to define the remaining research needs, and to inform the various stakeholders involved. The development and use of purpose-designed indicators to meet specific needs therefore remains a priority.

REFERENCES

Adriaanse A. *Environmental policy performance indicators. A study of the development of indicators for environmental policy in the Netherlands.* The Hague: Directorate-General for Environmental Protection, 1993.

Alexander L. *Health-Related Indicators for a Sustainable Community.* Lancaster: Environmental Epidemiology Research Unit, Lancaster University, 1994.

Briggs, DJ. Building a geographical information system in the European Community: the CORINE experience. In: *The added value of geographical information systems in public and environmental health* (M.J.C. de Lepper, H.K. Scholten and R.M. Stern, eds.). Dordrecht: Kluwer Academic Publishers, 1995a, pp. 299-314.

Briggs, D.J. Environmental statistics for environmental policy: data genealogy and quality. *Journal of Environmental Management* 1995b, 44, 39-54.

Carstairs V, Morris R. *Deprivation and health in Scotland.* Aberdeen: Aberdeen University Press, 1991.

Environment Canada *A Report on Canada's Progress Towards a National Set of Environmental Indicators.* SOE Report No. 91-1, Ottawa, Ministry of Environment, 1991.

FAO. *A framework for land evaluation.* FAO Soils Bulletin No. 32. Paris, FAO, 1976.

Hope C, Parker J. A pilot environmental index for the United Kingdom: results for the last decade. *Statistical Journal of the United Nations* 1991;8;85-107.

Inhaber, H. *Environmental indices.* London: John Wiley and Sons., 1976.

Jarman B. Underprivileged areas: validation and distribution of scores. *British Medical Journal* 1984;286:1705-9.

Kreisel WE. Representation of the environmental quality profile of a metropolitan area. *Environmental Monitoring and Assessment* 1984;4:15-33.

Local Government Management Board. *Sustainability Indicators - Guidance to Pilot Authorities.* London, Touche Ross, 1994.

Mullan RJ, Murthy LI. Occupational sentinel health events: An up-dated list for physician recognition and public health surveillance. *American Journal of Industrial Medicine* 1991; 19:775-99.

Newman, P.J. *Classification of Surface Water Quality. Review of Schemes Used in EC Mamber States.* Oxford, Heinemann Professional Publishingt, 1992.

OECD. OECD core set of indicators for environmental performance reviews. *Environmental Monograph* No 83, Paris: OECD, 1993.

Rothwell CJ, Hamilton CN, Leaverton PE. Identification of sentinel health events as indicators of environmental contamination. *Environmental Health Perspectives* 1991;94:261-3.

SCOPE (Scientific Committee on Problems of the Environment). Environmental Indicators: A systematic approach to measuring and reporting on the environment in the context of sustainable development. Paper presented at the workshop on Indicators of Sustainable Development for Decision-Making, Ghent, January 9-11, 1995.

Simonett, D.S. Considerations in integrating remote sensing and geographic information systems. In: *Building databases for global science.* (H. Mounsey, ed.). London, Taylor and Francis, 1988, 105-28.

Sustainable Seattle. *The Sustainable Seattle Indicators of Sustainable Community.* Seattle, Sustainable Seattle, 1993.

Swedish Environmental Protection Agency (SEPA). *Environment and Public Health. An Epidemiological Research Programme.* Solna: Swedish Environmental Protection Agency, 1993.

Townsend P. Deprivation *Journal of Social Policy.* 1987;16:125-46.

UNDP. *Human Development Report.* New York, Oxford University Press, 1990.

UNEP/RIVM. *An Overview of Environmental Indicators: State of the Art and Perspectives.* UNEP/EATR.94-01; RIVM/402001001, 1994.

UNEP/WHO. GEMS/Air — Global environmental monitoring system: A global programme for urban air quality monitoring and assessment. Doc. No. WHO/PEP 93.7; UNEP/GEMS/93.A.I. UNEP, Nairobi, 1993.

United Nations. *Agenda 21: Programme of Action for Sustainable Development.* New York: UN, 1993.

USEPA. *A Conceptual Framework to Support the Development and Use of Environmental Information* (draft). USEPA, 1994.

WHO. *Environmental Health Criteria* 27. Guidelines on studies in environmental epidemiology. Geneva: WHO, 1983.

WHO. *Air Quality Guidelines for Europe. WHO Regional Publications, European Series No. 23.* Copenhagen, WHO, 1987.

WHO. Joint UNEP/FAO/WHO food contamination monitoring programme. Doc. No. WHO/EHE/FOS/90.2. Geneva, WHO, 1990.

WHO. GEMS/Water 1990-2000: The challenge ahead. Doc. No. WHO/PEP/91.2. Geneva, WHO, 1991.

WHO. *Our Planet Our Health.* Report of the WHO Commission on Health and Environment. Geneva: WHO, 1992a.

WHO. *Twenty Steps for Developing a Healthy Cities Project.* Copenhagen, WHO, 1992b.

WHO. *Environmental Health Criteria* No. 134, Cadmium. Geneva: WHO, 1992c.

WHO. WHO consultation on the development and use of environmental health indicators in the management of environmental risks to human health. Dusseldorf, 15-18 December, 1992. Doc. No. WHO/EHE/93.3, 1993a.

WHO. Environment and health indicators for use with a health and environment geographic information system (HEGIS) for Europe. Report on a WHO consultation. Bilthoven 11-13 March, 1993. EUR/ICP/CEH 246, 1993b.

WHO. Informal consultation on Health and Environment Analysis for Decision-making (HEADLAMP) methods and field studies - Summary report. Doc. No. WHO/EHG/94.15. Geneva: WHO, 1994b.

WHO. *City Health Profiles. How to Report on Health in your City.* Copenhagen, WHO, 1995.

World Bank. Monitoring Environmental Progress: a report of work in progress. Washington: The World Bank, 1994.

World Commission on Environment and Development. *Our Common Future.* Oxford: Oxford University Press, 1987.

World Resources Institute. Environmental Indicators: A systematic approach to measuring and reporting on environmental policy performance in the context of sustainable development. New York, 1995.

World Wide Fund For Nature (WWF) and The New Economics Foundation. *Indicators for Sustainable Development.* London: WWF, 1994.

3

Exposure Assessment

T. Nurminen[a], *M. Nurminen*[a], *C. Corvalán*[b], *& D. Briggs*[c]

1. INTRODUCTION

The goal of environmental epidemiology is to understand the health effects of environmental factors that are outside the immediate control of the individual (Rothman, 1993). As such, it encompasses the processes and effects of exposures to physical and chemical pollutants not only in the open environment, but also in occupational environments, together with the study of the spread of infectious agents through environmental media such as the air, water and food. Moreover, psychosocial factors and the public's perception of environmental health risks are increasingly important in environmental epidemiology.

Typically, the exposures that are beyond individual control affect many people simultaneously. Measurement of individual exposures is thus difficult and costly. As a result, environmental epidemiological investigations often have to rely on the use of existing data, and to analyse these at the aggregate rather than individual level. It is also important to appreciate that epidemiologic studies require more than data on exposure and health. Equally important are data on other known or possible risk factors which may confound relationships with the health outcomes of interest. Environmental exposures often have small effects that may be masked or distorted by the effects of confounding. Observed health effects of air pollution, for example, may be confounded by risk factors such as smoking or occupational exposures. Socio-economic factors act as confounders for many environmental health effects. Moreover, the assessment of effect modification (i.e. the change of the strength of the association between exposure and health outcome according to some other factor) is important for generalizing observed exposure-effect relationships to other populations. These issues are dealt with in Chapter 4. In environmental epidemiology, problems connected with inference based on grouped data call for further methodological work. For example, by obtaining individual-level data on the exposure and certain covariates in samples of selected groups, it might be possible to determine the

[a]Finnish Institute of Occupational Health, Helsinki, Finland

[b]Office of Global and Integrated Environmental Health, WHO, Geneva, Switzerland

[c]Nene Centre for Research, Northampton, United Kingdom

limits of ecologic bias in estimating the health effects (Morgenstern and Thomas, 1993; Prentice and Thomas, 1993). The problems of ecologic studies are discussed in Chapter 5.

2. EXPOSURE PATTERNS AND PROCESSES

The assessment of exposure is clearly fundamental to environmental epidemiology, and methods of exposure assessment have consequently been the focus for much attention in the literature (e.g., ACGIH, 1989; AIHA, 1988; CEN, 1991; 1992; HSE, 1990; 1991; Hawkins et al., 1992; ISO, 1992; Rappaport and Smith, 1991). As discussed in Chapter 1, however, environmental exposures can occur in many different ways. Exposure may take place as a result of inhalation, ingestion or dermal absorption to pollutants which have been carried or stored in the air, water, food, biota (vegetation and animals) and soil. In many cases, exposure may occur simultaneously from many sources and through multiple routes. Pathways of exposure to lead, for example, include air pollution from traffic and industrial emissions, drinking water, food, tobacco smoking, dusts, paints and other industrially produced commodities and soil (IARC, 1982). Valid exposure assessment therefore typically requires detailed knowledge about the geographical distribution of the pollutants of concern, the temporal variations in pollution levels, and the processes of exposure.

2.1 Geographical Variations

The geography of environmental contamination is complex. Different pollutants may be derived from a wide range of different sources, including localised point sources (e.g. industrial chimneys), line sources (e.g. roads) and diffuse sources (such as agricultural activities). Release from any of these sources may also occur either through controlled pathways (e.g. from a stack or discharge pipe) or as fugitive emissions, which leak inadvertently into the environment. Once in the environment, they may be transferred by many different processes and pathways. On the way, they undergo a great variety of changes as the result of dilution, deposition, chemical reactions and physical decomposition. Rates of these processes depend upon the pollutant species and the environmental medium concerned. As a result, patterns of pollution differ markedly in their magnitude and extent. Some pollutants may be widely and relatively uniformly distributed, due either to the ubiquitous distribution of their source activities, or the effects of long-distance transport. Other pollutants show more localised patterns, reflecting the localised distribution of emission sources and the limited extent of transport. Atmospheric pollutants emitted primarily from tall stacks (e.g. sulphur dioxide from power stations and other major combustion plants), for example, may be widely dispersed. Nitrogen dioxide, which is derived primarily from low-level traffic sources, often shows marked variations even within an

individual street. Nitrate and phosphorus pollution of surface waters is extremely extensive. Organic pollution of drinking water, in contrast, commonly occurs at the level of a neighbourhood or household. Food contamination can be specific to a particular product and affect all population groups consuming the product, or it can be specific to a household or neighbourhood where food storage hygiene is locally inadequate. Exposure to electromagnetic fields can vary strikingly over short distances.

2.2 Temporal Variations

Temporal variations in pollution levels are equally important. Pollution levels typically show a number of different trends at different temporal scales. In many cases long-term trends exist, reflecting underlying changes in the rates of emission (e.g. as a result of technological or economic changes or due to policy intervention). Superimposed upon these there may be annual variations, reflecting year-to-year differences in climate or source activity. Many pollutants also show marked seasonal, weekly and diurnal patterns, due to cycles of activity and short-term climatic and other effects. Major, short-term pollution episodes may also occur as a result of sudden, accidental releases. Measurements of exposure will therefore vary according to both when sampling is carried out and the duration of individual measurements (the averaging time). There are also many different ways of expressing the exposure level - e.g. as the average, peak, percentile (95% and 98% are often used), frequency of exceedance of a specified level, or cumulative duration of exceedance. The time scale of interest and the specific indicator to be used will depend on the health outcome one wishes to study and existing aetiological knowledge about the exposure-effect process.

In order to model past or future concentrations, or to isolate the effects of specific pollution episodes, it may be necessary to unravel the effects of these different components of temporal variation. This is often extremely complex, for the different cycles are not easily identifiable and are often masked by considerable random variation in pollution levels. Time series analysis is often used for this purpose, but even this must be used with care since it involves a number of assumptions and decisions on the part of the user which may significantly affect the results.

2.3 Measurement Issues

Variations in individual absorption or metabolism of pollutants are also complex. Exposure assessment and dose estimation thus pose difficult problems for those investigating the health effects of environmental agents. As noted in Chapter 1, the term *exposure* refers both to the concentration of an agent at the boundary

between an individual and the environment and to the duration of contact between the two. *Dose*, in contrast, refers to the amount actually deposited or absorbed in the body over a given time period (Hatch and Thomas, 1993). *Internal dose* is the ideal measure from the scientific standpoint, but limits and standards set by health and safety legislation usually relate to external exposures. Occupational exposure to lead, for example, is regulated and monitored on the basis of blood lead levels in workers. Blood lead, however, is inadequate either for monitoring organic lead compounds or as an indicator of amounts of lead in target tissues and temporal variations of exposure levels (Kazantzis, 1988). While there is undoubtedly a need to improve externally derived measures of exposure, efforts are also needed to estimate internal dose using methods such as empirical dosimetric modelling, pharmacokinetic modelling and biologic markers (Hatch and Thomas, 1993).

The long lag time likely to occur between exposure and presumed health effect in many cases further exacerbates the difficulties of exposure assessment (Rothman, 1993). In these circumstances there is a need to link data on present-day health outcomes to data on past exposures. Estimation of past exposures, however, is often exceedingly difficult. Where good historical records are available, it may be possible to make generalised estimates of exposures, and examination of past patterns of pollution can provide a basis for modelling (Hatch and Thomas, 1993). Nevertheless, suitable historical data on exposures are often lacking. Changes in pollution levels, place of residence and lifestyle may also mean that it is not realistic to extrapolate back from recent data. For unrecorded and imperceptible exposures — such as electromagnetic fields, especially — retrospective evaluation can therefore be only approximate at best.

2.4 Pollutant Combinations

People are often exposed to different pollutants simultaneously. To isolate the effect of one requires that the others have been controlled for in the analysis. Exposure to these may occur at different locations (e.g. in the workplace and/or at home) and at different times. Thus, it may be necessary to establish different sampling regimes or to use different sources of information to obtain exposure estimates.

The full range of factors which may need to be examined in any particular study is therefore potentially large. It may include many different environmental pollutants (including hazardous chemicals, radioactivity, dusts and particulates), from many different anthropogenic sources (including energy production, industry, pesticide use, transportation etc) and natural sources (e.g. geological release of radon), released either continuously or sporadically, and either under controlled conditions (i.e. deliberate discharges) or accidentally. Data on these pollutants may need to be obtained either from monitoring sites within or around the study area, or through the use of modelling techniques. In the latter case, additional data may be needed on

levels of source activity (e.g. traffic density, industrial production), emission rates, meteorology and other factors which affect dispersion processes (e.g. topography). Different data sources, sampling regimes and analytical procedures may be needed for the different pollutants and sources involved.

Poor exposure data jeopardize the validity of linkage analysis.

3. SOURCES OF EXPOSURE DATA

3.1 Exposure Sampling Strategies

Information on geographical variations in pollution levels may be derived from a number of sources. Often the most useful are the results of monitoring exercises. Almost all countries now run routine monitoring networks for a wide range of pollutants, and networks in many countries are being extended. New sampling and analytical technologies are being developed, including the use of automatic samplers and remote sensing. The development of low-cost sampling devices (e.g. passive samplers) for an increasingly wide range of pollutants is also facilitating the use of purpose-designed surveys. Gradually, improved awareness about the spatial and temporal variations in pollution is contributing to improved sample designs, so that monitoring is being undertaken more effectively — for example, by sampling the micro environment where exposure principally occurs — including indoor environments (e.g., bedrooms and living rooms in the assessment of radon and electric and magnetic fields). The use of total exposure monitoring — in which all potentially relevant micro environments are sampled — also offers opportunities to improve exposure estimates (Hatch and Thomas, 1993). In addition, personal exposure monitoring is being incorporated to some extent into environmental health assessment. Questions that form the basis of any sampling strategy include: what to measure; how to sample; whose exposure should be measured; where to collect the sample; when to measure; how long to sample for; and how many measurements or readings to take (Gardiner, 1995).

Nevertheless, in many cases, it is not possible to obtain information directly on pollution levels for the locations or areas of interest. In these cases, models may have to be used to estimate exposures. Several approaches are available. Where suitable data exist, it may be possible to estimate pollution levels by interpolating data from nearby monitoring sites. With the development of geographical information systems (GIS), a wide range of interpolation and mapping methods have become available. This approach is normally only feasible, however, where the distances involved are relatively small and the spatial variation in pollution levels limited. Alternatively, it may be possible to

estimate concentrations in the areas of interest by using dispersion models. Again, a wide range of models have been developed over recent years, but their applicability is often limited by their relatively stringent data demands; most air pollution models, for example, require detailed data on emission sources and rates and meteorological conditions. Where neither of these methods are possible, it may be appropriate to use more empirical methods. Multiple regression techniques, for example, may be used to construct predictive equations based upon environmental factors thought to determine pollution levels.

Obtaining reliable estimates of exposure at the individual or group scale is nevertheless extremely difficult, especially where routinely collected data are being used and variations in concentration are localised. In these cases, the measurement stations may not be representative of the environment in which exposure occurs. Within a geographically defined population, considerable variations in exposure may also occur, reflecting local variations in pollution level and individual behavioural patterns. The application of a single exposure score to the entire group, based for example on results from monitoring stations, is therefore likely to be erroneous and must always be undertaken with care.

3.2 Routinely Collected Environment Data

Routine environmental monitoring provides one of the most important sources of exposure data. Most countries now undertake routine monitoring, and a number of international monitoring networks also operate (e.g. GEMS, the Global Environment Monitoring System). The advantages of routinely collected data are that they are likely to be relatively easily accessible (often through government departments) and widely available, to follow approved methods and to be available for a relatively long period of time. Nevertheless, routine data may not be optimal for exposure assessment and linkage with health data. Problems may include the relevance of the monitoring with regard to the population and the environmental health problems encountered, the frequency of the measurements, the spatial representativeness of the monitoring sites and the geographic and temporal completeness of the data. For example:

- environmental data may be collected in areas which do not correspond to where the main exposures occur, or to where people live;
- exposure data relevant to some important environmental health problems may not be collected;
- data may be collected on a weekly or monthly basis, when more frequent data would be preferable, or data may be recorded on a more frequent basis but only summary data made available; and

- data may be collected for certain periods of the year (e.g. when exposures are assumed to be higher), but the excluded data would be relevant for comparison purposes.

Clearly, not all the data required can be obtained from government departments. Therefore additional information will often have to be sought, for example from industry or private research establishments. Typical examples of the data which may be available from these sources include information on the types of pollution and waste treatment and control equipment, details of the manufacturing processes, raw materials usage, sales and data on emission rates. Difficulties with these data sources may include the confidentiality of the data, costs of data acquisition (there is an increasing tendency by many organisations to charge commercial rates for data), and lack of comparability.

3.3 Ad Hoc Data Sources

Further data may be obtained in some cases from results of previous, ad hoc surveys. These are often conducted as part of research investigations or as pilots for longer-term monitoring exercises. Large numbers of these studies have been carried out, especially in more developed countries. Commonly, they cover a restricted geographic area, but within these areas they often involve extremely detailed investigations. For this reason, they can be a rich source of environmental data. Problems may occur, however, in gaining access to results from ad hoc studies, for they may not be widely reported, and contacts with data holders may not be easy to arrange. Typically, also, they will not have been carried out specifically to investigate links between environment and health, so the survey design may not be optimal for such applications. In addition, the fact that the data are not routinely maintained means that they quickly become out of date. Problems of comparability may also occur, where there is a need to combine the data with results from other sources.

3.4 Purpose-Designed Surveys

In many instances, therefore, *purpose-designed surveys* would seem to provide the ideal source of data. These have the major advantage that they can be designed specifically to meet the needs of the study, so that the sampling framework, choice of exposure indicators and analytical techniques can all be optimised. In practice, however, they have two major drawbacks: they are likely to be costly and time-consuming. In optimising the survey design to meet the immediate needs of the study, comparability with other data sources may also be sacrificed. Furthermore, the short-term nature of most surveys of this type means that their results may rapidly become redundant. For these reasons, purpose-designed surveys should normally be undertaken only as a last resort —

when suitable data are not available from other, existing sources. In these circumstances, use of rapid survey techniques and low-cost sampling devices may help to minimise costs and time-delays (WHO, 1982).

3.5 Finding Environmental Data

The fact that environmental data are often collected not by official agencies but by private organisations and research groups means that searching for data can be a lengthy and frustrating task. This is especially so where data directories or metadatabases, listing available data sources, do not already exist. Even where directories are available, they may not be sufficiently informative, since they do not necessarily record details of data characteristics, such as the method of georeferencing, spatial resolution and averaging time, all of which may be crucial in determining the utility of the data for environment-health linkage studies. In the absence of such directories, data availability can often only be established through direct contacts with potential data holders and by careful literature searches.

4. ENVIRONMENTAL DATA QUALITY

4.1 Problems in Environmental Data

The complexity of the environment, the high costs of monitoring, and the technical limitations of many environmental monitoring techniques, mean that environmental data are subject to severe problems of quality. Major problems typically include:

1. gaps in data coverage and completeness due for example to:

    - equipment failure
    - detection limits (e.g. use of equipment which is unable to detect low concentrations of pollutants)
    - failure to report or analyse data
    - gaps in the sampling network
    - cessation of sampling programmes
    - disruptions such as war, strikes or storms

2. lack of data comparability, due for example to:

    - changes in measurement techniques
    - changes in sampling design
    - changes in analytical, classification or reporting methods
    - changes in the parameters measured

- administrative changes (e.g. in the administrative units for which data are collected)

3. bias and error, due for example to:

- non-representativeness in the sample design
- measurement error (e.g. poor detection)
- analytical or modelling error
- reporting or transcribing error
- aggregation error (e.g. rounding).

The effect of all these factors is to introduce considerable uncertainty into many environmental data sets (Briggs, 1995). In the case of atmospheric emissions, for example, it has been suggested that current techniques may have potential errors ranging from about 10% for SO_2 to 100% or more for volatile organic compounds, due primarily to uncertainties in the emission factors and source activity data used. Moreover, changes in the emission factors used mean that emissions data are often recalculated. In the UK, for example, there was a 40% change in the estimates of annual NO_2 emissions between 1983 and 1992, due to adjustments in methodology.

As noted previously, national air pollution networks are often too sparse to detect local variations in pollution concentrations. Similarly, many national monitoring networks for stream-water quality collect samples on only a few occasions each year, so they give only poor estimates of the annual pollution level and provide little or no data on short-term variations. Estimates of waste generation and collection are typically based on only the most limited monitoring and in any case face severe problems of how to classify and quantify waste materials. As a result, estimates may have margins of error considerably in excess of 100% (Briggs, 1995). For all these reasons, environmental data must be treated with considerable caution.

4.2 Quality Control

In the light of all the problems inherent in environmental data sources, quality control is of the utmost importance. Poor quality exposure data can totally undermine attempts to analyse linkages between environment and health. It is therefore vital to have good knowledge of the data collection procedures, so that the reliability of the data can be assessed (and, if necessary improved). This is particularly important where data were originally collected for purposes other than exposure assessment. Unfortunately, there is generally a lack of supporting information on the genealogy of environmental data sets. It is also often difficult or impossible to obtain independent measures of pollution or exposure against which to verify the data being used. As a result, it is often difficult in

practice to check the quality of the data, and there is an urgent need to establish standards for reporting and documenting data definitions and genealogy. In addition, the techniques available for quality assessment are as yet poorly developed. Equally important, therefore, is the development and application of improved methods for assessing and reporting data quality in environmental epidemiology (Hatch and Thomas, 1993).

It is particularly crucial to check the quality and consistency of information where data are obtained from different sources, for otherwise inherent inconsistencies may be overlooked. Among others, the following techniques can be used:

- constructing scattergrams to examine the relationship between exposure indicators and to search for obvious outliers
- visually comparing data with other, independently published sources
- statistical comparison of data from different sources
- mapping individual indicators or use of trend surface analysis techniques to look for discontinuities which coincide with the boundaries between different data sources.

While routine sources may provide a wide range of exposure data, it is often necessary to supplement these with data from other sources, or from purpose-designed rapid surveys. In all cases, special care is needed to check the consistency, quality and suitability of the exposure data in relation to the needs of linkage analysis.

4.3 Data Standards

If valid comparisons between countries or cities are to be made, it is evident that environmental data standards need to be improved. The health-related programmes of urban air quality, water quality, and food contamination, carried out under the *Global Environment Monitoring System (GEMS)*, have performed a valuable service in this respect by providing a framework for a standardized system of data collection for countries to follow. They have also provided advice on which exposures to monitor, and encouraged other countries to participate in this worldwide monitoring effort. In addition, the GEMS *Human Exposure Assessment Locations (HEAL)* program has provided resources directed to the collation of accurate and reliable data on human exposures (UNEP/WHO, 1993). Within Europe, both the European Environment Agency and Eurostat are also developing a major role in establishing standards and procedures for data

collection, in conjunction with other international agencies such as OECD and UN-ECE.

Nevertheless, the adoption and implementation of common standards is not always feasible. The historic investment which many countries (especially in the developed world) already have in monitoring systems, for example, may make them reluctant to change to new, international norms. Local or national priorities and circumstances may mean that standards developed elsewhere are not considered relevant. To identify and analyse local problems may require the use of specific methods and indicators. Ensuring comparability with other areas or countries, therefore, is not always appropriate. Nevertheless, much useful information may be lost if cross-comparisons between results from different studies cannot be made (e.g. in order to obtain improved estimates of exposure-effect relationships from a wider range of areas). Considerable care is therefore needed in designing studies or monitoring systems, so that the specific needs of the exercise can be met without jeopardising the wider relevance of the results.

REFERENCES

ACGIH. *Air Sampling Instruments for Evaluation of Atmospheric Contaminants, 7th Edition.* Cincinnati: American Conference of Governmental Industrial Hygienists, 1989.

AIHA. *Quality Assurance Manual for Industrial Hygiene Chemistry.* Prepared by the Analytical Chemistry Committee of the American Industrial Hygiene Association, 1988.

Briggs, DJ. Environmental statistics for environmental policy: data genealogy and quality. *Journal of Environmental Management* 1995; 44:35-52.

CEN. *General Requirements for the Performance of Procedures for Workplace Measurements - prEN 482.* Prepared by CEN Technical Committee 137, European Committee for Standardization, 1991.

CEN. *Workplace Atmospheres - Guidance for the Assessment of Exposure to Chemical Agents for Comparison with Limit Values and Measurement Strategy - prEN 689.* Prepared by CEN Technical Committee 137, European Committee for Standardization, 1992.

Gardiner, K. Needs of occupational exposure sampling strategies for compliance and epidemiology. *Occupational and Environmental Medicine* 1995;52:705-708.

Hatch M, Thomas D. Measurement issues in environmental epidemiology. *Environmental Health Perspectives Supplements* 1993;101(suppl 4):49-57.

Hawkins NC, Jayjock MA, Lynch J. A rationale and framework for establishing the quality of human exposure assessments. American Industrial Hygiene Association Journal 1992;53:34.

HSE. *General Methods for Sampling Airborne Gases and Vapours, MDHS 70, October 1990.* UK: Health and Safety Executive, 1990.

HSE. *Analytical Quality in Workplace Air Monitoring, MDHS 71, March 1991.* UK: Health and Safety Executive, 1991.

International Agency for Research on Cancer. *IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. Vol 23. Some*

metals and metallic compounds. Lyon: International Agency for Research on Cancer, 1982.

ISO. *Harmonized Proficiency Testing Protocol, ISO/REMCO No. 231 Revised, January 1992*. Geneva: International Organization for Standardization, 1992.

Kazantzis G. The use of blood in the biological monitoring of toxic metals. In: Clarkson T, Friberg L, Nordberg G, Sager P, eds. *Biological Monitoring of Toxic Metals*. New York and London: Plenum Press, 1988: 547-66.

Morgenstern H, Thomas D. Principles of study design in environmental epidemiology. *Environmental Health Perspectives Supplements* 1993;101(suppl 4):23-38.

Prentice RL, Thomas D. Environmental epidemiology: data analysis. *Environmental Health Perspectives Supplements* 1993;101(suppl 4):39-48.

Rappaport SM, Smith TJ, eds. *Exposure Assessment for Epidemiology and Hazard Control*. USA: American Conference of Governmental Industrial Hygienists, Lewis Publishers, Inc, 1991.

Rothman KJ. Methodologic frontiers in environmental epidemiology. *Environmental Health Perspectives Supplements* 1993;101(suppl 4):19-21.

United Nations Environment Programme (UNEP), World Health Organization (WHO). Guidance on Survey Design for Human Exposure Assessment Locations (HEAL) Studies. Nairobi: UNEP/WHO, 1993.

World Health Organization (WHO). Rapid Assessment of Sources of Air, Water, and Land Pollution. (WHO offset publication no. 62) Geneva: WHO, 1982.

Assessment of Health Effects

T. Nurminen[a], *M. Nurminen*[a], *C. Corvalán*[b], *& D. Briggs*[c]

1. INTRODUCTION

The health effects of pollutants found in the environment can be divided into two broad classes — acute (i.e. short-term) and chronic (i.e. long-term) effects. Each of these can range in severity from death to minor illness or discomfort. Microbiologically contaminated water or food, for example, can have an intense health effect a few hours after a short exposure, but with no detrimental long-term effects on health. Arsenic in water, on the other hand, may have a severe longer-term effect at low but constant exposure levels (e.g. cancer). Lead contamination provides an example of an environmental hazard which can have either acute or chronic effects. Thus, some pollutants may have an almost immediate effect after exposure and other substances may require accumulation in the target organ before causing any detectable adverse health effects. For some pollutants there may be a threshold level, below which no health effect is evident. For others there may be no threshold, and some effect may occur at even the lowest exposure levels. Moreover, some health outcomes may require a period of latency before the effect is observed.

People are not affected equally by the same environmental hazard. Substantial variations in sensitivity to an exposure may thus occur within a population. These differences may derive from a number of factors, including differences in characteristics of the individual. In this context, recent advances in the understanding of the role of genes has been particularly important, though problems may exist in determining whether the marker for sensitivity being examined is a measurement of the genotype itself, some host characteristic, or family history (Hatch and Thomas, 1993).

Age, nutritional status and state of general health are also important determinants of individual vulnerability. Exposure hazards for the normal "healthy" population, therefore, do not necessarily apply to all sectors of the population, and separate assessments may need to be made for particular high-risk groups such as infants and young children, the elderly, pregnant women and their foetuses, the nutritionally

[a]Finnish Institute of Occupational Health, Helsinki, Finland

[b]Office of Global and Integrated Environmental Health, WHO, Geneva, Switzerland

[c]Nene Centre for Research, Northampton, United Kingdom

deprived, individuals suffering from some disease (de Koning, 1987). Such groups can often be identified by assessing the degree of effect modification which occurs for each specific group compared to the "normal" population. It is especially important to identify these high-risk groups because they will usually be the first to experience adverse health outcomes as the level of the pollutant increases. A study in Romania, for example, showed higher values of lead in blood in children living near a lead smelter than in adults in the same area; these results indicated that biomonitoring should be extended at least to children in other parts of the concerned city (Verberk et al., 1992). A summary table of epidemiologic studies on the health effects of air pollution in children is given in the Annex to this Chapter.

Vulnerability of populations to hazards is also evident in the different abilities of individuals or groups to mitigate their exposure to, and effects of, environmental threats. For example, if microbiologically contaminated water leads to cases of morbidity, the effect of the contamination will depend in part on the ability of an individual or a group to gain access either to alternative water sources or to therapeutic treatment. Ability to cope with the effects of environmental hazards is very often limited by economic circumstances. Thus, while high rates of mortality in an area may be in part due to the existence of an environmental hazard, they are not necessarily a direct indication of pollution levels. Instead, the severity of the effect of the hazard may be more closely related to variations in the ability of individuals to protect themselves from exposure or to treat the effects of the hazard.

2. HEALTH ASSESSMENT

Adverse outcomes due to environmental exposures represent a broad spectrum of effects. They range from premature death of many individuals to premature death of an individual, severe acute illness or major disability, chronic debilitating disease, minor disability, temporary minor illness, discomfort, behavioural changes, temporary emotional effects and minor physiological change (de Koning, 1987).

Relatively few studies have shown associations between environmental pollutants and actual excess in deaths; even where it would otherwise have been expected, investigation usually has not revealed gross excess mortality (Lancet, 1992). The main exceptions are severe accidents or events which have resulted in release or accumulation of large amounts of toxic substances in the environment, leading to deaths due to poisoning. Long-term effects on mortality are invariably even more difficult to demonstrate. Typically, only a small subset of the population experiences high levels of exposure, and the doses received by the general population are so low that only vulnerable high-risk groups are severely affected. Any excess mortality due to a pollutant is therefore restricted to a small section of the population, and mortality across entire populations is a weak and insensitive indicator of environmental health effects in most situations (Landrigan, 1992). Therefore, whether

mortality is a reliable environmental health indicator, and if so for what groups, must be considered in the context of the particular circumstances.

Traditionally, concern about environmental hazards has tended to focus upon hazards believed to be contributing to excess mortality, in part because of the relative ease of obtaining mortality statistics. Nevertheless, it would clearly be beneficial to detect the effects of exposure much earlier in the process — at the stage of milder effects and before it results in death. One approach is to use data on morbidity. In the case of diseases for which formal registers exist, such as cancers (Swerdlow, 1992; Draper and Parkin, 1992), relatively sound morbidity data are widely available. Otherwise, obtaining suitable data poses severe problems, due to the inadequacies of many health surveillance and recording systems, and the inconsistencies inherent in the data. Disease occurrence, for example, may be measured in many different ways — as number of hospital admissions, length of hospitalisation, drug sales, medical consultations, days-off-work etc. All measure different components of morbidity and all are subject to substantial differences in reporting rates. Disease prevalence may be influenced by variations in the duration of the disease and survival rates. Incidence data are generally less easily accessed and can be subject to artificial variations in ascertainment (e.g. as a result of screening programs). In order to avoid dilution of weak associations through inclusion of irrelevant cases, therefore, it may be desirable to focus attention on subgroups of disease which — on the basis of prior observation — can be considered specifically responsive to the exposure of interest (Hatch and Thomas, 1993).

Various more subtle indicators of health outcome may also be sought, such as reproductive and developmental outcomes or premorbid changes in the state of health. Routinely collected data on these effects are rarely available, and reliable data on baseline rates and normal ranges for subclinical end-points are often lacking (Hatch and Thomas, 1993). Questionnaires can provide an effective means of obtaining data on perceived health, but severe problems may occur in obtaining unbiased response rates across all sectors of the population. Whether biochemical or physiological changes in individuals or complaints to local health authorities regarding nuisance factors in the environment are considered as indicators of adverse effects depends largely on the accepted concept of the term "state of health" (de Koning, 1987). Despite the understandable desire to use earlier indicators of health effect, therefore, serious problems remain in obtaining the relevant data.

> Mortality across entire populations is an insensitive environmental health indicator. In addition to mortality, therefore, effects should be sought at other levels of health damage.

3. HEALTH DATA

Health data are clearly of primary importance in environmental health studies. In the context of the HEADLAMP approach, they perform two main roles. Firstly, they provide indicators of the effects of known exposures to environmental pollution on human health. As such, data on health outcome — when linked to appropriate environmental data — can be used to assess or confirm exposure-effect relationships within the study area, or to quantify the contribution of specific exposures to total mortality or morbidity. Similarly, monitoring of health outcome can show the effects of changes in exposure, due for example to policy interventions or the adoption of new technologies. Additionally, they can provide an indication of the possible existence of previously undetected exposures. Thus, variations in health outcome may be used to infer the existence of underlying variations in exposure which need to be further investigated.

As does environmental data, health data may come from a variety of sources, including routine monitoring, ad hoc surveys and purpose-designed studies. These provide data on a variety of indicators, including *health status* (e.g. infant mortality, progress in child development, blood pressure), *disease* (morbidity, hospitalization, incidence or prevalence of different signs and symptoms) and *adverse effects* (e.g. premorbid changes in the state of health and complaints to local health authorities regarding nuisance factors in the environment).

Results from routine health monitoring programmes generally provide the most appropriate source, for they tend to be available on a continuous basis for the whole of the area concerned, relatively easily accessible (at least at an aggregate level), and standardized in terms of procedure. Routine monitoring of health is undertaken for a variety of purposes: to provide management information on the performance of the health service, to monitor trends and detect changes in health status, to provide an early warning about health problems, and to monitor the need for and effects of health policy. It is these requirements, rather than any explicit need to link the health data with information on the environment, which consequently determine the design of the monitoring systems. As a result, routine monitoring does not necessarily provide ideal data for environment-health linkage studies. Moreover, like all health data, routinely collected information may be subject to errors and inconsistencies in diagnosis, reporting and georeferencing.

3.1 Mortality Data

Data on *causes of death* are available in most developed countries and are the only health statistics for which comparatively long time series are available. Variations in diagnostic practice and coding will, however, affect the

comparability of death certificate information between different regions within a country or between countries. Cause-specific mortality data may also be subject to misclassification. Each year WHO, receives mortality data classified according to cause from thirty-seven developing countries and about the same number of developed countries. This information is readily available at WHO and published yearly in the World Health Statistics Annual Report. Of the developing countries, only twenty-two consider that the reporting of deaths is complete (WHO, 1987). Therefore, very few developing countries are in the position to monitor changes in causes of death on the basis of complete and reliable data.

Data on *infant mortality* are considered to be an indirect indicator of the level of health in the population. There are, however, a number of conceptual and practical problems with this indicator. A particular problem relates to differences in the definition of "infant death" for registration purposes in the first few days of life. The coverage of countries and areas in developing regions of the world in which registration of infant deaths is at least 90 percent complete is much less than for those reporting total population births and deaths (United Nations, 1985).

Many studies of environment-health relationships rely on time series analysis. These consequently require short-term (e.g. daily) counts of mortality. Daily mortality data are likely to be available in many countries, but perhaps not always in a form that is useful for computer analyses. Extra data entry or data processing may therefore be required.

3.2 Morbidity Data

Morbidity statistics are generally less readily available than mortality data even for developed countries. Typically, they are less complete and often refer only to specific subsections of the population. One exception to this is data on infectious diseases of significant public health importance. In most countries these must be recorded, and their reporting to a central health authority is often a legal requirement.

The accuracy of morbidity information depends on a number of factors, including the extent to which patients seek and obtain medical help, diagnosis practice and accuracy, the notification procedures, and treatment procedures. Variations in morbidity therefore do not necessarily reflect underlying differences in risk. When considering a small area, for example, it may be difficult to conclude whether a high prevalence of a disease is due to, say, a poor immunization rate or a good reporting of cases. Ongoing monitoring, instead of a cross-sectional assessment, is therefore desirable.

Disease registers are useful for obtaining incidence data on specific conditions. Most countries have registers of diseases, in particular of cancer. Other well organized registers include those of congenital malformations and mental disability. The utility of disease registers depends upon their level of completeness and quality of their records. Good registers may reach a completeness of 95% or greater, but there may be significant unevenness in the level of completeness between areas (Swerdlow, 1992). Unfortunately, independent data against which to assess the completeness of disease registers are rarely available, though indirect measures (such as mortality to registration ratios) may be used to indicate discrepancies (Muir and Waterhouse, 1987). A register of a terminal disease such as cancer may therefore not be considered complete until data from death certificates is used to complement those from referring hospitals and other regular sources. Other problems include duplicate registrations, differences in practice for dealing with multiple cancers, methods of georeferencing, and delays in registration (Swerdlow, 1992). Moreover, not all registers are yet fully computerized. A considerable investment of resources may thus be necessary to capture the data in a form suitable for analysis.

Annual data on *cancer incidence* are reported to the International Agency for Research on Cancer (IARC) from registries in participating countries. Although more than twenty developing countries report to IARC, data refer to population subsections and in many cases there is a question as to their reliability.

Some countries keep specific registers for certain diseases, such as myocardial infarctions or congenital malformations. In other cases, information on specific diseases is collected through purposely designed health surveys of representative populations or specific high-risk groups.

Information on *communicable diseases* is also available in many countries, and routine monitoring has played an important role in disease control in developed countries. Data on these diseases may be collected in a variety of ways, including mandatory notification, surveillance, sentinel networks and laboratory networks. The task of assessing the health impact of different communicable diseases on the population is made easier in many cases, because — in contrast to chemical and physical agents — the health effects tend to be very specific for a particular exposure (e.g. hepatitis caused by hepatitis virus).

Monitoring of *occupational diseases and accidents* has proved to be effective in their prevention. Most industrialized countries have consequently established monitoring programmes for occupationally exposed populations, while developing countries are in the early stages of implementing such programmes. In recent decades, increasing numbers of countries have also linked mortality data with occupation and place of residence. This has brought to light several associations with potential aetiologic factors, although in most cases subsequent epidemiological analyses have been required to confirm the relations. Linkage

of mortality data from health registers with exposure data can further enhance the detection of environmental risk factors. The effectiveness of such monitoring is increased with diseases specifically caused by environmental factors, such as pleural mesothelioma, lung cancer and asbestosis caused by the inhalation of asbestos dust. Similarly, the linkage of mortality and incidence data from cancer registers with information on occupation has provided a great deal of information on occupational cancers.

Additional sources of information exist in most countries which can be used for assessing disease and disability levels. These include hospital records, health service files, health insurance and physical payment systems, school records, work-day losses, and the sales of pharmaceutical agents. Although not ideal, these sources provide the basis for constructing indicators of certain aspects of health. Hospital morbidity data have the advantage of being detailed and fairly accurate, but detailed information is normally not coded; so, for many applications, data capture can be a time-consuming process. Multiple admissions are also not always easily detected, while differential use of health services is a well recognized problem. Another problem is the difficulty in determining the denominator population for the calculation of rates.

Sources of data such as hospital admission or discharge records, cancer registers and records of congenital malformations, therefore, do not on average meet the same levels of exhaustiveness and standardization as mortality statistics (except in the Nordic countries). The potential value of these systems is nevertheless considerable, since they offer the opportunity to detect and monitor health effects in advance of mortality. It is therefore extremely important to improve these systems by increasing their accuracy, completeness and accessibility.

Health surveys also provide a valuable source of morbidity data. Surveys are routinely performed in many countries, while special surveys may be undertaken to investigate specific health issues. The utility of the survey results depends to a great extent upon the survey design. Many surveys are targeted deliberately at particular sections of the population (e.g. high-risk groups) and thus do not provide data on the health of the general population. Surveys may also be designed to give only national data; sample sizes may then be too small to provide reliable estimates at the regional level.

Specific local or regional surveys may also be carried out to supplement existing data. A detailed survey of the region of interest, for example, may be the best means of obtaining detailed morbidity data for a city or region. A census of a region's hospital or any other health agency will provide information on the major reasons for service utilization.

The two main survey designs are the cross-sectional and longitudinal survey. A cross-sectional (prevalence) survey is probably the most practicable, as it

provides a picture of the population at one point in time, making it a rapid and inexpensive method. Longitudinal surveys collect information over time, providing a useful moving picture of the population (measuring change of health status), but at considerable expense and requiring a long duration.

The quality of the results again depends upon the sample design. Probability sampling is often the most reliable way of ensuring that the survey can provide valid sample-to-population inferences. If the region to be surveyed is very large, areas within the region can be randomly selected — using, for example, a multi-stage or stratified sampling technique. Determination of sample size is of great importance as it limits the precision of the survey estimates and thus constrains the analyses that can be legitimately carried out. The great advantage of a survey is that it can be designed to meet the specific needs of the study. Thus, it can be as detailed as necessary, and information on all the indicators of interest — e.g. morbidity, risk factors, and population characteristics — can be obtained simultaneously and within a consistent sampling framework. Although most surveys are designed to obtain data on morbidity, mortality data can be estimated by asking interviewees about deaths in the family. This is particularly useful, for example, in estimating infant mortality.

General guidance on survey methodology can be found in textbooks on sampling techniques (e.g. Cochran, 1960). A number of specialized treatises are also available on survey sampling methods for the assessment of human health (e.g. Lutz et al., 1992).

4. POPULATION AND COVARIATE DATA

Interpretation of patterns in health outcome cannot be reliably carried out without reference to the underlying population or to variations in those factors which may act as potential confounders to the relationship between environment and health. For these reasons, most studies of environmental health rely on the availability of data on population and covariates, such as social conditions and lifestyle. Moreover, processes such as in- and out-migration create major difficulties in interpreting exposure-health relations on either a temporal or spatial basis (Hatch and Thomas, 1993).

4.1 Population Data

Data on population numbers are essential for most environmental health studies. Expressed merely in absolute terms, data on health have little meaning, for variations are likely to depend more on differences in the size of the population than on any underlying differences in health. For most purposes, therefore, it

is more appropriate to express health outcome as rates — and this requires data on the population as a denominator.

In some cases, simple population totals (by gender) may suffice for this purpose. These data are normally readily available from censuses, at least at the national level. Of the 218 countries or areas from which the United Nations Statistical Office requests demographic data, only fifteen have not reported an official estimate since 1979. Nevertheless, most countries carry out complete censuses on about a ten-year cycle, so at any one time population statistics may be considerably out of date. Estimates are therefore generally made, based on population projections. Although these may be reasonably reliable at the national scale, considerable errors may develop over time at the small-area scale. Projections also tend to become less reliable with increasing time since the base census. Moreover, errors in enumeration are common in censuses, while significant differences may occur in the definitions of the resident population (e.g. in how transients are classified) between different countries. Because these errors and discrepancies often affect specific sections of the population disproportionately, population data for certain social or age-groups may be particularly vulnerable to uncertainty.

For many applications, data are needed not merely on total population, but on population subgroups (e.g. by age and gender). These are necessary, for example, where health effects are being studied within a specific age group (e.g. children), where disease rates may vary substantially between different age and gender, or where *time-trends* are being analyzed. For this purpose, *vital statistics* are ideally required. These provide a demographic profile of the population under study — essentially a count of persons cross-tabulated by age and sex and other personal characteristics. This information allows a computation of standardized rates as a basis for comparison both of the same population at other points in time, and with other populations.

To some extent, this information can be obtained from national censuses. While population by age is widely available for most developed countries, the number of developing countries for which reliable periodic estimates are available is much smaller than those with total population counts. Typically, data on population age structure are only available for census years, and the age classes used in different countries may differ, so that international comparisons may be difficult. Nevertheless, most countries also maintain some form of vital statistics which includes registration of births and deaths. Globally, reasonably complete registration of births and deaths occur in about 85-90 countries or areas (United Nations, 1985). This includes all the developed countries and about 40 developing countries. In about sixty developing countries, the registration of vital events is considered incomplete (WHO, 1987).

4.2 Confounder Data

As already noted, control of confounding is an important element of most ecological studies. Rarely are the relationships between environment and health simple and unitary; instead, they are usually affected by a variety of confounding variables, many of which are only partially known. Rarely, therefore, will interpretations of the linkages between environment and health be wholly valid unless allowance for potential confounding is made.

The confounder data required will clearly depend on the specific relationship being studied. A wide range of potential confounders may exist, including social factors (e.g. ethnic origin, occupation, housing condition, income, education), lifestyle (e.g. diet, smoking, drug use) and physical environment (e.g. exposure to other pollutants, climate). Obtaining data on these confounders is often one of the most difficult aspects of ecological studies. Some data may be available from routine sources such as censuses and lifestyle surveys, but the scope of these is often severely limited. Data may also be obtainable from attitudinal surveys and market research studies. With the growing opportunity to use such information for the targeting of advertising and direct sales operations, a growing number of databases are being compiled. They can provide useful information on a wide variety of social and lifestyle factors, including diet, income, housing status, smoking, household size and leisure patterns. They may, however, be relatively costly to acquire and data quality may be uncertain. In addition, the possibility exists to acquire data on confounders through purpose-designed surveys. As with acquisition of environmental or health data, these have the advantage of providing better control over the data collection process, and thus ensuring that the data specifically meet the needs of the study. Typically, however, they are expensive to conduct and may cause considerable delay.

Because of the limitations of data availability, it is often impracticable to obtain information on all the confounders of interest. In many situations, therefore, proxies need to be used, based on other, readily available, demographic or social statistics. Most covariates used in ecologic regressions are, in practice, either proxies or rather indirect or crude measures of the true confounder. The use of proxies, however, is clearly only valid where they do in fact provide a reliable surrogate for the confounder of concern. Unfortunately, this is not always the case — or, at least, the validity of the proxy is a matter of conjecture. In these circumstances, particular care is needed in interpreting the results.

Investigation of the lung cancer in cities provides an example. The causes of the higher incidence of lung cancer in many cities are insufficiently known, but are suspected to be related to smoking and socioeconomic status, among other factors. A study in Helsinki showed an apparent increase in cancer incidence

with increasing mean sulphur dioxide (SO_2) concentration (Pönkä et al., 1993). To interpret this correctly, however, clearly required the possible effects of confounding by smoking and other social factors to be taken into account. Information on smoking habits was not readily available so, instead, average education level was included as a covariate in the ecologic regression on the assumption that smoking levels were higher amongst the less-well educated. The analysis did, indeed, show a strong inverse association between education and cancer rate. Use of education level in the analysis thus helped to allow for some form of social confounding effect. Nevertheless, to interpret the results as evidence that *smoking* is related to lung cancer in the study area relies on the assumption that education level is a valid proxy for smoking rate.

The problem of controlling for confounders is further compounded by the potentially large number of confounders which may be of relevance, and the complex relationships which may exist between them. In other words, confounders do not necessarily act individually or in isolation, but may operate in unison. There is consequently a need to measure the multivariate (joint) distribution of the confounders; univariate distributions of the covariates or use of a simple confounder score may not suffice to achieve full control of confounding. Bobak and Leon (1992), for example, carried out an ecological study in the Czech Republic to test the hypothesis that atmospheric levels of pollution affect infant mortality risk. The socioeconomic data available included mean income, mean savings, mean number of persons per car, proportions of total births outside marriage, and legally induced abortions per 100 live births. While these allowed for control of a number of potential confounders, they were clearly not comprehensive, and allowance could not be made for potentially important confounders such as smoking, indoor pollution from heating or cooking, and family size. The potential also existed for interactions between the various confounders. As the investigators themselves acknowledged, therefore, an unknown amount of residual confounding may have been left unresolved.

> The problem of missing or inadequate information on confounding factors is especially serious in studies using aggregate data.

REFERENCES

Bobak M, Leon DA. Air pollution and infant mortality in the Czech Republic, 1986-88. *Lancet*, 1992; 340:1010-1014.

Cochran W. *Sampling Techniques*. New York: Wiley, 1960.

de Koning HW. *Setting Environmental Standards — Guidelines for Decision-making*. Geneva: WHO, 1987.

Draper GJ, Parlein DM. Cancer incidence data for children. In: Elliot P, Cuzick J, English D, Stern, R, eds. *Geographical and environmental epidemiology*. Oxford: Oxford University Press, 1992:63-71.

Hatch M, Thomas D. Measurement issues in environmental epidemiology. *Environmental Health Perspectives Supplement*, 1993;101(Suppl4):49-57.

Lancet. Environmental pollution: it kills trees, but does it kill people? (Editorial). *Lancet*, 1992;340:821-822.

Landrigan PJ. Environmental pollution and health. (Letter). *Lancet*, 1992;340:1220.

Lutz W, Chalmers J, Hepburn W, Lockerbie L. *Health and Community Surveys*. Vol 1-2. The MacMillan Press Ltd, International Epidemiological Association, World Health Organization, 1992.

Muir C, Waterhouse J. Comparability and quality of data: reliability of registration. In: Muir C, Waterhouse J, Mack T, Powell J and Whelan S, eds. *Cancer in five continents. Vol. V. IARC Scientific Publication No. 88.* Lyon: IARC, 1987;145-169.

Pönkä A, Pukkala E, Hakulinen T. Lung cancer and ambient air pollution in Helsinki. *Environment International* 1993;19:221-31.

Swerdlow AJ. Cancer incidence data for adults. In: Elliott P, Cuzick J, English D, Stern R, eds. *Geographical and environmental epidemiology*. Oxford: Oxford University Press, 1992: 51-62.

United Nations. *Handbook of Vital Statistics Systems and Methods*. Vol. 2: Review of National Practices, Series F, No. 35. New York: United Nations, 1985.

Verberk MM, de Wolf FA, Verplanke AJW. Environmental pollution and health. (Letter). *Lancet*, 1992;340:1221.

World Health Organization (WHO). *Health Monitoring in the Prevention of Diseases Caused by Environmental Factors*. Geneva: WHO and Commission of the European Communities, 1987.

Annex to Chapter 4 — Summary of epidemiologic studies of the health effects of air pollution in children. Page i of viii in table

Source Study location	Subjects	Exposure[a]	Health outcome[b]	Confounders controlled	Study design[c] Period	Association[d] found
Aubry et al., 1989 Northeast France	Children aged 9-12; 375 exposed, 523 unexposed	Complex industrial pollution (dust, SO_2, NO_x, hydrocarbons)	Rhinitis; school absenteeism	Parental smoking; use of coal fires; parental education	Cross-sectional	Higher prevalence in exposed
Bobak & Leon, 1992 Czech Republic	Infants from 46 of 85 districts	PM_{10}, SO_2, NO_x	Total and respiratory mortality	Socioeconomic characteristics of districts	Ecological 1986-88	Postneonatal mortality associated with PM_{10} increase; Respiratory mortality (highest to lowest quintile OR): 2.4 for PM_{10}; 3.9 for SO_2
Braun et al., 1989 Switzerland	1225 children aged 0-5, in 4 areas (2 urban, 1suburban, 1 rural)	NO_2, measured with personal samplers	Respiratory symptoms	Season; child's susceptibility to colds	Follow-up study (six weeks)	Association between outdoor NO_2 and respiratory symptoms
Castillejos et al. , 1992 Mexico City	143 7-9 yr old children, at 3 schools	Ambient 0_3	Respiratory function and acute respiratory symptoms		Follow-up study (six months)	O_3 associated with greater decrement in lung function for children with chronic phlegm
Dockery et al., 1989 Six US cities	Children from the Six Cities Study of Air Pollution and Health	TSP, PM_{15}, $PM_{2.5}$, $FS0_4$	Chronic cough, bronchitis, chest illness		Cross sectional 1980-81 school year	Health outcomes positively associated with all measures of particulate pollution.

Annex to Chapter 4 — Summary of epidemiologic studies of the health effects of air pollution in children. Page ii of viii in table

Source Study location	Subjects	Exposure[a]	Health outcome[b]	Confounders controlled	Study design[c] Period	Association[d] found
Forastiere et al., 1992, Italy: One industrial town, one rural area, and the city of Rome	2929 primary school children	Outdoor air pollution and passive smoking	Respiratory symptoms and illness		Cross sectional	OR for asthma: 1.4 for the industrial town and 1.3 for Rome (rural town as reference)
Goren & Hellerman, 1988. Ashod and Hadera, Israel	2nd and 5th grade schoolchildren	High and low air pollution areas	Respiratory symptoms and pulmonary diseases	Background variables	Cross sectional	OR: 1.5 for cough without cold; 2.7 for asthma (high pollution area)
Goren et al., 1990 Haifa Bay area	2334 2nd and 2000 5th grade schoolchildren	Low, medium and high air pollution areas	Respiratory symptoms	Background variables	Cross sectional, Spring 1984	OR: 1.4 for sputum with cold and 1.8 for sputum without cold (high pollution area as compared with low pollution area)

Annex to Chapter 4 — Summary of epidemiologic studies of the health effects of air pollution in children. Page iii of viii in table

Source Study location	Subjects	Exposure[a]	Health outcome[b]	Confounders controlled	Study design[c] Period	Association[d] found
Jaakkola et al., 1991. One polluted and two reference cities in Northern Finland	679 children 14-18 months; 759 children 6 years old	SO_2, particulates, NO_x, hydrogen sulphide (mainly industrial sources)	Respiratory infections	Potential confounders	Follow-up (12 month period, 1982)	OR: 2.0 in the younger group; 1.6 in the older group (polluted vs. less polluted).
Kinney et al., 1989. Kingston and Harriman, Tennessee	154 school children	O_3	Transient lung function decrease (FVC, FEV_{75}, MMEF, V_{max75})		Follow-up. Two-month period beginning February 1981	Health outcome associated with O_3 at levels below national standards
Kucerova et al., 1990. Slovakia	Children aged 7-14. 8973 in 1986; 9409 in 1987	Air pollution	Respiratory diseases		Cross-sectional	Higher annual incidence and higher mean duration in 11 contaminated areas

Annex to Chapter 4 — Summary of epidemiologic studies of the health effects of air pollution in children. Page iv of viii in table

Source Study location	Subjects	Exposure[a]	Health outcome[b]	Confounders controlled	Study design[c] Period	Association[d] found
Melia et al., 1981a. 28 areas of England and Scotland	Children aged 6-11 in 4 periods: 857, 1436, 2702 and 2036 children	Atmospheric smoke and SO_2	Change in number of respiratory conditions between examinations		Follow-up 1973-77	Change in health unrelated to pollution levels
Melia et al., 1981b. 19 areas of England and Scotland	4116 children aged 6-11	Atmospheric smoke and SO_2	Respiratory illness	Age, social class, population density, type of cooking fuel used at home, season	Cross-sectional 1975	Prevalence of respiratory conditions positively associated with levels of smoke
Mostardi et al., 1981. Akron, Ohio	Children in two schools (one next to industry)	SO_2, NO_2 monitored daily	Acute respiratory illness		Follow-up	Higher incidence of cough, runny nose and sore throat in the polluted area
Ong et al., 1991. Two districts in Hong Kong	Primary school children: 2009 from industrial area; 1837 from less polluted area	Exhaust emission from factories	Respiratory symptoms	Gender, age, socioeconomic factors, exposure to smoking	Cross-sectional April/May 1989	Higher prevalence of sore throat, evening cough, cough for more than 3 months, morning phlegm and wheezing in polluted area

Annex to Chapter 4 — Summary of epidemiologic studies of the health effects of air pollution in children. Page v of viii in table

Source Study location	Subjects	Exposure[a]	Health outcome[b]	Confounders controlled	Study design[c] Period	Association[d] found
Penna & Duchiade, 1991. Rio de Janeiro	Infants	Air pollution	Infant mortality from pneumonia	Area of residence, income	Ecological 1980	Association between average annual level of particulates and mortality from pneumonia
Pope & Dockery, 1992. Utah Valley	Symptomatic and asymptomatic samples of 5th and 6th grade school children	Respirable particulate pollution (PM_{10})	Acute respiratory symptoms		Follow-up Winter 1990-91	Association between the incidence of respiratory symptoms and PM_{10}, for both samples, but strongest for the symptomatic one
Roemer et al., 1993. Two nonindustrial towns in the Netherlands	73 children aged 6-12 with chronic respiratory symptoms	SO_2, NO_2, PM_{10} and black smoke (BS)	Acute respiratory symptoms	Ambient temperature	Follow-up Winter 1990-91	Association between PM_{10}, BS and SO_2 with the prevalence of wheeze and broncho-dilator use
Romieu et al., 1992. Mexico City	111 preschool children	O_3	Respiratory related school absenteeism	Demographic data, medical history, sources of indoor air pollution	Follow-up (Three months)	O_3 exposure associated with elevated risk of respiratory illness.

Annex to Chapter 4 — Summary of epidemiologic studies of the health effects of air pollution in children. Page vi of viii in table

Source Study location	Subjects	Exposure[a]	Health outcome[b]	Confounders controlled	Study design[c] Period	Association[d] found
Saldiva et al., 1994. São Paulo	Children under five years (excluding neonates)	Daily air pollution (SO_2,CO, PM_{10}, O_3, No_x)	Daily mortality due to respiratory diseases	Weather, seasonal effects	Ecological, May 1990 - April 1991	An association between mortality and the No_x levels
Schenker et al., 1986 Rural area, Western Pennsylvania	4071 children aged 6-11	SO_2 measured as 3-hr, 24-hr and annual mean, in low-, moderate- and high-pollution areas	Respiratory symptoms and pulmonary function levels	Known predictors	Cross-sectional; Spring of 1979	No significant association for SO_2 levels (highest exposure slightly above national standrds)
Schlipkoter et al., 1986. Two regions in Germany	Children	Two polluted areas, one with substantial emission reduction in the study period	Growth and bone maturation		Cross-sectional surveys, 1974 and 1984	No differences in height between areas and years. Bone age retardation (polluted area) at both surveys.
Schmitz-berger et al., 1993. European Alpine Region	1626 school age children	SO_2, NO_2, O_3 & infrared imaging & lichen mapping in three defined zones of exposure	Pulmonary function (FEV_1, FEF_{50}, FEF_{75}) and respiratory status	Age, sex, height, socioeconomic status, passive smoking	Cross-sectional	Decrements of FEV_1, FEF_{50} and FEF_{75} associated with areas of higher SO_2, NO_2, O_3. Higher prevalence of asthma in areas of increased O_3.

Annex to Chapter 4 — Summary of epidemiologic studies of the health effects of air pollution in children. Page vii of viii in table

Source Study location	Subjects	Exposure[a]	Health outcome[b]	Confounders controlled	Study design[c] Period	Association[d] found
Schwartz, 1989. USA	Children and youths aged 6-24 years	TSP, NO_2, O_3, SO_2	Lung function (FVC, FEV_1 and peak expiratory flow)	Age, height, race, sex, body mass, smoking, respiratory symptoms	Cross-sectional	FVC, FEV_1 and peak expiratory flow showed negative correlations with TSP, NO_2 and O_3.
Schwartz et al., 1991. Five German cities	Children	TSP, NO_2	Daily counts of children with croup or obstructive bronchitis	Seasonal confounding, weather factors	Follow-up (Two years)	TSP and NO_2 associated with croup cases. No association with obstructive bronchitis
Spinaci et al., 1985.	2385 School children from urban, peripheral urban and suburban areas	SO_2, TSP	Respiratory symptoms, pulmonary function	Household pollutants, smoking habits	Cross-sectional	Children from urban areas had lessened pulmonary function and higher prevalence of bronchial secretion with common colds
Stern et al, 1994. Ten Canadian rural communities	School children aged 7-11 years	Moderate and low exposure area of air pollution (sulfates, O_3)	Respiratory symptoms, pulmonary function	Age, sex, weight, height, parental smoking, gas cooking	Cross-sectional 1985-86	Decrement of FVC and FEV_1 associated with moderate exposures. No other effects observed.

Annex to Chapter 4 — Summary of epidemiologic studies of the health effects of air pollution in children. Page viii of viii in table

Source Study location	Subjects	Exposure[a]	Health outcome[b]	Confounders controlled	Study design[c] Period	Association[d] found
Ware et al., 1986. Six US cities	8380 white pre-adolescent children enrolled between 1974 and 1977	TSP, FSO_4, SO_2	Respiratory symptoms		Follow-up one year	Accross the cities: frequency of cough associated with mean concentration of all three pollutants during follow-up. Bronchitis and lower respiratory illness associated with mean particulate concentrations.
White et al, 1994. Atlanta	Black children aged 1-16 years	O_3	Asthma or reactive airway disease diagnosed in one public hospital		Ecological June 1990 - August 1990	Levels ≥0.11 ppm associated with higher frequency of cases
Wichmann et al., 1989. Germany	8420 children aged 6	Air pollution; traffic pollution (NO_2, NO, CO)	Respiratory diseases	Sex, parental education and nationality	Cross-sectional	Croup increased in streets with high traffic load. Asthma correlated with traffic related pollutants

[a] SO_2: sulphur dioxide; NO_x: nitrogen oxides; NO: nitrogen monoxide; NO_2: nitrous dioxide; O_3: ozone; TSP: total suspended particulates; PM_{15}, PM_{10}, $PM_{2.5}$: particulate matter less than 15, 10 and 2.5 microns; FSO_4: sulfate fraction of TSP; CO: carbon monoxide; BS: black smoke.

[b] FVC: forced vital capacity; MMEF: maximal mid-expiratory flow rate; V_{max75}: flow rate at 75% of expired FVC; FEV_1: forced expiratory volume in 1 second FEV_{75}, FEV_{50}: flow rates at 75 and 50% of vital capacity.

[c] Study design inferred, if not given.

[d] Major associations given. OR = Odds Ratio

REFERENCES

to Annex to Chapter 4

Aubry C, Teculescu D, Chau N, Viaggi MN, Pham QT, Manciaux M. [Respiratory health of children and atmospheric pollution. I. Respiratory symptoms]. *Revue des Maladies Respiratoires* 1989; 6(3):215-20.

Bobak M, Leon DA. Air pollution and infant mortality in the Czech Republic, 1986-88. *Lancet* 1992; 340(8826):1010-4.

Braun Fahrlander C, Ackermann Liebrich U, Wanner HU, Rutishauser M, Gnehm HE, Minder CE. [Effects of air pollutants on the respiratory system in young children]. *Schweizerische Medizinische Wochenschriffe* 1989; 119(41):1424-33.

Castillejos M, Gold DR, Dockery D, Tosteson T, Baum T, Speizer FE. Effects of ambient ozone on respiratory function and symptoms in Mexico City schoolchildren. *American Review of Respiratory Disease* 1992; 145(2 Pt 1) 276-82.

Dockery, DW, Speizer FE, Stram DO, Ware JH, Spengler JD, Ferris BG Jr. Effects of inhalable particles on the respiratory health of children. *American Reviw of Respiratory Disease* 1989; 139:587-94.

Forastiere F, Corbo GM, Michelozzi P, Pistelli R, Agabiti N, Brancato G, Ciappi G, Perucci CA. Effects of environment and passive smoking on the respiratory health of children. *International Journal of Epidemiology* 1992; 21(1):66-73.

Goren AI, Hellmann S. Prevalence of respiratory symptoms and diseases in schoolchildren living in a polluted and in a low polluted area in Israel. *Environmental Research* 1988; 45(1):28-37.

Goren AI, Hellmann S, Brenner S, Egoz N, Rishpon S. Prevalence of respiratory conditions among schoolchildren exposed to different levels of air pollutants in the Haifa Bay area, Israel. *Environmental Health Perspectives* 1990;89:225-31.

Jaakkola JJ, Paunio M, Virtanen M, Heinonen OP. Low-level air pollution and upper respiratory infections in children [see comments]. *American Journal of Public Health* 1991; 81(8):1060-3.

Kinney PL, Ware JH, Spengler JD, Dockery DW, Speizer FE, Ferris BG Jr. Short-term pulmonary function change in association with ozone levels. *American Review of Respiratory Disease* 1989; 139(1):56-61.

Kucerova A, Lipkova V, Liska J, Ursinyova M, Vanova R. [The effect of air pollution on the occurrence of respiratory tract diseases in children in Slovakia]. *Ceskoslovenska Pediatrie* 1990; 45(6):335-8.

Melia RJ, Florey CD, Chinn S. Respiratory illness in British schoolchildren and atmospheric smoke and sulphur dioxide 1973-7. II:Longitudinal findings. *J. Epidemiol. Community Health* 1981a;35(3):168-73.

Melia RJ, Florey CD, Swan AV. Respiratory illness in British schoolchildren and atmospheric smoke and sulphur dioxide 1973-7. I: cross-sectional findings. *Journal of Epidemiology and Community Health* 1981b; 35(3):161-7.

Mostardi RA, Woebkenberg NR, Ely DL, Conlon M, Atwood G. The University of Akron study on air pollution and human health effects II. Effects on acute respiratory illness. *Archives of Environmental Health* 1981; 36(5):250-5.

Ong SG, Liu J, Wong CM, Lam TH, Tam AY, Daniel L, Hedley AJ. Studies on the respiratory health of primary school children in urban communities of Hong Kong. *Science of the Total Environment* 1991; 106(1-2):121-35.

Penna ML, Duchiade MP. Air pollution and infant mortality from pneumonia in the Rio de Janeiro metropolitan area. *Bulletin of the Pan American Health Organization* 1991; 25(1):47-54.

Pope CA 3d, Dockery DW. Acute health effects of PM_{10} pollution on symptomatic and asymptomatic children. *American Review of Respiratory Disease* 1992; 145(5):1123-8.

Roemer W, Hoek G, Brunekreef B. Effect of ambient winter air pollution on respiratory health of children with chronic respiratory symptoms. *American Review of Respiratory Disease* 1993; 147(1):118-24.

Romieu I, Lugo MC, Velasco SR, Sanchez S, Meneses F, Hernandez M. Air pollution and school absenteeism among children in Mexico City. *American Journal of Epidemiology* 1992; 136(12):1524-31.

Saldiva PH, Lichtenfels AJ, Paiva PS, Barone IA, Martins MA, Massad E, Pereira JC, Xavier VP, Singer JM, Bohm GM. The association between air pollution and mortality due to respiratory diseases in children in São Paulo, Brazil: a preliminar report. *Environ Res* 1994;65:218-25.

Schenker MB, Vedal S, Batterman S, Samet J, Speizer FE. Health effects of air pollution due to coal combustion in the Chestnut Ridge region of Pennsylvania: cross-section survey of children. *Archives of Environmental Health* 1986; 41(2):104-8.

Schlipkoter HW, Rosicky B, Dolgner R, Pelech L. Growth and bone maturation in children from two regions of the F.R.G. differing in the degree of air pollution: results of the 1974 and 1984 surveys. *Journal of Hygiene Epidemiology Microbiology & Immunology* 1986; 30(4):353-8.

Schmitzberger R, Rhomberg K, Buchele H, Puchegger R, Schmitzberger Natzmer D, Kemmler G, Panosch B. Effects of air pollution on the respiratory tract of children. *Pediatric Pulmonology* 1993; 15(2):68-74.

Schwartz J. Lung function and chronic exposure to air pollution: a cross-sectional analysis of NHANES II. *Environmental Research* 1989; 50(2):309-21.

Schwartz J, Spix C, Wichmann HE, Malin E. Air pollution and acute respiratory illness in five German communities. *Environmental Research* 1991; 56(1):1-14.

Spinaci S, Arossa W, Bugiani M, Natale P, Bucca C, de Candussio G. The effects of air pollution on the respiratory health of children: a cross-sectional study. *Pediatric Pulmonology* 1985; 1(5):262-6.

Stern BR, Raizenne ME, Burnett RT, Jones L, Kearney J, Franklin CA. Air pollution and chidhood respiratory health: exposure to sulfate and ozone in ten Canadian rural communities. *Environ Res* 1994;66:125-42.

Ware JH, Ferris BG Jr, Dockery DW, Spengler JD, Stram DO, Speizer FE. Effects of ambient sulfur oxides and suspended particles on respiratory health of preadolescent children. *American Review of Respiratory Disease* 1986; 133(5):834-42.

White MC, Etzel RA, Wilcox WD, Lloyd C. Exacerbations of childhood asthma and ozone pollution in Atlanta. *Environ Res* 1994;65:56-68.

Wichmann HE, Hubner HR, Malin E, Kohler B, Hippke G, Fischer D, Bontemps M, Huenges R, Rebmann H, Walzer H, et al. [The significance of health risks caused by outdoor pollution, demonstrated by cross-sectional studies of pseudocroup in Baden-Wurttemberg]. *Offentliche Gesundheitswesen* 1989; 51(8-9):414-20.

Approaches to Linkage Analysis: Overview

M. Nurminen[a] *& D. Briggs*[b]

1. INTRODUCTION

Exploration of associations between environment and health is an integral part of environmental epidemiology, either in the search for previously unknown relationships between environment and health, or to test hypotheses about such relationships. *Linkage analysis* is an extension of this approach. It involves applying known exposure-response relationships, established in previous research and documented in the literature, to new empirical data as a basis for improved decision-making and policy support. In general, the health and environment data used for linkage analysis are derived from routine monitoring, though where necessary additional data may be collected from purpose-designed rapid surveys. In either case, the data often comprise series of data accrued over a long period of time, and are analysed in an aggregated form (e.g. at the small-area or regional level). As such, linkage analysis does not involve the direct use of individual records, though the data used may have been derived from individual-level sources, and individual-level analyses may be a valuable complement to linkage analysis. Instead, linkage analysis relies on methods for analysing grouped data.

In this context, it is important to emphasise that the aim of linkage analysis is not to discover new associations, nor to confirm suspected ones, but to use existing knowledge to assess the risks which exist, to identify need for action, to compare and evaluate the choices available, and to monitor and assess the effects of such actions. Linkage analysis thus requires the use of well-tested methods on reliable data, in the context of well-established knowledge about exposure-effect relationships. Results from linkage analysis should consequently be unbiased and agree with results that would have been obtained from more comprehensive ad hoc studies for which the statistical precision can be quantified.

[a] Finnish Institute of Occupational Health, Helsinki, Finland

[b] Nene Centre for Research, Northampton, United Kingdom

Many methods are available for this purpose. Three — *ecological analysis, time series analysis*, and *risk (or hazard) analysis* — will be outlined in this chapter Many tools are also available to support linkage analysis. One of the most important tools — and one which has received increasing attention in recent years — is *geographical information systems (GIS)*. Together, these methods and tools give an investigator with ingenuity countless opportunities to analyse and exploit existing data at greatly reduced cost. In the process, considerable value is likely to be added to the data, to knowledge about environment-health relationships in the area under study, and to the quality of decision-making.

Ecological analysis involves the investigation of group-level relationships between environment and health, by analysing spatial and/or temporal variations in exposure and health outcome. First used in sociology (Robinson, 1950), it has often been criticised for producing fallacious results. Particular concern has focused on the potential bias which may be introduced by aggregation of data, a problem which Selvin (1958) termed the *ecological bias* or *ecological fallacy*. Despite such theoretical shortcomings, however, ecological analysis has been widely used in environmental epidemiology, not least because it is relatively simple to perform, especially with the large, aggregated databases which are now available. For reasons of logistics and cost it may also be the only approach feasible where large population studies are required. Nevertheless, there has been a growing recognition that ecological or group-level associations are not necessarily consistent with those measured at the individual level (Greenland, 1992). Thus, much of the subsequent discussion of ecological methods has focused on how to identify, deal with or avoid the various biases involved, and how to quantify their effects compared to individual-level analyses For the future, more extensive use of the method may be anticipated, stimulated in part by the development of new statistical techniques and GIS.

Time series analysis is a similarly well-established method. It was developed to a large extent for econometric applications, but has since been adopted in a wide range of disciplines. It is typically used to investigate patterns in series of observations, as a basis for identifying and quantifying causal relationships. In the context of environmental epidemiology, it is often applied to long sequential observations, such as mortality statistics, data from morbidity registers (e.g. cancer registers, hospital discharge registers), or results from repeated health surveys. With simple data sets, it is a relatively simple method, and is supported by most proprietary statistical packages. Where temporal patterns are complex, however — and thus where relatively complex models need to be used to describe the time series — it can be computationally and statistically demanding, and can pose severe problems for both implementation and interpretation. In recent years, it has been extensively applied in studies of air pollution and health, and thus efforts have been made to formalize and standardise the techniques used (e.g. Katsouyanni et al., 1995). Moreover, as temporal data series are extended and improved, the opportunities for time-series analysis will

inevitably increase. Continued interest in the use of time-series analysis may thus be expected.

Risk analysis is the process of estimating the potential health effects of exposure to a hazard based upon existing information. As such, unlike the previous two methods, it does not explicitly involve bringing together data on environment and health, but uses available exposure data as a basis for quantifying potential health effects. Underlying such assessments, however, there must be known exposure-response relationships, usually derived from previous studies in other areas. There is, thus, an implicit linkage in the analysis. Note that the reverse process may also be envisaged — *hazard analysis* — in which health data are used to infer the levels of hazard (e.g. exposure) responsible for the observed health effects.

Application of risk analysis faces a number of problems. One is often the lack of information regarding the distribution and levels of pollution in the study area and the actual exposure of individuals to the pollutants. Another is uncertainty about the functional form of the exposure-response relation, and whether relationships established elsewhere fit the local conditions — for example, due to differences in background disease rate. Nevertheless, in many cases, risk analysis may be the only tool available for estimating the health outcomes of environmental pollution, due to the unavailability or poor quality of the health data.

Geographical information systems are computer systems for the capture, organization, storage, analysis and display of spatially referenced data. Developed initially for environmental applications in Canada and the USA, they have since been adopted world-wide in many areas of research, management and policy. Compared to some areas of application, their use in environmental epidemiology and health policy has been relatively slow, but they offer great potential for linkage analysis — for example in studying the incidence of disease events around emission sources, searching for disease clusters, or analysing relationships between pollution levels and health outcome (Briggs and Elliott, 1995). In the past, GIS technology has been expensive and more powerful systems especially have been complex and difficult to learn. Lack of suitably georeferenced data has also been a major constraint. In recent years, however, costs of both hardware and software have declined substantially, and a wide range of simpler introductory-level systems have emerged. At the same time, data availability has greatly improved. As a result, the use of GIS is rapidly expanding in both the developing and developed world.

In addition to the methods and tools outlined here, there are of course a wide range of more specialized approaches and techniques which may involve, or be relevant to, linkage analysis. Examples include the analysis of health clusters, exposure and disease mapping and studies of point-source exposures. Each of

these may have value in particular circumstances, but each also involves problems and dangers of which the investigator needs to be aware (see for example Rothman, 1990). Reference should therefore be made to relevant texts on methodology (e.g. Elliott et al., 1992).

> Linkage analysis applies known exposure-response relationships, established in previous research and documented in the literature, to new empirical data, as a basis for improved decision-making and policy support.

2. THE ECOLOGICAL METHOD

2.1 Introduction

The ecological method is an approach used in observational studies to detect and recognize patterns of disease occurrence across space and time, and to relate variations in the observed rates to environmental, social, behavioural and constitutional factors. The ecological design in epidemiology is also useful for the evaluation of the effects of interventions on risk factors — for example, to assess the effect of low-cholesterol diet on the rate of ischaemic heart disease.

Ecological analyses thus use aggregated or grouped data, rather than individual-level data, as the basic sampling unit of analysis. The grouping variate is usually a geographic region, although other factors such as ethnicity, socioeconomic class, time period, etc, could also be used. Ecological analyses of exposure-disease relationships are, however, subject to a number of biases. These include biases caused by model mis-specification, confounding, non-additivity of exposure and covariate effects (effect modification) and noncomparable standardization. Ecological correlations and rate estimates can be more sensitive to these sources of biases than individual-level estimates, because ecological estimates are based on extrapolations to unobserved individual-level data.

Given the availability of suitable exposure information, ecological analyses can be conducted in a number of different ways. One approach is to study the association between spatial variations in exposure and disease outcome in a single population at a given point in time. Alternatively, relationships between exposure and health outcome can be compared in two or more populations that differ in terms of their exposure. In either of these study designs, the data typically refer to a relatively short time span, and there are no multiple measurements over an extended time period. A third approach is to analyse time-trends within an ecological design. In this case, relationships between

exposure and health outcome are assessed by following changes in exposure and rates of disease within a single population over time.

2.2 Advantages and Disadvantages of Ecological Studies

The main advantage of the ecological approach is that it permits the study of very large populations (e.g. populations of entire countries), and thus allows the detection of relatively small increases in risk. The power of these studies is not a function of the size of the population studied, but of the large number of observations which are possible and the ability to use powerful methods of statistical analysis.

As noted earlier, ecological studies have remained popular in environmental epidemiology primarily because they are relatively easy and swift to conduct using existing databases. As a result, a well-designed ecological approach can serve as a cost-effective alternative to screening or monitoring of many health outcomes and environmental conditions. It should, however, be remembered that ecological studies often carry a number of hidden costs. Most studies, for example, rely on the use of data from existing monitoring networks which are themselves expensive to maintain, but whose costs are rarely charged to the ecological study. No doubt if these studies carried the full cost of the data they use they would seem less cost-effective. On the other hand, using routinely collected data as part of ecological studies does help to justify the costs of monitoring in the first place.

Ecological studies sometimes cover populations more markedly divergent in their exposures than those that can be readily obtained in studies of individuals. The ecological approach may also be useful for the investigation of clusters of disease in relatively small geographic areas.

The ecological design provides no information at all on the joint distribution of the exposure and disease variables at the individual level. All individuals within any "exposure class" are assumed to have experienced the same exposure. Thus there is no way of knowing from the ecological data whether individuals displaying the health outcome have in fact been exposed to the environmental risk factor, nor to what level they have been exposed. Only in exceptional cases, therefore, can valid inferences be made about individual-level exposure-disease relationships from ecological studies.

Routine data (e.g. hospital discharges or monitored pollution data) do not always meet the requirements of ecological research. In the case of health data, for example, an inappropriate system of disease classification may have been used. Equally, it may be difficult to define the population denominators (e.g. the catchment population of the hospital) which correspond to the health event

numerators. For less severe health events, such as acute asthmatic attacks, there may not be any records available at all. In the case of exposure data, there may be insufficient numbers of measurements of the pollutant(s) of concern, or measurements may not have been taken at locations which adequately characterize exposure. Commonly, health and environmental data have been collected on the basis of very different spatial frameworks. Health data, for example, are usually available for administrative units such as municipal health districts, municipalities, or provinces, whereas environmental pollutants and other exposures are usually available only for individual points or for areas which transect their boundaries. Considerable effort may thus be needed to convert the health and environmental data sets to comparable population subgroups. This may be done either by reallocating individuals to "pollution zones" based on their place of residence or, more commonly, by estimating pollution scores for each administrative area using mathematical models or spatial interpolation techniques.

2.3 Biases and Possibilities for Bias Correction

Unlike an individual-level study, an ecological study does not link individual disease events to individual exposure or covariate data, nor does it link individual exposure and covariate data. Instead, analysis is based upon the matching of aggregate or grouped data. Ecological studies are subject to a number of special biases as a result. In addition to the sources of bias ingrained in individual-level studies, for example, ecological estimates of effect can be biased by both *confounding* by group *(specification bias)* and *effect modification* by group. Covariates responsible for ecological bias in these ways may not even be confounders or effect modifiers at the individual level. *Aggregation bias* (or *cross-level bias)* may also be important in many cases. This refers to the incorrect estimates of exposure effects that derive from the analysis of data aggregated across study groups. Because the groups are typically heterogeneously exposed, cross-level bias is a more complex issue than that of simple confounding by group.

A further problem that has not been sufficiently considered in literature on ecological studies is that the ecological estimates of exposure are based on sample surveys and so are subject to sampling error. If the sampling error is not negligible, then exposure variables have standard errors which will bias the regression coefficients. If estimates of the standard errors are available from surveys, these may be incorporated to correct for this bias.

The susceptibility of ecological estimates to *measurement error* can also be an important source of uncertainty. Apart from basic demographic variables (such as sex and age), most variables used in ecological analyses are subject to substantial measurement error. The effects of this error are different for

ecological and individual-level studies. In part, this is because the samples used to estimate the distributions of the disease, exposure and covariate distributions for an ecological study are often disjointed. As a consequence, the measurement errors that arise may vary for different exposure classes, disease outcomes and covariates.

As a result of these various biases, estimation of exposure by ecological means may be affected by large *random errors* which may seriously disturb the analytical outcome. Independent *nondifferential misclassification* of an exposure indicator will usually result in a bias that is directed away from the no-effect hypothesis in ecological studies. By contrast, in individual-level studies the effect is in the opposite direction (i.e. towards the null).

Ecological studies in epidemiology typically deal with cause-specific mortality (and/or morbidity) rates rather than total mortality. Therefore, misclassification of disease outcome can be a source of severe bias. This bias can be far more important than the sampling variability of the disease outcome (the dependent variable within the regression analysis).

The most difficult sampling problem in ecological studies, however, often relates to measurement of potential confounders. Where existing databases are used, the availability of information on confounders is inevitably limited. Typically, the variables for which data are available do not include all the relevant covariates for the relationship under study, and only partial allowance can thus be made for confounding.

There is also a need in ecological studies for standardization of those variables that have a distribution not constant across the population. Published disease rates, for example, are invariably age-standardized, whereas exposure rates are seldom standardized. Since age is often associated with duration of exposure, regression of the disease rates on the exposure rates is biased, even if the precision of the rates used in the regression is high.

The main opportunity for bias prevention in ecological studies, as in epidemiologic studies in general, lies in the study design. If bias persists, there are only limited statistical methods available to reduce its effects at the analytical phase. These include influence analysis, sensitivity analysis, use of robust methods, and random-effect modelling. In general, the control of confounding in an ecological study is more demanding than in an individual-level study, because the measurement process for confounders is much more complicated. As with an individual-level study, an ecological approach also has the problem that the crude measurement or approximation of a confounder may be inadequate to achieve full control.

2.4 Conclusions and Recommendations

Ecological studies are based on a distinct methodological approach which sets them apart from individual-level studies. Either in planning ecological studies, or in critically evaluating the end-results of such studies, a number of specific factors thus have to be considered (see Greenland, 1992; and Morgenstern, 1982):

1. Ecological studies are much more sensitive to bias from model mis-specification than are results from individual-level studies.

2. Conditions for confounding differ in individual-level and ecological analyses.

3. In contrast to individual-level studies, independent and nondifferential misclassification of a dichotomous exposure usually leads to bias away from the null hypothesis in ecological studies.

4. In the *design of an ecological study* it is important to select areas with populations that:

   - are homogeneously exposed;
   - represent different extremes of exposure distribution;
   - are comparable with respect to covariate distributions; and
   - relate to the smallest sampling units possible.

5. In the *analysis of ecologic data* it is important to:

   - use weighted regression, instead of correlation, with weights proportional to the amount of information contained in each group;
   - include in the regression model all variables that are thought to be related to the grouping process;
   - examine multiple regression models with different and flexible structural forms beyond the standard linear form;
   - test the basic assumptions in the model;
   - conduct an influence analysis by examining the effect of deleting from the analysis various areas with unusual outcome, exposure, or covariate combinations;
   - conduct a sensitivity analysis of ecological estimates to misclassification;
   - take into account latency and induction periods separating causes and effects; and
   - give thorough consideration to the biases which are unique to such an analysis, as well as to those common to all epidemiological studies.

Whenever feasible, ecological studies using aggregated data should be supplemented by individual-level studies in a hybrid epidemiologic analysis.

3. TIME SERIES ANALYSIS

3.1 Introduction

Time series analysis (TSA) looks at the relation between observations recorded at consecutive, usually equally spaced, discrete time points. While TSA is also a regression method, it predicts the health outcome not from independent covariates, but from values of the outcome at previous points in time. The minimal requirements for TSA are the abilities to:

- plot the temporal series,
- derive new series (e.g. differenced series or smoothed series) and to plot these,
- examine scatter plots of time-lagged values,
- compute serial correlations, and
- display these graphically.

Current developments in graphical computing techniques for studying multidimensional relations will be valuable for TSA. Statistical computing aspects are especially important when the data sets used are large.

Three basic approaches to TSA exist, namely:

- *Poisson autoregression analysis* using *generalized estimating equations* (GEE);
- *Markov models* using *quasi-likelihood estimation* (QLE); and
- *Poisson risk function model* for time-stratified data using *maximum-likelihood estimation* (MLE).

It is beyond the scope and depth of this chapter to present the details of the statistics involved; it suffices here to outline in general terms how these models are applied in TSA.

3.2 Regression Models for TSA

The TSA model can be simply understood as a subclass of the *generalized linear model* (McCullagh & Nelder, 1983) in which the exposure effects are

multiplicative, the distribution of the errors is Poisson, and the link function is the natural logarithm. Thus, the model can be represented as:

$$\log[E(y_t)] = \mathbf{x}_t' \beta$$

where: $\mathbf{x}_t$ is the vector of covariates at time t, y_t is the count of observed outcomes at time t, and $E(y_t)$ denotes the expected count.

In order to account for the possibility of *overdispersion* and *autocorrelation,* the covariance matrix for the heath outcomes on the units of observation is assumed to have a special form; the regression parameters β are then estimated by the GEE (Liang and Zeger, 1986). This is because the form of the joint distribution of the time-dependent measurements is so complex as to be intractable; that is, it cannot provide useful and interpretable information.

Overdispersion in Poisson counts can arise for at least two reasons. First, the risk of an adverse outcome occurring to an individual may not be equal for all individuals, but depend on previous events that happened to that individual; that is, it varies over time. The second reason is that the risk may remain constant over time but not necessarily be equal for all individuals.

Markov models can also be applied for regression analysis of time series data (Zeger and Qaqish, 1988). As serial observations are unlikely to be independent, in the Markov models the expected response at a given time depends not only on the associated exposure variates and covariates but explicitly also on health outcomes at previous times. The regression coefficients can be estimated using the QLE approach (McCullagh and Nelder, 1983). Quasi-likelihood estimation allows one to estimate the regression relationship without full knowledge of the error distribution of the response variate.

There is a fundamental distinction between the GEE approach, which is a "pure" regression model with autocorrelated errors, and the QLE approach, which is a mixed regression-autoregression model. While these may be considered as alternatives, in general the coefficients in the two models are different. This is because, as in any regression equation, the interpretation of a parameter depends on what other variates are included in the model. It is also normally inappropriate to assume that the error in the exposure variates is negligible. When data on measurement errors is lacking, estimates should be obtained from independent survey samples. An advantage of the QLE approach over the GEE approach is that competing models can be compared directly with each other using a deviance statistic.

A particularly problematic aspect of studying temporal relations arises when there are sharp peaks present of similar frequency in both response and exposure series. For example, in an epidemiological study of daily death rate and

meteorological variates, seasonal fluctuations are likely to be present in all data sets. In other instances, long troughs or long-term trends may be present. TSA deals with this by studying the regressions separately in different seasons or periods.

Previous use of Poisson autoregression analysis models has generally been based on the assumption that the series is time-dependent. However, it is not clear either that the GEE approach or the QLE approach has advantages over simpler model building procedures sufficient to compensate for their greater statistical complexity. All the autoregressive methods involve complex and computer-intensive estimation procedures. A much simpler way of dealing with temporal data may be to adopt the working assumption that repeated observations from a unit are time-independent of one another. One can then proceed by dividing the study data into subgroups (strata) and fitting a Poisson risk function model to the time-stratified series. In this approach, the assumption of constant risk (or rate) ratio can be alleviated by including time-dependent covariates in the linear predictor. Computational demands can be reduced by using the *maximum likelihood estimation* methods available with existing software. Kuhn et al. (1994), for example, used this method for TSA and found that it compared favourably with the GEE approach.

3.3 Application of Poisson Regression for TSA

Given the potential use of Poisson regression to quantify time trends, it is worthwhile to consider some assumptions of the MLE method which may appear to be violated in the case of time-series data. Simple Poisson modelling requires that outcomes are independent. On first thought, it would seem untenable to assume that health outcomes occurring over time meet this requirement. For instance, the effects of social and environmental conditions are likely to persist at least in the short run. Poisson regression also assumes that the population subgroups are homogeneous with respect to the risk of adverse health outcome. This is another questionable assumption since the occurrence of, say, asthmatic or cardiac attacks do not occur at random but have predictable precursors and known patterns of risk.

In this context, two points need to be emphasized. The first is that the ordinary Poisson regression model requires that the study population meet the criteria of no overdispersion and heteroscedasticity conditional on the covariates. One effective way of removing overdispersion is to transform the data to a square root scale; this stabilizes the variance of the observed counts. The inclusion of time-dependent covariates may well result in *conditional independence* and help to define strata of homogeneous risk. Secondly, the Poisson regression allows the analysis of aggregate data to be comparable with the analytical methods used in cohort and case-base studies. Thus, although autoregressive TSA has been

promoted as the preferred method in analyses of sequential observations over long periods of time, it may equally be argued that Poisson regression provides a simple and viable alternative.

The minimum requirements — for epidemiological studies using aggregated temporal data (or time series studies) to be informative — are basically the same as those required for any valid and precise epidemiological study. Inadequacies in the data base, the sheer complexity of interactions among relevant variates, and other essentially non-methodological issues all contribute to the problem of inferring the dose-response relation between pollution and health. The HEADLAMP approach, however, avoids some of these problems by relying on established relationships, which are then applied at a local or national level to make inferences about excess risks. Thus it is not necessary to speculate on the biological mechanisms behind any associations found. Even within this framework, however, methodological difficulties may still arise. Appropriate application of statistical methods therefore needs to deal with the specific characteristics of time-series data on pollution and health.

A major drawback of a time-series design is the possible presence of unmeasured confounders. However, time-series studies of short-term effects, which use long series of short-interval data (e.g. days), often minimize such errors. An important feature in temporal studies using aggregated data is that the population being followed serves as its own control over time. Thus, possible confounders are only influential if they vary over the small time-intervals involved (e.g. from day to day). Such variations may certainly exist, for example in meterological data, but these can usually be accurately measured and easily taken into account in the analysis.

> While autoregressive models are often seen as the preferred method for advanced time series analysis, Poisson regression offers a simple and viable alternative.

4. RISK ANALYSIS

4.1 Introduction

Increasingly, authorities at the local, national and international level are faced with difficult decisions which involve weighing the social and economic benefits of technology against the health and environmental costs involved (McMichael, 1989). If these decisions are to be made on an informed basis, they require that health effects can be quantified. Often, indeed, some form of q*uantitative risk assessment* (QRA), is necessary for regulatory purposes. Moreover, because the

results of such assessments are often presented as a single number (for example, excess number of exposed disease cases), they give the appearance of scientific certainty and simplicity, both of which make the methods appealing to decision-makers. In practice, however, the ability to quantify the health effects of development is often limited and valid methods of risk assessment are both complex and uncertain. Methods of QRA, for example, are highly dependent on a series of assumptions and subjective choices which can have critical effects on the resulting risk estimates. Considerable care is therefore necessary in both using and interpreting results of QRA.

QRA can be defined as the application of a statistical relationship between exposure and the associated health outcome to assess either the health risk to a population or the exposure level associated with a given risk. Thus, two main types of quantitative risk estimate can be distinguished: *risk analysis (RA)*, which involves computation of 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 — and what is sometimes called *hazard analysis*, which involves calculation of the exposure/dose corresponding to a given level of risk — for example, the exposures estimated to cause adverse health outcomes in a certain percentage of exposed subjects.

QRA may be also applied at two different scales. *Individual risk* refers to the probability that an individual will develop a disease as a result of exposure in a specified time period. *Population risk* or disease burden refers to 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 that to a large population whose average exposure could be much lower.

4.2. Uses and Limitations of RA

Risk analysis is not a true linkage method in the sense that local health data are not utilized. Instead, it uses a predefined association between exposure and health outcome to determine the risk to an exposed population. The relationship between exposure and health is usually derived from independent studies, either within the study area or, more commonly, elsewhere.

The particular advantage of risk analysis is thus that it can be applied in areas where insufficient health outcome data are collected to allow the relationship between exposure and health to be locally determined. By the same token, risk analysis methods are the least resource-intensive, the easiest and the fastest to use of all the methods considered here. The success of the risk assessment process, however, depends on a number of issues such as the choice of the risk

prediction models and the adequacy of exposure assessment. All of these are subject to large uncertainties, though the exact form and magnitude of these problems vary depending on the particular context and purpose of the analysis.

> Risk analysis provides a relatively easy and rapid method of environmental health analysis, but the success of such assessments depends heavily on the validity of the risk prediction models and exposure estimates used.

One of the most important difficulties in QRA lies in obtaining reliable estimates of the exposure-response relationship. Results from epidemiological studies of one population cannot always be directly applied to others, due to differences in the range of exposures involved, in the methods of exposure estimation used, in the socio-economic contexts in which exposure occurs and in the baseline status of the populations concerned. A relationship for exposure to air pollution derived from a developed country or city, for example, is likely to underestimate the risks in developing countries, where the baseline health status is poorer (Ostro, 1994). Similarly, differences in the way in which exposure or health outcome are defined or measured in different areas (e.g. in the design of the pollution monitoring network, the specific definition of the pollutants measured, or in diagnosis) may make it difficult to transfer relationships from one area to another.

Particular care is also needed where the health outcome of concern is potentially related to more than one exposure. Both particulate matter and SO_2, for example, are known to contribute to respiratory diseases. In many areas, levels of the two pollutants are also highly correlated. When modelling the contribution of both, only one variable will remain statistically significant — the effect of the second will be subsumed within the first. When modelled separately, on the other hand, they may both show significant associations with health outcome. Summing these separate estimates of the effects will clearly exaggerate the estimated effect (e.g. the likely number of cases). Ideally, therefore, some measure of the combined effect should be obtained, by adjusting for the effect of the second exposure. In practice, this is often difficult, and in these cases a more conservative approach is to use only one measure of exposure — perhaps the one with the more complete data set.

A further source of uncertainty in QRA is the presence of population heterogeneity. In environmental health linkage analysis, risk factor data are usually collected and presented at high levels of aggregation. Aggregated risk estimates of this type can only be extrapolated back to the individual level if the population concerned is homogeneous. In reality, homogeneity within any

population rarely if ever exists. Unrecognized risk factors may be expected to subject different people to different background disease risks. As a result, individual risks may differ substantially from those implied by the aggregated data. Usually, variance estimators tend to be upwardly biased when risks are heterogeneous, rather than low. In undertaking a risk analysis, therefore, one should always check for hidden heterogeneity before presenting aggregate population statistics. If heterogeneity is discovered, then population risk estimates based on the aggregate data may be misleading. The populations should either be subdivided into more homogeneous subpopulations, or the statistics should be presented with due cautions for interpretation.

> Uncertainty in risk assessments should always be recognised, and estimates of environmental risks should always be accompanied by explicit estimates of the confidence interval or range estimate.

4.3 Presentation and Interpretation of Results

The results from any risk assessment clearly need to be communicated to the decision-maker in an appropriate form. This implies that the results are both clearly presented, yet also suitably qualified with regard to their reliability. The interpretation of the results, both by the risk assessor and the risk manager and later by the governmental and non-governmental organizations as well as the general public, may be critically dependent of the methods used to present the results. This is especially crucial in linking environment and health data since the decision-makers may not be well versed in the specialized statistical methods used. Moreover, there is the need to present the linkage results in such terms that they can be easily transformed to inputs for a societal or an individual cost-benefit analysis, or disseminated to other stakeholders (e.g. the public).

At present, there are no standardized procedures for analysing and presenting results from linkage analysis. To a large extent, this reflects the many different methods used to analyse the data, and the inherent differences in the data themselves. As a result, a standardized approach for the linkage analysis of environmental health data is often neither feasible nor necessary. It may not be feasible because of unresolvable differences in the data or methods available; it may be unnecessary because the study concerned does not involve comparisons across different areas or periods.

Standardization of methods is nevertheless beneficial insofar as it facilitates comparability. The diversity of analytical techniques so far applied in time series studies of air pollution and health, for example, has tended to hinder direct comparisons of the results, and made it difficult to derive general estimates of

exposure-effect relationships (e.g. from a meta-analysis). Lack of standardization also makes it difficult to verify the results of individual studies (e.g. by comparison with studies elsewhere) and reduces the opportunities to re-use the data at a later date. Standardization thus offers the possibility of obtaining added value from the data, and thus of improving the cost-effectiveness of data collection. One of the rare attempts to establish standardized procedures for time series analysis was the EU-funded APHEA project. This developed a standardized methodology to analyse data from 15 cities, representing a range of social, environmental and air pollution conditions across ten countries (Katsouyanni, 1995).

It is beyond the scope of this chapter to discuss how best to present the results of usual statistical analyses — they are well covered in many textbooks (e.g. Gore and Altman, 1982). Neither will the technicalities in quantifying human health risks be considered here — they have been described in books on risk assessment (e.g. Cox and Ricci, 1989). It is, however, useful to examine some of the general issues involved in the presentation of the results of linkage studies, as a basis for better informing the decision-maker.

The result of most interest to the health agency or risk manager in arriving at a decision is usually the quantitative estimate of risk. It is this risk estimate which provides the platform for subsequent policy action. In this context, two quantitative measures of risk are widely used:

- *increased individual risk* — the increased likelihood of an individual experiencing a specified health effect; and
- *disease burden* — the number of excess cases of the specified health effect ("body count").

Table 5.1, for example, shows the average worklife risk of lung cancer for an individual exposed to silica, while Table 5.2 shows the excess numbers of lung cancer in the exposed population for both 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 decremential risks. Thus, a risk-analyst might interpret the results of Table 5.2 as follows. Introduction of — and adherence to — an exposure standard of 0.2 $mg \cdot m^{-3}$ would produce a 12% reduction in the excess number of lung cancer cases. Alternatively, if the exposure standard was set at 0.1 $mg \cdot m^{-3}$, a 36% reduction would be predicted.

Table 5.1 Average risk of lung cancer for silica-exposed men employed from age 20 to 60 years. The predicted numbers and their associated 95% confidence levels are given separately for the currently prevailing exposure levels and for exposure levels up to 0.1 and 0.2 $mg \cdot m^{-3}$.

Exposure level	Estimated risk (%)	95% confidence interval (%)
Current	0.47	0.08-1.10
≤ 0.2 $mg \cdot m^{-3}$	0.42	0.08-0.90
≤ 0.1 $mg \cdot m^{-3}$	0.30	0.06-0.59

Source: Nurminen et al. (1992)

Table 5.2 Excess numbers of lung cancer in a dynamic population of 136,400 men exposed to silica in a 40-year follow-up period. The population was assumed to be stationary with respect to duration of exposure and age distribution. The predicted numbers and their associated 95% confidence intervals are given separately for the currently prevailing exposure levels and for exposure levels up to 0.1 and 0.2 $mg \cdot m^{-3}$

Exposure level	Estimated number	95% confidence interval
Current	630	120-1320
≤ 0.2 $mg \cdot m^{-3}$	550	110-1090
≤ 0.1 $mg \cdot m^{-3}$	410	90-780

Source: Nurminen et al. (1992)

The methods used for risk estimation inevitably give only approximate projections of risk, for they usually involve a myriad of assumptions, which cannot easily be verified. The presentation of simple point estimates of the expected risks and excess numbers thus tends to give a misleading impression of precision. Instead, it is important to provide clear information on both the

assumptions and limitations involved. Cox and Ricci (1989), for example, suggest the following guidelines for the presentation of risk estimates:

- 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 should be used to assess the effects on estimates of the key assumptions involved.

Linking environmental exposures to health outcomes is frequently achieved through the use of a regression model — for example, a multiple logistic regression. Whatever method is used, presentation of results after allowance for covariates should be in a form similar to that which would be used if no covariates were included in the risk function. Merely quoting the coefficients from the logistic model does not achieve this and is in any case artificial, since the logit transformation would not be necessary if there had 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. The analyst should, however, also provide 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).

> Whatever method is used for risk modelling, presentation of results after allowance for covariates should be in a form similar to that which would be used if no covariates were included.

A minor, yet more than cosmetic, point in presentation of results from risk analysis is the number of significant figures. In this context, the inherent precision of the results needs to be acknowledged. It is not sensible, for example, to give a result as "531.35 expected disease cases per year" when the

probable range is from zero to 1100. It might even be better not to give a single point estimate, but only to indicate the approximate confidence bounds. In presenting the results of a meta-analysis, the overall mean value can be shown along with the ranges for the lower and upper confidence limits.

Risk analyses frequently present information in terms of probability measures. Probability distributions can be difficult for a nonspecialist to interpret. While a plot of cumulative incidence rate (estimates of risk) allows one to read the median (and the percentiles of the distribution), the mean value cannot be determined from the plot. To avoid misinterpretations, therefore, it is important to present a plot of the cumulative distribution together with a graph of the incidence density curve, using the same horizontal scale, and to show also the mean risk on both curves (Ibrekk and Morgan, 1987).

To be of use for health policy making, epidemiologic data often need to be interpreted. Traditional epidemiology is mostly concerned with the increased incidence associated with exposure to a 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 *aetiologic fraction* — the proportion of the total incidence of the disease that can be attributed to that risk factor in the population. This indicates the proportion of incidence that could be prevented by the total elimination of that risk factor within the population.

In practice, prevention measures are rarely able to eliminate completely the prevalence of an environmental risk factor. As a result, a more useful measure is the *potential impact fraction* (Morgenstern and Bursic, 1982). This indicates the incidence that is avoided by a preventive intervention as a proportion of the incidence that would have occurred in that population without 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.

In the traditional epidemiologic literature, the term potential impact fraction is often used to imply an immediate elimination of excess risk after termination of exposure. In reality, this risk reduction may take many years to achieve, due to the lag effects involved. Ideally, therefore, estimates of effect should incorporate a time dimension. For this purpose, a methodology based on the preventive impact fraction has been developed (Gunning-Schepers, 1989). This comprises a computer simulation model, PREVENT (Gunning-Schepers et al., 1993), that can estimate the health benefits for a population of changes in risk-factor prevalence. Results are presented in graphical or tabular form and include: the intermediate output variates — aetiologic fraction, trend-impact fraction, and potential-impact fraction; and the 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. Indeed, in many cases, the effects are not expressed as real reductions in risk because of the demographic changes in the target population over time. This does not mean that prevention will have no beneficial effect. It does mean, however, that in order to see the effects it is important to show what would happen without the preventive intervention, and not merely to compare predicted effects with the current level of mortality. The potential utility of simulation models such as PREVENT in this respect lies in their ability to provide more precise quantification of effect estimates over time, and to take account of multiple risk factors and possible effects of demographic changes on the effects of intervention (Gunning-Schepers et al., 1993).

> The ideal measure of the health impact of a change in environmental exposure is the percentage of the incidence of a disease that would have occurred in the population without the change.

Although risk estimates produced by risk analysis have traditionally been used as the justifiable basis for regulating risks, the public's *perception of risk* is much broader than the "body counts" on which the quantitative risk assessments have focused. The public frequently misperceives risks because of the biases in the information to which they are exposed (e.g. the news media, government reports, industry reports). The public also perceives risk in a much wider context than that used in environmental epidemiology — perceptions reflect dread of the unknown, social and political impact, outrage and stigma. This difference in risk perception calls for two-way communication between the risk-analysts, risk-managers and other policy-makers, on the one hand, and the general public on the other (Morris, 1990). Useful guidelines and suggestions on how to communicate results of risk analysis to the public have been published by the U.S. Environment Protection Agency (Covello and Allen, 1988). These list "cardinal rules" for effective *risk communication*. In addition, a useful guide designed for industrial plant managers is available (Covello et al., 1988), which describes the technical information to be presented and provides guidelines for explaining risk-related numbers and risk comparisons.

5. GEOGRAPHIC INFORMATION SYSTEMS

5.1. Introduction

Analysing the links between environment and health is, by its very nature, a spatial problem. Levels of risk vary geographically in response to variations in

environmental conditions; health outcome and associated levels of need and health support vary as a consequence. Many of the questions facing the environmental epidemiologist and policy-maker are thus inherently geographical, and spatial analysis and mapping are vital components of their work. In research terms, they provide an important step in both the formulation and testing of hypotheses about links between environment and health. In policy terms — and thus in the context of HEADLAMP — they are a valuable means of directing policy to areas and problems of greatest need, and of monitoring policy performance and effects.

Until recently, spatial analysis and mapping in environmental health could only be carried out manually, or using relatively simple mapping packages. Over the last ten years, however, the capability for spatial data manipulation has been revolutionized by the development of *geographical information systems* (GIS). GIS can simply be described as systems for the collection, storage, manipulation and display of spatially-referenced data. As such, they have not only made mapping and many spatial analytical techniques much easier, but have also stimulated a wide range of new research into spatial operations and concepts that has greatly advanced our understanding of how to analyse and interpret spatial phenomena. GIS thus provide an increasingly important tool for environment-health linkage studies.

Establishing a GIS for environment-health applications is nevertheless far from easy. Since they were first developed in the 1970s, a wide variety of GIS have been developed. These differ in terms of their cost, functionality, complexity, power, underlying data models, and hardware and data requirements. The direct costs of GIS have declined rapidly in recent years, but the indirect costs — in terms of data acquisition, data cleaning and quality control, training and changes in operational structure of the organisation — can be substantial. In selecting and purchasing a GIS, therefore, particular care is needed to ensure that it can not only perform the tasks required with the types of data available, but that the system can be properly supported over time. This clearly implies the ability to foresee in advance both the uses and users of the GIS.

> In choosing a GIS it is important to ensure that the system can perform the tasks required, with the types of data available, and that the necessary resources are available to support and use the system effectively.

5.2. GIS Applications in Environmental Health

In the context of environment-health linkage studies, GIS offer a range of important capabilities. One of the most useful is the facility for efficient *map generation*. Digital data are captured either by digitizing or scanning, cleaned within the GIS to remove errors, and converted into map form. The wide range of display facilities available in GIS provide great flexibility in map representation and style, both as hard copy and on screen. Advanced search and retrieval facilities are also available, which allow maps to be interrogated interactively (e.g. to identify subsets of data based upon locational or other attributes). Many systems also provide limited statistical functionality, enabling results to be summarized in graphical or tabular form. Together, these make GIS highly influential systems for data presentation. This very persuasiveness is also a danger, however, for it can allow weaknesses in the underlying data or models to be easily disguised or overlooked. Maps are also subjective instruments; the message given by any map is dependent in part upon the way that the map has been compiled (e.g. map scale, projection, the class intervals or symbols and colour schemes used for mapping). For all these reasons, great care must be taken both in presenting and in interpreting results using GIS.

> GIS provide powerful and influential methods for data presentation. Because of their great potential for persuasion, however, it is essential to ensure that the results presented are based upon reliable data and well-validated models, and that the means of presentation are as objective as possible.

In many cases,, the generation of maps of environmental risk or health outcome involve more than simply the display of digital data. They also require some degree of spatial data transformation and analysis. One of the most useful techniques in this respect is *spatial interpolation*. This is often needed in two circumstances: to derive area coverages (i.e. continuous surfaces) from point data (e.g. to map air pollution based upon data from a network of monitoring stations); or to estimate conditions at unsampled sites (e.g. to estimate ambient pollution concentrations at the places of residence of cases and controls).

Modern GIS provide a range of spatial interpolation techniques, which can be classified in different ways. A common distinction is between proximal, local and global methods. Proximal interpolators (e.g. voronoi polygons) are the simplest in concept: they assume that the best estimate of the value at any unsampled point is provided by the nearest measured point. They thus assign the value of the nearest measured point to all intervening locations. Local interpolation methods fit regionalized functions through the points. They include

methods such as kriging and contouring. Global interpolation methods comprise those which fit a single mathematical function to all the data points — for example trend surface analysis. A distinction can also be made between mathematical interpolators and exact interpolators. The former fits smoothed surfaces through the data points. Exact interpolators fit surfaces which pass through the data points. The performance of the different interpolation methods depends upon a number of factors including the nature of the underlying spatial variation in the phenomenon under consideration and the sample density and distribution. In general, however, local methods of interpolation are to be favoured over global methods because they are more sensitive to local variations in the data and thus do not produce as much smoothing of the modelled surface.

Maps are a valuable means of presenting spatial data on environment and health, but the information they contain is not always immediately apparent. *Map analysis* is therefore an important step in the linkage of environment and health data. Its aim should be both to examine the integrity of the individual maps, and to determine whether any spatial variation or pattern actually exists. If no significant variation is discernible, there is clearly nothing to explain.

Two important approaches in this respect are the search for health *clusters* and *map-smoothing*. As yet, the use of GIS for map-smoothing has been limited — though Briggs et al. (1993) used map-smoothing techniques within a GIS to analyse infant mortality in Huddersfield, UK, and Elliott et al. (1995) used similar methods to examine variations in respiratory health in schoolchildren. The use of GIS for cluster searching has received rather more attention. Openshaw et al. (1987), for example, constructed what they referred to as a Geographical Analysis Machine (GAM) for cluster identification. This systematically constructed buffer zones around a fixed lattice of points in the study area. If the number of observed cases exceeded an expected number then a circle was drawn. Following repeated scanning with circles of different radii, the results were mapped, and locations which provided the focus for a large number of overlapping circles identified The method attracted considerable criticism, not least because it involves double-counting of individual cases and because of the difficulty of analysing the resulting maps, and a number of more robust approaches have thus been proposed (e.g. Besag & Newell 1991). These, however, have not yet been integrated into GIS.

GIS also clearly offer the opportunity for the spatial linkage and comparison of environmental and health data. The methods used depend to a large extent upon the measures of risk available. Where maps of pollution are available, it is possible to compare these directly with health outcome, using either overlay or point-in-polygon procedures. Where pollution or exposure has not been mapped, alternative indicators of risk may be used, such as the location of the emission

source. Analysing the links between pollution and health solely on the basis of distance from emission source is clearly an uncertain process. In general, stronger inferences can be drawn when a map is available of pollution level. In these circumstances, the pollution map may be used as an indicator of exposure, and compared to health outcome.

At its simplest, this may be achieved by *overlaying health and pollution maps*. From this, statistics can be compiled which compare the level of pollution with the standardized health outcome (e.g. SMR). A major difficulty of this approach, especially when applied to aggregated health data, is the mismatch which normally occurs between the spatial structure of the health data and pollution data. Overlay of health and pollution data — either in point or area form — also inevitably faces severe problems of confounding and effect modification (see Chapter 4.2.3). Unless these are taken into account, naive and misleading inferences may be drawn from simple map comparison.

As this implies, particular attention needs to be given to the quality of the data used in GIS. The problems lie not only in the source data themselves, but also in what happens to the data during GIS analysis. Linking different data sets to provide spatial coverages, or overlaying them to derive new information, for example, may generate complex and unseen error surfaces. Data quality control is thus of the utmost importance. The old adage of garbage-in-garbage-out is as true in GIS as in any other form of data analysis, but often less apparent because of the sophistication of the output, and the hidden complexity of the analytical operations involved.

Many of the operations performed by a GIS are run routinely with no transparency to the user. To integrate in one single output map the information coming from input maps from different sources each characterized by their own error structure can be hazardous and lead to error propagation. Other GIS operations, such as changes in the spatial resolution of a map or the discretization of continuous phenomena, are run routinely within a GIS context and yet can change substantially the results of any statistical analysis. Using data at too large a scale may simply be inconvenient; using data at too small a scale adds uncertainty and error to our analyses, and may generate false conclusions. Users therefore need to understand the spatial limitations of the data available in their GIS.

REFERENCES

Besag J, Newell J. The detection of clusters in rare diseases. *Journal of the Royal Statistical Society* 1991; A 154: 143-55.

Briggs DJ. GIS development for broad-scale policy applications: the lessons from CORINE. In: *Geographic Information 1991. The Yearbook of the Association for Geographic Information.* (J. Cadoux-Hudson and D.I. Heywood, eds.). London: Taylor and Francis, 1991. 113-20.

Briggs DJ, Elliot P. The use of geographical information systems in studies on environment and health. *World Health Statistics Quarterly* 1995;48:85-94.

Briggs DJ, Collins S, Elliott P, Martuzzi M. *Infant mortality in Huddersfield District Health Authority, 1982-1991. Report of a Study on behalf of West Yorkshire Health Authority.* Huddersfield: Institute of Environmental and Policy Analysis, University of Huddersfield and Environmental Epidemiology Unit, London School of Hygiene and Tropical Medicine, 1993.

Covello V, and Allen F. *Seven Cardinal Rules for Risk Communication.* Washington, D.C.: U.S. Environmental Protection Agency, 1988.

Covello VT, Sandman PM, Slovic P. *Risk Communication, Risk Statistics, and Risk Comparisons: A Manual for Plant Managers.* Washington, D.C.: Chemical Manufacturers Association, 1988.

Cox Jr. LA, Ricci PF. Risk, uncertainty, and causation: Quantifying human health risks. In: Paustenbach DJ, ed. *The Risk Assessment of Environmental and Human Health Hazards: A Textbook of Case Studies.* New York: Wiley, 1989. 125-56.

Elliott P, Briggs D, Lebret E, Gorynski P, Kriz B. Small area variations in air quality and health (the SAVIAH study): design and methods. (Abstract). *Epidemiology* 1995;6(4): S32.

Elliott P, Cuzick J, English D, Stern R, editors. Geographical and Environmental Epidemiology. Methods for Small-Area Studies. Oxford: Published on behalf of the WHO Regional Office for Europe by Oxford University Press, 1992.

Gore SM, Altman DG. *Statistics in Practice.* London: British Medical Association, 1982.

Greenland S. Divergent biases in ecologic and individual-level studies. *Statistics in Medicine* 1992;11:1209-23.

Gunning-Schepers LJ. The health benefits of prevention, a simulation approach. *Health Policy* 1989;12:1-256.

Gunning-Schepers LJ, Barendregt JJM, van der Maas PJ. PREVENT, a model to estimate health benefits of prevention. In: World Health Organization (WHO). Tools for Health Futures Research. Geneva: WHO, 1993:7-21.

Ibrekk H, Morgan MG. Graphical communication of uncertain quantities to nontechnical people. *Risk Analysis* 1987;7:519-29.

Katsouyanni K, Zmirou D, Spix C, Sunyer J, Schouten JP, Pönkä A, Anderson HR, Le Moullec Y, Wojtyniak B, Vigotti MA, Bacharova L. Short-term effects of air pollution on health: A European approach using epidemiologic time series data. (The APHEA project: Background, objectives, design). 1995.

Kuhn L, Davidson LL, Durkin MS. Use of Poisson regression and time series analysis for detecting changes over time in rates of child injury following a prevention program. *American Journal of Epidemiology* 1994;140:943-55.

Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13-22.

McCullagh P, Nelder JA. *Generalized Linear Models.* London: Chapman and Hall, 1983.

McMichael AJ. Setting environmental exposure standards: the role of the epidemiologist. *International Journal of Epidemiology* 1989;18:10-16.

Morgenstern H. Uses of ecologic analysis in epidemiologic research. *American Journal of Public Health* 1982;72:1336-44.

Morgenstern H, Bursic ES. A method for using epidemiologic data to estimate the potential impact of an intervention on the health status of a target population. *Journal of Community Health* 1982;7:292-309.

Morris S. *Cancer Risk Assessment. A Quantitative Approach.* New York: Marcel Decker, Inc., 1990.

Nurminen M. To use or not to use the odds ratio in epidemiologic analyses. *European Journal of Epidemiology* 1995;11:365-71.

Nurminen M, Corvalán C, Leigh J, Baker G. Prediction of silicosis and lung cancer in the Australian labor force exposed to silica. *Scandinavian Journal of Work Environment and Health* 1992;18:393-9.

Openshaw S, Charlton M, Wymer C, Craft A. A Mark 1 Geographical Analysis Machine for the automated analysis of point data sets. *International Journal of Geographical Information Systems* 1987; 1: 335-58.

Ostro B. Estimating the health effects of air pollutants. A method with an application to Jakarta. The World Bank. (Policy Research Working Paper 1301.) May, 1994.

Robinson WS. Ecological correlations and the behavior of individuals. American Sociological Review 1950;15:351-7.

Rothman KJ. A sobering start for the cluster busters' conference. *American Journal of Epidemiology* 1990;132, Suppl. 1;S6-13. (Keynote presentation at the National Conference on Clustering of Health Effects, Atlanta, GA, February 16, 1989.)

Selvin HC. Durkheim's "suicide" and problems of empirical research. *American Journal of Sociology* 1958;63:607-19.

Zeger Sl, Qaqish B. Markov regression models for time series: A quasi-likelihood approach. *Biometrics* 1988;44:1019-31.

Decision-making in Environmental Health

E. Schwartz[a] *& C. Corvalán*[b]

1. INTRODUCTION

Environmental health and epidemiology aims at preventing needless morbidity and mortality by protecting people from unnecessary exposure to environmental hazards. Unfortunately, despite the increasing knowledge about potentially harmful exposures that is embodied in a rich and expanding literature in environmental epidemiology, preventive action is often slow to materialise. The mismatch between knowledge and application or translation is often most acute in developing countries, where environmental and occupational exposures exceed national and international guideline levels by a considerable amount, yet where corrective action to control these problems is limited.

To reduce this growing deficit of action, results from environmental epidemiologic studies need to be translated more effectively and efficiently into public health practice. This requires that the epidemiologist provide the right type of information, and communicate it to the decision-maker in an easily understandable form. Better tools to help the decision-maker use the available epidemiological data also need to be developed.

Decision-making involves choosing among alternative ways of meeting objectives (Warner et al., 1984). Implicit in this definition is the notion that there are a number of alternatives, and that their effects can be measured or estimated. This, in turn, implies that there is adequate information to make an informed choice. Decision-making, however, is rarely about single objectives. Often a number of competing or conflicting objectives may exist. In the context of public health, for example, the objective may be not only to protect or improve health status, but also to maximize productivity and reduce costs. Moreover, there may be limited or inadequate information on the potential impact or costs of various policy alternatives — or even on what policy options are available. Together, this uncertainty and conflict often produces diverse conclusions about the "best alternative" when viewed by different observers.

HEADLAMP aims at providing the necessary tools for the management of environmental health problems. The main contribution of HEADLAMP to the

---

[a]Formerly, Office of Global and Integrated Environmental Health, World Health Organization, Geneva

[b]Office of Global and Integrated Environmental Health, World Health Organization, Geneva

decision-making process is in the creation and improvement of the information base required by those responsible for defining the policy actions needed to avoid or reduce adverse environmental health effects. Within this context, the development and use of environmental health indicators is fundamental to the decision-making process, for these provide the information on the state of, and trends in, environmental health that is essential for policy-related action and for monitoring the effects of policy implementation (WHO, 1994).

2. PUBLIC HEALTH PLANNING AND PRIORITIZING

Within the context of decision-making, an important distinction needs to be made between decisions (products) and the techniques (process) used to make them. Decisions are the final judgments regarding what actions should or should not be taken. Techniques refer to the methods used to reach a decision. These may take a variety of forms, including public debate, expert consensus conferences (e.g. Delphi techniques — Richey et al., 1985), algorithms or decision-trees (e.g. as part of a rule-making procedure), knowledge-based or expert systems, and mathematical modelling (Cross et al, 1995). Implicit in decision-making is a value system. For example, two common values systems focus on maximizing equity or utility. The egalitarian approach seeks to attain equal risks borne by all sectors of the population: on a global scale, this would require parity, for example, in the infant mortality rate among all nations; within a country, it would mean similar rates among different geographic or demographic groups. The utilitarian approach seeks to minimize risks or costs and, at the same time, maximize the benefits. Because this approach only places value on maximizing output per unit input, it is possible that inequalities in health between population groups may be produced or even accentuated.

The most common use of public health decision-making is in the process of planning and prioritizing public health programs. Planning involves three main steps: determining the current status of the programme (where we are now); determining the ultimate objectives of the programme (where do we want to be); and determining how we can get from where we are to where we want to be. Prioritizing involves valuing or ranking a set of possible programmatic alternatives; for example, choosing where to allocate resources, in what order to undertake a range of tasks, or which approach to adopt in any situation.

Epidemiology can contribute much of the information needed to accomplish national public health planning and priority setting. Epidemiological criteria, for example, are commonly used for evaluating programmes — for example, in terms of their effects on exposure or expected health outcome (e.g. prenatal care or infant mortality). In these cases, however, considerable care needs to be taken in selecting the epidemiological criteria used, since this choice is itself often value-laden (Schwartz, 1987). The use of relative outcome measures (e.g.

relative risk), for example, implies emphasis on equity considerations, for these will identify populations that experience higher risk than the comparison groups. Nevertheless, this approach may not necessarily produce the most effective use of resources if, for example, a disproportionate effort is needed to reduce risks in a relatively small high-risk group.

In contrast, use of a criterion such as the marginal number of avoidable cases or deaths seeks to maximize the impact of the intervention, providing the greatest good for the greatest number of individuals (a utilitarian goal). In many cases, more complex and subtle values may be inherent in the criteria used: potential years of life lost (PYLL), for example, used either as a relative or absolute measure, implicitly places greater weight on younger populations.

Each of these measures will clearly tend to produce different results in any given situation. As a result, the decisions taken are highly dependent on the epidemiological criteria used. This, in turn, means that the public health agenda will be a function of the epidemiological measures used to inform decision-making (Schwartz, 1987). The amount and type of information available is therefore a major driving force for policy.

3. COMMUNITY-BASED APPROACH VERSUS TARGETING

The goal of a community-based intervention is to shift the entire distribution of a risk factor (e.g., blood-lead level, hypertension) towards the desired target level (Figure 6.1). The extent to which any action will achieve this goal depends on:

- the shape of the dose-response curve,
- the likelihood that small decrements in exposure will produce concomitant changes in the outcome of individuals, and
- the marginal degree of preventability at various levels of exposure.

The existence of a non-linear dose-response curve or "threshold effect", for example, might yield very different beneficial results for a given intervention targeted at different levels of exposure. Equally, targeting the highest risk groups (those in the upper tail of the distribution curve) might be more difficult or more expensive than focusing on other segments of the distribution — for example, aiming a campaign to reduce smoking at long-term, "hard-core" smokers is likely to be less cost-effective than focusing an intervention on those smokers who recently took up the habit.

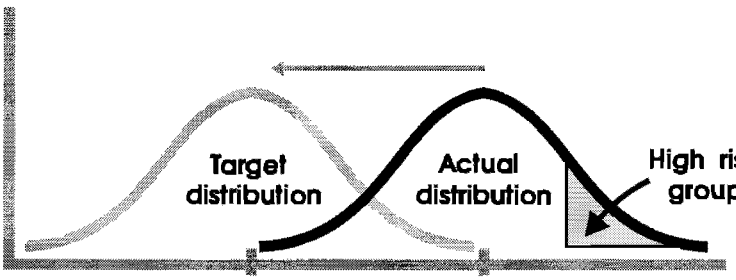


Fig. 6.1 The goal of a community based intervention

4. PARADIGMS AND TOOLS FOR DECISION-MAKING

Decisions are taken within both a social and a technological context. They are influenced by the belief systems of those who make decisions, and those for whom the decisions are made, and by the various external (e.g. political and cultural) pressures to which those involved may be subject. They may also be supported and guided by a wide range of techniques and tools.

An inherent belief in the primacy of the environment, for example, may mean that a decision-maker is willing to rely solely on exposure data, without reference to evidence of associated adverse human health outcomes. In this context, any increase in environmental pollution is likely to be seen as intrinsically bad, and adverse effects on health will be assumed. A commitment to the protection of scarce natural resources or to securing the environment for future generations, or simple distrust of science and modern technology — all may underlie such belief systems.

In contrast, many decisions are taken within a more strongly "accounting" or "optimization" frameworks. In these, risks and health gains, or costs and benefits, are seen as transferable commodities that can be set against each other. Various approaches have been developed within recent years based on this principle.

4.1 Risk Trading

Risk trading (or *risk substitution*) represents willingness to trade-off different risks in the search for an optimal or acceptable solution. This is based on

recognition that almost any decision is likely to produce a range of different, and often conflicting effects, that need to be compared. One example of the application of risk trading is provided by the effects of reducing the use of chemical preservatives in food or water. Prohibiting the use of nitrates as a food preservative because of their potential for increasing cancer risk, for example, must be weighed against the increased risk of *Clostridium botulinum*, a bacteria that causes a potentially fatal form of food poisoning. Similar examples occur in many other areas of public health and environmental management — for example, chlorination or fluoridation of water, use of pesticides, use of nuclear energy for power generation and others (de Koning, 1987).

In all cases, the tradeoffs involve some form of weighting of the different effects which are perceived. These weights are inevitably difficult to define, and tend to vary between different groups of people, and even between individuals, depending upon both personal and contextual factors — e.g. economic status, age, religious or other belief systems, level of education and past experience. The tradeoffs are thus far from universal, and they commonly engender conflict between the different stakeholders involved — for example, when interventions designed to maximize public health benefits require restricting an individual's right to privacy (e.g., compulsory vaccinations, use of seat belts, motorcycle helmets, STD contact tracing, etc).

The weights and preferences implied in risk trading are also difficult to assess. In recent years, methods of *contingent valuation* have been widely adopted in environmental science for this type of assessment. This involves asking respondents to evaluate (using some agreed scale or currency) alternative scenarios or outcomes. Similar preference methodologies can also be applied in the context of environmental health. People living near a smelter, for example, state their preferences between achieving lower environmental arsenic levels (by ceasing smelter operations, with potential adverse economic impact) or maintaining current levels in the face of risk of future lung cancer. In making the valuations, people are likely to be influenced not only by the direct environmental and health benefits on offer, but also by the potential economic and social effects — e.g. loss of employment and impacts on the local community.

In practice, however, these preference valuation methods are extremely difficult to apply in a rigorous and reproducible form, for responses depend to a great extent upon what information and supporting evidence is supplied to the respondent, how this information is presented, and even the emotional state of the person at the time. As a result, valuations may be highly unstable and open to manipulation. Moreover, the valuations obtained by such methods do not always seem to translate into personal action when the opportunity arises, implying that they do not give reliable measures of deep-seated preferences.

4.2 Cost-Benefit Analysis (CBA)

As its name implies, cost-benefit analysis is a process to weigh the benefits of an intervention against the costs by assigning monetary values to both. In the area of environmental health, it could, for example, be used to evaluate proposed measures to reduce air pollution. The costs of these measures -- in terms of both the direct costs of policy implementation and the indirect costs such as reduced industrial profits — would first be assessed. These would then be compared with the sum of the benefits (e.g. reduced costs of medical care and environmental remediation, improvements in productivity, avoided pain and suffering and environmental improvements. The ratio of the two would then be used to assess the viability of the policy, or to compare it with other possible strategies.

One example of the use of cost-benefit analysis is the recent study in the USA which examined the social benefits of reducing lead exposure (Schwartz, 1994). This estimated that a 1 μg/dl decrease in the mean blood-lead-level concentrations in children would produce at least $5 billion per year in benefits. Saved costs included medical costs and the costs for remedial education. A major estimated benefit was due to increased lifetime earnings reflected by increased IQs. Among adults, benefits included the reduction of lost wages due to hypertension, heart attacks, strokes and premature mortality. Similarly, Jarman (1994) estimated medical costs and absence from work due to road traffic in Oxford, UK totalled about £18 million/year.

Cost-benefit analysis has a long pedigree, and is widely used in many different policy areas. In many cases, it is a mandatory part of the policy process. Nevertheless, it faces many conceptual and practical problems, which may limit its acceptability in the area of environmental health. One of the most important is the difficulty in obtaining full estimates of costs and benefits. Clearly, if this is not done, the cost-benefit ratio is likely to be misleading, in that it is based upon only a partial view of the impacts involved. The problem is often most severe with regard to the benefits, many of which are intangible, remote and not easily converted to monetary values. Various methods have been developed to resolve this difficulty. One of the most widely used is *hedonic valuation*, which assesses people's "willingness to pay" for the benefits available (e.g. reduced risks or improved quality of life). These are normally based on surveys of the stakeholders involved. The extent to which such methods provide true estimates of the long-term social value of the benefits involved is, however, dubious, not least because many of the benefits may accrue to future generations or to sections of society (e.g. immigrants) not included in the survey. Moreover, these methods are only applicable to those costs or benefits which are both recognized by the analyst and are perceivable by the stakeholders concerned. In the case of new risks (e.g. health risks associated with biotechnology), lack of understanding and experience undoubtedly limits the ability of the public to

make meaningful valuations. For this reason, valuations may also be highly susceptible to influence and manipulation, either by the media or by the way in which the questions are posed. As noted earlier, the public tends to perceive risk in a very broad and often personal way, so that it is not easy to identify exactly what factors are being taken into account in any particular case. This leads to the potential for double-counting of costs and benefits (e.g. some impacts may be included in more than one set of valuations) and the omission of others.

A further problem with cost-benefit analysis is how to determine the "discount rate" for consequences (e.g. deaths) that are expected to occur in the future. Since many costs and benefits may be very long-lasting, the choice of discount rate can have an over-riding influence on the outcome of the assessment. In addition, cost-benefit analysis has a tendency to reduce the weight given to poorer sectors of society, since they are often able (and therefore willing) to pay only relatively small amounts for the potential benefits. For this reason, CBAs performed in countries at different stages of economic development may lead to quite different outcomes.

4.3 Cost-Effectiveness Analysis (CEA)

In view of the limitations of CBA, many people have tended to turn to cost-effectiveness analysis. It attempts to assess the marginal impact of a group of options: the "preferred" action would therefore be the one which provides the greatest effectiveness (e.g. in health status of a population) given the same investment of resources and effort. It thus defines a specific objective or goal which is considered worthwhile in social or other terms, and then compares different strategies for reaching that goal. In the context of environmental health, for example, it might be used to compare alternative forms of pollution control aimed at achieving a defined reduction in health risk; or, it might be used to assess the relative effectiveness of pollution control and awareness raising as a means of reducing exposure by a specified amount.

As such, cost-effectiveness analysis is clearly more limited in scope than CBA, and avoids many of the difficulties which CBA faces. Nevertheless, the very restrictiveness of cost-effectiveness analysis means that it is difficult to apply in complex situations, where different strategies may have different side-effects and secondary costs or benefits. For example, reducing emissions of an air pollutant may appear to be more cost-effective as a means of reducing exposure than an educational programme aimed at awareness raising. The latter, however, may lead to much more general health improvements due to wider public consciousness. Similarly, emission controls may have a wide range of additional non-health benefits — e.g. reduced environmental damage. Unless these are

considered, the wider effects of any action will not be taken into account, and decisions will tend to be made in too reductionist a manner.

4.4 Other Methods

A range of other methods are also available to support policy formulation and decision-making. These include distributional analysis, Bayes' rule, decision trees, sensitivity analysis, policy models, expert systems and decision support systems.

Distributional analysis involves the assessment of the distribution of costs and benefits within a community (de Koning, 1987). It aims to determine the extent to which these are balanced across the community, or whether some sections of society are subject to an unfair ratio of costs to benefits. The benefits of water treatment, for example, will not help those without access to piped water, and in many cases may even increase the risk of water contamination (as well as the costs) if water has to be purchased from vendors. Results of such an analysis may then be used to determine how the costs of the intervention should be shared, or whether additional action is required to help those who have not initially benefited.

Bayes' rule is based on the principle of prior and conditional probabilities, and has been used in clinical medicine to interpret the usefulness of diagnostic or screening tests in various populations. By the same token, Bayes' rule might provide guidance when determining the utility of conducting environmental exposure assessments or implementing environmental epidemiologic studies.

Decision trees can help identify all available choices and their potential outcomes by structuring a branching model of the alternatives. Using known or estimated probabilities of each option at each branch (node) of the tree, the relative utility of various strategies can be calculated. The strategy with the highest utility would be expected to provide the best choice. The main difficulty with this approach clearly lies in evaluating the utilities of the various choices in a comparable way. It is also a largely reductionist method, in that it does not easily allow for interactions or contingency between the various decisions — e.g. the utility of a decision at one level in the decision tree may be dependent upon both prior and subsequent decisions.

Sensitivity analysis refers to a wide range of methods which may be used to assess the robustness of any model or decision process. Typically, it involves reiterative assessment of the outcomes, subject to controlled modification of different elements or steps. It may be used in association with a decision tree, for example, to determine how sensitive the results are to decisions at specific points within the tree. It may also be used with more quantitative modelling

techniques, to assess the sensitivity of the model to the input data or to assumptions within the model.

Various *statistical models, expert systems* and *decision support systems* are also becoming widely available in relation to environmental health policy (Cross et al, 1995). The outcome of a policy can, for example, be modelled as a chain of events occurring over time. The sequence of events can be represented by a decision tree, each event having a given probability. The sequence can be repeated to determine how the impact changes over a long period of time. This type of model can be mathematically represented and analysed using *Markov chains* (Beck and Pauker, 1983). The software programme CANTROL, for example, has been developed to assist with public health decision-making in the area of cancer prevention and control (Eddy, 1986). Given various community demographic and baseline cancer incidence and mortality data, the programme estimates the potential impact of various levels of cancer screening or smoking prevalence. Expert systems work in a broadly similar way. Geraghty (1993), defined an expert system as a system which simulates:

> "the means by which a human expert tackles real-world problems using a set of rules, heuristics and inferences, programmed into a computer system"

The rules are normally derived from prior experience or from simulation models, and they are used to guide the decision-maker through the choices available, indicating at each stage the implications involved. In a full decision support system, these rules may be further supplemented by a range of supporting information (e.g. standards, guideline values, definitions), analytical facilities (e.g. simple programmes to compare the effects of different choices or to compute simple statistics) and display facilities (e.g. graphs and maps). Expert systems have been widely developed in the area of environmental policy (e.g. Geraghty, 1993), but have as yet attracted less attention in the area of environmental health. One example of a system which includes health considerations is the ISIS (Integrated System for Implementing Sustainability) which has been developed as a prototype to guide transport planning and policy (Hopkinson et al., 1994).

5. TOWARDS MORE EFFECTIVE DECISION-MAKING

5.1 Problems of Decision-making

As the preceding discussion has highlighted, the decision-making process is far from simple, and one in which numerous conflicts and uncertainties arise. One of the basic conflicts derives from the inexact nature of the process: while the public and politicians tend to expect swift and clear-cut solutions, the world

itself is extremely complex, often unpredictable and as yet poorly understood. As Steenberg (1989) has stated, there is no definable boundary between what is safe or hazardous, but rather a zone of uncertainty. In many cases, therefore, we can do no more than talk in terms of the "probability" of an effect being produced. Given the limited public understanding of statistical probabilities, such language is not always appropriate or readily accepted.

Decision-making is also bounded by a number of other constraints. Amongst these are the problems of:

- quantifying the extent to which prevention can be obtained;
- extrapolating from evidence derived at high doses to determine risk at lower doses;
- extrapolating from data derived from animal evidence to determine human risk;
- extrapolating from past or current data to future generations;
- allowing for variations in individual susceptibility;
- ensuring adequate control for all possible confounders;
- allowing for the effects of combinations of exposures and multiple routes of exposure;
- the unreliability of many of the models used, and the difficulties of model verification;
- gaps, inconsistencies and errors in many of the data used;
- determining true probabilities; and
- defining and valuing intangibles such as quality of life.

Setting clear guidelines to facilitate the decision-making process is therefore not a simple endeavour. All the items above are subject to interpretation, and even experts are likely to disagree regarding both the weight to allocate to each and the conclusions to which they point.

5.2 The Decision-making Process

Decision-making is a social and organic process. There is, as a result, no single model of the process which applies in all circumstances. The process varies substantially depending upon the organizational and administrative context, the issues and individuals involved, levels of knowledge, local perceptions and beliefs, and many other factors; for example:

- Value placed on health, human-life extension and environmental protection, concern for future generations
- Strength of data, extent of documentation
- Public understanding of data and perceptions (acceptability) of risk
- Costs of intervention: are they affordable?

- Leadership: ability to persuade/motivate, negotiate, resolve conflicting goals or competing interests
- Process that provides a forum for debate and permits input into public policy setting
- Emphasis on planning for the future, government responsibility for protecting public
- Degree of collaboration: government/business/non-government organizations
- Regulatory process
- Judicial process
- Seriousness of the outcome
- Involvement of Mass Media
- Targeted message for decision-makers

Figure 6.2, however, sets out a simple model of the decision process. Six main steps are defined: 1) stating the problem; 2) evaluating existing information; 3) identifying the need for new information; 4) identifying and evaluating options and alternatives; 5) making the decision, and 6) evaluating the impact of the decision.

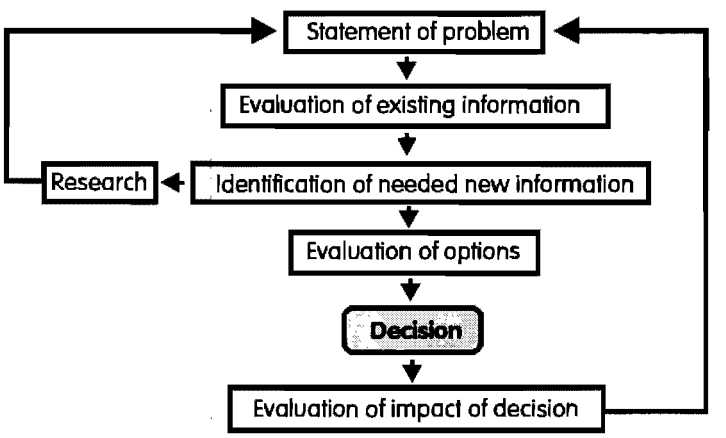


Fig. 6.2 Basic process of decision analysis
(Modified from Steensberg, 1989)

This concept of the decision process is a traditional one, and is information-driven. Information feeds the decision process — helping those concerned to identify the problems which need addressing and then to evaluate the options available.

Most decisions involve, and impinge on, a wide range of stakeholders and actors. These typically include:

- scientists — who may be involved in the initial research which identified the problem, and in helping to devise solutions
- business and industry — which may be implicated in the cause of the problem and may be partly responsible for implementing and financing solutions
- planners — who may be involved in translating general policies into local action, and in monitoring implementation
- the media — which may be involved in raising awareness about the problem and act as an unofficial watchdog on the actions taken
- politicians — who are charged with making the decisions
- the public — who in the end must accept, pay for and live with the results of the decisions made

Each of these groups is likely to have different agendas. Each of these will be moulded by a wide range of economic, professional, political and bureaucratic pressures. Consensus about the levels of risk involved, or about the relative merits of different policy actions, is therefore difficult to achieve (McMichael, 1991).

Nevertheless, the need to involve these various actors and stakeholders at all stages in the decision process should not be treated lightly. Some questions, for example, are unanswerable in strictly scientific terms because of gaps in our knowledge; in these cases, a dialogue with the community is essential in order to reach a mutually agreeable solution. Science can provide guidance but not provide all the answers. An open and participatory approach is more likely to make the results more credible and acceptable, to provide time for the community to consider in advance the technical concepts, and limitations and range of outcomes, and thus to allow decisions to be taken and implemented more effectively and speedily (Ozinoff and Boden, 1987).

In this context, de Koning (1987) notes five characteristics of an effective standard-setting process which can be applied generally to decision-making in the area of environmental health:

- Involve the major parties in the community, including politicians, citizen groups, industrial leaders and health officials. This should stimulate debate encompassing differing perspectives and values, leading to some compromises being made in both goals and methods, thus ensuring broad support in the society at large.

- Provide a mechanism through which technical and policy analysis can be generated, distributed and criticized.

- Provide a mechanism whereby the results of analyses can be presented to policy-makers and the other centres of interest in the society, to inform these groups of the costs, benefits, and impact of the proposals under consideration.

- Provide a mechanism for conflicting interests to be heard and discussed in a controlled manner, so that divergent opinions in the society can be aired and, as far a possible, accommodated in the implementation of the proposal.

- Provide a mechanism whereby the society can reach a decision and take useful action, even though such action may be less than what is "objectively" ideal.

5.3 The Role of Epidemiologists

These principles clearly provide a strong framework for effective decision-making in the area of environmental health. It is a framework within which epidemiologists must play an active and leading part. Epidemiologists are not only researchers but also practitioners of public health. As such, they have an ethical obligation to do more than conduct scientific inquiries and publish their results. They have specific obligations to the subjects they study (e.g. informed consent, maintaining confidentiality); to society more widely (e.g. to share their information and study results); to colleagues; to employers; and to research sponsors (Cook, 1991). For all of these, they need to ensure that they apply the following principles:

- Helping to make the problem explicit and clear;
- Implementing studies to answer relevant questions;
- Communicating and interpreting epidemiologic data for non-epidemiologists. The public has a right to know and the epidemiologist's responsibility is more than just publishing results in a journal;
- Being a good citizen and member of the community where they live and work. As Gordis (1991) has said, epidemiologists "should not hide under the mantle of scientific objectivity and detach ourselves from critical decision-making in public health".

6. CONCLUSIONS

Decision-making is not a simple process. A decision-maker must choose between competing alternatives, and may face uncertainties at every step. These difficulties, however, are no excuse for lack of action. There is a clear

mismatch between the sophistication in public health research and that of decision-making in public health and environmental issues (Schwartz, 1994). One way of rectifying this gap is to involve epidemiologists in the process of addressing the solutions to the problems they study. This would entail a change in both attitude and in training of environmental epidemiologists.

There is, sadly, commonly a gulf between the scientist and the decision-maker. Scientists are often reluctant to become embroiled in what has been called "the corridors of power" for fear of prejudicing their scientific objectivity (and simply because of lack of time). Those with decision-making responsibilities are not expected to be directly involved with the scientific technicalities behind the information they use. This cultural gap between science and decision-making is clearly not healthy. It requires both groups to attempt to close it. Scientists need to make more effort to convert their knowledge and the results of their research into a language and a form which the decision-maker can more clearly understand, though without reducing their scientific veracity. Decision-makers need to learn new ways of thinking and of evaluating information on health. They have a responsibility to understand both the value and the uncertainties and limitations of the information available to them.

Decision-making also requires the availability of better information and knowledge on the links between environment and health. Yet we cannot afford to delay while this information and knowledge is gathered, for while we wait suffering continues (Sandman, 1991). We must be prepared, therefore, to act with the data and methods we have. As Bradford Hill (1965) noted:

> *"All scientific work is incomplete — whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge that we already have, or to postpone the action that it appears to demand at a given time."*

It is within this context that HEADLAMP has been conceived. The aim of HEADLAMP is to help provide the tools needed to extract more information, more quickly, out of the data which already exist — and where adequate data are not available, to collect them speedily. Its aim is to improve the utility of the information gained by providing results in a form directly usable by the decision-maker. Its aim is to encourage epidemiologists and decision-makers to work more closely together, and for both to interact more openly with the public and other stakeholders concerned.

REFERENCES

Beck JR, Pauker SG. The Markov process in medical prognosis. *Medical Decision Making* 1983;3:419-57.

Cook R. Code of Ethics for Epidemiologists. *Journal of Clinical Epidemiology* 1991; 44 (Supplement 1):135S-139S.

Cross SS, Harrison RF, Kennedy RL. Introduction to Neural Networks. *Lancet* 1995; 346:1075-79.

de Koning HW. Setting environmental standards. *Guidelines for Decision-making*. Geneva: WHO, 1987.

Eddy DM. A computer-based model for designing cancer control strategies. *National Cancer Institute Monograph* No. 2, 1986.

Geraghty PJ. Environmental assessment and the application of expert systems: an overview. *Journal of Environmental Management* 1993;39:27-38.

Gordis L. Ethical and Professional Issues in the Changing Practice of Epidemiology. *Journal of Clinical Epidemiology* 1991;44(Supplement 1):9S-13S.

Harris JE. Environmental Policy Making: Act Now or Wait for More Information. In: National Research Council, "*Valuing Health Risks, Costs and Benefits for Environmental Decision Making*". Washington, D.C.: National Academy Press, (1990).

Hill AB. The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine* 1965;58:295-300.

Hopkinson P, Webber P, Briggs DJ. Developing an environmental management system for traffic pollution impacts. *Transportation Planning Systems* 1994;2:39-52.

Jarman M. Valuing wider environmental benefits from an urban traffic restraint package. *Transportation Planning Systems,* 1994;2: 23-37.

McMichael AJ. Setting environmental exposure standards: current concepts and controversies. *International Journal of Environmental Health Research* 1991;1:2-13.

Ozinoff D, Boden LI. Truth and Consequences: Health Agency Responses to Environmental Health Problems. *Science, Technology & Human Values* 1987;12:70-7.

Pauker SG & Kassirer JP. Decision Analysis. *New England Journal of Medicine* 1987;316:250-7.

Richey JS, Mar BW, Horner RR. The Delphi technique in environmental assessment. Parts 1 and 2. *Journal of Environmental Management* 1985;21:147-59.

Sandman P. Emerging Communication responsibilities of epidemiologists. *Journal of Clinical Epidemiology* 1991;44(Supplement 1):41S-50S.

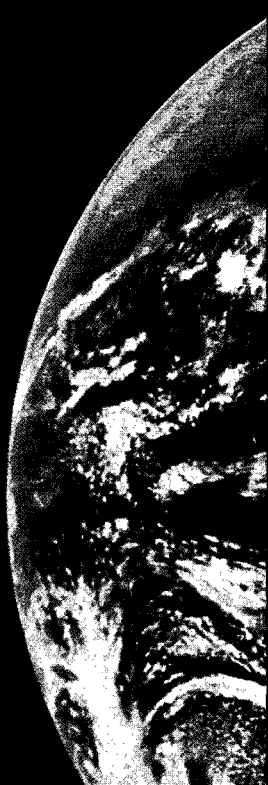
Schwartz E. Premature mortality in New Hampshire. *Morbidity Mortality Weekly Report* 1987;36:765-6.

Schwartz J. Societal Benefits of Reducing Lead Exposure. *Environmental Research* 1994; 66:105-24.

Steensberg J. *Environmental Decision Making: The Politics of Disease Prevention*. Copenhagen: Almqvist & Wiksell International, 1989.

Warner DM, Holloway DC, Grazier KL. *Decision Making and Control for Health Administration.*. Ann Arbor: Health Administration Press, 1984.

WHO. Informal consultation on Health and Environment Analysis for Decision-making (HEADLAMP) methods and field studies. Document No. WHO/EHG/94.15. Geneva: WHO, 1994.



Office of Global and Integrated Environmental Health
World Health Organization