
Chapter 1

COMPARATIVE QUANTIFICATION OF HEALTH RISKS: CONCEPTUAL FRAMEWORK AND METHODOLOGICAL ISSUES

CHRISTOPHER J.L. MURRAY, MAJID EZZATI,
ALAN D. LOPEZ, ANTHONY RODGERS AND
STEPHEN VANDER HOORN

1. INTRODUCTION

Detailed description of the level and distribution of diseases and injuries, and their causes are important inputs to strategies for improving population health. Data on disease or injury outcomes alone, such as death or hospitalization, tend to focus on the need for palliative or curative services. Reliable and comparable analysis of risks to health, on the other hand, is key for preventing disease and injury. A substantial body of work has focused on the quantification of causes of mortality, and more recently, the burden of disease (Murray and Lopez 1997; Preston 1976). Analysis of morbidity and mortality due to risk factors, however, has frequently been conducted in the context of methodological traditions of individual risk factors and in a limited number of settings (Kunzli et al. 2000; Leigh et al. 1999; McGinnis and Foege 1993; Peto et al. 1992; Single et al. 1999; Smith 2000; Smith et al. 1999; Willet 2002). The principal conclusions of this body of work are as follows:

- Causal attribution of morbidity and mortality to risk factors has been estimated relative to zero or some other constant level of population exposure. This single, constant baseline, although illustrating the total

magnitude of the risk, does not provide visions of population health under other alternative exposure distribution scenarios.

- Intermediate stages and interactions in the causal process have not been considered in the causal attribution calculations. As a result, attributable burden could be calculated only for those risk factor–disease combinations for which epidemiological studies had been conducted (often limited to individual risks).
- Causal attribution has often taken place using exposure and/or outcome at one point in time or over an arbitrary period of time (for notable exceptions see the works of Manton and colleagues [Manton et al. 1993b, 1994; Yashin et al. 1986] and Robins [Robins 1986, 1987, 1999a, 1999b; Robins and Greenland 1991; Robins et al. 1999]). Such “counting” of adverse events (such as death) has not been able to clearly distinguish between those cases that would not have occurred in the absence of the risk factor and those where occurrence would have been delayed. More generally, this approach is unable to consider the accumulated effects of time-varying exposure to a risk factor—in the form of years of life lost prematurely or lived with disability.
- The outcome has been morbidity or mortality due to specific disease(s) without conversion to a comparable unit, making comparison among different diseases and/or risk factors difficult.

To allow the assessment of risk factors in a unified framework while acknowledging risk-factor specific characteristics, the comparative risk assessment (CRA) module of the Global Burden of Disease (GBD) 2000 study is a systematic evaluation of the changes in population health which would result from modifying the population distribution of exposure to a risk factor or a group of risk factors (Murray and Lopez 1999). This unified framework for describing population exposure to risk factors and their consequences for population health is an important step in linking the growing interest in the causal determinants of health across a variety of public health disciplines from natural, physical, and medical sciences to the social sciences and humanities. In particular, in the CRA framework:

- The burden of disease due to the observed exposure distribution in a population is compared with the burden from a hypothetical distribution or series of distributions, rather than a single reference level such as the non-exposed population.
- Multiple stages in the causal network of interactions among risk factor(s) and disease outcome are considered to allow making inferences about combinations of risk factors for which epidemiological studies have not been conducted, including the joint effects of changes in multiple risk factors.

- The health loss due to risk factor(s) is calculated as a time-indexed “stream” of disease burden due to a time-indexed “stream” of exposure.
- The burden of disease and injury is converted into a summary measure of population health, which allows comparing fatal and non-fatal outcomes, also taking into account severity and duration.

It is important to emphasize that risk assessment, as defined above, is distinct from intervention analysis, whose purpose is to estimate the benefits of a given intervention or group of interventions in a specific population and at a specific time. Rather, risk assessment aims at mapping alternative population health scenarios to changes in distribution of exposure to risk factors over time, irrespective of whether exposure change is achievable using existing interventions. Therefore, while intervention analysis is a valuable input into cost-effectiveness studies, risk assessment contributes to assessing research and policy options for reducing disease burden by changing population exposure to risk factors.

Summary measures of population health (SMPH) and their use in burden of disease analysis are discussed elsewhere (Murray 1996; Murray et al. 2002). The next three sections of this chapter address the conceptual basis and methodological issues for the remaining three points above. We then discuss the sources and quantification of uncertainty.

2. CAUSAL ATTRIBUTION OF SMPH TO RISK FACTORS

Mathers et al. (2002) describe two traditions for causal attribution of health determinants, outcomes, or states: categorical attribution and counterfactual analysis. In categorical attribution, an event such as death is attributed to a single cause (such as a disease or risk factor) or group of causes according to a defined set of rules (hence 100% of the event is attributed to the single cause or group of causes). The International Classification of Disease system's (ICD) attribution of causes of death (WHO 1992) and attribution of some injuries to alcohol or occupational conditions are examples of categorical attribution. In counterfactual analysis, the contribution of one or a group of diseases, injuries or risk factors to a summary measure of population health is estimated by comparing the current or future levels of the summary measure with the levels that would be expected under some alternative hypothetical scenario, including the absence of or reduction in the disease(s) or risk factor(s) of interest. This hypothetical scenario is referred to as the counterfactual (see Maldonado and Greenland 2002 for a discussion of conceptual and methodological issues in the use of counterfactuals).

In theory, causal attribution of a summary measure to risk factors can be done using both categorical and counterfactual approaches. For

example, categorical attribution has been used in attribution of diseases and injuries to occupational risk factors in occupational health registries (Leigh et al. 1999) and attribution of motor vehicle accidents to alcohol consumption. In general however, categorical attribution of SMPH to risk factors overlooks the fact that many diseases have multiple causes (Rothman 1976). The epidemiological literature has commonly used the counterfactual approach for the attribution of a summary measure to a risk factor, and compared mortality or disability from the current distribution of exposure to the risk factor to that expected under an alternative exposure scenario.

The dominant counterfactual exposure distribution in these studies has been zero exposure for the whole population (or a fixed non-zero level where zero is not possible such as the case of blood pressure when defined as presence or absence of hypertension). The basic statistic obtained in this approach is the population attributable fraction (PAF) defined as the proportional reduction in disease or death that would occur if exposure to the risk factor were reduced to zero, *ceteris paribus* (Cole and MacMahon 1971; Eide and Heuch 2001; Greenland 1984; Levin 1953; MacMahon and Pugh 1970; Miettinen 1974; Ouellet et al. 1979; Rockhill et al. 1998; Uter and Pfahlberg 2001).¹ The attributable mortality, incidence or burden of disease due to the risk factor, AB , is then given as $AB = PAF \times B$ where B is the total burden of disease from a specific cause or group of causes affected by the risk factor with a relative risk of RR :

$$PAF = \frac{P(RR - 1)}{P(RR - 1) + 1} \quad (1a)$$

The exposed population may itself be divided into multiple categories based on the level or length of exposure, each with its own relative risk. With multiple (n) exposure categories, the PAF is given by the following generalized form:

$$PAF = \frac{\sum_{i=1}^n P_i(RR_i - 1)}{\sum_{i=1}^n P_i(RR_i - 1) + 1} \quad (1b)$$

Although choosing zero as the reference exposure may be useful for some purposes, it is a restricting assumption for others. The contribution of a risk factor to disease or death can alternatively be estimated by comparing the disease burden due to the observed exposure distribution in a population with that from another *distribution* (rather than a single reference level such as non-exposed) as described by the generalized

“potential impact fraction” equation (Drescher and Becher 1997; Eide and Heuch 2001; Walter 1980).

$$PIF = \frac{\int_{x=0}^m RR(x)P(x)dx - \int_{x=0}^m RR(x)P'(x)dx}{\int_{x=0}^m RR(x)P(x)dx} \quad (2a)$$

where $RR(x)$ is the relative risk at exposure level x , $P(x)$ is the population distribution of exposure, $P'(x)$ is the counterfactual distribution of exposure, and m the maximum exposure level. The first and second terms in the numerator of Equation 2a therefore represent the total exposure-weighted risk of mortality or disease in the population under current and counterfactual exposure distributions. The corresponding relationship when exposure is described as a discrete variable with n levels is given by:

$$PIF = \frac{\sum_{i=1}^n P_i RR_i - \sum_{i=1}^n P'_i RR_i}{\sum_{i=1}^n P_i RR_i} \quad (2b)$$

In addition to relaxing the assumption of the no-exposure group as the reference, analysis based on a broader range of distributions has the advantage of allowing multiple comparisons with multiple counterfactual scenarios. Equation 2a can be further generalized to consider counterfactual relative risks (i.e. relative risk may depend on other risks, new technology, medical services, etc.). For example the relative risk of injuries as a result of alcohol consumption may depend on road conditions and traffic law enforcement. Similarly, people employed in the same occupation may have different risks of occupational injuries because of different safety measures. Therefore, a more general form of Equation 2a is given by:

$$PIF = \frac{\int_{x=0}^m RR(x)P(x)dx - \int_{x=0}^m RR'(x)P'(x)dx}{\int_{x=0}^m RR(x)P(x)dx} \quad (2c)$$

2.1 COUNTERFACTUAL EXPOSURE DISTRIBUTIONS

Various criteria may determine the choice of the counterfactual exposure distributions. Greenland (2002) has discussed some of the criteria for the choice of counterfactuals, arguing that the counterfactuals should

be limited to actions that can be implemented (e.g. anti-smoking campaigns), rather than the effects of removing the outcomes targeted by those actions (e.g. smoking cessation) because, in practice, the implementation of counterfactuals for one risk factor or disease may affect other risks. The solution to Greenland's concern, however, is better analytical techniques for estimating joint risk factor effects, rather than abandoning non-intervention-based counterfactuals which, as argued by Mathers et al. (2002), is a limiting view. Estimating the contributions of risk factors to disease burden and the benefits of their removal, even in the absence of known interventions, can provide an understanding of their role in population health and visions of population health under different scenarios of risk factor exposure. This knowledge of risk factor effects can provide valuable input into public health policies and priorities, as well as research and development.

Murray and Lopez (1999) introduced a taxonomy of counterfactual exposure distributions that, in addition to identifying the size of risk, provides a mapping to policy implementation options. These categories include the exposure distributions corresponding to *theoretical* minimum risk, *plausible* minimum risk, *feasible* minimum risk and *cost-effective* minimum risk. Theoretical minimum risk refers to the exposure distribution that would result in the lowest population risk, irrespective of whether currently attainable in practice. Plausible minimum refers to a distribution which is imaginable, and feasible minimum is one that has been observed in some population. Finally, cost-effective minimum considers the cost of exposure reduction (through the set of known cost-effective interventions) as an additional criterion for choosing the alternative exposure scenario.

In addition to illustrating the total magnitude of disease burden due to a risk factor, the theoretical-minimum-risk distribution (or the current difference between theoretical and plausible or feasible risk levels) can guide research and development resources towards those risk factors for which the mechanisms of reduction (i.e. interventions) are currently underdeveloped. For example, if the reduction in the burden of disease due to improved medical injection safety is high and the methods for risk reduction are well-known, so that plausible/feasible and theoretical minima are identical, then current policy may have to be focused on the implementation of such methods. On the other hand, if there are large differences between plausible/feasible and theoretical minima risk levels for blood lipids or body mass index (BMI) (Powles and Day 2002), then research on reduction methods and their implementation should be encouraged. For this reason the total magnitude of the burden of disease due to a risk factor, as illustrated by the theoretical minimum, provides a tool for considering alternative visions of population health and setting research and implementation priorities.

Biological principles as well as considerations of equity would necessitate that, although the exposure distribution for theoretical minimum

risk may depend on age and sex, it should in general be independent of geographical region or population. Exceptions to this are, however, unavoidable. An example would be the case of alcohol consumption, which in limited quantities and when drunk in certain patterns has beneficial effects on cardiovascular mortality, but is always harmful for other diseases such as cancers and accidents (Puddey et al. 1999). In this case, the composition of the causes of death as well as drinking patterns in a region would determine the theoretical-minimum-risk distribution. In a population where cardiovascular diseases are a dominant cause of mortality, the theoretical-minimum-risk exposure distribution may be non-zero with moderate drinking patterns, whereas in a population with binge drinking and a large burden from injuries the theoretical minimum would be zero. Feasible and cost-effective distributions, on the other hand, may vary across populations based on the current distribution of the burden of disease and the resources and institutions available for exposure reduction.

The above categories of counterfactual exposure distributions are based on the burden of disease in the population as a whole. Counterfactual exposure distributions may also be considered based on other criteria. For example, a counterfactual distribution based on equity would be one in which the highest exposure group (or the group with the highest burden of disease) would be shifted towards low exposure values. Further, such equitable counterfactual distributions for each risk factor may themselves be categorized into theoretical (most equitable), plausible, feasible and cost-effective as described above. Similarly, a counterfactual distribution that focuses on the most susceptible groups in the population is one that gives additional weight to lowering the exposure of this group. Therefore, by permitting comparison of disease burden under multiple exposure distributions based on a range of criteria—including, but not limited to, implementation and cost, equity and research prioritization—relaxing the assumption of a constant exposure baseline provides an effective policy and planning tool.

2.2 EXPOSURE DISTRIBUTION FOR THEORETICAL MINIMUM RISK

In one taxonomy, risk factors such as those in the GBD project (Ezzati et al. 2002; see also the risk factor chapters in this book) can be broadly classified as physiological, behavioural, environmental and socio-economic. Some general principles that guide the choice of theoretical-minimum-risk exposure distribution for each category are:

1. *Physiological risk factors*: This group includes those factors that are physiological attributes of humans, such as blood pressure or blood lipids, and at some level result in increased risk. Since these factors are necessary to sustain life, their “exposure–response” relationship is J-shaped or U-shaped, and the theoretical-minimum-risk distribution is non-zero. For such risk factors, the choice of optimal exposure

needs to be based on empirical evidence from different scientific disciplines. For example, epidemiological research on blood pressure and cholesterol have illustrated a monotonically increasing dose–response relationship for mortality even at low levels of these risk factors (Chen et al. 1991; Eastern Stroke and Coronary Heart Disease Collaborative Research Group 1998; MacMahon et al. 1990; Prospective Studies Collaboration 1995). But, given the role of these factors in sustaining life, this relationship must flatten and reverse at some level. In the blood pressure and cholesterol assessment, a theoretical–minimum–risk exposure distribution with a mean of 115 mmHg for systolic blood pressure and 3.8 mmol/l for total cholesterol (each with a small standard deviation) were used (Ezzati et al. 2002). This distribution corresponds to the lowest levels at which the dose–response relationship has been characterized in meta-analyses of cohort studies (Chen et al. 1991; Eastern Stroke and Coronary Heart Disease Collaborative Research Group 1998; MacMahon et al. 1990; Prospective Studies Collaboration 1995). Further, these levels of blood pressure and cholesterol are consistent with levels seen in populations which have low levels of cardiovascular disease, such as the Yanomamo Indians (Carvalho et al. 1989) and rural populations in China (He et al. 1991a, 1991b), Papua New Guinea (Barnes 1965; Carvalho et al. 1989), and Africa (Mann et al. 1964). Although meta-analyses of randomized clinical trials have indicated that blood pressure and cholesterol levels may be lowered substantially with no adverse effects (LaRosa et al. 1999; Pignone et al. 2000), it is difficult to justify an optimal exposure distribution lower than that measured in population-based studies, since lower levels in individuals may be caused by factors such as pre-existing disease. Arguments from evolutionary biology would also support the choice of a lower bound on the optimal distribution based on historical survival of populations who are not substantially exposed to factors that raise blood pressure or cholesterol.

2. *Behavioural risk factors*: The exposure–response relationship for this group of risk factors may be monotonically increasing or J-shaped. For risk factors with a monotonic exposure–response relationship, such as smoking, the optimal exposure would be zero unless there are physical constraints that make zero risk unattainable. For example in the case of blood transfusion, there may be a lower bound on the safety of the blood supply process even using the best monitoring technology. With a J-shaped or U-shaped exposure–response relationship, the minimum risk would occur at the turning point of the exposure–response curve. An example of this is alcohol consumption in adult populations with high cardiovascular disease rates, since moderate consumption may result in a reduction in ischaemic heart disease (IHD) in some age groups (Corrao et al. 2000). With a

J-shaped exposure–response curve, similar to physiological risk factors, empirical evidence would have to be used to determine the theoretical minimum risk.

Finally, some behavioural risks are expressed as the absence of protective factors such as physical inactivity or low fruit and vegetable intake. In such cases, optimal exposure would be the level at which the benefits of these factors would no longer continue. With a monotonic exposure–response relationship or without detailed knowledge about a possible turning point, the theoretical-minimum-risk exposure distribution should be chosen based on empirical evidence about the highest theoretically sustainable levels of intake or exposure (for example very active life style or a purely vegetarian diet).

3. *Environmental risk factors*: The toxicity of most environmental risk factors is best described as a monotonically increasing function of exposure (potentially with some threshold). Therefore, the theoretical-minimum-risk exposure distribution for this group would be the lowest physically achievable level of exposure, such as background particulate matter concentration due to dust.
4. *Socioeconomic “risk factors”*: Socioeconomic status and factors—such as income (including levels and distribution) and associated levels of poverty and inequality, education, the existence of social support networks, etc.—are important determinants of health, often through their effects on other risk factors. The effects of each of these factors on health are, however, highly dependent on other socioeconomic variables as well as the policy context, including accessibility and effectiveness of health and welfare systems. For this reason, the theoretical-minimum-risk exposure distribution, even if meaningfully defined, is likely to change over time and space depending on a large number of other factors. Given this heterogeneity, the effects of socioeconomic variables are best assessed relative to counterfactual distributions defined based on policy and intervention options in specific times and settings, as discussed by Greenland (2002).

3. RISK QUANTIFICATION MODELS

Prediction implicitly assumes the use of a conceptual model which infers the value of the variable of interest at a point in time or space based on knowledge from a different time, or another location. Predictive models can be divided along a continuum between *aggregate* and *structural* categories. A completely aggregate model uses the previous trend of the variable of interest as the basis for predicting its future value. A structural model, on the other hand, identifies the components—and the relationships among them—of the “system” that determines the variable of interest. It then uses the knowledge of the system for predicting the value of

the variable of interest. Most predictive models lie between the two extremes and use a combination of aggregate and structural modelling.²

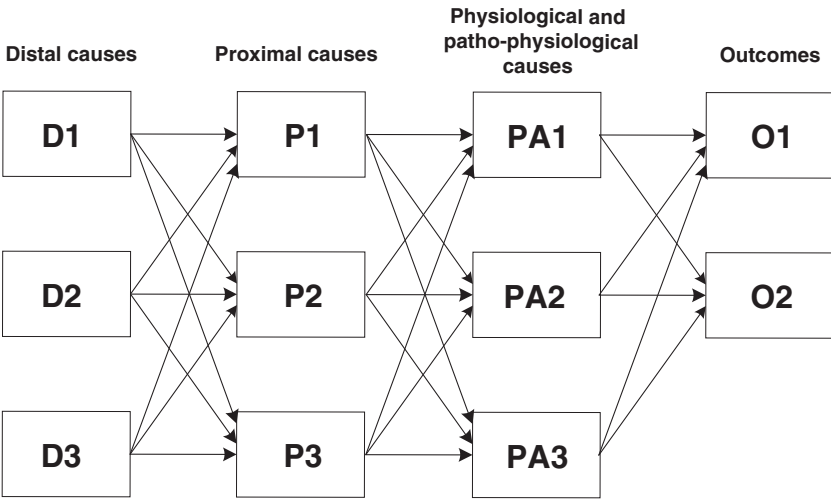
Consider for example predicting the future population of a city or the future ambient concentration of a pollutant. An aggregate model would extrapolate the historical levels to predict future values. Even in this case the model may include some structural elements. For example, the model may use a specific functional form—linear, exponential, quadratic or logarithmic—for extrapolation which involves an assumption about the underlying system. A structural model, in the case of population prediction would consider the age structure of the population, fertility (which itself may be modelled using data on education and family planning programmes), public health variables and rural–urban migration (which itself can be modelled using economic variables). In the case of air pollution, a structural model may consider demographic variables (themselves modelled as above), the structure of the economy (manufacturing, agriculture or service), the current manufacturing and transportation technology and effects of research and development on new technology, the demand for private vehicles, the price of energy and the atmospheric chemistry of pollution. Once again, in both examples the models may include some aggregation of variables by using historical trends to predict the future values of individual variables in the system, such as funding for family planning or research and development of new technologies.

The comparative advantage of structural and aggregate models lies in the balance between theoretical precision and data requirement. Structural models offer the potential for more robust predictions, especially when the underlying system is complex and highly sensitive to one or more of its components. In such cases, a shift in some of the system variables can introduce large changes in the outcome, which may be missed by extrapolation (such as the discovery of antibiotics and infectious disease trends or the change in tuberculosis mortality after the HIV epidemic). Aggregate models, on the other hand, require considerably less knowledge of the system components and the relationships among them. These models can therefore provide more reliable estimates when such information is not available, especially when the system is not very sensitive to inputs.

3.1 MODELS FOR RISK FACTOR–DISEASE RELATIONSHIP

Using the above aggregate, structural taxonomy, it is also possible to classify models that are used to predict changes in death or disease as a result of changes in exposure to underlying risk factors. Murray and Lopez (1999) described a “causal-web” which includes the various distal (such as socioeconomic), proximal (behavioural or environmental) and physiological and patho-physiological causes of disease, as shown in Figure 1.1. While different disciplinary traditions—from social sciences

Figure 1.1 Simplified schema for a causal-web illustrating various levels of disease causation

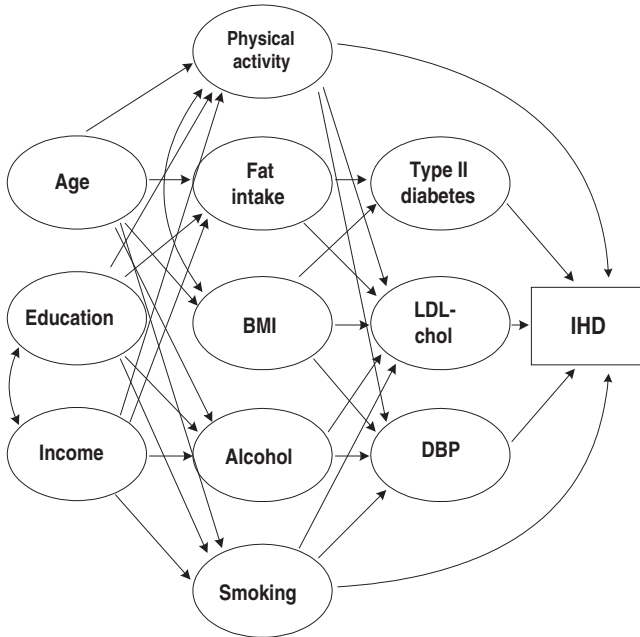


Note: Feedback from outcomes to preceding layers may also exist. For example, individuals or societies may modify their risk behaviour based on health outcomes.

and humanities, to the physical, natural and biomedical sciences—have focused on individual components or stages of these relationships, in a single multi-layer causal model with interactions the term “risk factor” can be used for any of the causal determinants of health (Mathers et al. 2002; Yerushalmy and Palmer 1959).³ For example, poverty, location of housing, lack of access to clean water and sanitation, and the existence of a specific pathogen in water can all be considered the causes of diarrhoeal diseases, providing a more complete framework for assessment of interventions and policy options. Similarly, education and occupation, diet, smoking, air pollution, physical activity, BMI and blood pressure are some of the risk factors at various levels of causality for cardiovascular diseases.

Compared to a causal-web, Equations 1 and 2 that use relative risk estimates from epidemiological methods (e.g. the Cox proportional hazard or other regression models) lie further towards aggregate modelling. In general, in such methods, relative risks are estimated so that they incorporate the aggregation of the various underlying relationship (ideally, but not always, controlling for the appropriate confounding variables)⁴ without considering intermediate relationships as separate causal stages. On the other hand, if specified and estimated correctly, considering the complete set of causal pathways which include multiple

Figure 1.2 A possible causal diagram based on established relationships for estimating the incidence of ischaemic heart disease



DBP Diastolic blood pressure.

Note: Other interactions may also be possible.

risk factors will allow making inferences about combinations of risk factors and risk factor levels for which direct epidemiological studies may not be available.

As discussed earlier, the appropriateness of the two approaches to estimation of attributable burden depends on the specific risk factor(s), outcomes and available data. For example, the relationship between smoking and lung cancer has been shown to be highly dependent on smoking intensity and duration which, with appropriate indicators of past smoking (Peto et al. 1992), can be readily estimated using the relative risk approach of Equations 1 and 2. Consider, on the other hand, the relationship among age, socioeconomic status and occupation, behavioural risk factors (such as smoking, alcohol consumption, diet, physical activity), physiological variables (such as blood pressure and cholesterol level) and IHD shown in Figure 1.2. Given the multiple complex interactions, IHD risk may be best predicted using a structural (causal-web) approach, especially when some risk factors vary simulta-

neously, such as smoking, alcohol and diet, requiring joint counterfactual distributions. Using a multi-risk model would also allow considering situations for which direct epidemiological studies may not have been conducted, such as the effects of physical activity on those people who have diets different from the study group or those who take medicine to lower blood pressure.

The health effects of global climate change provide another example where a structural approach to risk assessment may be appropriate. Economic activities (including manufacturing, agriculture and forest use, transportation and domestic energy use) affect the emissions of greenhouse gases (GHG). Changes in precipitation, temperature and other meteorological variables due to atmospheric GHG accumulation alter regional ecology, which in turn results in changes in agricultural productivity, quantity and quality of water, dynamics of disease vectors and other determinants of disease. All these effects are in turn modulated by local economic activities, land-use patterns and income (Patz et al. 2000; Reiter 2001; Rogers and Randolph 2000). A model based on the atmospheric physics/chemistry of GHG emissions and accumulation, climate models, plant and vector ecology and human activity might provide the optimal basis for the prediction of the health effects of climate change.⁵

SPECIFYING THE CAUSAL-WEB

Assuming for the moment no temporal dimension in the relationship between the different variables in the causal system (temporal aspects are discussed below), each layer of a causal-web may be characterized by the equation:

$$\mathbf{X}^n = f(\mathbf{B}(\mathbf{X}^{n-1}, \mathbf{X}^n), \mathbf{X}^{n-1}) \quad (3a)^6$$

where \mathbf{X}^n is the vector of the variables in the n th layer of the causal-web (which can be causal or output such as \mathbf{D} , \mathbf{P} , \mathbf{PA} , or \mathbf{O} using the notation of Figure 1.1); f is the functional form connecting the $(n-1)$ th layer to the n th layer; \mathbf{B} is a matrix of coefficients for f which itself may be dependent on the variables in the $(n-1)$ th and n th layers (\mathbf{X}^{n-1} and \mathbf{X}^n)⁷ (as well as time as we discuss below).

The attributable fraction of disease or mortality due to a single risk factor in the causal-web is then obtained by integrating the outcome (\mathbf{O}) over the current ($P(\mathbf{x})$) and counterfactual ($P'(\mathbf{x})$) population distributions of exposure, as for Equation 2.

$$AF = \frac{\int_{P(x)} \mathbf{O}(\mathbf{x}) - \int_{P'(x)} \mathbf{O}(\mathbf{x})}{\int_{P(x)} \mathbf{O}(\mathbf{x})} \quad (4)$$

3.2 JOINT RISK FACTOR CHANGES

The attributable fraction relationships described in Equations 1 and 2 are based on individual risk factors. Disease and mortality are however often affected by multiple, and at times correlated, risk factors (Rothman 1976; Walter 1980). Estimating the joint effects of multiple distal and proximal risks is particularly important because many factors act through other, intermediate, factors (Murray and Lopez 1999; Yerushalmy and Palmer 1959), or in combination with other risks. It is therefore important to consider how the burden of disease may change with simultaneous variations in multiple risk factors. Analysis of joint risk factor changes implicitly acknowledges that the disease causation mechanism involves multiple factors, and is therefore suited to a causal-web framework, with $P(\mathbf{x})$ and $P'(\mathbf{x})$ in Equation 4 being the joint distributions of the vector of risk factors, \mathbf{x} . Alternatively, when using Equations 1 or 2, knowledge of the distribution of *all* relevant risk factors and the relative risk for each risk factor, *estimated at the appropriate level of the remaining risk factors*,⁸ is required. Therefore, in Equation 2a, RR and P may represent joint risks and exposure distributions for multiple risk factors (Eide and Heuch 2001). In this case, the estimates from Equations 2a and 4 may in theory be identical.

ADDITIVITY OF ATTRIBUTABLE FRACTION

Many users of risk assessment desire information characterized by additive decomposition. In other words, users would like to know what fraction of the disease burden is related to any risk factor or group of risk factors, independent of the changes in other risk factors. As discussed by Mathers et al. (2002), additive decomposition is a property of categorical attribution and, in general, not of counterfactual attribution because many diseases are caused by the interaction of multiple risk factors acting simultaneously and therefore can be avoided by eliminating any of these factors (Rothman 1976; Rothman and Greenland 1998; Yerushalmy and Palmer 1959). Consider for example infant and child mortality due to acute respiratory infections (ARI), which are especially high among malnourished children, as a result of exposure to indoor smoke from solid fuels (Rice et al. 2000; Smith et al. 2000). In this case, removal of either risk factor can reduce mortality, some of which can therefore be attributed to both factors. Similarly the risk of mortality due to cardiovascular diseases among some of those who are exposed to smoking, low physical activity and poor diet may be reduced by elimination of any combination of these risk factors. Counterfactual causal attribution of disease and injury to individual risk factors does not normally allow additive decomposition and the sum of attributable fractions or burdens for a single disease due to multiple risk factors is therefore theoretically unbounded.

Although epidemiologically unavoidable and conceptually acceptable, the lack of additivity presents additional policy complexity and implies great caution is necessary when communicating and interpreting the estimates of attributable fraction and burden. With multiple attribution, the reduction of one risk factor would seem to make other, equally important risk factors potentially irrelevant from the perspective with a limited scope on quantification. At the same time multi-causality offers opportunities to tailor prevention based on availability and cost of interventions. It also necessitates the development of methods to quantify the effects of joint counterfactual distributions for multiple risk factors.

4. TEMPORAL DIMENSIONS OF THE RISK FACTOR–DISEASE RELATIONSHIP

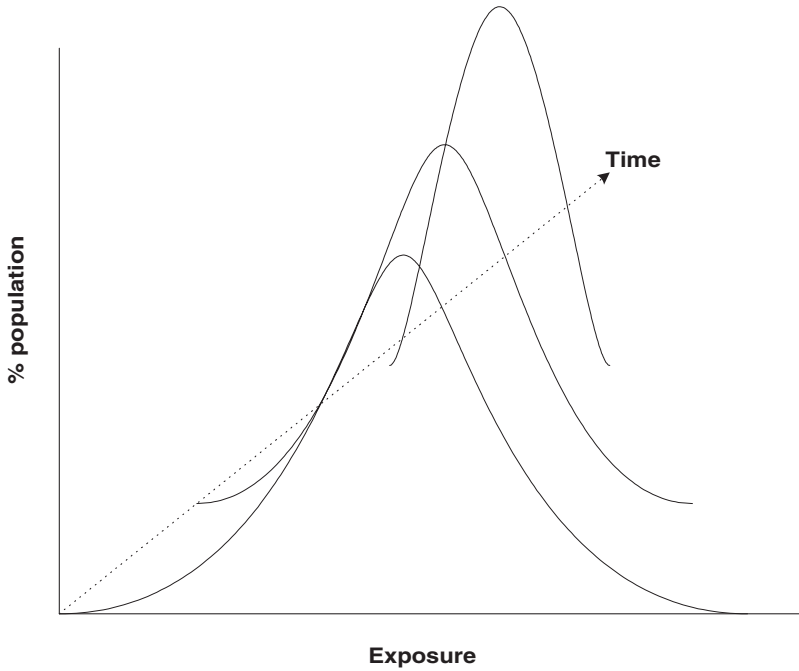
Both exposure to a risk factor and the health outcomes due to exposure include a time dimension. This can be described by a modified version of Equation 3 in which exposure and outcome as well as the model parameters (\mathbf{B}) are dependent on time. In the following two sections we consider the temporal characteristics of exposure and health outcomes, respectively.

4.1 TEMPORAL CHARACTERISTICS OF EXPOSURE

With the exception of acute hazards (e.g. injury risk factors) exposure to a risk factor affects disease over a time period. As a result, the distributional transition between any two exposure distributions includes a temporal dimension as illustrated schematically in Figure 1.3. The transition path is of little importance if exposure changes over a short time interval, especially relative to the time required for the effect of exposure on disease. Over long time periods, however, there is sufficient time for contributions from the intermediate exposure values, and the actual path of transition may be as important as the initial and final distributions in determining the disease burden associated with change in exposure. For example, the effects of reducing the prevalence of smoking or exposure to an occupational carcinogen by half in a population would be markedly different if the change takes place immediately, gradually over a twenty-year period or after twenty years. Therefore, the health effects of exposure to many risk factors depend on the complete profile of exposure over time, and may be further accompanied by a time lag from the period of exposure. Also, for some risk factors there may be complete or partial reversibility, with the role of past exposure gradually declining.

To capture the effects of exposure profiles over time, we begin by considering the role of temporal dimensions of exposure at the level of individuals (or groups of individuals with similar exposure) before considering the whole population.

Figure 1.3 A (three-dimensional) representation of a time-indexed distributional transition of population exposure to a risk factor, with a decreasing central tendency



Suppose that at time T the relative risk of a disease, RR , for individuals exposed to a risk factor (compared to the non-exposed group) depends on the complete *profile or stream of exposure* between time T_0 and T , denoted by $x(t)$, with some lag, L , between exposure and effect. Then, there is some function, $f(x)$, which can be used to describe the contribution of exposure at any point in time between T_0 and T to the relative risk (RR). In mathematical notation:

$$RR(x(t))|_{T_0}^T = RR \left(\int_{T_0}^T f(x(t-L)) dt \right) \quad (5)^9$$

The quantity $\int_{T_0}^T f(x(t-L)) dx$ is an *equivalent exposure*¹⁰ between T_0 and T and is dependent on: i) the profile of exposure (i.e. level of exposure at any point in time) described by $x(t)$; and ii) the contribution of previous exposure to current hazard characterized by $f(x)$, an *accumulative risk function*.¹¹ Some common forms for the accumulative risk function, $f(x)$, are given in Table 1.1.

Table 1.1 Possible forms for the accumulative risk function, $f(x)$

Accumulative risk function, $f(x)$	Interpretation	Relative risk	Potential example
$f(x) = \begin{cases} 1 & \text{if } t = T \\ 0 & \text{otherwise} \end{cases}$	RR depends only on current exposure, with no contribution from past exposure	$RR(x(t)) _0^T = RR(x(T))$	Instantaneous poisoning as a result of exposure to high levels of toxic chemicals; injuries or death in accidents due to binge drinking; infection with Hepatitis B or C as a result of an infected injection
2. $f(x) = 1$	RR depends on the accumulated exposure (or average exposure if normalized with respect to exposure time), without any effects from the temporal distribution of exposure	$RR(x(t)) _0^T = RR\left(\int_0^T x(t)dt\right)$	Cancer risk from lifetime exposure to carcinogens which have no threshold level
3. $f(x) = \begin{cases} \frac{1}{K} & \text{if } t > T - K \\ 0 & \text{otherwise} \end{cases}$	RR depends on current and past exposures. But the role of past exposure lasts for a limited time, K , and declines as a linear function of time	$RR(x(t)) _0^T = RR\left(\int_0^T \frac{(t-T+K)}{K} x(t)dt\right)$	Reduction in cardiovascular events after lowering blood pressure ^a
4. $f(x) = e^{\alpha(t-T)}$	RR depends on current and past exposures. But the role of past exposure decays as an exponential function of time	$RR(x(t)) _0^T = RR\left(\int_0^T e^{\alpha(t-T)} x(t)dt\right)$	Reduction in cardiovascular events after lowering blood pressure ^a

^a Example is applicable in illustrating gradual reduction in risk after exposure is reduced. The exact functional form of risk reduction may not necessarily be linear or exponential. For simplicity of notation, in all these cases we assume that: i) $L = 0$. Including a lag is straightforward and can be done by replacing t with $(t - L)$ in the corresponding formulas; and ii) there is no threshold for exposure. Including the threshold level is also straightforward using the $TRUE(x(z) \geq X)$ function. In scenarios 1, 3 and 4, where the effects of past exposure are absent or decline over time, risk reversibility can take place if exposure is reduced or removed. In scenario 1 there is immediate risk reversibility; in scenario 3, there is full reversibility after time K ; in scenario 4, risk reversibility asymptotically approaches 100%. In scenario 2 there is no risk reversibility and the effects of past exposure remain for an indefinite period.

The above framework can be extended from individuals to populations, by indexing the exposure profile ($x(t)$) to individuals (i.e. representing the exposure of the i th individual as $x_i(t)$) and considering how the distribution of exposure in the population evolves over time.¹² This in turn provides the population distributions of equivalent exposure (current or expected future and counterfactual) which form the basis of calculating attributable fractions (i.e. the terms in the numerator of Equations 2a, 2b or 4).

It is reasonable to assume that if the exposure of one individual is greater than that of another over the whole exposure period (i.e. tracking) (Foulkes and Davis 1981), the *equivalent* exposure of the former is also greater than the latter. In other words, the accumulative risk function, $f(x)$, has the following property:

$$\int_{T_0}^T f(x_i(t))dt > \int_{T_0}^T f(x_j(t))dt \quad \text{if } x_i(t) > x_j(t) \quad \forall t \in [T_0, T] \quad (6)$$

With this property, if the ordering of individuals in the exposure distribution remains unchanged over time (i.e. the rank-order correlation of individual exposures equals 1 between different points in time), the equivalent exposure will also have a distribution with the same ordering of individuals.

The method used by Peto et al. (1992) for estimating mortality due to smoking implicitly uses such a framework. It is well known that the accumulated hazards of smoking depend on a number of variables including the age at which smoking began, number of cigarettes smoked per day and cigarette type. Such data however are extremely rare. To overcome this problem, Peto et al. (1992) used the smoking impact ratio, *SIR*, which uses population lung cancer rates as a marker for accumulated hazard of smoking, to estimate the relative risk of the accumulated smoking exposure corresponding to the population. In the above notation:

$$RR(smoking(t))|_{T_0}^T = RR\left(\int_{T_0}^T f(smoking(t-L))dt\right) = RR(SIR(T))$$

The temporal profile of exposure for some risk factors may be more easily available than the range of indicators that are needed to estimate the accumulated hazards of smoking. For example, exposure to indoor smoke from solid fuels is likely to remain unchanged as long as household fuel and housing conditions remain the same. Therefore, estimating the effects of long-term exposure may require only knowledge of household fuel, housing and participation in cooking. Similarly, in the case of blood pressure, it is known that blood pressure follows a predictable age pattern (Tate et al. 1995; Yong et al. 1993), unless severely affected by a changes in social (stress), behavioral (diet or smoking)

or medical circumstances. In this case, the usual blood pressure of an individual reflects the history of the person's exposure. On the other hand, the patterns of fruit and vegetable consumption, smoking, or exposure to urban air pollution may change rapidly in countries with high rates of economic growth and urbanization, requiring more detailed data.

The above discussion is based on two implicit assumptions:

1. It considers the effects of exposure to a single risk factor over time. This approach may be appropriate for some risk factor–disease relationships (e.g. the effects of accumulated exposure to carcinogens with site-specific effects). But the single equivalent exposure cannot characterize other risk factor–disease relationships where risk factor interactions are important over time (e.g. physical activity, BMI, smoking and cardiovascular diseases). Extending this temporal dimension to multiple risk factors requires considering the accumulated effects of the vector of risk factors as well as their interactions. In this case, Equation 5 would be expressed in terms of the vector of risk factors of interest. Few epidemiological studies, however, have gathered the data needed for assessing accumulated interactive effects.
2. It considers exposure to each risk factor as an exogenous variable (i.e. intermediate exposure at any time, $x(t)$, is not affected by disease or other risk factors) whose accumulated effect can be captured in a single value using the risk accumulation function. For some risk factors, this assumption may not be valid since exposure to behavioural as well as environmental risk factors may be affected by knowledge of their current effects—individuals may change their diet or activity levels based on knowledge of their weight or blood pressure and governments may introduce regulations based on the level of various contaminants in air or water.¹³

Manton et al. (1993a, 1994) have relaxed these assumptions using a diffusion model for forecasting cardiovascular disease mortality in the United States of America. In this model, it is the change in the outcome at any time, t , that is modelled as a function of all the other variables in the system (i.e. other risk factors as well as outcomes) and their interactions. Using the notation of Equation 3:

$$d\mathbf{X}(t) = \mathbf{u}(\mathbf{X}(t), t)dt \quad (7)$$

where $\mathbf{u}(\mathbf{X}(t), t)$ is a drift term whose value depends on the current value of all the variables in the system as well as their interactions (and can be described by a functional form similar to that in Equation 3).¹⁴ Methods for estimation of such models using longitudinal data are discussed by Robins (1997, 1999b).

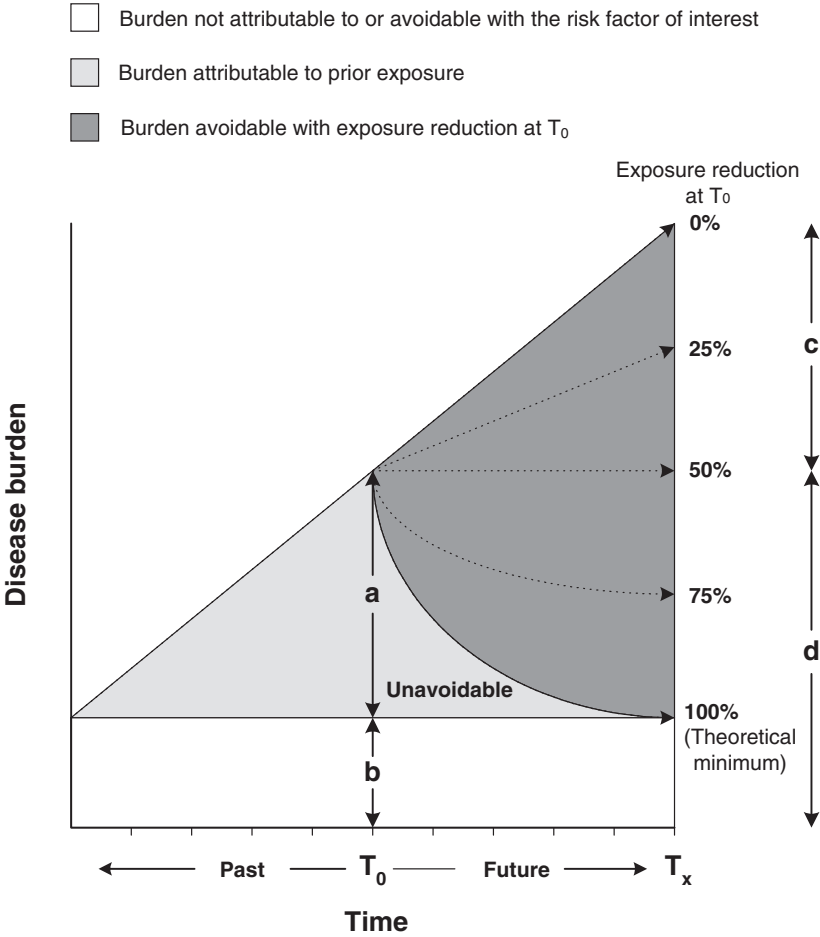
4.2 TEMPORAL CHARACTERISTICS OF HEALTH OUTCOMES

If the outcome variable used in causal attribution of disease and mortality to a risk factor only involves counting of adverse events (such as disease incidence or death), it is not possible to characterize those cases whose occurrence would have been delayed in the absence of the risk factor (Greenland and Robins 1988; Robins and Greenland 1989, 1991). A major shortcoming with this approach is that it does not take into account the accumulated effects of exposure—in the form of years of life lost prematurely or lived with disability. Parameterizing the above relationships by age (or birth cohort) would allow estimating the effects of exposure to a risk factor, not as an event without time dimension, but as an event at a certain age and time. More broadly, considering the time-indexed stream of health losses due to a risk factor requires using a time-based (and not event-based) SMPH.

Murray and Lopez (1999) have provided an additional temporal distinction for the burden of disease due to a risk factor by introducing the concepts of “attributable” and “avoidable” burden. Attributable burden is defined as the reduction in *current and/or future burden* of disease if *past exposure* to a risk factor had been equal to some counterfactual distribution. Avoidable burden is the reduction in the *future burden* of disease if the *current or future exposure* to a risk factor is reduced to a counterfactual distribution. Attributable and avoidable burden are shown graphically in Figure 1.4. While attributable burden is easier to measure and more certain, avoidable burden is more useful for policy purposes. The distinction between attributable and avoidable burden becomes less significant as the time between exposure to risk factor and effects on disease burden decreases. In this case, attributable burden is a good predictor of avoidable disease burden.

Figure 1.4 also illustrates a conceptual complexity in defining and estimating avoidable burden. Attributable burden is defined based on the difference between (accumulated) current exposure and a counterfactual. Measuring current exposure, while difficult and uncertain, is conceptually well defined. Avoidable burden, on the other hand, depends on the expectation of future exposure and a counterfactual, with the former being analogous to current exposure. Consider for example a population exposed to rising levels of air pollution or rates of obesity. In this setting, interventions that would maintain pollutant concentrations or BMI at their current levels would result in avoiding disease and mortality; they reduce exposure from what it would be in their absence. Therefore, avoidable burden (i.e. how much of future burden could be prevented) by definition requires estimates of future exposure (i.e. how much of future burden there is). Projecting future exposure in turn raises the need to provide a projection framework. To provide visions of public health under various intervention and policy scenarios we suggest that the future exposure level, with respect to which avoidable burden is

Figure 1.4 Attributable and avoidable burden



- a Disease burden at T_0 attributable to prior exposure.
- b Disease burden at T_0 not attributable to the risk factor of interest (caused by other factors only).
The burden not attributable to the risk factor of interest (white area) may be decreasing, constant or increasing over time. The constant case is shown in the figure.
Dashed arrows represent the path of burden after a reduction at T_0 .
- c Disease burden avoidable at T_x with a 50% exposure reduction at T_0 .
- d Remaining disease burden at T_x after a 50% reduction in risk factor exposure.

estimated, be the expectation of exposure if the current policy and technological context were to continue, referred to as the “business-as-usual” exposure trend. Therefore avoidable burden is the burden of disease averted due to reduction in exposure to a risk factor beyond its expected trends. We emphasize that with this definition, avoidable burden is the difference between two exposure scenarios: the expectation of future trends (business-as-usual) and a reduction with respect to this trend towards theoretical minimum.

4.3 CUMULATIVE VS PERIOD ESTIMATES OF ATTRIBUTABLE AND AVOIDABLE BURDEN

Although analytically inconsequential, the starting point and the duration of the time interval over which attributable or avoidable burden is reported has policy implications because reductions in various risk factors may provide health benefits that occur after short or long delays and last for different periods. Consider for example the health benefits of reductions in binge alcohol consumption, smoking and GHG emissions. Reducing binge drinking would result in immediate health benefits from a drop in alcohol-related accidents and injuries (as well as medium- and long-term benefits from reduction in other diseases). Lowering smoking will have some short- and medium-term benefits from reduction of acute respiratory diseases and cardiovascular disease as well as longer-term benefits from lowering cancers and chronic obstructive pulmonary disease (COPD). The benefits of policies that reduce climate change as a result of GHG emissions are likely to be heavily concentrated in the future.

Consider these examples of health benefits in terms of duration: the distribution of a drug that lowers blood pressure, or food aid to reduce malnutrition; and programmes that promote and sustain increased physical activity, the introduction of a new agricultural technology which results in higher food yields, or automotive technology which eliminates the use of leaded gasoline. While the benefits of all these interventions may be equally large and important for the current cohort, the first two actions have one-time health benefits (unless repeated) while the latter three are likely to last indefinitely.

The above discussion would suggest reporting the estimates of avoidable burden in multiple ways including both period (e.g. annual) and cumulative estimates, as well as over short and long time frames. The issue of future estimates and their policy relevance is further complicated by the growing uncertainty of estimates with increasing length of the estimation interval. Therefore, while it may be preferable to increase the prediction horizon, it is important to emphasize that long-term predictions are inherently more uncertain.

4.4 DISCOUNTING FUTURE RISK AND HEALTH EFFECTS

Individuals may discount consumption or welfare within their own lifespan and exhibit a preference for benefits today over the future. The theoretical and empirical arguments for and against individual discounting with specific emphasis on health, including the possibility of negative discount rates, are summarized elsewhere (Murray 1996; Murray and Acharya 2002) and are directly incorporated in the calculation of a summary measure of population health. In addition to individual discounting and discount rates, policies dealing with risk confront the issue of addressing benefits to different populations across time. As a result, these policies must address ethical and analytical dilemmas related to the valuation of current and future health and welfare, in the form of social discount rates (Kneese 1999). Discounting future risks, benefits and welfare has been a subject of great debate (Howarth 1996; Lind 1982; Portney and Weyant 1999; Schelling 1995; Toman 1999), motivating some economists to conclude that

maybe the idea of a unitary decision-maker—like an optimising individual or a wise and impartial adviser—is not very helpful when it comes to the choice of policies that will have distant-future effects about which one can now know hardly anything. Serious policy choice may then be a different animal, quite unlike individual saving and investment decisions. . . . “Responsibility” suggests something less personal (Solow 1999).

The arguments for and against discounting of future health and welfare, and their validity, have been discussed in detail elsewhere (Anand and Hansen 2002; Murray and Acharya 2002; Murray and Lopez 1996; Parfit 1984). According to one specific argument, “the disease eradication and health research paradox”, not discounting future health would imply investing all of society’s health resources in research programmes or programmes for disease eradication, which would result in an infinite stream of benefits, rather than in any programme that improves the health of the current generation. Such an excessive intergenerational “sacrifice” is a particularly powerful argument for discounting of future health (or more precisely for something that resembles discounting as we discuss shortly) (Parfit 1984). It is important to emphasize that this argument does not imply that future welfare or health is less valuable than current, but rather it uses discounting as a tool to avoid excessive sacrifice for the current generation, to the point of investing all resources in an infinite stream of future health. For this reason, Parfit (1984) argues that the issue of intergenerational distribution should be considered as an independent criterion, rather than explicitly discounting future benefits.

Koopmans (1960), Dasgupta and Mäler (1994) and Dasgupta et al. (1999), however, have shown that any preference-ordering defined over

a set of well-being paths over time can be represented by a numerical function with an *apparently* utilitarian form¹⁵ and therefore includes what *resembles* positive discounting of future well-being. We emphasize that this notion is simply a consequence of considering the paths of well-being (or temporal distributions), rather than a statement about the value of current or future welfare. With this formulation, Dasgupta and Mäler (1994) and Dasgupta et al. (1999) consider the implications of the choice of discount rate as a “derived notion”, as opposed to a value judgement. Dasgupta and Heal (1974) and Solow (1974b) have shown that if well-being is a result of consumption of an exhaustible resource, a zero discount rate would imply investing all available resources for the benefit of future generations, and hence no current consumption. This is because each unit sacrificed by the first generation would yield a finite loss to this generation, but an infinite stream of benefits to future generations (Arrow 1999) which, without discounting, would always be larger than the one-time sacrifice. Although the first generation cannot sacrifice everything,¹⁶ the logical conclusion of this situation would be that “given any investment [for future benefits], short of the entire income, a still greater investment would be preferred” (Arrow 1999), or a potentially excessive intergenerational sacrifice (Murray 1996).

On the other hand, a positive discount rate would imply that in the long run consumption of resources should become zero. In this case, however, the additional requirement that well-being should never fall below a certain threshold would in turn require downward adjustment of the discount rate (Solow 1974a). The stricter requirement of non-declining consumption and well-being would require a discount rate lower than the productivity of capital (Dasgupta and Mäler 1994).¹⁷ Based on these arguments, we suggest discounting of future attributable or avoidable disease burden due to risk factors, but with a low discount rate, to include the welfares of both current and future generations as described above.

5. UNCERTAINTY

Quantitative risk assessment is always affected by uncertainty about the existence, magnitude and distribution of risk (Graham et al. 1988). Quantitative analysis of uncertainty greatly adds to the usefulness of the results because it shows not only the “best-estimate” of the magnitude and distribution of exposure to a risk factor, and the resulting burden of disease, but also the range of potential outcomes.

5.1 SOURCES OF UNCERTAINTY IN ATTRIBUTABLE FRACTIONS

POPULATION DISTRIBUTION OF EXPOSURE

An important source of uncertainty in risk assessment is characterizing population distribution of exposure. Due to complexity and cost, for

most risk factors, exposure is measured only in small samples and in a limited number of settings. As discussed earlier, because a risk factor can be represented in different layers of causality, variables for which data are more readily available can be used as exposure proxies. The use of exposure proxies is sometimes also necessary because epidemiological studies have used such proxies in estimating hazard size. For example, anthropometric variables such as height-for-age or weight-for-age are used as indicators of childhood nutritional status; the presence of clean water sources or sanitary latrines as indicators of faecal-oral transmission of pathogens; concentration of particles as the measure of exposure to the various pollutants in ambient air, and so on. In addition to reduced data requirements, the use of such indirect indicators of exposure (or exposure scenarios [Kay et al. 2000]) may provide direct mapping to existing interventions. At the same time, these indicators, which are often more distal than actual exposure, do not capture the variability of exposure within each scenario, unless combined with other indicators which affect this variability (Ezzati et al. 2000). For example people using the same water source may experience different levels of faecal-oral transmission of pathogens due to different storage and hygiene behaviours. Therefore, the use of indirect exposure proxies results in additional uncertainty in exposure characterization.

Even with the choice of exposure proxies, extrapolation of exposure between different populations or age groups is often necessary. Such extrapolation (or spatial prediction) can be based on models as simple as using the average of subgroups with data for a whole population, or more complex prediction models. For example, urban air quality monitoring systems provide data on particle concentrations in some but not all cities in each region. Models to predict ambient concentrations of particulate matter based on energy consumption, number of vehicles and level of industrialization can be used to predict ambient air pollution levels for cities where data are not available. Similarly, the level of physical (in)activity in a population may be predicted from a model that uses rural-urban population distribution, income, education, distribution across occupational categories and available transportation modes in each geographical region. Each such extrapolation adds to the uncertainty of exposure distributions.

Finally, as we discussed earlier, for many risk factors, hazards are associated with accumulated effects of sustained exposure. Indicators of accumulated hazard for those risk factors with changing exposure such as smoking, urban air pollution or BMI, are needed but not always available. Further, few epidemiological studies have considered the role of a temporal profile of exposure on disease (see Peto 1986 for an example of an exception). Therefore even if longitudinal data on exposure prevalence were available, they could not always be used together with epidemiological studies that consider a single exposure variable—at the beginning or end of the follow-up period, for example. At the same time,

if the ordering of individuals in the exposure distribution remains unchanged over time (see above), risk estimates from epidemiological studies with similar ordering may be applicable, but result in additional uncertainty.¹⁸

RISK FACTOR–DISEASE RELATIONSHIP

At the most fundamental level, quantifying the hazards associated with exposure to a risk factor requires identifying the diseases and injuries that are caused by a risk factor. The criteria for establishing disease causality have been the subject of interest and debate for over a century (summarized in Evans 1976, 1978; Hill 1965; National Research Council 1994; Yerushalmy and Palmer 1959). Epidemiological studies have successfully provided the basis for establishing causality between some risk factor–disease pairs. For other risk factor–disease combinations, where the measurement of exposure or disease has been difficult or the delay between exposure and health effects is very long, observational or experimental epidemiology has had less success in establishing causality (Evans 1976; Robins 1999a). For this reason, epidemiological evidence must often be complemented with inferences from other disciplines such as toxicology, physiology, parasitology and, increasingly, biophysics, in establishing disease causation.

Even when causality is established, the magnitude of the hazard due to a risk factor needs to be quantified. Although the statistical issues around establishing causality and estimating the effect size are similar (lack of causality is equivalent to zero excess risk) (Robins 1999a; Robins and Greenland 2000), in practice, with knowledge from multiple disciplines in establishing causality, it is often the latter that is the source of increased uncertainty in risk assessment. For example, the collectivity of scientific knowledge from disciplines such as economics and behavioural sciences, vector biology, physiology and bio-mechanics and epidemiology would confirm the possibility that climate change or socioeconomic inequality would increase disease, or whether the relationships between occupational factors or physical inactivity and lower back pain are causal. At the same time, risk assessment would require estimating the hazard magnitude for each of these relationships. Therefore, the complexity of the causal relationship, or lack of detailed data, would shift the debate from causality to hazard size.

Epidemiological studies that quantify hazards are often conducted in a limited number of settings, with emphasis on estimating the average effect size in the whole study group. While the robustness of relative risk measures has been confirmed for more proximal factors in studies across populations (Eastern Stroke and Coronary Heart Disease Collaborative Research Group 1998; Law et al. 1994), their extrapolation is an important source of uncertainty for more distal risks (e.g. child sexual abuse) or for those whose effects are heterogeneous (e.g. alcohol and injuries vs alcohol and cancer) and has received less attention in the

epidemiological literature (Horton 2000). For some risk factors, it is likely that the magnitude of the hazard may depend on the levels of other variables (i.e. effect modifiers). Therefore, in extrapolating the results of individual epidemiological studies or meta-analyses, the very strength of the original study—applicability to the average person—would be the source of uncertainty if the population to whom the effect size is extrapolated has characteristics which would result in effect modification (Britton et al. 1999; Horwitz et al. 1990, 1996, 1998). The impact of alcohol drinking patterns on cardiovascular disease risk estimates (Britton and McKee 2000; Puddey et al. 1999) is an example of the importance of considering factors that modulate and modify hazards in risk extrapolation.

RISK FACTOR AND DISEASE CORRELATIONS

Because multiple risks and disease are correlated (e.g. higher malnutrition, unsafe water, sanitation and hygiene, indoor smoke and childhood mortality in poor rural households in developing countries; higher smoking, BMI and occupational risks in developed countries [Thun et al. 2000]), estimating attributable fractions would require stratified (e.g. by other risk factors) prevalence as well as disease data. Lack of stratified data is another source of uncertainty, in general leading to the underestimation of effects in the presence of positive risk factor correlations (Greenland 1984).

5.2 CHARACTERIZING AND QUANTIFYING UNCERTAINTY

Various taxonomies of uncertainty have been used in risk assessment (National Research Council 1994) including:

1. classification based on information type such as uncertainty in hazard identification, exposure assessment, exposure–response assessment, as discussed above;
2. classification based on uncertainty type such as randomness, true variability, and bias; and
3. classification based on the approach to handling uncertainty which divides uncertainty into *parameter uncertainty* and *model uncertainty*. Parameter uncertainty includes the uncertainty quantifiable using random-variable methods such as the uncertainty due to sampling and measurement error. Model uncertainty is due to gaps in scientific theory, measurement technology and data (National Research Council 1994). It includes uncertainty in the knowledge of causal relationships or of the form of the exposure–response relationship (threshold vs continuous, linear vs non-linear, etc.), the level of bias in measurement, etc. Defined broadly, model uncertainty also includes extrapolation of exposure or hazard from one population to another. Uncertainty in risk assessment is overwhelmingly dominated by model

uncertainty, which arises due to a lack of direct studies on exposure, hazard and background disease burden.

We distinguish between uncertainty, which is due to gaps in knowledge, methods or data, and variability, which is a real property of the world and itself may be known with certainty or with uncertainty. Variability can nonetheless be a source of uncertainty in the absence of population-specific data on exposure or the exposure–response relationship. For many risk factors, data on exposure distributions are available for a limited number of populations or demographic groups. The exposure distribution for other populations are then extrapolated from the available data based on some model. As discussed earlier, the extrapolation model may be as simple as using the population-weighted average of the existing data or more complex (based on a number of predictors). In such cases, the statistical uncertainty of the estimator (e.g. the 95% confidence interval of the mean or regression coefficients) is an underestimation of true uncertainty in predicted values due to the unexplained variability in the data. More complex models can increase the predictive power and therefore reduce uncertainty but even the most sophisticated models are unlikely to fully explain the variability of the data, resulting in residual uncertainty. Variability can also be a source of uncertainty in the estimation and extrapolation of exposure–response relationships or relative risks that are measured in a limited number of settings. In the presence of multiple estimates of hazard, it is common to use meta-analytical approaches to obtain an overall estimate. At the same time, the differences between various estimates may reflect true variability in effect size, especially if obtained from different populations, resulting in uncertainty in hazard estimates.

Parameter uncertainty can be readily included in quantitative analysis using random-variable statistical methods (Morgan and Henrion 1990). While we have discussed the various sources of uncertainty, the important issue of extrapolation of exposure and hazard using models requires new approaches to quantifying uncertainty in the presence of limited data. Quantitative analysis of model uncertainty, by definition, would require considering the uncertainty of the models and assumptions used (including assumptions about disease mechanism or data/parameter extrapolation) using the methods of Bayesian statistics.

6. CONCLUSIONS

We have described a framework for systematic quantification of the burden of disease due to risk factors that attempts to unify the growing interest in health risks across a number of health, physical and social sciences. The key attributes of the framework, along with the corresponding methodological issues that arise in its application, are:

- comparing the burden of disease due to the observed exposure distribution in a population with the burden from a hypothetical distribution or series of distributions, rather than a single reference level such as non-exposed;
- considering the multiple stages of causality and interactions between risk factor(s) and disease outcome to allow inferences about combinations of risk factors for which epidemiological studies have not been conducted, including the joint effects of changes in multiple risk factors;
- calculating the health loss due to risk factor(s) as a time-indexed “stream” of disease burden due to a time-indexed “stream” of exposure, including consideration of discounting; and
- describing the sources of uncertainty in the risk assessment process.

For each of the above aspects, we have outlined the important conceptual and methodological issues and their implications for risk assessment. While this framework provides a means for considering risk factors in different layers of causality, with multiple counterfactuals (Ezzati et al. 2002; see also the risk factor chapters in this book), its application is limited by the availability of data on risk factors and hazards (Powles and Day 2002). The availability and form of data on both exposure and hazard are often determined by disciplinary boundaries as well as measurement difficulties. Analysis of selected risks highlights data and monitoring needs for better quantification and intervention strategies, especially more detailed data on exposure, hazard accumulation over time, and heterogeneity of risk factor–disease relationships.

For more effective and affordable implementation of a prevention paradigm, policies, programmes and scientific research should acknowledge and take advantage of the interactive role of major risks to health, across and within causality layers. Despite the methodological complexity and empirical difficulties, especially in estimating time-based multi-risk exposures, this framework provides a consistent basis for better and more comparable information about the various causes of disease and injury. In the remaining chapters of this book, the burden of disease due to a diverse set of risk factors is assessed according to this conceptual framework, with a detailed description of data sources and risk factor-specific methodological issues.

ACKNOWLEDGEMENTS

This work has been sponsored by the National Institute of Aging Grant 1-PO1-AG17625. The participants in the various CRA project collaborators meetings and two reviewers provided valuable comments. J. Powles and J. Barendregt provided extensive comments on an earlier draft.

NOTES

- 1 As discussed by Greenland and Robins (1988), attributable fractions without a time dimension are not able to characterize those cases whose occurrence would have been delayed in the absence of exposure. The authors recommend the use of etiologic fractions with a time dimension to account for this shortcoming. Time-based measures are discussed in more detail below.
- 2 At the extreme, a structural model would attempt to use chemical or physical principles as the unit of analysis and modelling. This would of course be currently impossible in studying any system that involves population health.
- 3 The “driving force, pressure, state, exposure, effect” (DPSEE) model of Corvalan et al. (1999) does consider the multiple layers of causality. This model however focuses on the risk evolution process, which is less suitable for multi-risk factor interaction within and between layers. More complete discussions of causality and multiple causes are provided by Yerushalmy and Palmer (1959), Evans (1976, 1978) and Rothman and Greenland (Rothman 1976; Rothman and Greenland 1998).
- 4 The exception is those risk factors whose effects occur through intermediate variables which are themselves a risk factor for the outcome considered, such as the relationships between physical inactivity or obesity and IHD, which are mediated through blood pressure or cholesterol. In such cases, controlling for the intermediate risk factor would result in a bias (towards the null) in the estimation of total hazard of the distal factor (Greenland 1987).
- 5 There are no past studies on “climate change” as it is expected to take place in the future. For this reason, the relationship between climate change and health would always be based on a model which relates climate change to meteorological variables (e.g. temperature or rainfall). The relationship between these variables, disease vectors and disease could then be estimated from past data and vector biology (Craig et al. 1999; Martens et al. 1999; Rogers and Randolph 2000).
- 6 If the variables in the n th layer of the causal-web are affected directly by those in the $(n - 2)$ th layer in addition to the $(n - 1)$ th layer, or by variables within the n th layer itself (see Figure 1.2 for an example), Equation 3a can be expanded to include these links as well:

$$\mathbf{X}^n = f(\mathbf{B}_0(\mathbf{X}^n), \mathbf{X}^n; \mathbf{B}_1(\mathbf{X}^{n-1}, \mathbf{X}^n) \mathbf{X}^{n-1}; \mathbf{B}_2(\mathbf{X}^{n-2}, \mathbf{X}^n) \mathbf{X}^{n-2}) \quad (3b)$$

This can be extended to interactions across multiple causal layers, and in general any variable in the system can be affected by any other one as the concept of causal layer becomes more flexible.

- 7 In this case when some of the variables affect not only the other variables in the causal system but also the relationship(s) between variables, they are equivalent to “effect modification” in epidemiological literature (Rothman 1976). Graphically, in Figure 1.1 they would be represented as links (arrows) not between two variables but as links from one of the variables (the effect modifier) to another link in the system.
- 8 In other words the $RR(x)$ in Equations (2a) and (2b) are functions of the other covariates, referred to as effect modification earlier. Epidemiological

studies that stratify relative risks based on covariates other than age and sex are however rare.

- 9 The notation $|_{T_0}^T$ denotes “estimated between T_0 and T ”.
- 10 Equivalent exposure is an analytical concept and need not be physically realizable. In fact for many risk factors, such as carcinogens, where the effects are from life-long exposure, the equivalent exposure would be so high that its occurrence at a single instant would be impossible.
- 11 Further, if there is threshold, M , below which exposure has no effect:

$$RR(x(t))|_{T_0}^T = RR\left(\int_{T_0}^T f(x(t-L))(\text{TRUE}(x(t-L) \geq M))dt\right)$$

where

$$\text{TRUE}(x(t) \geq M) = \begin{cases} 1 & \text{if } x(t) \geq M \\ 0 & \text{if } x(t) < M \end{cases}$$

This framework can be easily modified to include cases where exposure has different effects below and above threshold by using $\text{TRUE}(x(t) \geq M)$ for the effect above the threshold and $\text{TRUE}(x(t) < M)$ for the effect below the threshold.

- 12 In this manner, the evolution of the exposure distribution is analytically similar to a “random process” in which a probability density function (PDF) describes a random variable which is a function of time. Exposure to a risk factor is not a random variable in the strict sense. But since a time-dependent exposure distribution has an accumulated distribution of 1.0, it has the same representation as a random process.
- 13 Robins (1999a) discusses this issue in the case of estimating the effects of a dynamic treatment regime whose dose is dependent on symptoms.
- 14 The diffusion model also includes a stochastic component to account for those interactive effects not described by the drift term.
- 15 The functional form is $\int_0^\infty W_t \alpha^t$ where $0 < \alpha < 1$ in Koopmans (1960) and $\int_0^\infty W_t \exp(-\delta t)$ where $\delta > 0$ in Dasgupta and Mäler (1994) and Dasgupta et al. (1999); α and δ are the social rate of discount.
- 16 The last unit of sacrifice will have infinite marginal utility therefore matching the future infinite stream of benefits.
- 17 These two additional constraints are external to the economic efficiency arguments as defined by maximizing aggregate welfare. In fact, these additional constraints of minimum acceptable or non-declining welfare result in an “inefficient” outcome in order to achieve better distribution across generations (Montgomery 1999). See Weitzman (1999) for another argument for the choice of lower discount rates.
- 18 If exposure is sustained for a longer time in the risk assessment population than in the study population and if the whole exposure period contributes to hazards, this would result in an underestimation of risk (and vice versa). For example, in many cohorts in current epidemiological studies, BMI

increased when the subjects were in their twenties or thirties. There is however increasing child and adolescent overweight or obesity in many regions of the world. If this continues into adult life, the hazards may be higher than subjects in the current study cohorts.

REFERENCES

- Anand S, Hansen K (2002) Age weights and discounting in disability-adjusted life years: a critical review. In: *Summary measures of population health*. Murray CJL, Salomon J, Mathers CD, Lopez AD, Lozano R, eds. World Health Organization, Geneva.
- Arrow KJ (1999) Discounting, morality, and gaming. In: *Discounting and intergenerational equity*. Portney PR, Weyant JP, eds. Resources for the Future, Washington, DC.
- Barnes R (1965) Comparisons of blood pressure and blood cholesterol levels in New Guineans and Australians. *Medical Journal of Australia*, 1:611–617.
- Britton A, McKee M (2000) The relationship between alcohol and cardiovascular disease in eastern Europe: explaining the paradox. *Journal of Epidemiology and Community Health*, 54:328–332.
- Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C (1999) Threats to applicability of randomised trials: exclusion and selective participation. *Journal of Health Services Research and Policy*, 4:112–121.
- Carvalho JJ, Baruzzi RG, Howard PF et al. (1989) Blood pressure in four remote populations in the Intersalt study. *Hypertension*, 14:238–246.
- Chen Z, Peto R, Collins R, MacMahon S, Lu J, Li W (1991) Serum cholesterol concentration and coronary heart disease in population with low cholesterol concentrations. *British Medical Journal*, 303:276–282.
- Cole P, MacMahon B (1971) Attributable risk percent in case-control studies. *British Journal of Preventive and Social Medicine*, 25:242–244.
- Corrao G, Rubbiati L, Bagnardi V, Zambon A, Poikolainen K (2000) Alcohol and coronary heart disease: a meta-analysis. *Addiction*, 95:1505–1523.
- Corvalan CF, Kjellstrom T, Smith KR (1999) Health, environment, and sustainable development: identifying links and interactions to promote action. *Epidemiology*, 10:656–660.
- Craig MH, Snow RW, le Sueur D (1999) A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitology Today*, 15:105–111.
- Dasgupta P, Heal GM (1974) The optimal depletion of exhaustible resources. *Review of Economic Studies*, 41.
- Dasgupta P, Mäler K-G (1994) *Poverty, institutions, and the environmental resource base* (World Bank Environment Paper No. 9.) World Bank, Washington, DC.
- Dasgupta P, Mäler K-G, Barrett S (1999) Intergenerational equity, social discount rates, and global warming. In: *Discounting and intergenerational equity*. Portney PR, Weyant JP, eds. Resources for the Future, Washington, DC.

- Drescher K, Becher H (1997) Estimating the generalized impact fraction from case-control data. *Biometrics*, 53:1170–1176.
- Eastern Stroke and Coronary Heart Disease Collaborative Research Group (1998) Blood pressure, cholesterol, and stroke in eastern Asia. *The Lancet*, 352:1801–1807.
- Eide GE, Heuch I (2001) Attributable fractions: fundamental concepts and their visualization. *Statistical Methods in Medical Research*, 10:159–193.
- Evans AS (1976) Causation and disease: the Henle-Koch postulates revisited. *Yale Journal of Biology and Medicine*, 49:175–195.
- Evans AS (1978) Causation and disease: a chronological journey. *American Journal of Epidemiology*, 108:249–258.
- Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJL, Comparative Risk Assessment Collaborative Group (2002) Selected major risk factors and global and regional burden of disease. *The Lancet*, 360:1347–1360.
- Ezzati M, Saleh H, Kammen DM (2000) The contributions of emissions and spatial microenvironments to exposure to indoor air pollution from biomass combustion in Kenya. *Environmental Health Perspectives*, 108:833–839.
- Foulkes MA, Davis CE (1981) An index of tracking for longitudinal data. *Biometrics*, 37:439–446.
- Graham JD, Green LC, Roberts MJ (1988) *In search of safety: chemicals and cancer risk*. Harvard University Press, Cambridge.
- Greenland S (1984) Bias in methods for deriving standardized morbidity ratio and attributable fraction estimates. *Statistics in Medicine*, 3:131–141.
- Greenland S (1987) Quantitative methods in the review of epidemiologic literature. *Epidemiologic Reviews*, 9:1–30.
- Greenland S (2002) Causality theory for policy uses of epidemiological measures. In: *Summary measures of population health: concepts, ethics, measurement and applications*. Murray CJL, Salomon J, Mathers CD, Lopez AD, eds. World Health Organization, Geneva.
- Greenland S, Robins JM (1988) Conceptual problems in the definition and interpretation of attributable fractions. *American Journal of Epidemiology*, 128:1185–1197.
- He J, Klag MJ, Whelton PK et al. (1991a) Migration, blood pressure pattern, and hypertension: the Yi migrant study. *American Journal of Epidemiology*, 134:1085–1101.
- He J, Tang YC, Tell GS, Mo PS, He GQ (1991b) Effect of migration on blood pressure: the Yi people study. *Epidemiology*, 2:88–97.
- Hill AB (1965) The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, 58:295–300.
- Horton R (2000) Common sense and figures: the rhetoric of validity in medicine (Bradford Hill memorial lecture 1999). *Statistics in Medicine*, 19:3149–3164.

- Horwitz RI, Singer BH, Makuch RW, Viscoli CM (1998) Clinical versus statistical considerations in the design and analysis of clinical research. *Journal of Clinical Epidemiology*, 51:305–307.
- Horwitz RI, Singer BH, Viscoli CM, Makuch RW (1996) Can treatment that is helpful on average be harmful to some patients? A study of the conflicting information needs of clinical inquiry and drug regulation. *Journal of Clinical Epidemiology*, 49:395–400.
- Horwitz RI, Viscoli CM, Clemens JD, Sadock RT (1990) Developing improved observational methods for evaluating therapeutic evidence. *American Journal of Medicine*, 89:630–638.
- Howarth RH (1996) Climate change and overlapping generations. *Contemporary Economic Policy*, 14:100–111.
- Kay D, Prüss AM, Corvalan CF (2000) *Methodology for assessment of environmental burden of disease*. ISEE session on environmental burden of disease. World Health Organization, Geneva.
- Kneese AV (1999) The Faustian bargain. In: *The RFF reader in environment and resource management*. Oates WE, ed. Resources for the Future, Washington, DC.
- Koopmans TC (1960) Stationary ordinal utility and impatience. *Econometrica*, 28:287–309.
- Kunzli N, Kaiser R, Medina S et al. (2000) Public-health impact of outdoor and traffic-related air pollution: a European assessment. *The Lancet*, 356:795–801.
- LaRosa JC, He J, Vupputuri S (1999) Effect of statins on risk of coronary disease: a meta-analysis of randomized controlled trials. *Journal of American Medical Association*, 282:2340–2346.
- Law MR, Wald NJ, Thompson SG (1994) By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? *British Medical Journal*, 308:367–373.
- Leigh J, Macaskill P, Kuosma E, Mandryk J (1999) Global burden of disease and injury due to occupational factors. *Epidemiology*, 10:626–631.
- Levin ML (1953) The occurrence of lung cancer in man. *Acta Unio Internationalis Contra Cancrum*, 9:531–541.
- Lind RC, ed. (1982) *Discounting for time and risk in energy policy*, Johns Hopkins University Press for Resources for the Future, Baltimore, MD.
- MacMahon B, Pugh TF (1970) *Epidemiology: principles and method*. Little Brown and Company, Boston, MA.
- MacMahon S, Peto R, Cutler J et al. (1990) Blood pressure, stroke, and coronary heart disease. Part I, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *The Lancet*, 335:765–774.
- Maldonado G, Greenland S (2002) Estimating causal effects. *International Journal of Epidemiology*, 31:422–429.

- Mann GV, Shaffer RD, Anderson RS, Sandstead HH (1964) Cardiovascular disease in the Masai. *Journal of Atherosclerosis Research*, 4:289–312.
- Manton KG, Dowd JE, Stallard E (1993a) The effects of risk factors on male and female cardiovascular risks in middle and late age. In: *Forecasting the health of elderly populations*. Manton KG, Singer BH, Suzman RM, eds. Springer-Verlag, New York.
- Manton KG, Singer BH, Suzman RM (1993b) *Forecasting the health of elderly populations*. Springer-Verlag, New York.
- Manton KG, Stallard E, Singer BH (1994) Methods for projecting the future size and health status of the US elderly populations. In: *Studies in the economics of aging*. Wise DA, ed. University of Chicago Press, Chicago, IL.
- Martens P, Kovats RS, Nijhof S et al. (1999) Climate change and future populations at risk of malaria. *Global Environmental Change-Human and Policy Dimensions*, 9:S89–S107.
- Mathers CD, Ezzati M, Lopez AD, Murray CJL, Rodgers AD (2002) Causal decomposition of summary measures of population health. In: *Summary measures of population health: concepts, ethics, measurement and applications*. Murray CJL, Salomon J, Mathers CD, Lopez AD, eds. World Health Organization, Geneva.
- McGinnis JM, Foege WH (1993) Actual causes of death in the United States. *Journal of American Medical Association*, 270:2207–2212.
- Miettinen OS (1974) Proportion of disease caused or prevented by a given exposure, trait or intervention. *American Journal of Epidemiology*, 99: 325–332.
- Montgomery WD (1999) A market-based discount rate: comments on Bradford. In: *Discounting and intergenerational equity*. Portney PR, Weyant JP, eds. Resources for the Future, Washington, DC.
- Morgan MG, Henrion M (1990) *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*, Cambridge University Press, Cambridge.
- Murray CJL (1996) Rethinking DALYs. In: *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Global Burden of Disease and Injury, Vol. 1. Murray CJL, Lopez AD, eds. Harvard School of Public Health on behalf of WHO, Cambridge, MA.
- Murray CJL, Acharya A (2002) Age weights and discounting in health gaps reconsidered. In: *Summary measures of population health: concepts, ethics, measurement and applications*. Murray CJL, Salomon J, Mathers CD, Lopez AD, eds. World Health Organization, Geneva.
- Murray CJL, Lopez AD, eds. (1996) *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Global Burden of Disease and Injury, Vol. 1. Harvard School of Public Health on behalf of WHO, Cambridge, MA.
- Murray CJL, Lopez AD (1997) Global mortality, disability, and the contribution of risk factors: global burden of disease study. *The Lancet*, 349:1436–1442.

- Murray CJL, Lopez AD (1999) On the comparable quantification of health risks: lessons from the global burden of disease. *Epidemiology*, **10**:594–605.
- Murray CJL, Salomon J, Mathers CD, Lopez AD, eds. (2002) *Summary measures of population health: concepts, ethics, measurement and applications*. World Health Organization, Geneva.
- National Research Council (1994) *Science and judgment in risk assessment*. National Academy Press, Washington, DC.
- Ouellet BL, Romeder JM, Lance JM (1979) Premature mortality attributable to smoking and hazardous drinking in Canada. *American Journal of Epidemiology*, **109**:451–463.
- Parfit D (1984) *Reasons and persons*. Oxford University Press, New York.
- Patz JA, McGeehin MA, Bernard SM et al. (2000) The potential health impacts of climate variability and change for the United States: executive summary of the report of the health sector of the US national assessment. *Environmental Health Perspectives*, **108**:367–376.
- Peto R (1986) Influence of dose and duration of smoking on lung cancer rates. In: *Tobacco: a major international health hazard*. Zaridze D, Peto R, eds. (IARC Scientific Publication No. 7.) International Agency for Research on Cancer, Lyon.
- Peto R, Lopez AD, Boreham J, Thun M, Heath C, Jr. (1992) Mortality from tobacco in developed countries. *The Lancet*, **339**:1268–1278.
- Pignone M, Phillips C, Mulrow C (2000) Use of lipid lowering drugs for primary prevention of coronary heart disease: meta-analysis of randomised trials. *British Medical Journal*, **321**:983–986.
- Portney PR, Weyant JP, eds. (1999) *Discounting and intergenerational equity*. Resources for the Future, Washington, DC.
- Powles J, Day N (2002) Interpreting the global burden of disease. *The Lancet*, **360**:1342–1343.
- Preston SH (1976) *Mortality patterns in national populations: with special reference to recorded causes of death*. Academic Press, New York.
- Prospective Studies Collaboration (1995) Cholesterol, diastolic blood pressure, and stroke: 13 000 strokes in 45 000 people in 45 prospective cohorts. *The Lancet*, **346**:1647–1653.
- Puddey IB, Rakic V, Dimmitt SB, Beilin LJ (1999) Influence of pattern of drinking on cardiovascular disease and cardiovascular risk factors: a review. *Addiction*, **94**:649–663.
- Reiter P (2001) Climate change and mosquito-borne diseases. *Environmental Health Perspectives*, **109**:S141–161.
- Rice AL, Sacco L, Hyder A, Black RE (2000) Malnutrition as an underlying cause of childhood deaths associated with infectious diseases in developing countries. *Bulletin of the World Health Organization*, **108**:367–376.
- Robins JM (1986) A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survival effect. *Mathematical Modelling*, **7**:1393–1512.

- Robins JM (1987) Addendum to “a new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survival effect”. *Computers and Mathematics with Applications*, **14**:917–921.
- Robins JM (1997) Causal inference from complex longitudinal data. In: *Latent variable modeling and applications to causality. Lecture notes in statistics (120)*. Berkane M, ed. Springer-Verlag, New York.
- Robins JM (1999a) Association, causation, and marginal structural models. *Synthese*, **121**:151–179.
- Robins JM (1999b) Marginal structural models versus structural nested models as tools for causal inference. In: *Statistical models in epidemiology*. Halloran E, ed. Springer-Verlag, New York.
- Robins JM, Greenland S (1989) Estimability and estimation of excess and etiological fractions. *Statistics in Medicine*, **8**:845–859.
- Robins JM, Greenland S (1991) Estimability and estimation of expected years of life lost due to a hazardous exposure. *Statistics in Medicine*, **10**:79–93.
- Robins JM, Greenland S (2000) Comment on “causal inference without counterfactuals” by Dawid AP. *Journal of American Statistical Association—Theory and Methods*, **95**:477–482.
- Robins JM, Greenland S, Fu-Chang H (1999) Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome. *Journal of American Statistical Association*, **94**:687–700.
- Rockhill B, Newman B, Weinberg C (1998) Use and misuse of population attributable fractions. *American Journal of Public Health*, **88**:15–19.
- Rogers DJ, Randolph SE (2000) The global spread of malaria in a future, warmer world. *Science*, **289**:1763–1766.
- Rothman KJ (1976) Causes. *American Journal of Epidemiology*, **104**:587–592.
- Rothman KJ, Greenland S (1998) *Modern epidemiology*. Lippincott-Raven, Philadelphia, PA.
- Schelling TC (1995) Intergenerational discounting. *Energy Policy*, **23**:395–401.
- Single E, Robson L, Rehm J, Xie X (1999) Morbidity and mortality attributable to alcohol, tobacco, and illicit drug use in Canada. *American Journal of Public Health*, **89**:385–390.
- Smith KR (2000) The national burden of disease from indoor air pollution in India. *Proceedings of the National Academy of Sciences*, **97**:13286–13293.
- Smith KR, Corvalan CF, Kjellstrom T (1999) How much global ill health is attributable to environmental factors. *Epidemiology*, **10**:573–584.
- Smith KR, Samet JM, Romieu I, Bruce N (2000) Indoor air pollution in developing countries and acute lower respiratory infections in children. *Thorax*, **55**:518–532.
- Solow RM (1974a) The economics of resources or the resources of economics. *The American Economic Review*, **64**:1–14.

- Solow RM (1974b) Intergenerational equity and exhaustible resources. *The Review of Economic Studies*, 41:29–45.
- Solow RM (1999) Foreword. In: *Discounting and intergenerational equity*. Portney PR, Weyant JP, eds. Resources for the Future, Washington, DC.
- Tate RB, Manfreda J, Krahn AD, Cuddy TE (1995) Tracking of blood pressure over a 40-year period in the University of Manitoba follow-up study, 1948–1988. *American Journal of Epidemiology*, 142:946–954.
- Thun MJ, Apicella LF, Henley SJ (2000) Smoking vs other risk factors as the cause of smoking-attributable mortality: confounding in the courtroom. *Journal of American Medical Association*, 284:706–712.
- Toman MA (1999) Reconciling philosophy and economics in long-term discounting. In: *Discounting and intergenerational equity*. Portney PR, Weyant JP, eds. Resources for the Future, Washington, DC.
- Uter W, Pfahlberg A (2001) The application of methods to quantify attributable risk in medical practice. *Statistical Methods in Medical Research*, 10: 231–237.
- Walter SD (1980) Prevention of multifactorial disease. *American Journal of Epidemiology*, 112:409–416.
- Weitzman ML (1999) Just keep discounting, but ... In: *Discounting and intergenerational equity*. Portney PR, Weyant JP, eds. Resources for the Future, Washington, DC.
- Willet WC (2002) Balancing life-style and genomics research for disease prevention. *Science*, 296:695–698.
- WHO (1992) *International statistical classification of disease and related health problems. Tenth edition*. World Health Organization, Geneva.
- Yashin AI, Manton KG, Stallard E (1986) Dependent competing risks: a stochastic process model. *Journal of Mathematical Biology*, 24:119–140.
- Yerushalmy J, Palmer CE (1959) On the methodology of investigations of etiologic factors in chronic diseases. *Journal of Chronic Disease*, 108:27–40.
- Yong LC, Kuller LH, Rutan G, Bunker C (1993) Longitudinal study of blood pressure: changes and determinants from adolescence to middle age. The Dormont High School follow-up study, 1957–1963 to 1989–1990. *American Journal of Epidemiology*, 138:973–983.