Chapter 2

THEORETICAL PERSPECTIVES ON
DETERMINANTS OF MORTALITY

2.1 Introduction

This chapter sets out the theoretical framework on which the analyses are based and involves socioeconomic and time dimensions. The socioeconomic dimension of the framework relates to the main goal of the research: to identify socioeconomic and other factors that are associated with mortality. These factors do not occur randomly and seldom operate alone: mortality risks are determined by causal chains consisting of a series of factors operating successively over time. Therefore, a knowledge of the mechanisms that are behind mortality trends and variations is essential for the prediction of future mortality trends (Kunst et al., 1999; Willekens, 1990). Such factors include health exposures that range from very specific behavioural risk factors (e.g. smoking) to factors that affect health indirectly (e.g. education), and in which the indirect factors are also related to the direct factors (e.g. education to smoking).

The second theoretical dimension arises from the principle that whatever the risk factor, the outcome will always come later. With the perspective that about 90% of mortality occurs after the age of 45, the latency period of a risk factor becomes important in considering the study of mortality differences as well as for projection purposes. In regard to adult mortality this period stretches from gestation to later adult life (Kuh and Ben-Shlomo, 1997). In ecological analysis this latency effect can to some extent be taken into account by calculating time lags for each variable included in the analysis provided there are long time series of data.

Exposure to health determinants occurs throughout life, and may be specific to the individual (e.g. biological and behavioural factors) or shared by those in the same cohort or time period (e.g. the educational and health care systems). As noted previously, the health effects of most of these determinants is not instantaneous. The two elements, the different types of health-related exposures and the time component, are essential aspects of what is referred to as the life course approach. However, while within demographic research the life course is an accepted framework (Van Wissen and Dykstra, 1999), it remains difficult to apply in ecological studies, but one important contribution
of the life course approach for this thesis is that it provides an opportunity to set out all the relevant variables that are associated with mortality.

The next section provides several possible explanations for social and economic mortality differences, and is followed by a more detailed outline of established direct and indirect factors that may influence the likelihood of succumbing to a particular disease or that contribute to mortality in general. The penultimate section of this chapter discusses why the problem of ecological fallacy – that is, making inferences about the individual from macro-level data – is not a major concern for this thesis. The chapter concludes with a discussion on how the life course theory can be applied in macrodemographic research and how it fits in conceptually.

2.2 Explanations of socioeconomic mortality differences

In recent years, several well-recognised studies have found evidence of a negative association between socioeconomic status (SES) – that is, the position on the social ladder according to level of education, income, type of occupation or a combination these – and mortality (see Van de Mheen and Mackenbach, 1990; Balarajan and McDowall, 1988; Davey Smith et al., 1998b; Kunst et al., 1997; Marmot et al., 1991; Marmot and McDowall, 1986). Empirical studies of socioeconomic health differences (SEHD) have generally paid most attention to the influence of social position and related differential exposures in adulthood (Davey Smith, 1997). However, research on the effect of unfavourable social environments in early life on health in adulthood is by no means new (see ibid. for references). One of the more recent propositions is Barker’s "foetal origins" hypothesis, which proposes that early life stages, such as the environment during foetal and infant life, programme people from socioeconomically unfavourable backgrounds to be at an elevated risk of adult ischaemic heart disease (IHD) (Barker, 1995). Several others have taken a more holistic approach by emphasising a greater range of biological and social experiences in childhood, adolescence and early adulthood, which combine to influence adult disease risk (e.g., Kuh and Ben-Shlomo, 1997; Martelin, 1994; Stronks, 1997). This framework of reasoning is usually referred to as the life course approach and, according to this approach, both exposures that act throughout the life course and interactions between exposures that occur at various stages of life, contribute to disease risk (Davey Smith, 1997). This implies that socioeconomic factors affect the risk of disease throughout life, since health-damaging exposures and health-enhancing opportunities are socially patterned. This suggests that an individual’s response to health-damaging exposures or health-enhancing opportunities – which may modify their impact or alter the risk of future exposures – will be powerfully affected by the social and economic experience of the individual (Kuh et al., 1997).

In the existing literature, there are several conceptual frameworks that attempt to explain how inequalities in terms of SES are related to inequalities in morbidity and/or mortality which only
focus on adult ages (e.g. Martikainen, 1990; Kunst, 1997; Cavelaars, 1998). Other, more sophisticated, frameworks include other life stages and also non-socioeconomic factors (e.g. Martelin, 1994; Stronks, 1997). However, in general, they depict similar relationships: mortality is caused by both direct and indirect determinants of disease. Figure 2.1 is a simplified illustration of this, and serves as the framework for this thesis. Implicit in the model is that mortality is explained by causal mechanisms, through which mortality is either directly or indirectly determined. According to the causation mechanism, a person’s SES affects their health, not directly, but through more specific determinants of health and disease (Davey Smith et al., 1994) which are called proximate risk indicators in the model. The “proximate” part of this term reflects direct associations between determinants and mortality: there are no intermediary factors in the relationship between a determinant and mortality. “Risk” refers to aspects of personal behaviour or lifestyle (e.g. smoking, dietary habits, the consumption of drugs and alcohol, physical activity), material factors (e.g. working and housing conditions), psychosocial stress-related factors (e.g. stressful events such as losing a job or the death of a spouse), or an inborn or inherited characteristic which, based on epidemiological evidence, is known to be associated with undesirable health-related condition(s) (Last, 1988; Stronks, 1997). Throughout the text, the term “indicator” is used rather than “factor” because not all links between a proximate risk factor and a disease are causal. According to Wolleswinkel-van den Bosch (1998), a change in a proximate risk indicator (or “proximate determinant” as she called it) is expected to be directly related to mortality change, while indirect or distal determinants can initiate changes in the proximate risk indicators that could then lead to mortality decline. A person’s SES is an example of such a determinant. Although distal determinants are only indirectly associated with mortality change, they are more stable than the proximate variables and can, in some cases, be more easily perceived as a condition for mortality change than as a determinant of mortality change. Other distal determinants include macroeconomic conditions (e.g. Gross Domestic Product (GDP)), political institutions, sociocultural factors (e.g. religion), health care factors (e.g. effectiveness of the health care system) and ecological setting (e.g. climate). However, contrary to Wolleswinkel-van den Bosch, I refer to these latter distal factors throughout the thesis as contextual factors, because these factors are collective in nature, as for instance the market value of the goods and services produced by one individual has little influence on the GDP of a country. However, as the model implies, in the context of adult mortality, contextual factors can instigate a direct change in the proximate risk indicators, as well as through both early life factors and adult socioeconomic status. In addition, there are also associations within each group of factors (e.g. between a political system and GDP, or between education and income). With the exception of the contextual factors, it is assumed that the strength of the relationship between the determinants and mortality will often vary according to sociodemographic characteristics. As examples, childhood illnesses are generally less of a life threat among recent cohorts than they were 50 years ago, and women experience an increase in the risk of IHD after menopause as oestrogen is no longer produced (Meeter and Witteman, 1997).
The conceptual framework also includes causes of death because the strength and direction of the relationship between a proximate risk indicator and mortality will vary according to the cause of death. For instance, whereas alcohol has no known causal association with respiratory system disease mortality, about 85% of the deaths from chronic liver disease and cirrhosis of the liver (LDC) have been attributed to an excess alcohol consumption (Nizard and Muñoz-Perez, 1994). Conversely, alcohol appears to protect against IHD when consumed in moderate amounts (Renaud and de Lorgeril, 1992). Furthermore, certain causes of death have one or two very specific proximate risk indicators that cause the onset of the disease (e.g. lung cancer and smoking), while other diseases, such as IHD, have a larger range of important (some interrelated) proximate risk indicators (e.g. a family history of the disease, physical inactivity, obesity, high blood pressure, high serum cholesterol and smoking). Further, each indicator has its own latency period for a particular cause of death. This can be seen, for instance, by the pattern of lung cancer deaths, which began to drop in many parts of Europe 30 years after smoking rates began to decline, while a similar fall in IHD deaths occurred after just 15 years (Ruwaard and Kramers, 1993). Variations also exist in the time lags of the indirect effects of the distal determinants on cause-specific mortality. What it all boils down to is that studying only all-cause mortality differences between countries or over time makes little sense, because variations or changes in total mortality are caused by differences or changes in the cause-of-death structure (see also Chapter 4). Therefore, given the aetiological differences in causes of death, differences or changes in the cause-of-death structure imply more rapid differences or changes in the prevalence of those proximate risk indicators that were instigated by differences or changes in distal determinants, and these are the main factors of interest in this thesis. Studying mortality by cause of death therefore allows one to gain initial insights into possible reasons for mortality differences, since each cause of death has its own
aetiology (although there will be some overlap when analysing larger groups of causes). In turn, the knowledge of the aetiology of the specific diseases enables a better estimate to be made of the impact of the proximate and distal determinants of the specific causes of death by considering the differences in magnitudes and latency periods.

Besides causation, mortality differences between socioeconomic groups may also be the result of selection effects. For example, if one believes that health has an effect on one’s socioeconomic position, people with poor health are less likely to move up the social ladder and are more likely to descend than those in good health (Stronks, 1997). This is particularly the case with childhood illnesses as a child unwillingly gains less knowledge due to school absence, and an illness during childhood also seems to have a bearing on health status later in life which may contribute to class differences in health among adults (Lundberg, 1988). Closely related to the selection explanation is the explanation of inequalities in health due to genetic predisposition, which argues that people with a genetic illness will eventually move down the social ladder, resulting in an uneven distribution of genetic material among the population. In other words, SEHD is partly the outcome of differences in the distribution of genetic factors in the different segments of the population. A fourth explanation is statistical artefact, which assumes that established inequalities in health between socioeconomic groups are biased by the research methods and the instruments used (Stronks, 1997). For example, results may be biased by the way that mortality is measured, as was shown in a paper on social class differences in mortality in the United Kingdom (Davey Smith et al., 1994). However, evidence suggests that statistical artefact, genetic predisposition and the selection mechanism do not make a significant contribution to the explanation of SEHD or provide reliable estimates (Stronks, 1997). Therefore, in this thesis, they will not be considered further.

2.3 Socioeconomic determinants of adult mortality from the life course perspective

As noted earlier, the life course may be regarded as combining interacting biological and social elements. The biological development of an individual takes place within a social context which structures life chances such that advantages and disadvantages tend to cluster cross-sectionally and accumulate longitudinally. For example, during a particular period of life, working and living conditions may both be health promoting (e.g. a well-paid non-hazardous job, which permits one to reside in good-quality housing and have a varied diet). At the same time, this position is likely to have been preceded by, and will be succeeded by, similar advantages in other phases of life such as a good education and a secure pension. These social processes interact with health in a number of ways and the relationships may be direct and disease-specific, or more general (Blane, 1999).
In other words, the life course approach does not only consider adult life factors, but also social experiences in childhood, adolescence and early adulthood which, together with biological experiences in the foetal and post-natal period, combine to influence adult disease risk. However, “what appears to be important is not any one factor which has a major long-term influence on health, but a number of comparatively small differences which become linked into a chain of disadvantage” (ibid.). Nevertheless, it has been suggested that while poor childhood growth and development, and adverse early environmental conditions, may increase the likelihood of acquiring a chronic disease, other lifetime exposures may contribute more to the underlying causal factor (Ben-Shlomo and Davey Smith, 1991).

2.3.1 Early life stages

In the industrialised world, the majority of people die at an adult age from chronic diseases that are usually associated with adult risk factors, particularly adult life style. However, research has also shown that certain early life factors are associated with later adult risk factors or disease, particularly those associated with so-called “critical periods” where exposures have long lasting or lifelong effects on the structure or function of organs, tissues and body systems, and which are not modified in any dramatic way by later experience. This is also known as “biological programming” and is the basis of Barker’s “foetal origins of adult disease” hypothesis (Ben-Shlomo and Kuh, 2002). To give an example, under-nutrition during middle or late gestation has been shown to raise the risk of IHD during adulthood as a result of programming blood pressure, fibrinogen concentration, factor VIII concentration and glucose intolerance (Paneth and Susser, 1995). Known markers of lack of nutrients or oxygen at particular stages of gestation are small birth size, disproportional head size, low height and weight. These reflect (sometimes permanent) adaptations that the foetus makes to sustain its development (Barker, 1995). Associations between reduced early growth, cardiovascular risk factors and IHD mortality are found in various populations and remain after controlling for socioeconomic factors and measures of adult lifestyle (Fall et al., 1995a, 1995b; Leon et al., 1998; Leon and Ben-Shlomo, 1997). However, the same patterns are not apparent world-wide, for example, in developing countries such as China babies tend to be proportionally small at birth but IHD is a rare event (Barker, 1995). In addition to cardiovascular diseases, there is evidence regarding a causal role between early life factors and adult morbidity and mortality in terms of respiratory system diseases, particularly chronic bronchitis which may be caused by repeated childhood infections. With other diseases, associations have usually ‘disappeared’ after controlling for factors such as deprivation and social class (Ben-Shlomo and Davey Smith, 1991). Further, in addition to foetal and infant body size and childhood infectious diseases, various sociodemographic and socioeconomic characteristics of the family and some aspects of the maternal characteristics
before and during pregnancy may also augment the risk of adult chronic diseases (Kuh and Ben-Shlomo, 1997).

2.3.2 Social pathways between childhood and adult health

The social environment during childhood also has an impact on the susceptibility to acquiring a disease, because socioeconomic factors act across the life course, affecting health and risk of death at any age (see for example, Martelin, 1994, Stronks, 1997, Davey Smith et al., 1997, 1998a). In particular, mortality in late middle age and early old age appears to be associated with various phases of the life course trajectories (Blane, 1999). For example, in one British study, adverse social and material circumstances in childhood (measured by the father’s occupational class) had a specific influence on mortality from strokes and stomach cancer in adulthood, and was not due to the continuity of the social disadvantage throughout life (Davey Smith et al., 1998a). In other words, the negative associations that were found between socioeconomic factors during adulthood and mortality were, to a certain extent, an effect of social pathways between the childhood socioeconomic environment and the adult socioeconomic environment. In particular, adolescence and the early adulthood period are important life stages as many decisions are then made about education, occupation and marriage. These decisions have a considerable impact on an individual’s life trajectory, and may therefore be associated with different disease risks since health-damaging exposures or health-enhancing opportunities are socially patterned (Kuh and Ben-Shlomo, 1997).

The socioeconomic environments during childhood and adulthood are highly correlated. The former may be indicated by the family background, such as parental education, father’s social class and other parental or household characteristics. These factors are to a large extent responsible for the educational opportunities and attainment during adolescence and early adulthood, which have important influences as the level of education is a powerful predictor of adult occupation and subsequent income. Children from more favoured family backgrounds have a much better chance of achieving a high social status in adult life, and therefore the influence of parents plays an important role in the development of the social capital of their offspring. Affluent parents will often provide financial backing and social contacts that will help prepare their children for a similar class position and capacity for earning. Other characteristics of the social environment (school and peers, workplace and peers) also help to shape an individual’s social capital, which will, in turn, affect behaviour. The social environment is also influenced by the socioeconomic and cultural characteristics of the time and place (Kuh et al., 1997).

---

6 Social capital comprises cognitive and social skills, coping strategies, self-esteem, attitudes and values. A family’s social capital also includes relationships between family members (Kuh et al., 1997).
When entering early adulthood, a number of sources of risk for later health, such as a greater risk of exposure and vulnerability to psychosocial stress, and behaviours, such as cigarette smoking, are acquired. On the other hand, "in late adolescence or early adult life, key life transitions provide an opportunity to discontinue adolescent behaviour patterns. For example, the acquisition of educational qualifications is associated with healthier diets, less smoking and more exercise in adult life, and the formation of stable partnerships and families moderates heavy drinking" (Kuh et al., 1997). Additional sources of risk are incurred in adulthood due to the constraints and opportunities afforded by one’s adult socioeconomic position, which is again the outcome of earlier socioeconomic processes. Little is known about whether a good start to life protects one against the impact of poor adult circumstances, but there is evidence that behavioural risk factors are more strongly associated with socioeconomic position in adulthood than in childhood (Blane et al., 1996). Nevertheless, health-related behaviours might have a greater impact on health in the non-manual classes because of the risk accumulation from other factors that play a greater role in the manual classes, such as their higher risk of exposure to a range of physical hazards in their residential and occupational environments, while the low status attached to occupying a position at or near the bottom of a social hierarchy is in itself a source of stress (Kuh et al., 1997).

Moreover, social and economic experiences determine to some extent an individual’s possible responses to health-damaging exposures or health-enhancing opportunities which may modify their impact or alter the risk of future exposures. The positive or adverse effects of the socioeconomic environment are hypothesised as being due to cumulative exposure over the entire life course: although some development stages may be more sensitive, there are no critical periods of development during childhood or adolescence (Kuh et al., 1997).

### 2.3.3 Socioeconomic factors in adult life

The majority of deaths occur in old age. Adult diseases are usually associated with adult risk factors, where socioeconomic factors belong to distal determinants that affect mortality through more-proximate determinants of disease. Accordingly, a change in a factor at the distal level may initiate change in a proximate risk indicator. In Figure 2.1, both contextual and adult life factors in the form of socioeconomic factors are regarded as distal determinants of disease.

Social inequalities exist in society because both material and non-material resources and rewards are unequally distributed. This inequality can be portrayed as a system of social stratification, in which people attain positions according to their educational attainment, occupation and income. This position is usually referred to as socioeconomic status or SES for short (Kunst, 1997). According to the life course approach, one’s adult SES is largely determined during adolescence and early adulthood, and there is only limited evidence of intragenerational mobility (i.e. a change in social status during adult life; Davey Smith et al., 1994), although shifts in the labour market have been
common during the recent economic transition in Eastern Europe (Večerník and Matějů, 1999). Employment status is a fourth socioeconomic indicator that is known to be associated with mortality: being unemployed, particularly in the long-term, raises the mortality risk (Martikainen, 1990). Further, the lower socioeconomic classes are more likely to experience unemployment than those who are higher up the social ladder, particularly during economic recessions (Bartley and Owen, 1996).

In recent years, many studies have been devoted to associations between socioeconomic factors and mortality. The fact that socioeconomic inequalities in mortality exist has been of concern to policy makers as these differences were considered to be unnecessary, unfair and avoidable (Kunst et al., 1999). Several studies have searched for explanations for the socioeconomic differences in mortality by analysing the simultaneous effects of socioeconomic status, lifestyle, and material factors on health. Results have tended to show that a substantial part of the observed socioeconomic inequalities in health are due to the differential distribution of behavioural factors and material factors in adulthood across the different socioeconomic groups (e.g. Marmot et al., 1991; Davey Smith et al., 1994; Stronks, 1997; Schrijvers et al., 1998). In other words, the higher up the socioeconomic ladder, the more favourable the material conditions and behavioural patterns are, and the lower the mortality. Behavioural factors, such as alcohol consumption, smoking and physical activity, are often partly embedded in material factors, such as income, adverse housing conditions and overcrowding. People who engage in health damaging behaviour such as smoking may do so in order to cope with adverse material circumstances. (Davey Smith et al., 1994).

As education, occupation, income and employment status are all very much interrelated, the patterns of behaviour and material factors are similar between the various socioeconomic factors. For example, a person’s socioeconomic background, familial environment and personal characteristics may all influence the attainment, or not, of a high level of education and a high status job, and a job that is high on the occupational ladder is usually secure and well-paid. Nevertheless, as differences have also been observed in the relationships between education, occupation, income, employment status and mortality they should be studied separately. The differences occur because they are related to differing kinds of resources and rewards (Kunst, 1997). The independent contribution of each of the four factors to SEHD is briefly described below:

2.3.3.1 Education
The influence of education on health has been depicted as indirect and positive. Reasons as to why health and mortality variations exist between different levels of education include:

i. the acquisition of knowledge regarding health damaging behaviours;
ii. the increased ability to benefit from new information;
iii. the ability to optimise the use of health services;
iv. the development of time preferences favourable to health maintenance;
v. an increasing willingness to invest in human capital;
vi. the promotion of the psychological attributes of high self esteem and self efficacy and a sense of mastery and control (Davey Smith et al., 1998b, Kunst, 1997).

The beneficial effects on health of these resources are assumed to work through more proximate determinants of health, such as less smoking and drinking, and the ability to cope with stress, which implies that a rise in the educational level of a population will have a favourable influence on future trends of mortality.

2.3.3.2 Occupation
Among the different types of socioeconomic factor, occupation has probably been the most investigated in epidemiology in recent years. While no standard form of social stratification is applied, a frequent distinction that is made is between manual and non-manual labour (e.g. Kunst et al., 1998a). Other studies distinguish additional occupational classes. For example, the Registrar General’s social class scheme that is used in official statistics in Britain has five levels, from professional to unskilled occupations, whereby the third social class, skilled occupations, has been split into non-manual and manual (OPCS, 1978). The well-known Whitehall studies uses five different social grades of civil servants (e.g. Marmot et al., 1984). In some instances, similarities in social standards and behaviour are also taken into account, as in the Erikson-Goldthorpe-Portocarero scheme (Vågerö and Erikson, 1997). Each of these studies, as well as others (e.g. Valkonen et al., 1993a; Davey Smith et al., 1998b; Kunst, 1997) confirm the existence of an occupational mortality gradient, i.e. the lower on the social scale, the higher the mortality. When farmers were included in such studies, they usually showed lower levels of all-cause mortality than manual workers, but higher levels than the white collar workers in a given period (e.g. Valkonen et al., 1993a).

As noted earlier, the origins of manual workers being at a higher risk of dying than non-manual workers precede entering the workforce. The attainment of a high occupational class, for example, is influenced through education and/or more directly by a person’s socioeconomic background, family environment and personal characteristics. Any subsequent material and non-material rewards may protect a person against exposure to a wide range of risk factors for disease, including unhealthy housing and working conditions, psychosocial stress, and unhealthy consumption patterns (Kunst, 1997).

The socioeconomic differences that exist in exposure are translated into socioeconomic differences in specific causes of death, although they are not necessarily consistent in direction. For example, in the case of lung cancer, early Finnish studies showed higher mortality rates among upper non-manual workers when compared to manual workers (Valkonen, 1993a), while later studies revealed an opposite trend (Valkonen et al., 1993). Such a reversal in social gradient has yet to occur in all
developed countries, for example in Portugal lung cancer mortality rates among men aged 45-59 are currently similar among different occupational classes (Kunst, 1997), while in the case of women, particularly in southern Europe, smoking prevalence rates remain highest among the higher occupational classes (Cavelaars, 1998).

What should also be taken into account is that the effect of one particular proximate risk indicator only explains a certain amount of the SEHD that is observed. For instance, results from the Whitehall I study showed that an administrator who smoked twenty cigarettes a day had a lower risk of lung cancer mortality than a lower grade civil servant who smoked the same amount, and that a social gradient in IHD also existed among non-smokers. This is probably because the general susceptibility to ill health operates on both the biological and social levels. At the social level, lower grade workers are more exposed to a range of specific influences on health that includes smoking, but also poverty, stressful life events, and the psychosocial work environment. Biologically, lower grade workers are shorter in height, influenced by both genes and childhood environment, and this has been shown to be a predictor of adult mortality independent of employment grade (Marmot et al., 1984; Marmot, 1995; Davey Smith et al., 1998c; Karasek and Theorell, 1990).

Part of the observed social class differences in mortality is also due to occupational hazards that tend to be greater among manual, rather than non-manual, jobs. A very specific example is from England & Wales where, in 1970-72, mortality from diseases of the respiratory system among male miners and quarrymen aged 15-64 was about 2.5 times higher than the national average (OPCS, 1978). A more recent Finnish study concluded that although inequalities associated with occupational category and class were more attributable to varying levels of education and income, the elimination of adverse working conditions would have reduced the number of all circulatory system disease deaths by 8% and those from myocardial infarction and stroke deaths by 10% and 18% respectively (Virtanen and Notkola, 2002).

Occupational mortality differences do not only exist in market economies, but also in former socialist countries despite the smaller differences in income, status and other attributes between high and low socioeconomic groups. In fact, the mortality range is larger, or at least as large, in Eastern European countries as in Western countries. This suggests that “some other features of the former socialist society have overshadowed the positive effects, if any, of egalitarian socioeconomic, health care and other policies” (Kunst, 1997). Given the larger proportion of industrial workers in Eastern Europe, this is likely to be an important factor in explaining mortality differences between Western and Eastern Europe.

2.3.3.3 Personal income

It has frequently been suggested that the inverse association between social class and health reflects the causal effect of income (Stronks, 1997). Income is one of the material rewards of occupation, and reflects the material standard of living since families or persons with higher incomes can spend more money on items such as food, accommodation and clothing, which may lead to better health
(Van Poppel, 1991). In a study conducted by Statistics Netherlands (Kardaun and Glerum, 1995),
deads that occurred in 1990 were linked to 1989 income data from the Netherlands Tax Department
(Belastingdienst). Deaths at ages below 20 and over 89 were excluded, leaving a total of 110,000
death for the analysis. Results showed that the male death rate was higher for those with a low
income compared to those with a high income. Among women, the relationship between income
and death was ‘n’-shaped, i.e. middle income women had higher death rates than both low- and
high-income earners. This relationship largely applied to most of the causes of death investigated
(see Table 2.1). The strongest associations between income and mortality were found for CRB,
OHD and lung cancer among men; and IHD and traffic accidents among women, although the
association generally disappeared or diminished with age. The male results are similar to a US
follow-up study in which, during a 16-year period from 1973-75 more than 300,000 white men
were followed (Davey Smith et al., 1996). Results showed that for the same causes of death as listed
here for the Dutch study, an increased income led to lower death rates and that adjusting for risk
factors did not diminish the effect by much.

<table>
<thead>
<tr>
<th>Cause of death</th>
<th>Men (linear)</th>
<th>Men (quadratic)</th>
<th>Women (linear)</th>
<th>Women (quadratic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause mortality</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Stomach cancer</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Colon cancer</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Breast cancer (women only)</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Remaining cancer</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Ischaemic heart disease (IHD)</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Other heart diseases (OHD)</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Cerebrovascular disease (CRB)</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Traffic accidents</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Suicide</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>


Note: a lack of a sign indicates that no association was observed (significance set at the 95% level). A negative quadratic
term implies that the observed association takes on an ‘n’ shape, i.e. the highest death rates were found for middle
incomes; the opposite applies with a positive quadratic. A negative/positive linear association implies that those with the
highest wages observed lower/higher mortality than those with the lowest wages.

2.3.3.4 Employment status

Mortality rates have been found to be lower for wage earners than for those who are unemployed for
many of the major causes of death, particularly accidents and violence (Martikainen, 1990). In this
Finnish study, results changed little after controlling for occupation, education, marital status, use of
reimbursable medicines, and the number of sick allowance days. One possible explanation that was
given was that unemployment itself causes excess mortality (causal explanation), either directly or
through intervening social factors (e.g. changes in lifestyle, social cohesion, income and other material factors, and increased stress). Alternatively, the apparent interdependence between unemployment and mortality may be spurious if some background variables cause both unemployment and mortality (selection hypothesis). For example, poor health can in itself cause unemployment, either directly or through poor educational attainment. However, as only marginal groups were over-represented among the unemployed, the effect of selection could only be responsible for a small part of the increased mortality among the unemployed. An important finding was that the effects of unemployment on mortality increased with duration, suggesting a possible change in behaviour and psychosocial stress in people who become unemployed. The negative health effects of unemployment would thus be minimised if a quick return was made to the workforce.

2.3.4 Mechanisms in adult life: proximate risk indicators

Differences in mortality patterns between social groups and regions have in the past been partly explained by differences in behaviour, material factors, psychosocial stress, the use of health care, and biological factors. How this is thought to occur is explained in more detail below.

2.3.4.1 Behaviour factors / lifestyle

*Cigarette smoking,* in particular, has been the focus of much media attention and medical concern. In countries where cigarette smoking has been common for many decades, tobacco now accounts for a substantial proportion of premature deaths. In fact, throughout middle age the death rates of current cigarette smokers (most of whom have smoked regularly throughout adult life) have been reported as being more than twice the rate of those who had never smoked regularly. In the US, tobacco was responsible for about 20% of all US deaths in 1985, while about half of all regular cigarette smokers would eventually, in either middle or old age, be killed by their habit (Peto et al., 1992).

With different smoking patterns existing between men and women, among different cohorts, lower and higher social classes and between countries, part of the variations in mortality between such groups have been explained by smoking. Or (2000) noted that variations across countries in the evolution of tobacco consumption gave rise to different results in terms of their impacts on changes in health outcomes over time. For example, while between 1970 and 1992 the number of potential years of life lost (PYLL) in Greece declined by about 48%, increased tobacco consumption was responsible for a 14% increase in PYLL. In comparison, the PYLL declined by 56% in the UK, of which 9% was due to a reduction in tobacco consumption.

In terms of social class and gender patterns in smoking, international differences and changes over time are marked. For example, Portugal is still one of the few western countries where people with a
high social status smoke more than those with a lower social status. Conversely, northern European countries are considered furthest advanced in the smoking epidemic, as smoking prevalence rates have been declining for a long time, most evidently among the higher social classes. This has resulted in relatively steep social gradients in terms of smoking in such countries (Cavelaars, 1998). Many types of specific diseases have been linked with smoking. In terms of total number of deaths, the largest contribution comes from IHD followed by lung cancer. Other causes of death where smoking plays a contributing factor include chronic obstructive pulmonary disease (COPD); peptic ulcers; duodenal ulcers; cancer of the oral cavity, pharynx, oesophagus, pancreas, cervix, bladder and kidney; accidents caused by fire and flames; and neonatal deaths (Nizard and Nuñoz-Perez, 1994; Peto et al., 1992; Balarajan and McDowall, 1988).

Another frequently studied health-related behaviour is the consumption of alcohol. Depending on the nature of consumption, alcohol has been found to have both positive and negative effects on the risk of premature adult mortality. When taken regularly and in moderate amounts, alcohol consumption reduces the risk of IHD, and it appears that it is the alcohol itself, rather than the type of drink, that reduces the risk of IHD, and particularly when it is consumed during meals. The reason is that it is absorbed slowly, prolonging the protective effect of alcohol against platelet aggregation, a risk factor for IHD (Renaud and de Lorgeril, 1992; Rimm, et al., 1996; Kiechl et al., 1998). On the other hand, in a prospective population-based study in Finland, beer bingeing (regular sessions of heavy beer drinking) has been linked with increased risk of death, independent of the total average consumption of alcohol, and after controlling for known behavioural, psychosocial, or biological risk factors (Kauhanen et al., 1997). In addition to the obvious external causes of death being over-represented among binge drinkers, this increase also pertained to circulatory system diseases (in particular IHD). A strong positive association has also been observed with mortality from CRB among men who drink 35 or more units per week (Hart et al., 1999). Other causes of death that have been linked with alcohol consumption include LDC; oesophageal cancer; cancers of the oral cavity, pharynx and larynx; cancer of the liver and alcohol dependence syndrome (Nizard and Nuñoz-Perez, 1994; Balarajan and McDowall, 1988). Alcohol-related causes of death, to some extent, explain the socioeconomic differentials in mortality. In a study by Mäkelä et al. (1997), alcohol related mortality\(^7\) accounted for 14% of the adult mortality differentials among men, and 4% among women, between manual workers and upper non-manual employees. The role of alcohol in the socioeconomic differences was modest in deaths from diseases but substantial in accidental and violent deaths, where it accounted for 50% of the difference in the mortality rates between upper non-manual employees and manual workers in accidental and violent mortality.

\(^7\) Cause of death includes a reference to alcohol in the death certificate.
Epidemiologists have frequently studied associations between dietary factors and mortality, especially due to IHD. According to one study (Ulbricht and Southgate, 1991), seven dietary factors either promote or protect against the development of IHD subsequent to the onset of atherosclerosis (obstruction of blood flow). The factors that promote the development of IHD include foods that contain saturated fatty acids (SFA) since these raise serum cholesterol levels in the blood, promoting atherosclerosis. Unsaturated fatty acids (UFA), dietary fibre and antioxidants all reduce the risk of IHD (ibid.; Nyyssönen et al., 1997). However, not all countries show the same relationship between a high intake of SFA and high levels of IHD mortality. In France, for example, the mortality rate from IHD is much lower than in other industrialised countries such as the US and UK, even though the intakes of SFA and concentrations of serum cholesterol and other risk factors of IHD, including blood pressure, body-mass index, and cigarette smoking are similar (at least for men). This finding is known as the French paradox (Renaud and De Lorgeril, 1992). Even though no adequate explanation has been produced as to why the situation in France is unique, when investigating the French diet more closely, it seems that it is varied, consisting of small amounts of butter and a high consumption of bread, vegetables, fruit, cheese, vegetable fat and wine. As already discussed, the last item has been particularly associated with a reduced risk of IHD. In Spain, regional differences and increases in fruit consumption have been linked to regional differences and decreases in CRB mortality (Artalejo et al., 1998). Imbalanced diets or glutinous diets have been linked to several types of cancer, including colorectal cancer (Frankel et al., 1998), and higher levels of mortality from oesophageal cancers (Balarajan and McDowall, 1988). On the other hand, dietary restriction reduces the likelihood of carcinogenesis due to the delay in cell proliferation during development, as well as cell renewal in most organs at later stages (Frankel et al., 1998).

2.3.4.2 Material factors
Socioeconomic gradients in health are associated with both social position and different material circumstances. The most important material conditions are housing conditions and working conditions. Poor housing conditions, such as cold and damp dwellings, a lack of central heating and running water, and overcrowding can lead to increased risks of ill-health. This is because they are related to the easy transmission of, predominantly airborne, infectious diseases. Sanitary conditions, another indicator of housing conditions, are related more to diseases of the digestive system (Wolleswinkel-van den Bosch, 1998). As these material conditions are difficult to measure on a large scale, alternative variables can be used that are simpler to collect and process such as housing tenure. Housing tenure is also an indicator of SES, and has proven to be an effective discriminator of mortality (Filakti and Fox, 1995). They produced similar results using another indicator of SES that has been used in the past, namely household access to a car.
Working conditions may also have a direct influence on health. In particular, manual workers such as construction workers are exposed to occupational hazards that may result in fatalities, while non-
manual workers are often exposed to potential non-fatal chronic injuries such as repetitive strain injury (RSI).

2.3.4.3 Psychosocial stress
The psychosocial wellbeing of a person is perhaps most strongly influenced by life events: changes in the lives of individuals, such as retirement, loss of income and the death of a close family member. Age plays a role in the effect of life events on the psychosocial wellbeing, with the elderly being less influenced than younger individuals, even when events such as the death of a spouse occur at unexpected times. The reason why older people are more tolerant is that they expect such life events to happen and are therefore psychologically prepared for their occurrence (Joung, 1996). Women are also better able to cope with such events (Blazer, 1980), while people with higher education, better physical functioning and greater social support also have a higher rate of adjustment (Nieboer, 1997). The death of a spouse is seen as being the life event with the highest risk factor for the development of physical and psychiatric illness, followed by divorce (Blazer, 1980).

In terms of SEHD, it is perhaps not surprising that psychosocial resources, including optimism, coping style, a sense of mastery or control and social support, may partially mediate the relationship between SES and health, and to varying degrees these resources appear to be differentially distributed by social class (Taylor and Seeman, 1999). While interrelated pathways between psychosocial factors and disease include health-related traditional risk factors such as smoking, diet, alcohol consumption, or physical activity, psychosocial factors may also cause direct acute or chronic pathological changes (Hemmingway and Marmot, 1999). The Whitehall II studies, for instance, showed that traditional risk factors could only explain 25% of the health differences between high and low class civil servants, as differences in control over the work situation, work satisfaction and social support were also important (Marmot et al., 1991). Similarly, it is also thought that the poor socioeconomic situation in Eastern Europe has an effect on adult mortality through psychosocial factors rather than because of absolute material deprivation (Bobak and Marmot, 1996; Kristenson et al., 1998). In the latter of these studies, standard cardiovascular risk factors failed to explain the diverging IHD mortality trends, and found an unfavourable cluster of psychosocial risk factors for IHD in 50-year-old Lithuanian men compared with 50-year-old Swedish men.

2.3.4.4 Health care factors (individual level)
Although health care factors are discussed later in this chapter from a contextual perspective, at the individual level there are also health care factors that can contribute to mortality differences between populations. There is plenty of evidence of social inequalities in access to health services. For instance, Tod et al. (2001) studied the British South Yorkshire coalfields and found that communities with the greatest needs had lower referral rates and uptake of services than the more
affluent communities. They identified five categories of factors that influenced the use of health care there:

i. structural factors (e.g. access to transport);

ii. personal factors (e.g. the fear of illness and denial of symptoms) and social and cultural factors (e.g. social mores and expectation, such as an unwillingness to depend on doctors or family);

iii. past experiences and expectations (e.g. negative attitudes of health professionals and previous problems of accessing health care);

iv. diagnostic confusion (e.g. delay of medical help as symptoms were attributed to industrially related lung disease instead of heart problems);

v. knowledge and awareness.

Although this regional example cannot be generalised for all of Europe, there are other examples of differences in the utilisation and quality of health care, such as the distinction between rural-urban (Pereboom et al., 1989) or by ethnicity (Cooper et al., 1998), and it is clear that the composition of the population according to socioeconomic status (and its cultural derivatives) and the cultural and medical knowledge of general practitioners and other medical personnel all affect the utilisation of health care by individuals, which in turn affects the average health of a population.

2.3.4.5 Biological factors
As noted earlier in this chapter, adult risk factors that independently affect adult disease may have been biologically programmed during gestation or early infancy (Barker, 1995). Biological risk factors, however, are also acquired and accumulated during childhood, adolescence and early adulthood, and may exacerbate the risk of adult disease. For instance, a positive association between energy intake during early life and later mortality from non-smoking related cancers was found in the British Boyd Orr cohort study after adjusting for household expenditure, social class, number of children and deprivation. It was suggested that dietary restriction would retard cell proliferation during development as well as cell renewal in most organs at later stages, thus potentially reducing the chance of genetic changes arising from carcinogens or active oxygen species (Frankel et al., 1998). Other commonly used biological risk factors associated with SEHD, especially cardiovascular diseases, include serum cholesterol (Ebrahim et al., 1998; Hart et al., 2000), blood pressure (McCarron et al., 2000; Everson et al., 1997) and body mass index (in particular combined with low birth weight) (Eriksson et al., 1999; Frankel et al., 1996).

It is not difficult to accept that biological pathways are closely related with social pathways. For instance, the results of one Finnish study provided evidence that the greatest accumulation of adverse biological and other risk factors was among those who were born and remained socioeconomically disadvantaged over their entire life course. Moreover, evidence also supported the view that those who were upwardly mobile accumulated fewer risks, but that this socioeconomic
climb did not completely obliterate the effects of poor socioeconomic origins in childhood on a higher level of adult risk factors (Salonen, 1988; Lynch, 2001).

2.3.5 **Sociodemographic factors**

The effects of early life stages, adult life SES and proximate risk indicators on health are to some extent associated with sociodemographic characteristics. The best example is sex, because it is predetermined at birth and has its own independent influences during various stages of the life course. Similarly, there are health risks associated with specific ages, such as cot deaths for infants and dementia for the aged. Age is also synonymous with cohort as, for instance, a thirty-two year old today has had different life course experiences than a thirty-two year old fifty years ago. The third type of sociodemographic factor that is expanded upon below is marital status since, on average, married people live longer than unmarried ones.

2.3.5.1 **Age, cohort, period**

Death is nothing more than the result of biological processes that are expressed in the biological, functional and mental status of a person (Martelin, 1994); and age is inherent in these biological processes. In the absence of disease, genuine biological ageing affects everyone and is manifested to the outside world in the form of age-related physiological changes of the organs and/or the entire human organism. Examples of this include the deterioration of vision, hearing, lung function, memory, bone structure; and an increase in blood pressure, cholesterol and blood sugar levels. All this implies disturbances in the proper functioning of various organismal systems, but there is no common programme in the functional and constitutional decline of the human body that occurs with increasing age: the decrements appear and progress independently in each individual.

Death rates are relatively high in the first years of life, after which they decrease through later childhood. Among males the rate rises abruptly in late adolescence and continues to rise thereafter, whilst for females there is a more gradual increase. Biological or intrinsic ageing is only one factor that influences the age at death since cohort effects also occur with people of different ages and those who occupy different roles being differentially exposed to, and influenced by, particular social and economic changes (Ben-Shlomo and Kuh, 2002). One of the best known examples is the effect of the 1944-45 Dutch ‘hunger winter’ on girls in their early years of puberty. The scarcity of food, particularly containing fat, led to extremely low breast cancer rates among this cohort many years later (Dirx et al., 1999). Period effects can have an impact on most or all of the population. For instance, the post-1984 period in the Russian Federation is marked by large swings in life expectancy, with dramatic improvements during Gorbachev’s anti-alcohol campaign (1985-1987), but a sharp deterioration from 1991 until the mid-1990s coinciding with Russia’s erratic attempts at economic reform and a period of great macroeconomic instability (Brainerd and Varavikova, 2001; Shkolnikov and Nemtsov, 1997). One point to bear in mind is that most cohort and period effects
that contribute to mortality differences are not instantaneous, and are more likely to be observed years of decades later. While, in individual-level studies, the history of exposure can be taken into account directly, in population-level studies only rough estimates of latency periods can be made (see also Chapter 4).

2.3.5.2 Sex and gender
In all European countries, life expectancy at all ages is higher for women than for men. During the life course, sex differences in life expectancy are divided between biological (sex) and behavioural (gender) factors (Verbrugge, 1983; Waldron, 1986). During gestation, more male than female foetuses are spontaneously aborted or stillborn and in all Western societies higher male mortality continues to be marked during the first six months of life. This is possibly due to sex differences in chromosomal structures and a slower maturing of boys’ lungs from the effects of testosterone (Waldron, 1986). During adult life, there are several sex-specific causes pertaining to the male and female reproductive organs that are responsible for a considerable number of deaths each year, the principal ones being breast and ovarian cancer among women and prostate cancer among men. Mortality from sex-specific causes of death is more common among women (WHO, 2001).

On the other hand, most of the excess male mortality is due to IHD, for which smoking and poor dietary habits are two of the most important behavioural risk factors. As a result, during young and middle age, high blood pressure, high levels of cholesterol and overweight occur more frequently among men. However, above the age of 55, these risk factors actually occur more frequently among women, and the increase in the occurrence of IHD among women after middle age suggests that the lack of oestrogen production after the menopause play a role in this. A reduction of 30 to 50% in IHD has been observed among women treated with hormonal supplements for menopausal and other complaints (Meeter and Witteman 1997; see also Blatt Kalben, 2000). Male-female differences in IHD mortality have also been attributed to the extent of previously beneficial medical care, perception of symptoms, and the choice of and obedience to treatment (Deeg and Pot 1997). Women also tend to drink less alcohol, coinciding with fewer deaths from LDC and external causes. Gender roles of women have been changing during recent decades, from traditional role patterns to lifestyles that are more liberal and with an increase in their labour participation (Annandale and Hunt, 2000). The latter does not seem to increase the risk of heart disease among women, whereas social isolation does (Meeter and Witteman, 1997). However, the most obvious change in behaviour has been the increase in female smoking at a time when rates among men have been declining (OECD, 2001). As a consequence, most European countries are now showing a decline in sex-differences in mortality from certain smoking-related diseases, particularly from lung cancer (WHO, 2001). How ongoing social changes in men’s and women’s lives in the domains of education and work, household and the family, leisure and consumption are associated with changing differentials in male-female mortality is explored further in Chapter 7.
Lastly, it should be noted that when older men develop a disease it is more often fatal than is the case among older women, and that women tend to have a greater chance of acquiring less fatal health problems such as depression, joint complaints and other limitations of the body.

2.3.5.3 Marital status
Research has consistently shown that, after controlling for initial health status, married people live longer and have fewer health problems than those who are unmarried, divorced or widowed. Relative differences are largest for men and for both sexes at younger ages. Causes of death that contribute to the mortality differences by marital status include external causes, LDC, diabetes, lung cancer, COPD and non-IHD, although there are some differences internationally. It is assumed that the effect on health of marital status is mediated by psychosocial factors, material circumstances and health behaviour (Hu and Goldman, 1990; Joung, 1996).

Marital status groups differ in both their exposure and vulnerability to psychosocial stress. For example, the loss of a beloved person is an important source of stress in itself, and divorce (or marital separation) can evoke stress through feelings of failure, lowered self-esteem, and a sense of incompetence. Among both the widowed and the divorced there are many possible concurrent changes, such as lowered income, changes in parental responsibilities, forced move to other housing or the loss of familiar activities and habit systems, all of which contribute to the total stress experienced. There are also differences in psychological stress between marital status groups that are caused by mechanisms other than the stressful character of events such as bereavement and divorce. Negative societal attitudes towards certain marital status can also be a source of stress. Although societal attitudes towards alternatives to marriage have become more liberal in recent years, prejudices and stereotyped images of never-married and divorced people still exist (Joung, 1996).

The different material circumstances experienced by the various marital status groups also lead to potential differences in health, as people who share a household are able to benefit from economies of scale, such as in the purchase and use of housing and other goods and services. A spouse is also likely to provide social support, which might ameliorate the negative health effects of stress. When a marriage dissolves, the woman tends to be financially worse off than the man, especially when the husband is the sole wage-earner. However, divorced men are also materially disadvantaged compared to married men as a result of material arrangements made with their former partner, such as alimony and the division of the goods (ibid.).

Differences between marital status groups in health behaviour are also seen as an intermediary of the association between marital status and health. Evidence has shown that fewer married than unmarried people engage in health-damaging behaviour such as smoking and excessive drinking. Married people are also more likely to pursue preventative medical services and follow any required prolonged and methodological course of treatment (Joung et al., 1996).
2.3.6 Contextual factors

In the theoretical framework, contextual factors are considered as extrinsic factors that may indirectly affect health. They include macroeconomic conditions, cultural factors, the health care system and ecological factors. Although the individual has little influence on specific contextual factors such as the economic production of a country or the amount of air pollution, in some areas such as religion one can make certain choices. This does, however, depend on a combination of a social acceptance of individualism and personal characteristics (e.g. the ability to withstand the social pressure to conform).

2.3.6.1 Macroeconomic condition

The term macroeconomic condition refers to the economic indicators of a geographic area as a whole, rather than of individuals. As Figure 2.1 shows, macroeconomic factors can have a direct or an indirect effect on proximate risk indicators that are directly associated with health. Therefore, a changing state of the economy will at some later point affect the health of a population. For instance, an increase in government expenditure on public health care will lead to a healthier population (Van Poppel, 1991). The most commonly used macroeconomic indicator is per capita gross domestic product (GDPc) which is a measure of the average economic output per person of a country and is sometimes referred to as the average level of income or wealth in a country. When reference is made to a region, per capita Gross Regional Product is used. Not surprisingly, from a world perspective, a higher GDPc results in longer and healthier lives, in which a growth in income has more impact on poor populations because additional resources allow more of the basic necessities to be bought, particularly food and shelter. The World Bank estimated that the effect of doubling income from $1000 to $2000 in purchasing power parity (PPP – used to adjust for different price levels) corresponds with a life expectancy increase of 11 years. However, a doubling of PPP from $4000 to $8000 only results in a gain of four years (Hilderink, 2000). Moreover, after a certain point of economic development, there are indications that further improvements in material standards will have little influence on health (Wilkinson, 1997), or that confounding factors obscure the mortality lowering effects of high living standards (Mackenbach and Looman, 1994). Rather than absolute living standards, Wilkinson (1995), based on several pieces of evidence, proposed that mortality in developed countries is more affected by relative living standards. He first showed that mortality was related more closely to relative income within countries than to differences in absolute income between countries. Secondly, the countries with the lowest mortality rates were those with the smallest income differentials and, thirdly, most of the long-term rise in life expectancy was unrelated to long term economic growth rates. One explanation that was given for the association between greater income equality and better health was that this tends to improve social cohesion and reduce the social divisions while, conversely, poverty is associated with social exclusion. Moreover, social processes are linked with biological processes such as psychosocial stress. In addition to the
fact that relative deprivation causes psychosocial effects on health; increased death rates from accidents, violence and alcohol-related causes also seem to be particularly closely related to wider income inequality.

Although this income inequality hypothesis has influenced a wide range of authoritative reports emanating across the world, it is refuted by several researchers. After a careful review of the evidence provided by Wilkinson in a number of papers, Judge (1995) found little support of the income inequality hypothesis. Besides several reservations about the actual data that were used, it was noted that no reference was made to the fact that the coefficients associated with per capita income were not significant, whereas most of the coefficients associated with measures of family income were. This was an important finding as the data on family income were used without adjusting for differences in typical family size between countries. In other words, since support for the income inequality hypothesis seems to depend on the choice of income distribution indicator, then a clear rationale for selecting one measure over another is essential. Rather than seeking to establish the association between income inequality and mortality, it may be more important to pursue the causal pathways, such as psychosocial characteristics, which are strongly patterned by income (Kaplan, 1995). In a study by Kaplan et al. (1996), US States with the greatest inequality in incomes showed higher rates of violence, more disability, more people without health insurance, less investment in education and literacy, and poorer educational outcomes. However, according to Gravelle (1998), care should be taken when interpreting results of this type of aggregate study as different associations may be found at the level of the individual (see also Section 2.4). Part of Chapter 4 is devoted to the debate between absolute and relative income in relation to how these affect differences and changes in mortality.

Besides the absolute and relative living standards in a region that may affect mortality, there are other macroeconomic factors that could confound this association. For example, population density, agricultural and industrial employment (both as a percentage of total employment) are all associated with the level of socioeconomic development of a region (Mackenbach and Looman, 1994). As the service sector adds, on average, more value to an economy than the agricultural or industrial sectors, and each type of occupation has an independent effect on mortality through more specific behavioural and environmental factors (Fox et al., 1984), changes or regional differences in the composition of the employment structure is expected to be indirectly associated with changes or differences in mortality. Although part of this indirect association is likely to originate from a cohort effect (e.g. older, less educated, agricultural and industrial manual workers leaving the workforce and being replaced by younger, educated non-manual workers), intragenerational mobility has been common during the recent economic transition in Eastern Europe. If the results of the French study by Cambois (2002) were to be applied to the Eastern European context, it would be expected that age-standardised mortality levels would decline as people moved upward in occupational class, in
particular because there was little evidence that health deteriorated for those who moved down the occupational ladder.

One of the consequences of the political and economic reforms in Eastern Europe has been the privatisation of state-owned enterprises and the re-orientation of production and services towards demand. This has not only caused changes in the employment structure, but also widespread unemployment as well as new opportunities. The effect of the economic transition should not be underestimated. For example, in one regional study that aimed to identify aspects of socioeconomic change that were associated with the steep decline in life expectancy in Russia between 1990 and 1994, it was found that ‘labour turnover’ (sum of job gains and losses as a percentage of average employment) yielded a higher positive association with mortality than a fall in income or an increase in crime rates (Walberg et al., 1998).

2.3.6.2 Cultural factors
Culture may be defined as “a complex whole of learned behaviour and its products in a social context” (Hammel, 1990). It includes language, customs and conventions by which human behaviour can be distinguished from that of other primates, and focuses on the collective nature of social behaviour. Cultural factors have attracted much less attention in international reports on mortality differences than economic factors. They are often ignored because they are supposed to confound the relationship between economic factors and mortality. However, culture plays an important role in shaping the behaviour of social actors, while the behaviour of actors, in return, redefines culture. Culture identifies a social grouping whose values and norms may explain why communities or persons living under apparently identical conditions, but differing in language or tradition, often behave very differently demographically. Culture may also explain why the population of a geographic region or linguistic area continues to behave demographically in much the same way over time even though economic conditions change, or why demographic differentials between populations persist even though demographic changes are similar over time (Hammel, 1990). Nevertheless, although the concept of culture is not something that can be measured directly, several cultural indicators can be used to identify different populations such as language, ethnicity, religion and post-materialism.

Language is an important part of culture, as one identifies a country or region by the language that is spoken there. Although language does not directly influence mortality, clear contrasts exist in cause-specific mortality patterns between adjacent countries speaking different languages. As country-specific macroeconomic and healthcare factors are likely contributors to this, the use of this culture proxy is more convincing when language regions within the same country exhibit different mortality patterns, as is the case between the Flemish and French speaking regions of Belgium. Here, while differences narrowed slightly after controlling for potential confounding factors at the individual-
level such as age and various socioeconomic characteristics, the essential pattern remained (Gadeyne and Deboosere, 2002). Moreover, since the Second World War, the geographical pattern of mortality in Belgium has changed little (Mackenbach et al., 1991).

As a result of international migration, arrays of various ethnic populations have established themselves in many European countries. These ethnic minorities are often distinguishable through their different cultural characteristics, as well as their lower SES. It is not surprising that mortality ratios for migrants are therefore often different to those of the general population. When trying to ascertain underlying reasons for these differences, there seems to be no uniform pattern as various factors may be of importance depending on the context. For example, health status is less likely to be influential on the decision to migrate to a country in close proximity than to a distant country where there are restrictions on immigration, as results from Wild and Mckeigue (1997) showed for Scottish and Irish migrants to England & Wales. In this study, social class did not account for the higher level of mortality among Irish immigrants, but rather behavioural factors such as smoking were blamed since, within each social class, their mortality was higher than the national average. In general, ethnic minorities from developing countries have higher rates of mortality than the host population because they bring particular disease risks with them from their country of origin, such as adverse environmental conditions including deprivation, poor working conditions, lack of social support and an unhealthy life style. Given the cumulative nature of such risks, they could have long term effects on immigrants and may persist or develop in the host country (ibid.). Some of these risks also influence infant mortality. For instance, in the Netherlands, the IMR among non-western ethnic minorities was 20% higher than for the western population, even after controlling for economic factors (Van Duin, 2002). However, mortality rates are not always worse among non-western immigrants. While a higher than average proportion of Caribbean immigrants to England & Wales were manual workers, they experienced lower than average all-cause mortality. This was mainly attributable to there being very few IHD deaths, a trend that was preserved from 1971 to 1991. This was in contrast with the expectation that the mortality of an immigrant population would change within one or two generations to levels that approximated more closely the disease risks of the host country (Wild and Mckeigue, 1997; Van Poppel, 1990). It was not clear why the Caribbean and West African immigrants to England & Wales both had lower levels of IHD and excess mortality from CRB and hypertensive diseases, since the two migrant groups had not shared a common environment for the past 300 years. However, it does suggest a possible genetic explanation for the susceptibility to hypertension in people of west-African descent (Wild and Mckeigue, 1997).

Religion is a cultural indicator that is known to influence the way people live as it addresses itself to all of man’s major uncertainties. Most religions include rules and activities meant to ward off sickness and death, such as abstaining from smoking, drinking and certain types of food. In
particular, one is reminded of Weber’s classical analysis of the Protestant way of life that is supposed to be characterised by rationality and self-discipline (Weber, 1969; Jarvis and Northcott, 1987; Van Poppel, 1991; Mackenbach et al., 1991). More recent studies on the Seventh-Day Adventists and the Mormons in the United States have revealed that their devout lifestyles have resulted in lower rates of mortality from cancer and circulatory system disease than the population as a whole (Manton et al. 1991; Enstrom, 1989). Other positive effects on health of religion include the likelihood of a reduction in the impact of stressful life events and the creation of supportive social networks (Jarvis and Northcott, 1987).

In the Netherlands, religion has played a role in explaining regional differences in mortality. According to a study by Mackenbach et al. (1991), approximately 90% of the excess mortality in the southern part of the country in the 1980s was due to cardiovascular diseases, mainly IHD and CRB. Results of their regression analysis showed that this excess mortality was primarily related to the high percentage of Roman Catholics in this part of the country, and additionally to a slightly lower income. Other causes of death that were linked to Roman Catholicism include lung cancer, arterial disease and chronic non-specific lung disease. Survey data also showed that these associations were partly due to higher prevalence of smoking among Roman Catholics. The association between geographical mortality patterns and the percentage of Roman Catholics in the population has attenuated since the 1950s, and regional mortality rates are converging, which suggests a reduction in the lifestyle differences between the population groups. As Jarvis and Northcott (1987) pointed out, in order to measure the effects of religion on various aspects of health and mortality, it is important to distinguish among religious affiliations, levels of participation or attendance, and strength of faith (beliefs and adherence, etc.).

Urbanisation has been used in previous epidemiological studies as a cultural or social factor. Research has shown, for example, that, at least in the Netherlands, a high level of urbanisation is related to factors such as atheism, a left-wing political orientation, a high proportion of single people, one-parent families and divorcees, and fewer elderly people (Mackenbach and Kunst (1995): factors which may have a negative influence on health. In terms of mortality, they discovered that the geographic variation in the onset of a decline of deaths from IHD within the Netherlands started in the 1970s in the highly urbanised areas that were regarded as the cultural centres of the country and gradually diffused to more ‘peripheral’ regions, although socioeconomic factors may have also played an important role in this diffusion process. Despite this, levels of life expectancy in the Netherlands have been more favourable in rural areas than in the most urbanised areas of the country, with mortality in urban areas being particularly high with respect to cancer (except of the stomach), arterial diseases and respiratory diseases, and diseases in those parts of the body that are aetiologically related to behavioural factors such as smoking, alcohol consumption and sexual promiscuity (Mackenbach and Kunst, 1995; Schouten et al., 1996; Mackenbach and Verkleij, 1997). Less densely populated areas also have a reduced probability of contamination, as
there are fewer personal contacts, more ventilation, better quality housing, less pollution and fewer industries associated with major occupational hazards. Conversely, the availability and accessibility of various forms of public health services and medical technology is positively associated with urbanisation (Van Poppel, 1991).

In western society, much has been said and written about the negative effects of certain types of behaviour such as smoking. According to Inglehart (1997), as people put more emphasis on self-actualisation they could be expected to reflect more on how one’s behaviour can affect one’s physiological and psychological well-being. This would imply that regions that are influenced by the ‘culture of post-materialism’ would have better health practices than regions where materialistic values still dominate. Tabeau et al. (1999) tested Inglehart’s hypothesis and discovered that in those sub-national regions, in seven Western European countries, where relatively many people held post-materialistic views the mortality rates were indeed more favourable.

2.3.6.3 Health care factors (contextual level)

In addition to improvements in living standards, education and other previously mentioned factors, health care factors have also been important for the improvements in life expectancy in the twentieth century (Manton et al., 1991; Wolleswinkel-van den Bosch, 1998), although the contribution of health care to the health of a population has been subject to some debate. The best-known argument has come from McKeown who showed that three-quarters of the decline in mortality in England & Wales between 1841 and 1971 was due to a reduction in deaths from infectious diseases, and that three-quarters of this reduction had preceded the widespread introduction of immunisation and antibiotics. He argued that the main influences on health had been nutrition, the environment and behaviour, a view that is shared by many public health professionals. Others have argued that health care has the potential to damage the health of patients from the side effects of prescribed drugs, hospital-acquired infections and poorly performed surgery (see McKee, 1999 for references). However, today there is a growing body of evidence that health care has a demonstrable positive effect on health at a population level, with many formerly fatal conditions now susceptible to treatment. For instance, the declining rates of cancer mortality in the US in recent years has been partially attributed to improvements in medical care as cancers are diagnosed earlier in their course leading to an improved survival rate (Cole and Rodu, 1996). One method that has been used to measure the effectiveness of health care is to calculate the mortality rate from causes of death that are usually amenable to medical intervention, such as tuberculosis, cancer and respiratory diseases among children, appendicitis, Hodgkin’s disease, cervical cancer, maternal and perinatal mortality, and CRB (Mackenbach et al., 1988; McKee, 1999). Analysing these so-called avoidable causes of death would enable regions to be identified as being suspected of having less effective health care, as well as monitor changes over time. For example, in one Dutch study, reductions in avoidable mortality (due to 35 causes of death) between 1950 and 1984 added a total of 2.90 and 3.95 years
respectively to the life expectancy of men and women in the Netherlands (Mackenbach et al., 1988). It should be noted that, in a review of published work reporting on mortality from conditions amenable to medical intervention, Mackenbach et al. (1990) concluded that declines in mortality might not be entirely due to better and greater supply of health care as socioeconomic conditions have also improved, while a number of amenable conditions saw ‘spontaneous’ declines in incidence. Geographical and social differences in access to health care may also confound the association between health care and mortality (McKee, 1999), and this is difficult to measure at the population level. For geographical differences, the share of government financing in total health expenditure can be considered as an indirect approximation, but since this share has remained fairly constant over the past two decades, its contribution to the decline in premature mortality in industrialised countries has been negligible (Or, 2000).

2.3.6.4 Ecological factors
Regional mortality differences within European countries have also been partly attributed to differences in temperature (Donaldson et al., 1998a, 1998b), precipitation (Valkonen and Notkola, 1977; Balarajan and McDowall, 1988), sunshine (Balarajan and McDowall, 1988), soil composition (Valkonen and Notkola, 1977) and air pollution (Balarajan and McDowall, 1988). Although air quality has improved in many countries over recent decades, the effects of air pollution on health is still an important issue. One study concluded that up to 3% of all-cause, all-age mortality in the Czech Republic was caused by air pollution, which translated into 9% of the mortality gap between this country and Western Europe (Bobak and Feachem, 1995). Within countries, differences are also found, as Balarajan and McDowall (1988) suggested that in the industrial north and northwest of Great Britain air pollution contributed to the higher mortality rates from chronic bronchitis and lung cancer compared with the rest of the country. Short-term effects of pollution are also known. According to another European study, a rise in the concentrations of particulate matter and sulphur dioxide over a 24-hour period, was associated with a rise in mortality (Katsouyanni et al., 1997). The consistency in the results across 12 cities, although the effect was larger in Western Europe than in Eastern Europe8, with differing climate, topography, and sources of pollution supported a causal association. The effects of both pollutants were stronger during the summer, and were mutually independent. Although the biological mechanisms through which exposure to particulate matter increases mortality is not well understood, it is thought that fine particles penetrate into the respiratory region, and that the effect of sulphur dioxide seems to be limited to patients with asthma or bronchitis (Katsouyanni et al., 1997). Emissions from motor vehicles may also trigger circulatory disorders (Poloniecki et al., 1997).

8 Most of the mortality differences between Eastern and Western Europe are due to circulatory system diseases and external causes (Bobak and Marmot, 1996). For this reason, one possible criticism of this study is that all-cause mortality was analysed, rather than only causes of death such as lung cancer and respiratory system diseases where pollution is one of the most important risk factors.
2.4 Ecological fallacy

In the previous section, a large number of epidemiological studies were used to identify factors known to have caused mortality differentials between different groups of people, with a particular focus on socioeconomic factors. Many of these studies were conducted using individual level data in which mortality was linked to a range of personal characteristics. In the studies undertaken for this thesis, the units of analysis are countries and sub-national administrative regions, and entail aggregated-level data such as the national average income or alcohol consumption. Care should be taken when inferring relationships at the individual level from associations between variables at the population level, particularly when the relationship is non-linear at the individual level since in that case the result at the population level may be a statistical artefact. This is known as the problem of ecological fallacy (Valkonen, 1993b; Gravelle, 1998; Elford and Ben-Shlomo, 1997). To give an example, one study found negative correlations between population health and income inequality when aggregating the individual data but, once the analysis was conducted at the individual level using individual incomes, the area’s income inequality had no effect (Fiscella and Franks, 1997). It is possible that the importance of factors in determining disease at the individual level may be different from that at the population level because, within populations, the variation in an exposure may be far less than between populations, and/or at an individual level the relationship between exposure and outcome may be weaker due to the multifactorial nature of disease causation (Elford and Ben-Shlomo, 1997). Nevertheless:

"the failure to replicate an individual-based study at an ecological level does not negate the causal importance of that variable. Assuming that this is not merely a measurement error problem, it either suggests that the other variables are more important in explaining ecological variations or that there may be more important interactions between different risk factors that are more marked at the ecological rather than the individual level" (Elford and Ben-Shlomo, 1997).

Fortunately, the potential problem of ecological fallacy is not very relevant in the analyses presented here because the main contribution of this thesis is the quantification of the sensitivity of socioeconomic and other factors on mortality for future mortality scenarios. Moreover, even though the analyses were performed on aggregated data, most of the variables were selected on the basis of their well-documented associations with mortality at the individual level. Therefore, while all results are expected to be generally in line with individual-level findings, the failure to replicate an individual-based study at the ecological level suggests that the factor of interest may not be important at the population level, and this is an important result for scenario making – even if part of the result is due to statistical artefact. Due to this greater uncertainty with ecological analysis, the significant variables in the model outcomes obtained for this thesis suggest associations rather than causal relationships. A related conceptual issue that ought to be explained is that when one or more
variable in a model alleviate, augment or change the association between the investigated variable and mortality, it is considered to be a confounding factor; and this should not be confused with epidemiological confounding that refers to the failure of a crude (or partially adjusted) association to properly reflect the magnitude of an exposed effect because of differences in the distribution of extraneous risk factors among exposed and unexposed individuals (Greenland and Morgenstern, 1989).

2.5 Conclusions

This theoretical chapter has reviewed factors known to be associated with mortality. In the SEHD studies, there is consistent evidence that those with lower education, income or occupational status suffer worse health. However, this is not just a matter of education etc., but a combination of different life trajectories that were, initially, responsible for differences in educational attainment etc., and subsequently for differences in the proximate risk indicators that led to SEHD. For this reason, one comes to realise that the life course should be acknowledged as a conceptual framework. The problem is how to conceptualise this in ecological research.

The life course approach advances our understanding of both disease aetiology and the production of socioeconomic differentials in health and it was therefore used as a basis in selecting variables for the analyses presented in Chapters 4-7. From a life course perspective the range of factors that determine an individual’s risk of death at adult age is large: from one’s father’s occupation to the food that one eats, from the critical period during gestation to the lifelong exposure to air pollutants – the list is endless. For the purpose of this study most of these factors cannot be used, or at least not in the form that they have been applied in individual-level studies. It is therefore necessary to look for alternatives. Moreover, while it might be worthwhile investigating the effect of single events on mortality differences between countries or over time (for example, war, famine, economic crisis), it would be too difficult to integrate critical periods into mortality projections. New insights have also emerged suggesting that each phase of life appears to be capable of adding its own protection or disadvantage (although the weightings might vary for different diseases) as some damage can be reversed or even eradicated later in life by a change in behaviour (Blane, 1999). For instance, people who stop smoking, even well into middle age, avoid most of the subsequent risk of lung cancer, and stopping before middle age avoids more than 90% of the risks attributable to tobacco (Peto et al., 2000).

Incorporating the concept of risk accumulation, that is defined as “exposures or insults [that] gradually accumulate [throughout the life course] through episodes of illness, adverse environmental conditions and behaviours increasing the risk of chronic disease and mortality” (Kuh and Ben-Shlomo, 1997), into macro-level mortality analysis appears to be more convincing.
and plausible than that of critical periods. Although the citation by Kuh and Ben-Shlomo (1997) only pertained to individuals, it would seem that different levels of accumulation effects may also cause mortality differences between populations when we consider that the divorced, the less well educated, manual classes, the poor, and men consistently have higher mortality rates than the married, the highly educated, non-manual class, the rich, and women. Therefore, by deduction, a society with higher proportion divorcees, few educated people, a large industrial sector and poverty should experience worse health. As a result, compositional factors become important in the geographical patterning of health, over and above contextual factors such as GDP and government health expenditure. But while these compositional factors will have some independent effect on geographical mortality differences, or over time as the composition of a population also changes (for instance younger cohorts are better educated than older cohorts), socioeconomic factors will affect mortality particularly through proximate risk indicators such as smoking and alcohol consumption. However, this is where the interpretation of results from ecological research becomes problematic, because the composition of the population according to occupation, etc., cannot be split into, for instance, smokers and non-smokers, which would allow us to ascertain if smoking levels are declining because there are fewer industrial workers or if they occur at all social levels equally. We can therefore not infer any causal relationships when one variable influences the coefficient of another, although we may suggest that smoking ‘confounds’ the association between, in this case occupation, and mortality. Given the strong influence of both smoking and occupation on mortality and the fact that the two factors are invariably linked, it is nevertheless pertinent that both are included in mortality analysis.

While individual-level determinants of mortality have formed the basis for the selection of aggregate-level indicators, in order to ascertain their effect on international mortality differences and changes over time, ecological analysis can be justified in its own right by considering the effect of contextual factors, such as environmental pollution, that can only be measured at the population level. There are also factors that have a different impact at the population level than at the level of the individual, as is the case with income. GDP, i.e. national level income, benefits the population’s health through taxation paying for health services, health education, health research, better work and road safety, etc., areas in which an individual’s income has little influence (you might be able to buy a safer car but the roads remain the same).

The life course approach illustrates that “studies with data on socioeconomic circumstances at only one stage of life are inadequate for fully elucidating the contribution of socioeconomic factors to health and mortality risk” (Davey Smith et al., 1997). Given that mortality differences, whether socioeconomic, geographic, according to sex or otherwise, do not occur overnight but are the result of the accumulation of exposures to various risk factors that have been accrued during the period from before birth to well into adulthood, the life course approach is essential in improving current
national-level mortality projections: that is, the main goal of this research project. This becomes even more apparent if we consider that about 90% of natural deaths occur after the age of 45. How then can we take this aspect of the life course into account in ecological research? One way is by using lagged values of the relevant exogenous variables. This is particularly applicable to the Western Europe situation, as improvements in health have been slow over time, and current international differences in mortality are small. For this reason, many of the factors that are associated with these changes cannot be contemporaneous effects, but must be the result of many years of differential exposure to various risk factors. The situation in Eastern Europe is somewhat different, which in itself makes it a very special case. Given the large fluctuations in mortality rates there, it seems that many people must be very vulnerable to the changing economic climate. However, even in the case of Eastern Europe, the latency effects of some factors such as education still need to be considered.

Although this method of lagging variables to accommodate life course effects at the population level is rather crude, it is nevertheless superior to a model which only includes contemporaneous effects. It is also more logical to assume that a change in the magnitude of an exposure will, at some time in the future, bring about a change in health which will occur at both the individual – and the macro-level. Chapter 4, in dealing with the European analyses, explains in more detail how different latency effects can be estimated.

To conclude, some summarising remarks:

i. The integration or life course approach, as a theoretical framework in ecological research, has most value when mortality is analysed by cause of death because causes of death have, by definition, their own unique aetiologies since the effects of proximate risk indicators will vary for different causes of death, not only in terms of the strength of effect but also the latency period. These are important factors to consider when mortality is modelled on the basis of exogenous variables.

ii. As parental SES is an early life influence in several chronic diseases, in particular IHD, the consequent health effects are destined to continue for a number of years, even if the lower classes were all to increase their status. However, overall intergenerational health effects have contributed little to the explanation of SEHD (Blane et al., 1993; cited in Davey Smith et al., 1994).

iii. The compositional factors that will be used in the subsequent analyses are particularly related to adult SES. One of the questions to be answered is to what extent the composition of the population by occupational category (primary, secondary or tertiary sector) is relevant in explaining international mortality differences, given that various occupational groups have been characterised by different propensities to certain types of health-related behaviour and other factors.
iv. While specific aspects of the life course cannot be tested in the type of macro-research presented in this thesis, certain explanatory variables enable some of its concepts to be incorporated. For instance, given the early life factors identified by Barker’s foetal origin thesis, it might be useful to include labour-related variables as a way of accommodating the parents’ SES in analyses of infant and childhood mortality. Similarly, as poor growth and development, and adverse early environmental conditions, are positively associated with social class, children of lower SES are more likely to develop chronic diseases during adulthood. Referring back to the first comment, while this may apply to IHD; mortality from lung cancer, other cancers, as well as from accidents and violence, were predominantly influenced by risk factors that were related to social circumstances in adulthood (Davey Smith et al., 1998a). In other words, it might be, data permitting, useful to include estimates of socioeconomic circumstances, both say 40 years ago as well as 5 years ago, in order to estimate IHD mortality among 45-64 year olds. Conversely, a lag of just one year might be sufficient to estimate mortality from accidents and violence.

v. It seems obvious that time series data are needed to test for the effects of economic and other variables on mortality. As mortality is, in most cases, a process taking many years, cross-sectional analyses would be of little value.