The persistence of health inequalities in modern welfare states: The explanation of a paradox

Johan P. Mackenbach
Department of Public Health, Erasmus MC, P.O. Box 2040, 3000 CA Rotterdam, Netherlands

A B S T R A C T

The persistence of socioeconomic inequalities in health, even in the highly developed ‘welfare states’ of Western Europe, is one of the great disappointments of public health. Health inequalities have not only persisted while welfare states were being built up, but on some measures have even widened, and are not smaller in European countries with more generous welfare arrangements. This paper attempts to identify potential explanations for this paradox, by reviewing nine modern ‘theories’ of the explanation of health inequalities. The theories reviewed are: mathematical artifact, fundamental causes, life course perspective, social selection, personal characteristics, neo-materialism, psychosocial factors, diffusion of innovations, and cultural capital.

Based on these theories it is hypothesized that three circumstances may help to explain the persistence of health inequalities despite attenuation of inequalities in material conditions by the welfare state: (1) inequalities in access to material and immaterial resources have not been eliminated by the welfare state, and are still substantial; (2) due to greater intergenerational mobility, the composition of lower socioeconomic groups has become more homogeneous with regard to personal characteristics associated with ill-health; and (3) due to a change in epidemiological regime, in which consumption behavior became the most important determinant of ill-health, the marginal benefits of the immaterial resources to which a higher social position gives access have increased.

Further research is necessary to test these hypotheses. If they are correct, the persistence of health inequalities in modern European welfare states can partly be seen as a failure of these welfare states to implement more radical redistribution measures, and partly as a form of ‘bad luck’ related to concurrent developments that have changed the composition of socioeconomic groups and made health inequalities more sensitive to immaterial factors. It is argued that normative evaluations of health inequalities should take these explanations into account, and that a direct attack on the personal, psychosocial and cultural determinants of health inequalities may be necessary to achieve a substantial reduction of health inequalities.

Introduction

The persistence of socioeconomic inequalities in health is one of the great disappointments of public health. All countries, including those ranking high on indices of economic prosperity and human development, have systematic inequalities in mortality and morbidity between citizens with a higher and a lower socioeconomic position, as indicated by education, occupation, income or wealth. These health inequalities are often substantial, and usually amount to between 5 and 10 years difference in average life expectancy at birth, and between 10 and 20 years difference in disability-free life expectancy (Commission on Social Determinants of Health, 2008; Mackenbach et al., 2008; Sihvonen et al., 1998).

This also applies to the highly developed ‘welfare states’ of Western Europe. All Western European countries have created extensive arrangements aiming to reduce socioeconomic inequality and its various consequences. With notable variations all these ‘welfare regimes’ include measures to redistribute income (e.g. by progressive taxation and social security) and a range of collectively financed provisions (e.g. public housing, education, health care, access to culture and leisure facilities) (Esping-Andersen, 1990; Ferrara, 1996).

There is good evidence that welfare policies have contributed to a reduction of inequalities in income, housing quality, health care access and other social and economic outcomes (Esping-Andersen,
and sometimes also on an absolute scale, has been reported for four decades, a widening of inequalities in mortality, on a relative also in other countries with available data. During the last three or four decades, a widening of inequalities in mortality must have occurred in its early stages. The widening of mortality inequalities has continued into the 21st century, not only in Britain (Fair Society, Healthy Lives, 2010) but also in other countries with available data. During the last three or four decades, a widening of inequalities in mortality, on a relative and sometimes also on an absolute scale, has been reported for many Western European countries (Fawcett et al., 2005; Mackenbach et al. 2003; Strand et al., 2010; Valkonen et al., 2000). This widening started well before the welfare reforms (with cutbacks in provisions and entitlements) of the 1990s (Fawcett et al., 2005; Strand et al., 2010; Valkonen et al., 2000; Wilkinson, 1989).

This paradoxical situation is made even more puzzling by the lack of association between the extent or intensity of welfare policies in a country on the one hand, and the magnitude of its health inequalities on the other hand. Comparative studies have found that socioeconomic inequalities in mortality and morbidity are not smaller in countries with relatively universal and generous welfare policies (e.g. the Nordic countries) than they are in other countries (e.g. the United Kingdom with its more liberal welfare regime, or Southern European countries with their more family-based welfare arrangements). This was first observed for the 1980s (Mackenbach et al., 1997) and then confirmed for the 1990s (Mackenbach et al., 2008) and 2000s (Eikemo, Huisman, Bambra, & Kunst, 2008; Eikemo, Bambra, Joyce, & Dahl, 2008), and applies both to relative and absolute measures of health inequalities. Recent reviews have concluded that the association between welfare regimes and health inequalities is inconsistent (Beckfield & Krieger, 2009; Muntaner et al., 2011).

Table 1 illustrates this for absolute inequalities in mortality in a sample of Western European countries. Within Western Europe, both relative and absolute inequalities in total mortality tend to be smallest in the South, e.g. in the regions of Turin and Madrid. This is primarily due to small inequalities in mortality from cardiovascular disease (among men and women) and cancer (among women) (Mackenbach et al., 2008). In the North of Europe, Sweden tends to have relatively small inequalities in mortality, but Norway's inequalities in mortality are considerably larger than those seen in the South or in England/Wales, partly because of large inequalities in smoking-related causes of death (Mackenbach et al., 2008; Van der Heyden et al., 2009).

In this paper I will explore the explanation of this paradox. What explains the persistence and even widening of socioeconomic inequalities in health in the highly developed welfare states of Western Europe, and the lack of association between generousness of welfare policies and the magnitude of health inequalities? Although this question has been discussed before (Bambra, 2011; Dahl et al., 2006: Eikemo & Huijts, 2009; Hurrelmann et al., 2011), clear hypotheses have not emerged. I will review current scientific theories on health inequalities, and evaluate what elements they could contribute to an explanation of the paradox. My approach resembles that of Bambra (2011), but while the latter study's main conclusion is that “[existing] theories provide little insight into the issue” and that this “public health puzzle highlights the limitations of existing theories” (p. 744), I will argue that some of these theories do suggest plausible hypotheses on the explanation of the paradox. I will also review a wider range of potentially relevant theories. The conclusions of the analysis will then be used to discuss some policy implications, starting with the question of what these explanations would imply for the normative evaluation of health inequalities in modern welfare states. The main purpose of this paper is to generate hypotheses, and to provide a starting point for further scientific discussion.

Current theories and how they may help to explain the paradox

The three components of social stratification

Socioeconomic inequalities in health ultimately derive from social inequality. An analysis of why health inequalities persist should therefore be based on an understanding of social stratification. According to general sociological theory, systems of social stratification are made up of three components, each of which can vary between societies and over time: (1) mobility mechanisms that “sort” individuals into social strata, (2) allocation rules that distribute resources to social strata, and (3) social processes that render some resources more valuable than others (Grusky, 2010).

These three distinct components can also be used to identify the general mechanisms underlying health inequalities. The magnitude of health inequalities in a society will then be a function of: (1) social mobility, and the resulting differences between social strata in personal characteristics of their individual members, (2) resource distribution, and the resulting differences between social

Table 1

<table>
<thead>
<tr>
<th>All causes</th>
<th>Cancer-related causes</th>
<th>Cardiovascular disease</th>
<th>Injury-related causes</th>
<th>Alcohol-related causes</th>
<th>Smoking-related causes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>625</td>
<td>90</td>
<td>309</td>
<td>52</td>
<td>50</td>
</tr>
<tr>
<td>Norway</td>
<td>980</td>
<td>169</td>
<td>434</td>
<td>70</td>
<td>62</td>
</tr>
<tr>
<td>England/W</td>
<td>862</td>
<td>225</td>
<td>401</td>
<td>19</td>
<td>28</td>
</tr>
<tr>
<td>France</td>
<td>1044</td>
<td>333</td>
<td>232</td>
<td>109</td>
<td>196</td>
</tr>
<tr>
<td>Italy (Turin)</td>
<td>639</td>
<td>232</td>
<td>140</td>
<td>23</td>
<td>63</td>
</tr>
<tr>
<td>Spain (Madrid)</td>
<td>530</td>
<td>181</td>
<td>38</td>
<td>26</td>
<td>73</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>381</td>
<td>73</td>
<td>172</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>Norway</td>
<td>518</td>
<td>103</td>
<td>239</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>England/W</td>
<td>462</td>
<td>111</td>
<td>236</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>France</td>
<td>375</td>
<td>50</td>
<td>130</td>
<td>36</td>
<td>39</td>
</tr>
<tr>
<td>Italy (Turin)</td>
<td>197</td>
<td>15</td>
<td>94</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Spain (Madrid)</td>
<td>51</td>
<td>76</td>
<td>56</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td><strong>Source:</strong></td>
<td>Mackenbach et al., 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Injury Alcohol-related causes Smoking-related causes

Fig. 1. Mortality by occupational class among men in England and Wales, ca. 1920—ca. 1980. Source: modified data from Pamuk (1985).
strata in access to material and immaterial resources, and (3) resource benefits, i.e. the value of the resources for the avoidance of health problems that are prevalent in that society. Variations over time or between countries in any of these factors may give rise to variations in the magnitude of health inequalities.

Over the last two decades, the persistence and widening of health inequalities in Western Europe, despite advances in material well-being and welfare policies, has given rise to a rich body of empirically grounded literature. Box 1 gives an overview of nine ‘theories’ of the explanation of health inequalities that are

<table>
<thead>
<tr>
<th>Focus of theory</th>
<th>Main proponents</th>
<th>Short description</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematical artifact</td>
<td>Scanlan⁷</td>
<td>Increasing relative inequalities in health outcomes are inevitable when the over-all level of the outcome falls, and persistence of health inequalities is an artifact of the focus on relative inequalities in negative outcomes.</td>
<td>Relative inequalities in mortality tend to be higher when over-all mortality is lower, but this is not a mathematical necessity. Paradox also applies to absolute inequalities in mortality.</td>
</tr>
<tr>
<td>Fundamental causes</td>
<td>Link⁶</td>
<td>Socioeconomic position involves access to resources which can be used to avoid disease risks or to minimize the consequences of disease once it occurs, regardless of what the current profile of diseases and known risks happens to be.</td>
<td>Reformulates the problem without identifying the specific pathways linking socioeconomic position and health. However, refocusing attention on fundamental aspects of social stratification is useful.</td>
</tr>
<tr>
<td>Life course perspective</td>
<td>Wadsworth⁶</td>
<td>Health at adult ages is partly determined by exposure to biological and social factors at the start of life, and the roots of health inequalities may therefore lie in inequalities experienced in the womb and during childhood and adolescence.</td>
<td>May explain why health inequalities at adult ages respond with long delays only to more equal living conditions. However, there is no evidence that health inequalities are smaller in generations exposed to more extensive welfare arrangements.</td>
</tr>
<tr>
<td>Social selection</td>
<td>Black report⁶</td>
<td>In modern societies, people are socially mobile, and are sorted into social classes on the basis of health (‘direct health selection’) or health determinants (‘indirect health selection’).</td>
<td>Evidence for ‘direct’ health selection mainly limited to income-health relationship. ‘Indirect’ health selection is difficult to measure, but may explain paradox if it has increased over time or is associated with welfare policies. Empirical support is growing. May provide pathway for ‘social selection’ theory, and may explain paradox if relative importance of ‘personal characteristics’ for health has increased over time.</td>
</tr>
<tr>
<td>Personal characteristics</td>
<td>Batty⁷</td>
<td>In modern societies, socioeconomic position is strongly associated with personality, cognitive ability and other personal characteristics which affect health.</td>
<td>Persistence of inequalities in material resources is well documented, and availability of material resources still affects health, but cannot explain trends over time or geographical patterns within Western Europe.</td>
</tr>
<tr>
<td>‘Neo-material’ factors</td>
<td>Lynch⁸</td>
<td>Inequalities in material resources, both at the individual and community level, are still universal, and lead to accumulation over the life course of exposures and experiences which affect health.</td>
<td>Persistence of inequalities in exposure to psychosocial stress is well documented, and psychosocial stress does affect health, but cannot explain trends over time or geographical patterns within Western Europe.</td>
</tr>
<tr>
<td>Psychosocial pathways</td>
<td>Marmot⁹</td>
<td>Socioeconomic position is strongly associated with psychosocial stress, e.g. through variations in exposure to demand-control imbalance or through relative deprivation.</td>
<td>Persistence of inequalities in exposure to psychosocial stress is well documented, and psychosocial stress does affect health, but cannot explain trends over time or geographical patterns within Western Europe.</td>
</tr>
<tr>
<td>Diffusion of innovations</td>
<td>Rogers⁹</td>
<td>Increasing inequalities in health outcomes result from a faster rate of improvement in higher socioeconomic groups, which is due to earlier adoption of new behaviors and earlier uptake of new interventions.</td>
<td>Supported by a lot of evidence, but theory does not identify the specific pathways linking socioeconomic position and adoption of new behaviors or uptake of new interventions.</td>
</tr>
<tr>
<td>Cultural capital</td>
<td>Bourdieu⁸</td>
<td>In modern societies, socioeconomic position is still strongly associated with cultural factors such as normative beliefs and knowledge on health risks, which strongly affect health because the latter is largely determined by lifestyle.</td>
<td>Empirical support is limited, but may provide specific pathway for ‘diffusion of innovations’ theory. May explain paradox if relative importance of ‘cultural capital’ for health has increased over time.</td>
</tr>
</tbody>
</table>

⁸ Vageró & Erikson, 1997.  
¹⁰ Phelan, Link, & Tehrani, 2010.  
¹³ Mackenbach, 2001a.  
¹⁸ Rogers, 1962.  
¹⁹ Victoria, 2000.  
²¹ Abel, 2008.
potentially relevant for the explanation of the paradox. Not all of these ‘theories’ may fully deserve this label — the notion of ‘mathematical artifact’ is a rather narrow account only, and the ‘life course perspective’ could perhaps better be labeled a broad conceptual scheme — but for convenience we will consistently use the term ‘theory’ to refer to each of them. None of these theories was developed to explain the paradox that is central to this analysis, but I will assess them for their usefulness in explaining the paradox, by deriving hypotheses that are compatible with the observations of persisting and even widening health inequalities, also in highly developed welfare states. Please note that I do not intend to provide a general assessment of the validity of these theories. None of these theories is mutually exclusive, and they may therefore operate simultaneously and reinforce each other. With the exception of three very general theories (‘mathematical artifact’, ‘fundamental causes’, and ‘life course perspective’) each relates to one of the three components of health stratification mentioned above.

It is important to note that, in order for a theory to contribute to the explanation of the paradox, the strength of the proposed mechanism or pathway must have become larger over time, and/or be larger in countries with more extensive welfare arrangements. As modern European welfare states have to some extent been successful in reducing inequalities in access to material resources, one or more of the other components must have counteracted that beneficial effect.

Three general theories

The theory of ‘mathematical artifact’ suggests that increasing relative inequalities in health outcomes are inevitable when the over-all level of the outcome falls (Scanlan, 2001, 2006; Vagerø & Erikson, 1997). In intuitive terms, the idea is that when the background risk of mortality or morbidity is lower (as it tends to be in more recent time-periods, and in more advanced welfare states), it is ‘easier’ for determinants to produce a high relative risk (even though the absolute risk difference declines). There is indeed an association between the average frequency of health problems in a population and the level of the relative risk for socioeconomic status: relative risks for mortality and morbidity tend to be higher when average mortality and morbidity are lower (Eikemo et al., 2009). However, it has been shown that larger inequalities are not a mathematical necessity when over-all health improves (Houweling et al., 2007), and an association between average frequency and relative risk does not prove causality. Furthermore, absolute inequalities in mortality have also increased in several countries (Mackenbach et al., 2003; Strand et al., 2010; Valkonen et al., 2000), and their pattern of between-country variation is largely similar to that of relative inequalities in mortality (Mackenbach et al., 2008), so that it is difficult to see how this theory could explain the paradox.

The theory of ‘fundamental causes’ stipulates that it is the social forces underlying social stratification that ultimately cause health inequalities, and not exposure to the proximal risk factors, which are usually studied by social epidemiologists (like smoking, psychosocial stress, working conditions etc.). According to this theory, the persistence of health inequalities in different time-periods and different national conditions is due to the fact that a person’s socioeconomic status provides him or her with “flexible resources”. These include knowledge, money, power, prestige, and beneficial social connections that can be used to avoid disease risks or to minimize the consequences of disease once it occurs regardless of the prevailing circumstances. The association between socioeconomic status and health is reproduced over time via the replacement of intervening mechanisms, and as opportunities for avoiding disease continue to expand so will health inequalities continue to exist (Link & Phelan, 1995; Phelan et al., 2004, 2010). This theory is an elegant reformulation of the problem, and the paradox may actually be seen as an example of the workings of this theory. But it does not provide a specific explanation of the paradox. What are the new intervening mechanisms that replace the mechanisms eliminated or attenuated by the welfare state? Or are health inequalities perhaps reproduced in highly developed welfare states because people with higher socioeconomic status make better use of these welfare resources? We need other theories to provide more specific explanations.

The ‘life course perspective’ is another useful and increasingly popular approach to the explanation of health inequalities. This approach is based on the observation that health at adult ages is partly determined by experiences in early life, both biological and social. ‘Biological programming’ of the fetus has been hypothesized to increase vulnerability to chronic diseases like diabetes and cardiovascular disease (Barker et al., 1989), and unfavorable social and health conditions in childhood may be the starting-points of pathways leading into both health and social disadvantage in adulthood (Wadsworth, 1997). Some of these processes have been shown to extend over several generations (Modin & Fritzell, 2009). While this may explain a long delay between exposure to modern welfare arrangements and a reduction of health inequalities among adults, it does not explain the fact that even among generations who were born during the welfare state there is no association between welfare policies and the magnitude of health inequalities (Bambara et al., 2010). Furthermore, long delays are not inevitable as both mortality and morbidity often respond quickly and dramati-
cally to changing social conditions (Vaupel et al., 2003). In any case, as with ‘fundamental causes’ more specific explanations will be needed.

Two theories focusing on social mobility and the composition of social strata

The theory of ‘social selection’ suggests that health inequalities result from health-related selection during social mobility. Health problems may lead to downward social mobility (thereby creating ‘direct health selection’) (Inequalities in Health, 1982), and upward mobility is more likely for those with personal characteristics conducive to good health (thereby creating ‘indirect health selection’) (West, 1991). Direct health selection is equivalent with reverse causality (health problems leading to low socioeconomic status, instead of vice versa), while indirect health selection can also be thought of as confounding (by third factors which influence both health and socioeconomic status). This theory would help to explain the paradox if social mobility has increased over time, or if social mobility has become more selective with regard to health-relevant characteristics, particularly in countries with generous welfare arrangements.

Intergenerational social mobility has increased slowly but systematically in all high-income countries since the Second World War, both with regard to educational achievement and occupational class (Breen, 2004; Ganzeboom, 2007; Ganzeboom et al., 1989; Heath, 1981; Raad voor Maatschappelijke Ontwikkeling, 2011). International-comparative studies have not found consistent differences between countries in rates of intergenerational social mobility, but countries with well-developed welfare policies such as the Nordic countries usually also have egalitarian education policies and substantial social mobility (Erikson & Goldthorpe, 1992). As a result, the scope for both ‘direct’ and ‘indirect health selection’ is likely to have become larger over time and to be larger in more advanced welfare states.

While mobility reduces the degree of intergenerational transmission of social (dis)advantage in a society, and therefore may
the health effects of social stratification, it also increases opportunities for selection into higher social positions on the basis of personal characteristics, such as cognitive ability and personality profiles. To a certain degree modern societies have become ‘meritocracies’, in which educational and occupational achievement are no longer dependent on family background but on personal talent and effort (Arrow, 2000). To the extent that these personal characteristics are important for health, e.g. because they determine health-related behavior, inequalities in health could have increased as a result.

A number of studies have shown that cognitive ability and the Big Five personality traits statistically explain part of the gap in health between socioeconomic groups (Batty et al., 2006; Hart et al., 2004; Judge et al., 1999; Nabi et al., 2008). Socioeconomic inequalities in these personal characteristics may partly arise from differences in early environment, because the comparatively harsh living conditions of families of low socioeconomic status increase family stress and hamper family investments in children, which could harm their personality and cognitive ability (Mackenbach, 2010a), partly from selection during social mobility. While inequalities in early environment are likely to have become smaller over time, and to be smaller in countries with more extensive welfare arrangements, opportunities for social selection on the basis of these personal characteristics may well have become larger, as noted above.

Two theories focusing on the distribution of specific resources

The ‘neo-materialist theory’ emphasizes the fact that despite increases in average prosperity and some redistribution inequalities in access to material resources are still universal, and still generate health inequalities (Davey Smith et al., 1994; Lynch et al., 2000). There is indeed no doubt that inequalities in material advantage are still substantial everywhere, even in countries with relatively small income inequalities (Luxembourg Income Study, 2011). The welfare state, and its redistribution of lifetime welfare through taxation, cash transfers and non-cash benefits, does attenuate material inequalities, but what remains is still substantial (Ter Rele, 2007). Over the past decades, even ‘diseases of affluence’ have become associated with material disadvantage, because the latter partly determines health-related behaviors, such as leisure-time physical exercise, diet etc. (Van Oort et al., 2005).

However, while this theory may provide a plausible explanation for the persistence of health inequalities, it is unlikely to explain their widening. Over the 20th century income inequality has declined in Western European countries, partly as a result of welfare policies, and the same applies to wealth inequalities (Nolan & Lenski, 2004). It is only since the 1980s or 1990s that income and wealth inequalities have tended to rise again, partly as a result of welfare reform, partly as a result of other influences (changes in household composition, globalization etc.) (OECD, 2008). Inequalities in mortality started to rise well before that (Pamuk, 1985; Strand et al., 2010; Valkonen et al., 2000). Also, within Western Europe there is no clear association between the magnitude of income inequalities and the magnitude of socioeconomic inequalities in morbidity or mortality (Mackenbach et al., 1997).

Similarly, the ‘psychosocial theory’, which emphasizes the role of psychosocial stress, lack of social support, and sense of control (Marmot, 2004; Wilkinson, 2005), cannot plausibly explain the widening of health inequalities, although the unequal distribution of these psychosocial factors may contribute to their persistence. Despite the welfare state, considerable differences in power and prestige have continued to exist. People with a higher level of education and income have a much stronger sense of control over their own lives, and this is linked to healthier behavior, and lower rates of morbidity and mortality (Bosma et al., 1999). The perception of material inequality, and in particular of one’s own ‘relative deprivation’, may have a direct effect on a person’s health via psychosocial stress mechanisms (Wilkinson & Pickett, 2009). The welfare state may have blurred some of the inequalities in psychosocial stress, but the middle classes have also benefited from the welfare state, for example because it has reduced the psychosocial stress of job insecurity even among the employed (Sjoberg, 2010).

However, in order to explain the widening of health inequalities one would have to assume that inequalities in exposure to psychosocial risk factors have increased over time, and this is unlikely, if only because material hardship in lower socioeconomic groups must have been a powerful source of psychosocial stress in the past. It has sometimes been argued that advanced welfare states may raise unrealistic expectations of a better life among people with a lower socioeconomic position, and therefore induce higher levels of frustration and stress (Dahl et al., 2006; Wilkinson & Pickett, 2009), but this is largely speculative.

Two theories focusing on the value of resources for health gain

Studies of the ‘diffusion of innovations’ have observed that people with a higher socioeconomic position often tend to be early adopters of new behaviors, only later to be followed by those with a lower social position (Rogers, 1962). This theory provides a plausible explanation for a (temporary) widening of health inequalities when major improvements in population health occur that are mediated by behavior change. Many Western European countries, particularly those in the North and West, have entered a new phase of the epidemiological transition, in which the ‘diseases of affluence’ (e.g. cardiovascular disease, some cancers, injuries), which dominated the third stage, are rapidly being pushed back (Olshansky & Ault, 1986; Omran, 1971). This is partly the result of health care interventions and partly the result of behavior change, such as quitting smoking and adopting a more healthy diet (Capewell et al., 1999; Goldman & Cook, 1984). As predicted by this theory, these behavior changes tend to follow a trajectory through populations in which those with a higher social position adopt the new behavior first (Lopez et al., 1994; Sobal & Stunkard, 1989). As a result, this dynamic phase is characterized by large and widening inequalities in health behaviors, which in turn lead to large and widening inequalities in mortality (Cavelaars et al., 2000; Roskam et al., 2010). Within Western Europe, countries in the South happen to be in a later stage of this transition, with inequalities in health behaviors only recently arising and inequalities in ‘diseases of affluence’ still being small (Avendaño et al., 2006; Mackenbach et al., 2008; Van der Heyden et al., 2009).

The ‘inverse equity’ hypothesis, which postulates that “new interventions will initially reach those of higher socioeconomic status and only later affect the poor [...] which results in an early increase in inequality ratios” (Victor et al., 2000) is based on a similar line of reasoning, but focuses on the emergence of inequalities in the use of preventive or curative health care during periods of rapid health improvement. These inequalities may be due to differences in both access and uptake, and may explain part of the socioeconomic inequalities in mortality from conditions which have become amenable to medical intervention in Western Europe (Stirbu et al., 2010).

It is as if, during the most dynamic phases of health improvement, the marginal benefits of a higher social position for health temporarily become larger. Economic resources may play a role: in the early stages of their implementation new interventions are often expensive, and being rich may temporarily become very important for preserving health. But in the case of behavior change,
cultural resources may also play a role. The theory of ‘cultural capital’ explains inequalities in consumption behavior from differences in attitude, knowledge and competency between socioeconomic groups, which are transmitted across generations. These differences partly arise from the need for ‘social distinction’: people in higher socioeconomic groups behave differently to show off their social position, but doing this successfully requires a lot of ‘cultural capital’ (Bourdieu, 1984). It is unlikely that, during the last decades, inequalities in ‘cultural capital’ have increased, but the need for ‘social distinction’ on the basis of health-related behavior may have increased due to the decline in opportunities to distinguish oneself by outward signs of material prosperity. More importantly, the relevance of cultural capital for health is likely to have increased with the rise and subsequent decline of the consumption-related diseases which happen to be prevalent in highly developed welfare states (Abel, 2008).

Discussion

Synthesis

Only some of the reviewed theories can provide a plausible and specific explanation for the persistence and even widening of health inequalities in the advanced welfare states of Western Europe. The ‘mathematical artifact’, ‘fundamental causes’ and ‘life course’ theories reframe the problem in different (and potentially useful) terms but do not provide a specific explanation for the paradox. ‘Neo-materialist’ and ‘psychosocial’ theories help us to understand why health inequalities persist, i.e. because inequalities in access to material and immaterial resources persist. If modern welfare states had abolished the economic and social structures that produce unequal lives, health inequalities would probably have largely disappeared. But these theories do not explain their widening, or the fact that health inequalities do not tend to be smaller in countries with more generous welfare policies.

The other four theories hold more promise, and can be used to derive two general hypotheses to address the paradox. The first hypothesis is that the lower social strata have become more exclusively composed of individuals with personal characteristics that increase the risks of ill-health. This is the result of decades of upward intergenerational social mobility, which may have increased opportunities for social selection and may have made the lower social groups more homogeneous with regard to personal characteristics like low cognitive ability and less favorable personality profiles. The increase of intergenerational social mobility is primarily due to changes in the economy that have led to an expansion of higher education, but to the extent that welfare policies have contributed to making the education system more merit-based, they may paradoxically have contributed to a widening of health inequalities.

The second hypothesis is that advanced welfare states happen to be further along in their epidemiological development, and have now reached the fourth stage of the epidemiological transition in which health improvement depends to a large extent on behavior change. This may have increased the importance of immaterial factors like cultural capital (and personal characteristics like cognitive ability) for health — factors that may be as strongly (cultural capital) or even more strongly (personal characteristics) socially differentiated than they were before, because they have largely been left untouched by the welfare state. To the extent that welfare policies have contributed to making an ‘affluent lifestyle’ widely affordable, they may again paradoxically have contributed to a widening of health inequalities.

That inequalities in access to material and immaterial resources have not been eliminated could be seen as a partial failure of the welfare state: had more radical redistribution measures been taken, then social inequalities might have been reduced more effectively, and health inequalities might have declined as well. The welfare state, however, has never had such revolutionary goals. On the contrary, its rise reflected a compromise between the interests of employers and employees, between the laboring and the middle classes, and between the political ideologies representing these interests (Hicks, 1999). It therefore should not come as a surprise that its redistributive effects were modest.

It is also not surprising that the welfare state did not have effective solutions for widening inequalities in ‘diseases of affluence’. The welfare state’s ‘bad luck’ was that, while it had been invented to remediate poverty and its consequences, it was soon overtaken by the rise and subsequent fall of ‘diseases of affluence’, against the social patterning of which the old arsenal was largely ineffective. One poignant illustration is tobacco control: countries that have made stronger tobacco control efforts tend to have larger socioeconomic inequalities in smoking (Schaap et al., 2008).

As explained in the introduction, the approach followed in this paper was similar to that found in Bambra’s article (2011). Bambra reviewed six theories: ‘artifact’, ‘health selection’, ‘cultural-behavioural’, ‘materialist’, ‘psychosocial’, ‘life course’. Although they were sometimes conceptualized differently, these theories roughly correspond with what I have labeled ‘mathematical artifact’, ‘social selection’, ‘diffusion of innovations’, ‘neo-material factors’, ‘psychosocial pathways’ and ‘life course’. Bambra did see some value in the ‘cultural-behavioural’ approach for explaining the paradox, but her over-all conclusion was that these theories “provide little insight”, mainly because most of them “to a greater or lesser extent expect health inequalities to be smaller in the Scandinavian countries” (Bambra, 2011, p. 743). The analysis reported here reviewed three additional theories (‘fundamental causes’, ‘personal characteristics’, ‘cultural capital’), and comes to rather different conclusions. It articulates two hypotheses based on, respectively, a combination of the ‘social selection’ and ‘personal characteristics’ theories, and a combination of the ‘cultural capital’ and ‘diffusion of innovations’ theories. It also emphasizes the importance of the ‘neo-materialist’ and ‘psychosocial’ theories for understanding why health inequalities persist (because the welfare state has far from eliminated inequalities in access to material and immaterial resources).

Limitations

As stated in the Introduction, this analysis had a hypothesis-generating character, and does not pretend to provide final answers. Its limitations include the theoretical nature of the analysis: although most of the nine theories have substantial empirical support, empirical analyses that try to explain the persistence and widening of health inequalities in advanced welfare states from these theories are almost non-existent. I suggest that further research on the paradoxical relationship between welfare policies and health inequalities should focus on the specific hypotheses that can be derived from these theories. This will require considerable efforts in data mining and/or data collection, because internationally and/or historically comparable data on the composition of social strata or on the social distribution of immaterial resources are difficult to find. The hypothesis on the changing composition of social strata can be empirically tested by comparing socioeconomic inequalities in, for example, personality profiles and cognitive abilities between different birth cohorts. The hypothesis on the rising importance of immaterial resources can be empirically tested by comparing the contribution of socioeconomic inequalities in indicators of cultural capital to the explanation of inequalities in health behavior and health outcomes.
It is likely that further research would also benefit from unpacking welfare regimes into specific components: some Nordic welfare policies may reduce inequalities in health, while others are neutral or may even exacerbate them. The same applies to measures of socioeconomic status. This analysis was limited in the sense that I have used a global concept of socioeconomic status, ignoring the fact that education, occupational class, income etc. have partly different associations with mortality and morbidity. The various mechanisms that were reviewed may also have different relevance for each of these indicators of socioeconomic position: for example, selection effects (reverse causality) are likely to be more important for income than for education. Further research will benefit from a higher degree of specificity with regard to the social determinants as well.

**Policy implications**

Can health inequalities as they are currently found in Western Europe still be regarded as ‘unfair’, and do widening health inequalities imply that they are becoming more unfair? Such questions immediately arise when one considers the hypotheses suggested above. Contrary to what social epidemiologists usually assume (Commission on Social Determinants of Health, 2008), it is not obvious that health inequalities are ‘unfair’. Health inequalities are certainly ‘tragic’; it is sad that people who have less of everything also tend to live shorter lives and spend more years in ill-health, and it is disappointing that after all that has been done to reduce social inequalities and their consequences, health inequalities have not been eliminated. But because health is not a ‘good’ that lends itself for redistribution, the unfairness of health inequalities does not automatically follow from their existence, however tragic this may be. Normative evaluations need to focus on how health inequalities arise: it is the processes by which, or the circumstances from which they arise that should be evaluated for their distributive fairness (Daniels et al., 1999; Deaton, 2011; Whitehead, 1991). The analysis presented in this paper of how health inequalities arise therefore also raises potentially important normative questions.

If the hypotheses presented in this paper are correct, then current health inequalities arise from partly different mechanisms as compared to those observed before the rise of the modern European welfare state, some 50 years ago (with variations between countries). Nowadays, health inequalities arise in a society with considerable social mobility, and whether one stays poor when one is born poor may have become partly dependent on personal characteristics like cognitive ability and personality factors. Some of the involuntary causes of health inequalities have been taken away by extensive welfare arrangements, and although patterns of oppression and exploitation can still be distinguished these have to some extent been softened by democratization and equal rights movements. Health inequalities between social strata partly arise from differences in consumption behavior, which may partly be determined by cultural factors and personal characteristics, and not only by material disadvantage. Health inequalities have recently widened, not necessarily because of widening inequalities in access to material or even immaterial resources, but possibly also because of changes in the composition of social strata and because the marginal benefits of a higher social position are larger in times of rapid epidemiological change.

Even those (like the present author) who believe that current inequalities in income, status and power are too large to be justified by whatever differences there are between individuals in merit or effort, and must therefore partly result from subtle and less subtle forms of exploitation, and who believe that social inequalities in consumption behavior do not result from the exercise of free choice but from differences in material and immaterial living conditions, may ask themselves whether if the hypotheses presented here are correct, current health inequalities are perhaps less unfair than those observed 50 years ago – even though they are larger. Can health inequalities that result from differences in composition of social strata be considered unfair, even if they are a by-product of social policies promoting equal educational opportunities? Can health inequalities that result from differences in immaterial resources (such as cultural capital), which do not directly result from forms of oppression or exploitation, and which do not easily lend themselves for redistribution between social strata, be considered unfair? Can health inequalities that are a temporary side-effect of over-all progress, from which everyone in the end will probably benefit, be considered unfair?

But even if current health inequalities would not unambiguously and completely be considered unfair, they are still likely to be seen as partly unfair, and unfairness is not the only legitimate reason for reducing them. One may also find them tragic, and want to reduce them because they reinforce social inequalities by depleting the smaller resources of people in lower socioeconomic groups, and by preventing them to reinforce their social position. Other reasons include the huge population health losses due to health inequalities, and the economic costs of health inequalities (e.g. through lost labor productivity and costs of health care and social security) (Mackenbach et al., 2011).

It is no wonder, therefore, that there have been many pleas for reducing or even eliminating health inequalities, and that several European countries have developed and sometimes implemented national programs to tackle health inequalities (Independent Inquiry into Inequalities in Health, 1998; Mackenbach et al., 2003; Mackenbach & Stronks, 2002; National Action Plan to Reduce Health Inequalities 2008-2011; National strategy to reduce social inequalities in health, 2007; Public Health Policy Report, 2005). These attempts have so far, however, been insufficiently effective. The most powerful attempt at tackling health inequalities so far, the English strategy of successive Labor governments in the period 1997–2010, has produced a few positive intermediate results but has failed to reduce health inequalities. This was due to a lack of policies of proven effectiveness and a lack of political determination to pursue sufficiently radical solutions (Mackenbach, 2010b, 2011).

If the hypotheses presented in this paper are correct, neither of these two circumstances is surprising. According to these hypotheses, modern European welfare states have been ineffective against health inequalities, partly because of a failure to implement more radical redistribution measures, partly because of concurrent developments which have changed the composition of socioeconomic groups and made health inequalities more sensitive to immaterial factors. In that case, a substantial reduction of health inequalities can only be achieved with more radical redistribution measures, partly because of a failure to implement more radical redistribution measures, and/or a direct attack on the personal, psychosocial and cultural determinants of health inequalities. As long as there is insufficient political support for the first, and as long as the second is unfeasible because of a lack of effective interventions, those who want to reduce health inequalities will have to be satisfied with small steps forward.

**Acknowledgment**

The author would like to thank Dr Mauricio Avendano, Dr Terje Eikemo and Dr Frank van Lenthe for their comments on a previous version of this paper. The valuable comments of the anonymous reviewers have also helped to further improve this paper and clarify its reasoning.