

Social inequality in chronic disease outcomes

The role of differential exposure and vulnerability to health behaviours

Helene Nordahl

This review has been accepted as a thesis together with four previously published papers by University of Copenhagen 6th March 2014 and defended on 28th May 2014

Tutors: Merete Osler, Birgitte Lidegaard Frederiksen & Naja Hulvej Rod

Official opponents: Johan Hallqvist, Jørn Olsen & Lene Theil Skovgaard

Correspondence: Helene Nordahl, Section of Social Medicine, Department of Public Health, University of Copenhagen, Denmark

E-mail: helenenordahl@gmail.com

Dan Med J 2014;61(11)B4943

THE THESIS IS BASED ON THE FOUR STUDIES

1. Nordahl H, Hvidtfeldt, Diderichsen F, Rod NH, Osler M, Frederiksen BL, Prescott E, Tjønneland A, Lange T, Keiding N, Andersen PK, Andersen I. Cohort Profile: The Social Inequality in Cancer Cohort Study. *Int J Epidemiol*. first published online February 17, 2014 doi:10.1093/ije/dyu003
2. Nordahl H, Rod NH, Frederiksen B, Andersen I, Lange T, Diderichsen F, Prescott E, Overvad K, Osler M. Education and risk of coronary heart disease: assessment of mediation by behavioral risk factors using the additive hazards model. *Eur J Epidemiol*. 2013;28:149-57
3. Nordahl H, Osler M, Frederiksen BL, Andersen I, Diderichsen F, Prescott E, Overvad K, Rod NH. Combined effects of socioeconomic position, smoking and hypertension on risk of ischemic and hemorrhagic stroke. *Stroke*. 2014;45:2582-7
4. Nordahl H, Lange T, Osler M, Diderichsen F, Andersen I, Prescott E, Tjønneland A, Frederiksen BL, Rod NH. Education and cause-specific mortality: the mediating role of differential exposures and vulnerability of behavioral risk factors. *Epidemiology*, 2014;25:389-396

INTRODUCTION

Socioeconomic differences in morbidity and mortality, particularly across educational groups, are widening even in the Nordic welfare states (1-5). Taking Denmark as an example, the difference in life expectancy at age 30 between lower and higher educational groups have increased from less than five years to more than six years during the past three decades (6). Such differences in life expectancy imply adverse health consequences at the individual level and could lead to expenses at the societal level (7).

Chronic diseases such as heart disease, cerebrovascular disease, chronic obstructive lung disease, and lung cancer are some of the illnesses that contribute the most to social inequality and the burden of disease in Europe (7;8). From a public health perspective it seems unfeasible to modify the fundamental social stratification in order to tackle social inequality in these chronic diseases. It appears more relevant to know, what fraction of this inequality could be eliminated by modifying well-known risk factors (9). Health behaviours such as smoking, excessive consumption of alcohol, physical inactivity, certain dietary patterns, and obesity are well-established modifiable risk factors for common chronic diseases (10-15). Although many of these risk factors are unequally distributed across socioeconomic groups, the pathways through which socioeconomic factors affect morbidity and mortality are not fully understood and may vary in time, place and type of outcome (16;17).

Differential exposures to behavioural risk factors, particularly smoking and obesity, have been shown to play an important mediating role on the social inequality in cardiovascular disease (18-21). However, little is known about the extent to which socioeconomic factors interact with health behaviour, creating subgroups that are more vulnerable than others. From a public health perspective it appears obvious to identify in which population subgroups specific interventions or preventive strategies could prevent most cases (22). To further understand how socioeconomic factors (e.g., education) combines with behavioural risk factors to affect chronic disease outcomes, the research summarised in this thesis is a practical implementation of contemporary statistical methodology. In this implementation it is possible to regard behavioural risk factors not only to as mediators but also the role of their interaction with socioeconomic factors.

AIM OF THE THESIS

The overall aim of the thesis was to address how behavioural risk factors have contributed to the social inequality in heart disease, cerebrovascular disease, and cause-specific mortality.

The thesis was organized around two research questions:

How did differential exposure to behavioural risk factors mediate the effect of socioeconomic position on chronic disease outcomes in Denmark?

To what extent did behavioural risk factors interact with the effect of socioeconomic position on chronic disease outcomes - creating differential vulnerability in Denmark?

To answer these research questions the objectives were:

1) To establish a sufficiently large study population in which power demanding questions of mechanisms underlying social inequalities in chronic disease outcomes can be investigated by pooling and harmonising prospective data from existing cohort studies in Denmark (Paper 1)

2) To examine smoking, body mass index and physical inactivity as potential mediators of the relationship between education and coronary heart disease. Further, to compare the results from a new approach to mediation analysis with a conventional approach (Paper 2)

3) To examine interaction between education, smoking, and hypertension in relation to ischemic and haemorrhagic stroke (Paper 3)

4) To examine the mediating and interacting role of smoking, body mass index, physical inactivity, and alcohol intake in the relationship between education and cause-specific mortality (Paper 4)

BACKGROUND

Social inequality in chronic diseases and health behaviour

Despite substantial declines over the past decades, cardiovascular diseases, in particular heart disease and cerebrovascular disease are among the leading causes of death in Europe (2;8;17). These illnesses are also great contributors to the social inequality in burden of disease, measured as life years lost due to premature death and disability (7). Cancer, especially tobacco-related cancers, also seems to contribute to the pattern of social inequalities in diseases (23-25). Studies have suggested that lung, stomach, and rectum cancer are more common in lower socioeconomic groups, whereas breast and colon cancer have the opposite social gradient (26-31). Within recent decades, there has been an emerging interest in the rising social inequality in respiratory diseases, particularly chronic obstructive pulmonary disease (8;32). In Denmark, socioeconomic position is strongly and consis-

tently associated with lung function and subsequent hospital admission for chronic obstructive pulmonary disease (33).

In most westernized countries reducing “unhealthy lifestyle” are a key target for general improvements of the population’s health and quality of life (34-38). Although the proportion of everyday smokers has decreased in the general population, socioeconomic differences in smoking are widening (39-41). Taken Denmark as an example smoking is about three times as common among lower compared to higher socioeconomic groups (7). For obesity (BMI ≥ 30), both the proportion and the social inequality are increasing, while for physical inactivity and consumption of alcohol, the picture is more complicated (42;43). For example, people with low education often have physically demanding jobs, while they are less physically active in their leisure time, whereas the opposite pattern is observed for people with high education (7). In Denmark alcohol consumption is almost equally distributed across socioeconomic groups, although with a slightly higher proportion of heavy drinkers among women with high socioeconomic position (43;44).

Mechanisms underlying social inequality in chronic diseases

In this thesis the investigation of pathways leading from socioeconomic position to chronic disease outcomes is founded in a slightly modified version of the conceptual framework developed by Diderichsen and Hallqvist (45), depicted in Figure 1.

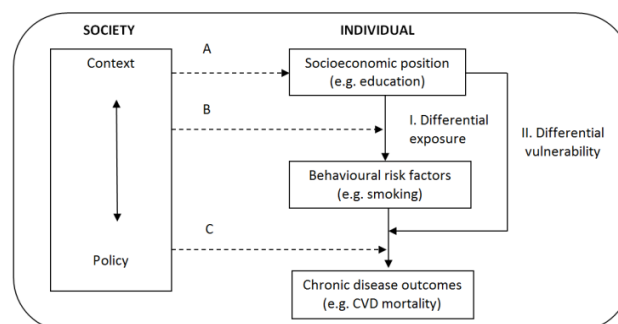


Figure 1

A conceptual framework for studying the mechanisms (I and II) related to social inequality in health and associated policy entry points (A-C). Modified Source: Whitehead, Burstrom & Diderichsen (60).

This framework considers how a person’s socioeconomic position (defined by the educational level, occupation class or income, for example) influences its exposure to specific patterns of behavioural risk factors. This mechanism (I) is referred to as *differential exposure*. Throughout their lives people are exposed to a number of risk factors that vary between socioeconomic groups by type, amount and duration. Other things being equal, these exposures can potentially mediate the effect of socioeconomic position on chronic disease outcome. For example, it might be that the knowledge and skills attained through education affect a person’s cognitive functioning and ability to be more receptive to health education messages and choose healthy lifestyles, which in turn could lead to low risk of ill health (46). In addition, the framework considers how potential mediators might interact with each other

and with socioeconomic position. This mechanism (II) is referred to as *differential vulnerability*. The effect of socioeconomic position on heart disease may, for example, be dependent on smoking. Since people with low socioeconomic position will often be exposed to smoking, these two risk factors may act together in causing heart disease and produce higher vulnerability to the effect of smoking among people with low socioeconomic position (47). Finally, the framework also illustrates how policy, as part of the social context, may influence on the pathways between socioeconomic position and chronic disease outcomes at three entry points (48): (A) influencing social stratification, (B) decreasing exposure, and (C) decreasing vulnerability.

Table 1 summarizes previous studies on behavioural risk factors as potential mediators of the relationship between socioeconomic position (primarily defined by education) and cardiovascular disease and cause-specific mortality. Briefly, some studies have demonstrated that health behaviours explain a substantial part of the relative socioeconomic differences in cardiovascular disease and mortality (18-21;49-53), but others have concluded that health behaviours cannot account for such inequalities (54-58). Most of these studies quantified how differential exposure to health behaviours mediated the inequality in relative terms (as the ratio between groups), and only a few studies (56;57;59) offered such quantification in absolute terms (as differences between groups). None of the previous evaluations have accommodated the possibility that the mediated effect may require the joint operation of exposure and mediator. Thus, no insight has been given into the extent to which socioeconomic position interacts with health behaviours - creating differential vulnerability.

MATERIAL AND METHODS

The Social Inequality in Cancer cohort study

To provide a large study population with a wide age distribution and long follow-up the Social Inequality in Cancer (SIC) cohort study was initiated in 2011 by pooling and harmonising prospective data from seven existing cohort studies: the Copenhagen City Heart study (63;64); the Diet, Cancer and Health study (65); five selected cohorts from the Research Centre for Prevention and Health (66); MONICA I-III, the 1936-cohort, and the Inter99 study. For enrolment in the SIC cohort the inclusion criteria were: a population-based study with data on behavioural and biological risk factors for sub-types of cancer and cardiovascular diseases initiated after 1980 (socioeconomic information drawn from the central registries was only available after January 1980). Although cancer was the initial focus of the SIC cohort, it contained plentiful data which enabled using it for addressing other purposes such as subtypes of cardiovascular diseases and cause-specific mortality, which was the objective of this thesis.

The SIC cohort is described in details in Paper 1. Briefly, participants had been randomly selected from the general population. They all participated in health examinations, blood samples, and self-administered questionnaires. The work of this thesis should

be seen in the context of the great effort that has gone into pooling and harmonising these data. The distribution of selected harmonised measures and characteristics of participants from each of the seven cohort studies pooled into the SIC cohort are presented in Table 2. The seven cohorts had quite different study entry and sample size. The earliest cohort, the Copenhagen City Heart study, was initiated in 1981, whereas the latest cohort, the Inter99 study, was initiated in 1999. The largest cohort, the Diet, Cancer and Health study, constituted more than 70 % of the pooled study population. However, the other cohorts combined provided almost half of the person-years at risk, because most of them had longer follow-up than the Diet, Cancer and Health study. The distribution of the harmonised behavioural risk factors varied across the cohort studies. For example, the proportion of current smokers varied from 58% to 36%.

Measurements of health behaviour

Information on health behaviour was based on data from self-administered questionnaires and health examinations. The participants were asked about their *smoking behaviour* in terms of smoking status and current level of smoking. These questions were combined into a four-category smoking variable: Never smoker; Former smoker; Smoker of 1-15 g/day; Smoker of >16 g/day. The questionnaire data on *physical activity* in leisure time from the seven cohort studies were harmonised into a four-category variable about the participants' usual physical activity per week within the past year: Sedentary (less than 2 hours of light activity e.g. housekeeping, walking, bicycling); Light physical activity (about 2-4 hours of light activity); Moderate physical activity (more than 4 hours of light activity or 2-4 hours of high-level activity e.g. heavy gardening, running, swimming); High physical activity (more than 4 hours of high-level activity). Participants were also asked about their *weekly alcohol intake* with separate items about beer, wine, and spirits. The level of alcohol intake was calculated as average gram alcohol per day, and harmonised into a six-category alcohol variable (<1; 1-7; 8-14; 15-21; 22-28; ≥29 drinks per week). The participants also had health examinations with anthropometric measurements collected by trained nurses, and *body mass index* (BMI) was calculated as weight divided by height squared and harmonised into a four-category variable: Underweight (BMI of 18.5 kg/m² or less); Average weight (BMI of 18.6-24.9 kg/m²); Overweight (BMI of 25.0-29.9 kg/m²); Obese (BMI of 30 kg/m² or more). Blood pressure was measured according to the WHO guidelines. A four-category variable for *systolic blood pressure* was constructed (<120; 120-139; 140-159; >160 mmHg). *Hypertension* was defined as systolic blood pressure ≥140 mm Hg.

Table 1

Summary of studies on behavioural risk factors as potential mediators of the relationship between education and cardiovascular disease and cause-specific mortality

First author Country Year	Study design Study population Follow-up	SES indicators	Behavioural risk factors	Outcome	Confounders	Statistical analysis	Results
Albert (49) US 2006	Randomized placebo controlled trial 22688 female health professionals aged 45+ 10 years follow-up	Education income	BMI, smoking, hypertension, diabetes, LDL and HDL cholesterol, triglycerides, hor- mone use, family history of MI, alcohol, physical activity	CVD incidence	Age, race	Cox propor- tional hazard model	HR of CVD associated with the highest vs. lowest level of educa- tion changed from 0.5 (0.3-0.7) to 0.8 (0.5-1.2) when adjusting for all the included behavioural risk factors. Smoking and waist circumference were the strong- est mediators.
Beauchamp (18) Australia 2009	Cohort study 41514 men and women aged 27-80 9.4 years follow-up	Education	Smoking, fruit and vegetables, physical activity, alcohol, saturated fat intake	CVD mortality	Age, sex, country of birth	Cox propor- tional hazard model	HR of CVD mortality associated with the lowest vs. highest level of education changed from 1.7 (1.1-2.5) to 1.2 (0.8-1.8) when adjusting for all the included behavioural risk factors. Smoking were the strongest mediators.
Ernstsen (50) Norway 2010	Cohort study 44128 men and women aged 30+ 9.1 years follow-up	Education	Smoking, physical activity, alcohol	CHD mortality	Age, any limiting long- standing illness	Cox propor- tional hazard	Health behaviour accounted for 25% of the relative difference between primary and tertiary education level in CHD mortality among women and 53 % in men.
Harald (54) Finland 2006	Cohort study 19272 men and women aged 35-64 5 years follow-up	Occupation education income	Smoking, alcohol, physical activity, cholesterol, blood pressure, BMI,	CHD incidence	Age, baseline year, area	Cox propor- tional hazard model	No clear differences between educational or income groups were observed. Traditional CVD risk factors explained 31% of the relative difference between male manual workers and upper-level employees in CHD. Smoking was the strongest mediator.
Kerr (55) UK 2010	Meta-analysis (12 cohort or case- control studies	Socio-economic status: regardless of choice of indicator	At least one vascular risk factor (blood pressure, smoking, diabetes, lipids, atrial fibrillation history of vascular disease, BMI, physical activity)	Stroke incidence		Random effects model	HR of stroke associated with lower vs. higher SES changed from 1.7 (1.5-1.9) to 1.3 (1.2-1.5) when adjusting classic vascular risk factor. Smoking was the strongest mediator, followed by obesity, physical activity, and hypertension.
Kershaw (19) Netherlands 2013	Cohort study 15067 men and women aged 20-65 11.5 years follow-up	Education	Smoking, alcohol, diet, physical activity, BMI, hypertension, diabetes, hypercho- lesterolemia	CHD incidence	Age, sex, marital status	Path- analysis	The examined risk factors ac- counted for 57% of the relative difference in low vs. high educa- tion in CHD. The proportions mediated were for smoking 27%, obesity 10%, and physical inactiv- ity 6%.
Khang (56) South Korea 2009	Cohort study 8366 men and women aged 30+ 7 years follow-up	Education occupation	Smoking, alcohol, physical activity	All-cause mortality	Survey year, sex, age	Cox propor- tional hazard model and Binomial model	Health behaviour accounted for 15% of the relative difference between the lowest vs. highest educational level in all-cause mortality. However the absolute explanatory power reached 84%.
Kilander (51) Sweden 2001	Cohort study 2301 men aged 50 25 years follow-up	Education	BMI, smoking, blood pressure, choles- terol, triglycerides, glucose, physical activity, body height, serum fatty and antioxidants	CVD and cancer mortality	Age	Cox propor- tional hazard model	When comparing the lower to higher educational groups before and after adjustment for all examined risk factors, the HR for CVD mortality changed from 1.7 (1.2-2.4) to 1.0 (0.7-1.5) and for Cancer mortality from 1.9 (1.2- 3.1) to 1.5 (0.9-2.5).

Table 1 (continued)

First author Country Year	Study design Study population Follow-up	SES indicators	Behavioural risk factors	Outcome	Confounders	Statistical analysis	Results
Laaksonen (20) Finland 2007	Cohort study 60000 men and women aged 25-64 11.9 years follow-up	Education	Smoking, alcohol, physical activity, vegetable use, fat on bread, coffee drinking, weight	CVD, CHD, and stroke mortality	Age, study year, pre-existing chronic disease	Cox propor- tional hazard model	Health behaviours accounted 54% of the relative difference between primary and higher education in CVD mortality among in men and 22% in women. Smoking, vegetable use and physical activity were the strongest mediators.
Lantz (53) US 2010	Cohort study 3617 men and women 19 years follow-up	Education, income	Smoking, alcohol, physical activity, BMI	All-cause mortality	Age, sex, race, urban city	Cox propor- tional hazard model	HR of mortality associated with lower vs. higher educational level changed from 1.4 (1.1-1.9) to 1.2 (0.9-1.7) when adjusting behav- ioural risk factor. Further, educa- tion indirectly influenced mortal- ity through its strong association with income.
Lynch (57) Finland 2006	Cohort study 2682 men aged 42, 48, 54, 60 10.5 years follow-up	Education	Smoking, choles- terol, blood pres- sure, diabetes	CHD incidence	Age	Cox propor- tional hazard model, calculation of the PAR	Classic vascular risk factors reduced relative social inequality by 24%. In a low risk population free from classic vascular risk factors absolute social inequality reduced by 72%.
McFadden (61) UK 2008	Cohort study 22486 men and women aged 39-79 10 years follow-up	Occupation education	Smoking and BMI	All-cause, CVD, and cancer mortality	Age	Cox propor- tional hazard model	When comparing men with high vs. low education before and after adjustment for smoking and BMI, the HR for CVD mortality changed from 0.6 (0.4-0.9) to 0.7 (0.5-1.0) and for Cancer mortality from 0.7 (0.5-0.9) to 0.8 (0.6-1.1). No association was found in women. Further, education was not associated with mortality independently of occupation.
Nandi (59) US 2014	Cohort study 8037 men and women aged 51-61 10 years follow-up	SES measure based on education occupation and income	Smoking, alcohol, BMI, physical activity	All-cause Mortality	Age, sex, health status	Invers probability- weighted mediation models on risk ratio and risk differ- ence scale	Health behaviours reduced relative social inequality by 68% and the absolute social inequal- ity by 51%.
Strand (21) Norway 2004	Cohort study 66200 men and women aged 35-49 23.6 years follow-up	Education	Smoking, physical activity, BMI, choles- terol, blood pressure	CHD and CVD mortality	Age	Cox propor- tional hazard model	HR of CHD mortality associated with low vs. high educational group changed in men from 1.33(1.18 to 1.50) to 1.03 (0.91 to 1.17) when adjusting for all the included behavioural risk factors. For women the HR change from 1.72 (1.23 to 2.41) to 1.24 (0.88 to 1.75). Smoking, cholesterol, blood pressure were the strong- est mediators.
Stringhini (62) UK and France 2011	Cohort studies 9771 (UK) and 17760 (France) participants age 30-55 17-20 years follow- up	Education occupation income	Smoking, alcohol, diet, physical activity	All-cause Mortality	Age, sex	Cox propor- tional hazard model	Health behaviours accounted for 56% in the UK cohort and 17% in the French cohort of the relative difference in mortality between low vs. high education.
Wamala (52) Sweden 1998	Case-control study 292 cases and 292 age-matched controls women aged <65	Education	Smoking, physical activity BMI, total cholesterol, blood pressure, haemo- static factors	CHD incidence	Age	Logistic regression	Health behaviour accounted for 48% of the relative difference between low vs. high educational attainment in CHD.

CVD= cardiovascular disease, CHD= Coronary Heart Disease, HR= Hazard Ratio, PAR=population attributable risk

Table 2

Characteristics of participants from the seven cohorts pooled into the Social Inequality in Cancer (SIC) cohort study

	CCHS	1936-COHORT	MONICA I	MONICA II	MONICA III	DCH	INTER99	Pooled SIC
Study entry, years	1981-1983	1981-1982	1982-1984	1986-1987	1991-1992	1993-1997	1999-2001	1981-2001
Sample size, No.	12,693	991	3780	1417	2024	55,806	6295	83,006
Response rate, %	61	81	79	75	73	36	53	N.A.
Age at study entry, Mean (range)	56 (20-98)	45 (45-45)	45 (30-60)	45 (30-60)	50 (30-70)	56 (50-65)	45 (30-60)	54 (20-98)
Women, %	55	53	49	50	50	52	51	52
Death from all-causes, %	59	25	31	22	25	12	3	20
Person-years at risk ^a	154,910	24,281	83,528	28,094	29,928	642,056	56,417	1,019,215
Low education, %	29	36	36	36	34	28	24	28
Current smokers, %	58	49	56	48	46	36	36	41
Obese, %	13	8	10	9	13	15	18	14
Low physical activity, %	17	32	28	29	25	16	22	18
High alcohol intake, %	13	19	14	31	12	24	15	20
Systolic blood pressure, mmHg, Median (5%-95%)	138 (109-180)	120 (102-143)	122 (101-155)	119 (96-153)	123 (100-161)	138 (109-176)	129 (105-161)	136 (107-175)

CCHS=Copenhagen City Heart Study. MONICA (I, II, III)=Multinational MONitoring of trends and determinants in CARDiovascular disease. DCH=Diet Cancer and Health study. INTER99=Randomized non-pharmacological INTERvention study for prevention of ischaemic heart disease. SIC=Social Inequality in Cancer cohort study.

^aPerson-years at risk from date of study entry until date of emigration, death from all-causes or end of follow-up (31st December 2009)

Assessing education and outcomes by linkage to register data

Since 1968 the Central Population Registry has provided every citizen in Denmark with a unique number for personal identification. This identification number is the key linking participant data in the SIC cohort to data from national registries.

Educational attainment

The Integrated Database for Labour Market Research provided information on individual highest attained education. The level of education was measured one year before study entry as a three-category variable. 'Low education' was defined as primary (grade 1 to 6) and lower secondary (grade 7-9/10) education. 'Medium education' was defined as upper secondary, vocational or technical education as well as short-cycle higher non-university pro-

grammes (~11-14 years of education). 'High education' was

defined as medium-cycle university or non-university programmes as well as long-cycle university programmes (≥15 years of education).

Follow-up and outcomes

Participants were followed by linkage to the Danish National Registry of Patients providing information about discharge diagnoses from hospital admissions and causes of deaths from the Causes of Death Registry using the International Classification of Diseases (ICD), 8th Revision from 1981 to 1994 and the 10th version thereafter. The research in this thesis focused on the following outcomes:

- Incident cases of coronary heart disease (ICD8: 410-4; ICD10: I20-I25) in paper 2
- Incident cases of ischemic stroke (ICD8: 433 to 434; ICD10:I63), haemorrhagic stroke (ICD8:430 to 431; ICD10:I60 to I61), and unspecified stroke (ICD8: 436; ICD10: I64) in paper 3
- Mortality defined as all-cause, cardiovascular (ICD8: 390–458; ICD10: I00–I99, G45), cancer (ICD8: 140–209; ICD10: C00–C97), and respiratory (ICD8: 460–519; ICD10: J00–J99) mortality in paper 4

Follow-up was assigned from the date of study entry until date of diagnose (as specified above), death, emigration, or to the end of follow-up (31th December 2009). Fewer than 0.1% was lost to follow-up due to emigration. The mean follow-up time was approximately 14 years.

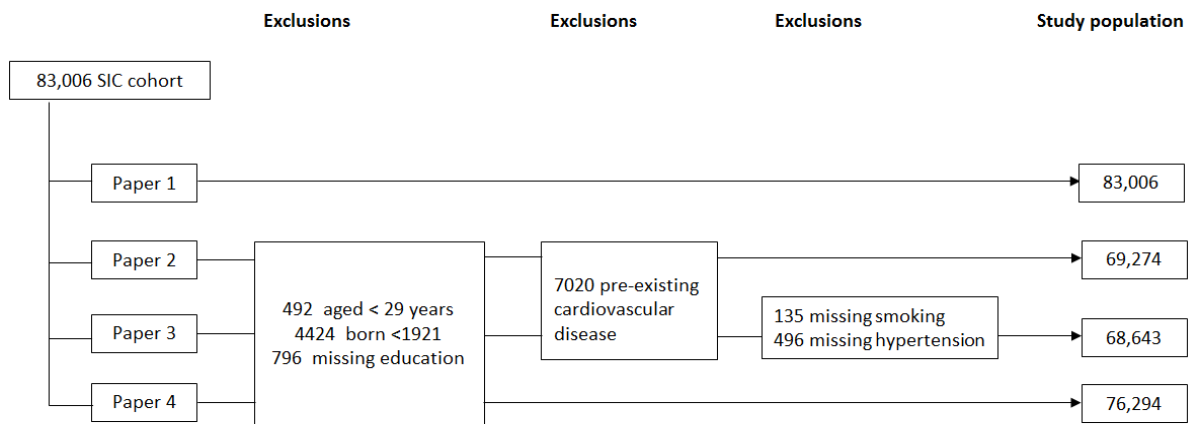


Figure 2
Flow chart of exclusions and study populations in Paper 1 to 4

Study population

The study population was defined differently in the four studies depending on the specific analyses. Figure 2 gives an overview of exclusions and study population in each of the studies. Paper 1 presented the complete SIC cohort with a total of 83,006 men and women aged 20 to 93 years. In Paper 2, 3, and 4 several exclusions were made. To obtain a robust measure of highest attained education participants aged 29 or below were excluded assuming that some of these participants had not yet reached their final level of education. Participants born before January 1921 were also excluded due to invalid information on education. Additionally, participants with missing information on education for unspecified reasons were excluded. Based on these exclusions the study population comprised 76,294 participants born between 1921 through 1970. In addition, all participants with pre-existing (prior to study entry) cardiovascular disease were excluded in Paper 2 and 3 leaving 69,274 participants for the analysis. Please note that for Paper 2 the size of the study population displayed in Figure 2 are somewhat different from the one reported in the published article. A detailed clarification of this inconsistency can be found in the published erratum of Paper 2.

Graphical presentation

Directed acyclic graphs have been used in this thesis as a tool for summarizing assumptions of direct and indirect effect and to identify potential confounding (67;68). As an example, the graph presented in Figure 3 represents the assumed relationship between education and coronary heart disease (CHD) based on a comprehensive review of the existing literature on variables affecting this relationship. The graph illustrates that education could be associated with CHD through other pathway than health behaviour, for example through psychosocial factors such as stress, life events, lack of social support (69-71). Further, it shows that in addition to the major confounders (age, sex and cohort) of this relationship, other variable such as early life circumstances and residential area should be included to adequately control for confounding.

Statistical methods

In recent years important work on developing methods for mediation and interaction analysis for survival outcomes have been initiated by Lange et al. (72;73) and Rod et al. (74). These contemporary methods used, the additive hazards model, in a counterfactual framework (75;76), which allowed for definitions of direct and indirect effect, even in settings with interactions (77). Applied to an example from Paper 4, this model yielded an estimate of the absolute change in the mortality rate when comparing low to high education (as the reference group). The estimate of this *total effect* was interpreted as the number of extra deaths per 100,000 person-years lived at risk in the low compared to high educational group. The total effect was separated into natural direct and indirect effects.

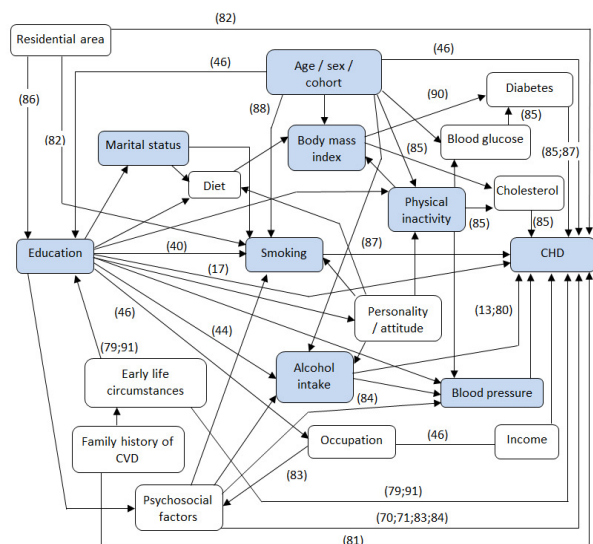


Figure 3
Graphical presentation of the assumed relationship between education and coronary heart disease (CHD) including both measured (blue shading) and unmeasured variables. Round brackets indicate references to existing literature on variables affecting this relationship.

The natural *direct effect* compared the average mortality rate in the low versus the high educational group when, in both groups, the level of smoking (as the potential mediator) was what it would have been in the high educational group. The estimate of this direct effect was interpreted as the number of extra deaths per 100,000 person-years lived at risk in the low compared to high educational group attributed to the direct path or to other mediators.

The natural *indirect effect* compared the average mortality rate in the low educational group, when the level of smoking was what it would have been in the low versus high educational group. The estimate of the indirect effect was interpreted as the number of extra deaths per 100,000 person-years lived at risk in the low compared to high educational group attributed to mediation through smoking.

In the additive hazards model the magnitude of interaction (defined as deviation from additivity of absolute effects (22)) between education and a behavioural risk factor was directly assessed by including a product term of these variables. Again, applied to the example from Paper 4, the interaction would be nonzero when the average mortality rate of being in the low educational group and the level of smoking was what it would have been in the low educational group (i.e., both the exposure and the mediator were present) differed from the sum of the average mortality rates of having only the exposure or the mediator present.

The additive hazards model was applied for slightly different purposes in this thesis:

In Paper 2, the additive hazards model was applied to quantify the extent to which the association of education (exposure) with coronary heart disease (outcome) was mediated by smoking, physical inactivity, and body mass index after accounting for measured confounders (i.e., sex, age and cohort). The first step was to fit a linear regression model to each mediator (measured as binary or continuous variables) conditioning on education and confounders (age and cohort). Separate analyses were made for men and women. The second step was to fit an additive hazards model using age as the underlying time scale to onset of coronary heart disease conditioning on education, each of the mediators and confounders. The indirect effect through either smoking, physical inactivity or body mass index was given by the product of the parameter estimates for education on the mediator (from the linear regression in step one) and the parameter estimate for the mediator on coronary heart disease (from the additive hazards model in step two). The direct effect of exposure was given directly from the additive hazards model. Finally, the total effect was found as the sum of the direct and indirect effects. For the direct effect, 95% confidence limits were readily available from the additive hazards model, while limits for the indirect and total effects were computed by bootstrap (using 100,000 replications).

In Paper 3, the additive hazards model was applied in order to estimate the magnitude of interaction and combined effects of the three exposures; low education, smoking, and hypertension on onset of stroke (outcome) after accounting for measured confounders (i.e., sex, age and cohort). The additive interaction was directly assessed by fitting an additive hazards models using age as the underlying time scale to onset of stroke conditioning on education, smoking, hypertension and confounders, and in addition including a product term of the exposures.

In Paper 4, the additive hazards model was applied to quantify the extent to which the association of education (exposure) with mortality (outcome) was mediated by smoking, alcohol intake, physical inactivity and body mass index and simultaneously estimate the magnitude of interaction between education and the mediators, after accounting for measured confounders (i.e., sex, age and cohort). Initially a multinomial logistic-regression model was fitted to each mediator (measured as categorical variables), conditioning on education and baseline confounders. Separate analyses were made for men and women. Then a new data set was constructed where a new variable (education*) corresponds to the value of education relative to the indirect path. Then weights were computed by applying fitted multinomial logistic-regression models using education* and education respectively. Finally, a three-way decomposition of total effect into three components as suggested by VanderWeele (78): 1) the average pure direct effect, 2) the average pure indirect effect, and 3) the product of an additive interaction between the exposure and the mediator on the outcome, and the average effect of the exposure on the mediator – referred to as the average mediated interaction was obtained by fitting an additive hazards model to cause-specific mortality including only education and education* (and their product term) as covariates weighted by the weights from step two and confounders. For the direct effect, 95% confidence limits were readily available from the additive hazards model, while limits for the mediated interaction, the indirect effect and the total effect were computed by bootstrap (using 10,000 replications).

Table 2**Educational difference in health behaviours among 76,294 men and women enrolled in the Social Inequality in Cancer cohort study**

	MEN			WOMEN			
	Total	Educational level			Educational level		
	N=76294	Low (n=8954)	Medium (n=19239)	High (n=8195)	Low (n=14552)	Medium (n=18197)	High (n=7157)
Tobacco smoking							
Never smoker	25149	18	25	32	34	44	46
Former smoker	20314	28	31	36	20	22	27
Smoker of 1-15 g/day	15887	20	17	15	29	23	19
Heavy smoker of ≥ 16 g/day	14784	35	26	17	17	12	9
No. missing	160						
Alcohol intake							
<1 drinks/week	4917	6	3	2	14	6	5
1-7 drinks/week	32887	33	30	26	58	56	51
8-14 drinks/week	17333	23	27	28	17	21	23
15-21 drinks/week	7834	12	12	14	6	9	12
22-28 drinks/week	6070	9	12	15	3	5	6
≥ 29 drinks/week	6873	17	16	15	2	3	3
No. missing	380						
Leisure-time physical activity							
Sedentary	13268	18	15	16	21	18	16
Light activity	30689	38	37	34	44	45	43
Moderate activity	25460	35	36	38	30	30	34
High activity	6676	9	11	12	6	7	8
No. Missing	201						
Body mass index							
≤ 18.5 kg/m ² - Underweight	801	<1	<1	<1	2	2	2
18.6-24.9 kg/m ² - Normal weight	34244	32	36	45	47	54	62
25.0-29.9 kg/m ² - Overweight	30304	48	49	46	34	33	28
≥ 30 kg/m ² - Obese	10833	20	15	9	17	12	9
No. missing	112						
Systolic Blood pressure							
<120 mmHg	15180	14	14	17	24	26	32
120-140 mmHg	31319	42	42	45	39	39	41
141-160 mmHg	20334	30	30	28	25	24	19
>160 mmHg	8924	14	13	10	12	11	8
No. missing	537						

RESULTS

Educational differences in health behaviours

Clear educational gradients in the prevalence of health behaviours, particularly smoking and body mass index, were observed among the 76,294 participants age 30 to 70 years, as presented in Table 3. Among men, the prevalence of heavy smoking in the three educational levels varied from 17% in the high, 26% in the medium to 35% in the low educational level. Correspondingly, the prevalence of obesity varied from 9% in the high, 15% in the medium to 20% in the low. Similar patterns were observed among women, although the prevalence and variation were less pronounced. In addition, those with low education had slightly less favourable levels of physical activity and blood pressure, whereas those with high education had the highest levels of alcohol intake.

Educational differences in chronic disease outcomes

During the follow-up time a total of 12,340 participants died: 5310 from cancer; 3182 from cardiovascular disease; 975 from respiratory disease; 2873 from other causes. In addition, among participants free of pre-existing cardiovascular disease there were 7,461 incident cases of coronary heart disease and 4,389 incident cases of stroke.

Clear educational differences were observed for all the chronic disease outcomes reported in this thesis, as summarised by the diagram in Figure 4. In general the differences when comparing low to high level of education were more pronounced in men than women. As expected, given the higher risks of deaths in older age, educational differences were more pronounced for deaths per 100,000 person-years lived at risk over age 65 than under age 65. In men, the largest educational difference in mortality was observed in death from cardiovascular diseases. In women, the largest educational difference in mortality was observed in death from cancer. In both men and women, educa-

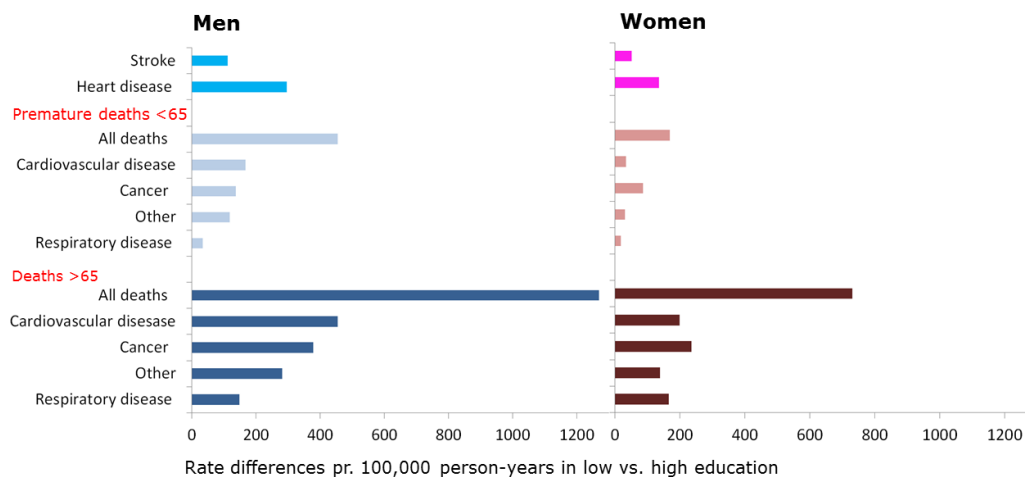


Figure 4

Educational differences in the chronic disease outcomes per 100,000 person-years for men and women comparing low to high level of education, age and cohort adjusted

tional differences were substantially larger for incidence of coronary heart disease than for stroke.

Education and coronary heart disease: Mediation by behavioural risk factors

The mediating role of smoking, body mass index, and physical inactivity on the association between education and incidence of coronary heart disease was investigated in 69,274 men and women free of pre-existing cardiovascular disease. Low compared to high level of education was associated with 296 (95% confidence interval= 209 to 383) extra cases per 100,000 person-years at risk of coronary heart disease in men, and 135 (80 to 190) in women. Body mass index (five-unit increment) was the strongest mediator, with 71 (57 to 85) cases in men and 25 (17 to 33) cases in women ascribed to this pathway. Further, 27 (19 to 35) cases in men and 17 (12 to 22) cases in women could be ascribed to the pathway through smoking (defined as being current smoker or ex-smoker). The effect of physical inactivity (defined as less than 4 hours of light activity per week) was negligible. The observed mediated effects derived from the additive hazards model were moderately stronger than the mediated effects derived from the cox proportional hazards models. Taking obesity as an example, using the additive hazards model, the proportion mediated was 24 % in men and 19 % in women; the corresponding estimates from the cox proportional hazards model were 7% and 5 %, respectively.

Education and stroke: Interaction with behavioural risk factors

Interaction between education, smoking and hypertension in relation to incidence of ischemic and haemorrhagic stroke was examined in 68,634 participants from the SIC cohort, who were free of pre-existing cardiovascular disease and had full information on both smoking and hypertension. The combined effect of low education and current smoking was more than expected by the sum of their separate effects on ischemic stroke but negligible for hemorrhagic stroke. This was most pronounced among men,

where the separate effect of having low education was associated with 42 (-5, 90) extra cases per 100,000 person-years at risk of ischemic stroke and the separate effect of being current smoker was associated with 112 (38, 186) extra cases. However, the combination of current smoking and low education was associated with 289 (238, 340) extra cases per 100,000 person-years at risk of ischemic stroke. Thus, 134 (49, 219) extra cases per 100 000 person-years at risk of ischemic stroke could be ascribed to the interaction between smoking and education. We could not confirm evidence of similar patterns among women. There was no clear evidence of interaction with respect to the combination of low education and hypertension on risk of ischemic stroke (men, $P=0.89$ women, $P=0.05$) or hemorrhagic stroke ($P=0.53$). However, the combined effect of current smoking and hypertension was more than expected by the sum of their separate effects on ischemic and hemorrhagic stroke. This was most pronounced among women, where 178 (103, 253) extra cases per 100,000 person-years at risk of ischemic stroke could be ascribed to the interaction between smoking and hypertension.

Education and mortality: Mediation and interaction by behavioural risk factors

The association between education and cause-specific mortality and the mediating role of smoking, body mass index, physical activity, and alcohol intake was analysed in 76,294 men and women from the SIC cohort. A three-way effect decomposition of the total effect into a direct effect, an indirect effect, and a mediated interaction was used to simultaneously regard the behavioural risk factors as intermediates and clarify the role of their interaction with education. Smoking (defined in a 4-category variable) was the strongest mediator on the educational differences in death from cardiovascular disease, cancer, and respiratory disease. The mediated effect of smoking was most pronounced for cancer deaths among men, were 160 (95% confidence interval= 110 to 209) extra deaths per 100,000 person-years lived at risk over age 65 in the low compared to high

educational group could be attributed to the pure indirect effect through smoking. In addition, 69 (6 to 132) deaths could be attributed to the mediated interaction between education and smoking. When combining the pure indirect effect and the mediated interaction the overall proportion mediated through smoking was 42% + 18% = 60%. The mediated effect through body mass index (defined in a 4-category variable) was stronger among men than women, and varied for the different causes of death. Among men, comparing low to high level of education gave a total effect of 461 (344 to 579) extra deaths from cardiovascular disease per 100,000 person-years lived at risk after age 65, and the proportion mediated through body mass index was 15% (6% to 26%). An interaction between education and body mass index was observed for cancer mortality. In this case, the proportion mediated through body mass index require a combination of the pure indirect effect (16% [4% to 33%]) and the mediated interaction (-31% [-56% to -15%]). The mediating effects through physical activity (defined in a 4-category variable) and alcohol intake (defined in a 6-category variable) were negligible.

Synthesis of findings

To enhance comparability and provide a clearer overview of the finding in the summarised papers, the mediated effect of smoking on the association between education and incidence of coronary heart disease and stroke have been re-analysed and reported in accordance with the approach used in Paper 4. That was; using a three-way effect decomposition, a 4-category smoking variable, a 3-category education variable, age as underlying time scale, adjusting for cohort, and separate analysis for men and women. The diagram in Figure 5 presents these new results for coronary heart disease and stroke as well as selected results for cause-specific mortality reported in Paper 4. It is shown how a substantial proportion of the total effect of education on chronic disease outcomes were attributable to differential exposure to smoking (indicated by the indirect effect of education mediated through smoking). Additionally, in many of these cases a considerable part of the mediated effect was also attributable to differential vulnerability (indicated by the mediated education-by-smoking interaction).

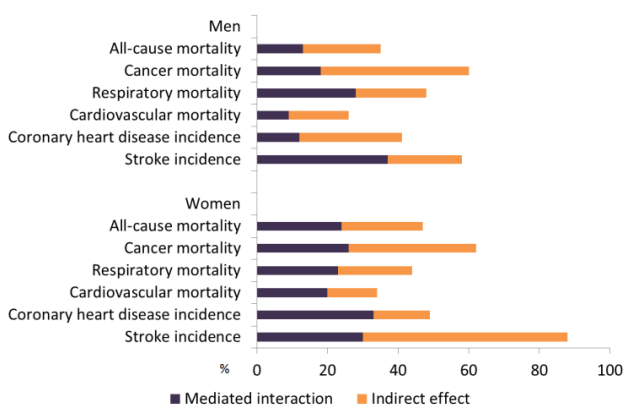


Figure 5
Proportion mediated by smoking on the relation between education

(comparing low to high) and chronic disease outcomes among men and women, age and cohort adjusted

Smoking had the strongest mediating effect on incidence of Stroke, with an overall proportion mediated of 88% (58% + 30%) among women. For incidence of coronary heart disease the overall proportion mediated through smoking was 41% (29% + 12%) among men and 49% (16% + 33%) among women. These re-analysed results for coronary heart disease show a stronger mediating effect of the 4-category smoking variable than of the dichotomized smoking variable reported in Paper 2.

Auxiliary analyses

For the purpose of this thesis auxiliary analyses were performed in order to exemplify the variations in study-specific effects of education and smoking on death from all-causes after age 65. As shown in Figure 6, comparing low to high level of education among women, the estimate from the Inter99 study indicating the opposite educational difference in all-cause mortality than the pooled estimate. However, the confidence intervals were wide, and there were no clear evidence of interaction between education and cohort study ($P > 0.05$). Further, the pooled estimate was not affected by potential calendar effects when stratifying by birth cohort (5-year intervals). Similar pattern was observed among men (data not shown)

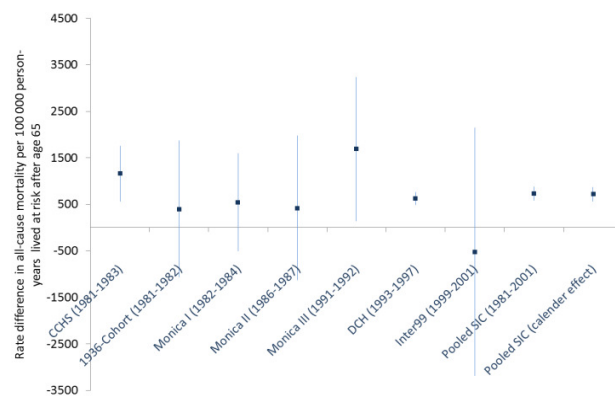


Figure 6
Study-specific effects of education on death from all-causes after age 65 among women, comparing low to high level of education, age-adjusted

As shown in Figure 7, comparing current to never smokers among women, the estimates from the earlier cohorts (Copenhagen City Heart study, 1936-cohort, and the Monica studies I, II, III) were substantially higher than the pooled estimate. Additionally, an interaction between smoking and cohort study was observed ($P < 0.05$). However, the pooled estimate was not affected by potential calendar effects when stratifying by birth cohort (5-year intervals). Similar pattern was observed among men (data not shown).

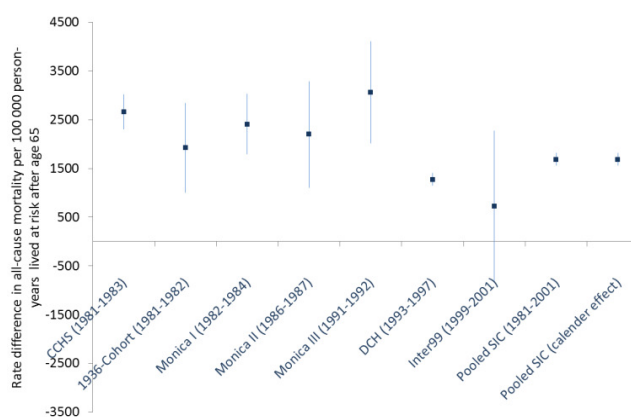


Figure 7
Study-specific effects of smoking on death from all-causes after age 65 among women, comparing current to never smokers, age-adjusted

Table 4 presents the mediating effect of smoking on the association between education and all-cause mortality after age 65 in the earlier and later cohorts. When comparing low to high level of education, the overall proportion mediated through smoking was 31% (23% + 8%) among men from the earlier cohorts and 37% (22% + 15%) among men from the later cohorts. Correspondingly, among women the proportion mediated was 51% (33% + 18%) and 40% (22% + 18%).

DISCUSSION

Discussion of main findings

The results in this thesis can be seen as an attempt to quantify two central mechanisms for understanding social inequality in chronic disease outcomes; differential exposures and differential vulnerability to behavioural risk factors. In paper 2, smoking and body mass index partially mediated the observed educational

differences in incidence of coronary heart disease. This result indicated that some of the social inequality in coronary heart disease might have been enhanced by differential exposure to behavioural risk factors (i.e. smoking and obesity). The concept of differential vulnerability was quantified, in statistical terms, analogous to assessing additive interaction (i.e., deviation from additivity of absolute effect). Thus, the interaction between education and smoking on incidence of ischemic stroke observed in Paper 3 indicated that participants, particularly men, with low level of education were more vulnerable to the effect of smoking than those with high level of education. Finally, Paper 4 revealed that behavioural risk factors, primarily smoking, explained a considerable part of the educational differences in cause-specific mortality. In particular, this paper added important knowledge about the substantial part of the mediated effect, which was due to interaction between education and smoking.

Previous studies investigating the extent to which health behaviour contribute to social inequality in chronic disease, particularly cardiovascular disease and mortality, have found substantively different results across various contexts. A very illustrative example on this matter was recently given in a study by Stringhini and colleagues (62). They found that health behaviours were strong predictors of mortality in two European cohorts, the British Whitehall II and the French GAZEL study, but the social characterisation of these behaviours was considerably different. Thus, the health behaviours were unequally important mediators of the social inequality in mortality in these countries. In the study population of this thesis the social disparities were substantially more pronounced for smoking and obesity than for physical inactivity and high alcohol intake, which explain the greater contribution of those behavioural risk factors to the social inequality in incidence

Table 3

Rate difference in extra deaths per 100,000 person-years lived over age 65 years by educational level (decomposition of total effects into direct, indirect, and mediated interaction effects of smoking) among men and women from the early and later cohorts enrolled in the SIC cohort.

All-cause mortality after age 65 ^a	Earlier cohorts ^b (1981-1992)		Later cohorts ^c (1993-2001)	
	Low vs. High education	Proportion mediated by smoking	Low vs. High education	Proportion mediated by smoking
	RD (95% CI)	% (95% CI)	RD (95% CI)	% (95% CI)
Men				
Total effect	1780 (1202-2359)		1103 (888-1318)	
Direct effect	1212 (634-1791)		685 (483-888)	
Indirect effect	417 (233-602)	23 (12-41)	247 (175-319)	22 (15-32)
Mediated Interaction	151 (-75-377)	8 (-5-22)	170 (68-273)	15 (6-24)
Women				
Total effect	1040 (590-1491)		621 (481-761)	
Direct effect	507 (56-958)		374 (241-507)	
Indirect effect	341 (159-525)	33 (13-71)	136 (83-190)	22 (12-35)
Mediated interaction	191 (0-384)	18 (0-42)	111 (47-174)	18 (8-27)

RD= rate difference per 100,000 person-years at risk lived after age 65

CI= confidence interval

^aAll analyses are adjusted for age and cohort study

^bEarlier cohorts: Copenhagen City Heart study, 1936-cohort and Monica (I, II, III)

^cLater cohorts: Diet Cancer and Health study and Inter99

of coronary heart disease and cause-specific mortality. This finding was similar to what Stringhini and colleagues found for the British Whitehall II cohort, whereas for the French GAZEL cohort none of the behaviours were considerable socially skewed and therefore contributed less to social inequality in mortality.

The mixed results in previous studies may also reflect diverse methodological approaches. For example, some studies (18;21;50;52;57) have used education as a unique socioeconomic indicator, whereas other studies (53;54;61) have examined the effects of education while simultaneously adjusting for occupation and/or income. In addition, composite indicators, which capture several aspects of socioeconomic position, have been investigated (59). Furthermore, multiple versions and levels of the different health behaviours have been applied, and the strategies for including these as potential mediators have been inconsistent. Some studies (18;20;21;49;51;54;61) have investigated the effect of each mediator one at a time, while others have only reported the simultaneous effect of the included mediators (50;52;56;57). For studies with an analogous approach to the one used in the thesis (i.e., investigating education as a unique socioeconomic indicator and the including each mediator one at a time) the conclusions were also somewhat similar to the results presented here; behavioural risk factors explained some but not all the observed educational differences in incidence of coronary heart disease (19), death from all-causes (62), cardiovascular disease (18;20;21) and cancer (51), and smoking was the strongest mediator.

The previous research in this field relied on the conventional regression approach (sometimes referred to as the change-in-estimate method²) of adding a mediator as a covariate to the cox proportional hazards model and then calculation of how much the coefficient for education changed (92;93). Thus, the proportion of the educational difference eliminated by for example smoking was expressed on the excess relative risk scale, and therefore quantified relative inequality independent of the prevalence of the smoking. In this thesis, the use of the additive hazards model as a natural effects model proposed by Lange et al. (73) offered complementary information for the assessment of social inequality in chronic disease outcomes by estimating excess number of cases attributable to a given health behaviour. Further, the implementation of the three-way effect decomposition proposed by VanderWeele (78) to simultaneously quantify two central mechanisms underlying social inequality in health, differential exposure and differential vulnerability to health behaviour was in fact a new contribution to the existing literature. When analysing how much a set of mediators contribute to social inequality in chronic disease and mortality it would be reasonable to assume that public health decision makers might expect that epidemiologists quantify the unique share of the inequality that could be eliminated by addressing each of the examined behavioural risk factors. This question would receive a relatively simple answer in the absence of interactions between the exposure and the media-

tors, while a slightly more complex answer in the presence of interaction (94). Although, previous studies have compared their findings derived from a risk ratio model and risk difference model, none of them reported whether there were interactions between socioeconomic position and behavioural risk factors (56;57;59).

Methodological considerations

Pooling of existing cohort studies

The pooling of seven existing cohort studies generated a large study population with long follow-up. This enhanced statistical power to examine behavioural risk factors as potential mediators and their interaction with socioeconomic position on different chronic disease outcomes. However, some methodological challenges evolved from this procedure. The distributions of educational levels and behavioural risk factors varied across the different cohort studies pooled together (Table 2). Thus, the underlying assumption about homogeneity across study populations and follow-up periods were unlikely to be completely satisfied (68). A number of study characteristics such as study location, length of follow-up, year of study entry, and birth cohorts could be potential sources of the observed heterogeneity (95;96). Taken educational level as an example, since people usually attain their education in young adulthood, education is less likely to change over time and be influenced by lifestyle or disease (reverse causation), as opposed to other much more dynamic socioeconomic indicators such as income and wealth (89). However long-term comparison could be difficult by the fact that the educational system and actual composition of educational groups do change over the years. In relation to the SIC cohort, it seems feasible to expect that having lower secondary education would be a relatively high level of education, especially for women in the oldest birth cohort – while for the women in the younger birth cohorts, this would be considered as a low level of education. Furthermore, it was shown from the auxiliary analyses in this thesis that the effect of education on all-cause mortality varied across cohort studies (Figure 6). A variety of statistical modelling techniques have been proposed to account for such existing heterogeneity (97). Recently, approaches which combine all individual participant data in a single meta-analysis and incorporate random effects to allow for heterogeneity have been suggested (98). However, such a random effect was not directly applicable to the techniques presented in this thesis. Instead, thorough sensitivity analyses of study-specific effects have been performed in Paper 2, 3, and 4, for example by including exposure-by-cohort study interaction, comparing pooled estimates by systematically removing each individual study one at a time, reducing length of follow-up time, as well as adjusting for cohort study and birth cohorts (calendar effect). The conclusions were robust to these procedures.

Harmonisation of survey data and misclassification

The retrospective harmonisation of the already collected data on health behaviours was established through a series of consensus meetings bringing together researchers with expertise in using data from the included cohort studies. Although, the various

cohort studies had used different self-administered questionnaires where phrasing of questions and time periods of responses varied, it was possible to maintain sufficient consistency in the synthesised data set.

The measurements of health behaviours were harmonised into continuous and categorical variables making it possible to model these potential mediators using dose information. However, in Paper 2 the mediated effect of smoking was estimated using a dichotomised measure of smoking status defined as 'never smoker' vs. 'smoker or ex-smoker'. This was done because, at that time, the technique used to estimate the direct and indirect effect could not handle categorical variables. Consequently, the chosen cut off misclassified ex-smokers as smokers. As pointed out by Ogburn and VanderWeele (99); if a mediator is dichotomised, the natural indirect effect can be biased either upward or downward. However, in the presented context it seems that the use of a dichotomised smoking variable have underestimated the mediated effect of smoking, as the results from the reanalysis using a four-category smoking variable suggested a much stronger mediated effect of smoking (Figure 5).

Another concern was that smoking behaviour may differ across socioeconomic groups (e.g., varying proportions of heavy smokers, ex-smokers, etc.) and therefore the observed interaction between education and smoking in Paper 3 and 4 could have been a consequence of crude categorization of the smoking variable. However, the four-category smoking variable was constructed as a combination of past smoking status, as well as, the current level of smoking. Furthermore, the sensitivity analysis performed in Paper 3 and 4 showed that the observed interactions were robust to different cut-off point of heavy smoking.

Health behaviours were measured at a unique time point. However, it is likely that health behaviours may influence chronic disease outcomes in accumulative fashion over the life-course. For example, it has been documented that smoking has a dose-response and cumulative effect on mortality and also that risk wanes after quitting (100). In Denmark, the prevalence of daily smoker dropped from 44% in 1987 to 27% in 2005 (101). Due to the long follow-up in the SIC cohort, it seems feasible to expect that some of the participants, who were smokers at study entry, might have quit smoking along the way. Actually, the proportion of current smokers was higher among participants from the earlier cohorts than from the later cohorts (Table 2). Consequently, some of the participants, particularly those from the earlier cohorts, could have been misclassified as 'false-positive' smoker. Intuitively, it seems reasonable to expect that such measurement error weaken the association between smoking and mortality, for example, which would bias the mediated effects towards the null. However, the auxiliary analyses performed in this thesis, showed that the effect of smoking on all-cause mortality was stronger in the earlier than later cohorts (Figure 7). Further, the proportion mediated through smoking on the association between education

and all-cause mortality was slightly lower in the earlier than later cohorts among men, whereas the opposite pattern was found among women (Table 4). Thus, these analyses did not offer any insight of the direction of the potential bias. It could be assumed that the misclassification was differential with the exposure (i.e., education), because in previous research on trends in smoking in Western Europe, greater declines have been observed among higher educated, particularly women, compared with their less educated counterparts (88). In scenarios of such differential measurement error sensitivity analysis could offer interesting implication for the direction of bias of the mediated effect (102). Unfortunately, the available formulas were not directly applicable for situations with exposure and mediator interaction.

Quality of register data and potential misclassification

For decades, and hopefully also in the future, the legislation in Denmark and in most of the Nordic countries has allowed researchers to carry out studies linking data from various registries by means of the personal identification number. This provided a unique opportunity to link the data on health behaviour from the SIC cohort to high-quality data on the participants' level of education and chronic disease outcomes from nationwide registries.

Generally, the information on highest attained education was of high quality from 1981 and onwards, because it was based on annually collections on information about passed exams from all educational institutions (103). However, for those who had already attained their education before 1981, information was obtained from national census data of 1970. More than 90% of the participants in the SIC cohort were born before 1950, and these birth cohorts would have been above 30 years in 1981. Thus, they would be expected to have attained their education at that time. Although, participation in the 1970 census was compulsory, the data were reported in self-administered questionnaires filled out by the head of the household and information bias might have led to misclassification. Since a validation study has reported major irregularities in the educational data for people born in before 1921, these birth cohorts were discarded (104). Further, in order to enhance a robust measure of education people below 30 years at study entry were excluded, as they might not have attained their highest educational level.

The quality of the data on causes of death relied on the accuracy of the reporting and the coding practices. The reporting of causes of death on the death certificate might have changed within the study period (1981-2009). For example, within the last ten years, higher mortality rates from chronic obstructive pulmonary disease have been observed in many countries, which could be ascribed to enhanced diagnostics and focus on this condition (105). However, as indicated by Jensen et al. (106) misclassification of death from chronic obstructive pulmonary disease could still be a potential source of bias. The validity of the registration of causes of death could have been threatened by lower autopsy rates and stricter criteria for reporting e.g. coronary heart dis-

ease. However, the validity of diagnosis of coronary heart disease in the Danish National Registry of Patients and as a cause of death in the Causes of Death Registry have shown to be relatively high (107). Furthermore, the potential misclassification of the outcomes in this thesis was assumed to be non-differential, since the measurement errors were probably the same across educational groups.

Selection bias

The linkage to central and administrative registries gave access to nearly complete follow-up, with less than 0.1% lost to follow-up due to emigration. Thus, selection bias due to lost to follow-up was considered a minor problem. However, in Paper 4 selection bias could be expected, as level of education might have affected the probability of being selected into the study. To be recruited into the SIC cohort participants must be alive at study entry and therefore have survived on average 54 years (SD = 7) since birth. Because persons with low education tend to have higher mortality rates than those with high education (even at early ages), and because the analysis was restricted to participants who were still alive at study entry, it is likely that the educational differences in mortality were underestimated.

Confounding

To obtain valid estimates of direct effects, indirect effects and mediated interactions it had to be assumed that there were no unmeasured confounders for; 1) the exposure-outcome, 2) the exposure-mediator, and 3) the mediator-outcome relationship. In addition, it had to be assumed that no effect of the exposure confounded the mediator-outcome relation (no exposure-dependent confounders) (77). These were very restrictive assumptions and unlikely to have been completely satisfied in this set up.

Although, control was made for major confounders such as age, sex and cohort it seems unlikely that these were even close to being sufficient to avoid confounding. A recent American study by Nandi et al. (79) reported evidence of a direct effect of early life socioeconomic position on risk of heart disease and stroke in adult life through pathways other than adult socioeconomic position. In addition, another American study by Loucks et al. (108) showed that other early life risk factors such as childhood intelligence, parental mental illness, childhood chronic health conditions explained a considerable part of the observed association between education and risk of coronary heart disease. Thus, early life circumstances could be potential confounders of the relationships with education, as represented by the graph of Figure 3. It was, however, not possible to apply such life course perspective in this thesis as information on early life circumstances was not available in the SIC cohort. However, inter-sibling analysis in a recent Danish register-based study by Søndergaard et al. (109) showed that the inverse education-mortality relationships were only marginally attenuated after control for a range of potential sibling-specific confounders of serious health conditions

in early life. As was recently addressed by VanderWeele (110), in his commentary regarding Paper 2, a systematic and formal assessment of the potential impact of unmeasured confounders could be extremely valuable. Sensitivity analysis could offer insight of the likely direction in which the effects are potentially biased due to unmeasured confounding. Unfortunately, the presented techniques do not yet directly cover sensitivity analysis for situations with exposure and mediator interaction.

In Paper 4, the measure of pre-existing morbidity included having a diagnosis of cardiovascular disease, cancer or respiratory disease prior to study entry. Since education was related to disease incidence and subsequent death, it could also be related to pre-existing morbidity. Thus, this morbidity might operate as mediators of the association between education and mortality. Excluding participants with pre-existing morbidity would dilute the actual effect of education. However, pre-existing morbidity could contemporaneously be a potential confounder of the association between the smoking-mortality relationships, for example, as the participants with pre-existing morbidity might have undergone changes in health behaviour (e.g. quit smoking) prior to study entry. The presence of such a confounder generally undermines the identifiability of the natural indirect and direct effects (111). To address this issue, sensitivity analyses of all-cause mortality conditioning on exposure relative to the direct and the indirect path, confounders and in addition pre-existing morbidity was obtained (see eTable 14 in the Appendix of Paper 4). Controlling for pre-existing morbidity did not substantively change our findings.

Intertwined pathways

Paper 2 and 4 addressed a number of health behaviours as potential mediators. These mediators were investigated separately and thereby it was assumed that each mediator represented independent and non-intertwined causal pathways (112). This assumption was likely too strong. Each of the behavioural risk factors may lie on the pathway between education and outcome, and these factors may also be related to one another, either through direct pathways or through other common causes of the mediators. For example, people might have coherent attitude towards various health behaviours. This attitude could be reflected in a "healthy" or "unhealthy" lifestyle. In Paper 4, this issue was addressed, modelling the mediators concurrently. When comparing the summarised mediated effects of the separate mediator variables with the overall mediated effect of the combined mediator variable, it was shown that the difference, which corresponds to the potential amount of double counting due to intertwined causal pathways between the mediators, was marginal (see eTable 13 in the Appendix of Paper 4). These differences were by no means ignorable and indicated that the separate mediator effects may be slightly overestimated due to potential intertwined causal pathways. Thus, in light of the notion of "unhealthy" lifestyle proposed above a plausible explanation

would be that each mediator modelled separately also captures a share of the effect of the other omitted mediators.

Representativeness

All participants have been randomly selected from the general population, but as participation rates have decreased over the years the representativeness of the SIC cohort can be questioned. Participation studies of the cohorts pooled into the SIC cohort have shown a tendency to higher participation rate with higher level of education (64-66). In addition, participants had lower mortality rates than non-participants (66;113). This indicated that non-participants had a high risk profile (i.e., low socioeconomic position, smokers, low physical activity, overweight) and poor health. It could be expected that the exposure contrast in educational levels and health behaviours observed in this study population were underrated compared to the general population, because those with severe social problems and extreme lifestyle factors declined participation. Generalization of results for participants to the general population should be made with caution. However, representativeness will always be a historical concept because it is time- and place-specific (114;115). For example, when the aim was to study the contribution of behavioural risk factors to the social inequality in mortality, the patterns observed referred to a population of deceased.

CONCLUSION AND IMPLICATIONS FOR FUTURE RESEARCH AND PUBLIC HEALTH POLICY

The research summarised in this thesis is a practical implementation of contemporary statistical methodology, which offers the opportunity to quantify the two central mechanisms – differential exposure and differential vulnerability – simultaneously. This is, an important contribution to the existing literature, because these two mechanisms for understanding inequality in health are not mutually exclusive. The results indicate that research on social inequality in chronic disease outcomes should regard not only that the smoking prevalence is higher in lower socioeconomic groups (due to differential exposure) but also that health consequences of being a smoker can be worse in this subgroups (due to differential vulnerability).

Where does this thesis bring epidemiological research in this field? It seems clear that there are a number of data and modelling challenges to address in future research. Sensitivity analyses to overcome limitations inherent in data are rapidly being proposed in these years (116). This emphasises the need for better data, particularly repeated measures of health behaviour, to study the accumulation of behavioural risk over time. On the modelling side, strategies to study the independent contribution of behaviours and at the same time the joint contribution of two behaviours (in case of interaction) could help prioritize interventions. Finally, it is certainly an open field of research to develop methods in which mediated effects can be identified even when an intermediate variable is affected by the exposure and that in turn confounds the mediator-outcome relationship (117).

The findings in this thesis have substantial implications for public health policy. An immediate implication is that it would be a relevant strategy to reduce social inequality in chronic disease outcomes by challenging the disparities in unhealthy behaviours, particularly smoking. A number of structural interventions that could address differential exposure and vulnerability to smoking are currently being evaluated such as increasing the tax on tobacco products to reduced tobacco's availability, establishing tobacco-free environments, providing knowledge on tobacco's adverse effects and on tobacco control resources and tools (34;118;119). If such structural interventions could yield the same reduction of smokers across socioeconomic groups then it would have the strongest effect on morbidity and mortality among the lower socioeconomic groups. Similarly, targeted interventions (e.g., cessation services designed to target the disadvantaged smokers) with a given reduction in smoking among those in the lower socioeconomic groups would yield a greater reduction in morbidity and mortality than interventions targeting the higher socioeconomic groups. An exclusive focus on interventions targeting the disadvantaged smokers is based on the hope that social inequality in health can be addressed without changing the fundamental social stratification. However, policies focusing on behavioural mediators without addressing socioeconomic disparities might be ineffective. As Chaix and colleagues (116) recently exemplified; smoking may be seen as a way to cope with the stress and adversity associated with socioeconomic disadvantage. Thus, addressing smoking without its socioeconomic determinants would mean leaving people in poverty without one of the strategies to cope with it.

ACKNOWLEDGEMENTS

I thank my supervisors Merete Olser, Birgitte Lidegaard Frederiksen and Naja Hulvaj Rod as well as the collaborators behind the Social Inequality in Cancer cohort study: the Copenhagen City Heart study, the Diet, Cancer and Health study, and the cohorts at the Research Center for Prevention and Health.

SOURCES OF FUNDING

This work was supported by the Danish Cancer Society, Commission of Social Inequality in Cancer [grant number SU08004]

CONFLICTS OF INTERESTS

I have no conflicts of interest

SUMMARY

Socioeconomic differences in morbidity and mortality, particularly across educational groups, are widening. Differential exposures to behavioural risk factors have been shown to play an important mediating role on the social inequality in chronic diseases such as heart disease, cerebrovascular disease, chronic obstructive pulmonary disease, and lung cancer. However, much less attention has been given to the potential role of interaction, where the same level of exposure to a behavioural risk factor has different

effect across socioeconomic groups, creating subgroups that are more vulnerable than others.

In this thesis, Paper 1 describes the unique cohort consortium which was established by pooling and harmonising prospective data from existing cohort studies in Denmark. This consortium generated a large study population with long follow-up sufficient to study power demanding questions of mechanisms underlying social inequalities in chronic disease outcomes. In Paper 2 on incidence of coronary heart disease, smoking and body mass index partially mediated the observed educational differences. This result suggested that some of the social inequality in coronary heart disease may be enhanced by differential exposure to behavioural risk factors (i.e. smoking and obesity). In Paper 3 on incidence of stroke, an observed interaction between education and smoking indicated that participants, particularly men, with low level of education may be more vulnerable to the effect of smoking than those with high level of education in terms of ischemic stroke. Finally, Paper 4 revealed that behavioural risk factors, primarily smoking, explained a considerable part of the educational differences in cause-specific mortality. Further, this paper added important knowledge about the considerable part of the mediated effect, which could be due to interaction between education and smoking.

In conclusion, the research in this thesis is a practical implementation of contemporary statistical methodology, the additive hazards models, in which the potential role of behavioural risk factors can be regarded not only as mediation but also as interaction with the effect of socioeconomic position on chronic disease outcomes. The results support that two central mechanisms, differential exposure and differential vulnerability to behavioural risk factors, particularly smoking; have contributed substantially to the social inequality in chronic disease outcomes in Denmark. These mechanisms are not mutually exclusive and should be regarded simultaneously. However, the findings could be non-causal associations due to, for instance, psychosocial or environmental factors. Nevertheless, research on social inequality in chronic disease outcomes should regard not only that the smoking prevalence is higher in lower socioeconomic groups (differential exposure), but also that health consequences of being a smoker seems to be worse in these subgroups (differential vulnerability).

REFERENCES

- 1 Huisman M, Kunst AE, Bopp M, Borgan JK, Borrell C, Costa G, et al. Educational inequalities in cause-specific mortality in middle-aged and older men and women in eight western European populations. *Lancet* 2005 Feb 5;365(9458):493-500.
- 2 Mackenbach JP, Stirbu I, Roskam AJ, Schaap MM, Menvielle G, Leinsalu M, et al. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 2008 Jun 5;358(23):2468-81.
- 3 Eikemo TA, Huisman M, Bambra C, Kunst AE. Health inequalities according to educational level in different welfare regimes: a comparison of 23 European countries. *Social Health Illn* 2008 May;30(4):565-82.
- 4 Mackenbach JP. The persistence of health inequalities in modern welfare states: the explanation of a paradox. *Soc Sci Med* 2012 Aug;75(4):761-9.
- 5 Mackenbach JP, Bos V, Andersen O, Cardano M, Costa G, Harding S, et al. Widening socioeconomic inequalities in mortality in six Western European countries. *Int J Epidemiol* 2003 Oct;32(5):830-7.
- 6 Bronnum-Hansen H, Baadsgaard M. Widening social inequality in life expectancy in Denmark. A register-based study on social composition and mortality trends for the Danish population. *BMC Public Health* 2012;12:994.
- 7 Diderichsen F, Andersen I, Manuel C, Andersen AM, Bach E, Baadsgaard M, et al. Health inequality - determinants and policies. *Scand J Public Health* 2012 Nov;40(8 Suppl):12-105.
- 8 Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006 May 27;367(9524):1747-57.
- 9 Hafeman DM, Schwartz S. Opening the Black Box: a motivation for the assessment of mediation. *Int J Epidemiol* 2009 Jun;38(3):838-45.
- 10 World cancer research fund/American institut for cancer research. Food, nutrition, physical activity, and prevention of cancer. A global perspective. Washington DC: AICR; 2007.
- 11 Corrao G, Bagnardi V, Zambon A, La VC. A meta-analysis of alcohol consumption and the risk of 15 diseases. *Prev Med* 2004 May;38(5):613-9.
- 12 Kaplan GA, Keil JE. Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation* 1993 Oct;88(4 Pt 1):1973-98.
- 13 Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. *BMJ* 2011;342:d671.
- 14 Rose G. Strategy of prevention: lessons from cardiovascular disease. *Br Med J (Clin Res Ed)* 1981 Jun 6;282(6279):1847-51.
- 15 Shaper AG, Phillips AN, Pocock SJ, Walker M, Macfarlane PW. Risk factors for stroke in middle aged British men. *BMJ* 1991 May 11;302(6785):1111-5.

- 16 Rose G, Marmot MG. Social class and coronary heart disease. *Br Heart J* 1981 Jan;45(1):13-9.
- 17 Harper S, Lynch J, Smith GD. Social determinants and the decline of cardiovascular diseases: understanding the links. *Annu Rev Public Health* 2011 Apr 21;32:39-69.
- 18 Beauchamp A, Peeters A, Wolfe R, Turrell G, Harriss LR, Giles GG, et al. Inequalities in cardiovascular disease mortality: the role of behavioural, physiological and social risk factors. *J Epidemiol Community Health* 2010 Jun;64(6):542-8.
- 19 Kershaw KN, Droomers M, Robinson WR, Carnethon MR, Daviglus ML, Monique Verschuren WM. Quantifying the contributions of behavioral and biological risk factors to socioeconomic disparities in coronary heart disease incidence: the MORGEN study. *Eur J Epidemiol* 2013 Oct;28(10):807-14.
- 20 Laaksonen M, Talala K, Martelin T, Rahkonen O, Roos E, Helakorpi S, et al. Health behaviours as explanations for educational level differences in cardiovascular and all-cause mortality: a follow-up of 60 000 men and women over 23 years. *Eur J Public Health* 2008 Feb;18(1):38-43.
- 21 Strand BH, Tverdal A. Can cardiovascular risk factors and lifestyle explain the educational inequalities in mortality from ischaemic heart disease and from other heart diseases? 26 year follow up of 50,000 Norwegian men and women. *J Epidemiol Community Health* 2004 Aug;58(8):705-9.
- 22 Hallqvist J, Ahlbom A, Diderichsen F, Reuterwall C. How to evaluate interaction between causes: a review of practices in cardiovascular epidemiology. *J Intern Med* 1996 May;239(5):377-82.
- 23 Dalton SO, Schuz J, Engholm G, Johansen C, Kjaer SK, Steding-Jessen M, et al. Social inequality in incidence of and survival from cancer in a population-based study in Denmark, 1994-2003: Summary of findings. *Eur J Cancer* 2008 Sep;44(14):2074-85.
- 24 Parkin DM, Bray FI, Devesa SS. Cancer burden in the year 2000. The global picture. *Eur J Cancer* 2001 Oct;37 Suppl 8:S4-66.
- 25 Pisani P, Bray F, Parkin DM. Estimates of the world-wide prevalence of cancer for 25 sites in the adult population. *Int J Cancer* 2002 Jan 1;97(1):72-81.
- 26 Aarts MJ, Lemmens VE, Louwman MW, Kunst AE, Coebergh JW. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. *Eur J Cancer* 2010 Oct;46(15):2681-95.
- 27 Carlsen K, Hoybye MT, Dalton SO, Tjonneland A. Social inequality and incidence of and survival from breast cancer in a population-based study in Denmark, 1994-2003. *Eur J Cancer* 2008 Sep;44(14):1996-2002.
- 28 Dalton SO, Steding-Jessen M, Engholm G, Schuz J, Olsen JH. Social inequality and incidence of and survival from lung cancer in a population-based study in Denmark, 1994-2003. *Eur J Cancer* 2008 Sep;44(14):1989-95.
- 29 Dano H, Andersen O, Ewertz M, Petersen JH, Lynge E. Socioeconomic status and breast cancer in Denmark. *Int J Epidemiol* 2003 Apr;32(2):218-24.
- 30 Egeberg R, Halkjaer J, Rottmann N, Hansen L, Holten I. Social inequality and incidence of and survival from cancers of the colon and rectum in a population-based study in Denmark, 1994-2003. *Eur J Cancer* 2008 Sep;44(14):1978-88.
- 31 Sidorchuk A, Agardh EE, Aremu O, Hallqvist J, Allebeck P, Moradi T. Socioeconomic differences in lung cancer incidence: a systematic review and meta-analysis. *Cancer Causes Control* 2009 May;20(4):459-71.
- 32 Prescott E, Vestbo J. Socioeconomic status and chronic obstructive pulmonary disease. *Thorax* 1999 Aug;54(8):737-41.
- 33 Prescott E, Lange P, Vestbo J. Socioeconomic status, lung function and admission to hospital for COPD: results from the Copenhagen City Heart Study. *Eur Respir J* 1999 May;13(5):1109-14.
- 34 Blas E, Kurup AS. Equity, social determinants and public health programmes. Switzerland: World Health Association; 2010.
- 35 Marmot M, Atkinson T, Bell J, Black C, Broadfoot P, Cumberlege J. Fair society, healthy lives. The Marmot review. Strategic review of health inequalities in England post-2010. 2010.
- 36 Government of Denmark. Regeringens folkesundhedsprogram 1999-2008. Et handlingsorienteret program for sundere rammer i hverdagen. 1999.
- 37 Government of Denmark. Sund hele livet - de nationale mål og strategier for folkesundheden 2002-10. 2002.
- 38 Government of Denmark. Sundhedspakken 09. 2009.
- 39 Lopez AD, Collishaw NE, Piha T. A descriptive model of the cigarette epidemic in developed countries. *Tobacco Control* 1994 Sep 1;3(3):242.
- 40 Osler M, Holstein B, Avlund K, Damsgaard MT, Rasmussen NK. Socioeconomic position and smoking behaviour in Danish adults. *Scand J Public Health* 2001 Mar;29(1):32-9.

- 41 Huisman M, Kunst AE, Mackenbach JP. Educational inequalities in smoking among men and women aged 16 years and older in 11 European countries. *Tob Control* 2005 Apr;14(2):106-13.
- 42 Nordahl H, Glümer C, Jørgensen T. *Udviklingen i risiko faktorer for hjerte-karsygdom i vestegnskommunerne 1978-2006*. Glostrup: Research Centre for Prevention and Health. Capital Region of Denmark; 2008.
- 43 Kjølner M, Juel K, Kamper-Jørgensen F. *Folkesundhedsrapporten Danmark 2007*. Copenhagen: National Institute of Public Health; 2007.
- 44 Osler M, Jørgensen T, Davidsen M, Gronbaek M, Bronnum-Hansen H, Madsen M, et al. Socioeconomic status and trends in alcohol drinking in the Danish MONICA population, 1982-92. *Scand J Public Health* 2001 Mar;29(1):40-3.
- 45 Diderichsen F, Hallqvist J. Social inequality in health: Some methodological considerations for the study of social position and social context. In: Arve-Páres B, editor. *Inequality in Health - a Swedish perspective*. Stockholm: Swedish Council for Social Research; 1998. p. 25-39.
- 46 Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey SG. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health* 2006 Jan;60(1):7-12.
- 47 Mutsers R, Jager KJ, Zoccali C, Dekker FW. The effect of joint exposures: examining the presence of interaction. *Kidney Int* 2009 Apr;75(7):677-81.
- 48 Diderichsen F, Evans T, Whitehead M. The social basis of disparities in health. In: Evans T, Whitehead M, Diderichsen F, Bhuiya A, Wirth M, editors. *Challenging inequities in health. From ethics to action*. New York: Oxford University Press; 2001. p. 12-23.
- 49 Albert MA, Glynn RJ, Buring J, Ridker PM. Impact of traditional and novel risk factors on the relationship between socioeconomic status and incident cardiovascular events. *Circulation* 2006 Dec 12;114(24):2619-26.
- 50 Ernstsen L, Bjerkeset O, Krokstad S. Educational inequalities in ischaemic heart disease mortality in 44,000 Norwegian women and men: the influence of psychosocial and behavioural factors. The HUNT Study. *Scand J Public Health* 2010 Nov;38(7):678-85.
- 51 Kilander L, Berglund L, Boberg M, Vessby B, Lithell H. Education, lifestyle factors and mortality from cardiovascular disease and cancer. A 25-year follow-up of Swedish 50-year-old men. *Int J Epidemiol* 2001 Oct;30(5):1119-26.
- 52 Wamala SP, Mittleman MA, Schenck-Gustafsson K, Orth-Gomer K. Potential explanations for the educational gradient in coronary heart disease: a population-based case-control study of Swedish women. *Am J Public Health* 1999 Mar;89(3):315-21.
- 53 Lantz PM, Golberstein E, House JS, Morenoff J. Socioeconomic and behavioral risk factors for mortality in a national 19-year prospective study of U.S. adults. *Soc Sci Med* 2010 May;70(10):1558-66.
- 54 Harald K, Pajunen P, Jousilahti P, Koskinen S, Vartiainen E, Salomaa V. Modifiable risk factors have an impact on socioeconomic differences in coronary heart disease events. *Scand Cardiovasc J* 2006 Apr;40(2):87-95.
- 55 Kerr GD, Slavin H, Clark D, Coupar F, Langhorne P, Stott DJ. Do vascular risk factors explain the association between socioeconomic status and stroke incidence: a meta-analysis. *Cerebrovasc Dis* 2011;31(1):57-63.
- 56 Khang YH, Lynch JW, Yang S, Harper S, Yun SC, Jung-Choi K, et al. The contribution of material, psychosocial, and behavioral factors in explaining educational and occupational mortality inequalities in a nationally representative sample of South Koreans: relative and absolute perspectives. *Soc Sci Med* 2009 Mar;68(5):858-66.
- 57 Lynch J, Davey SG, Harper S, Bainbridge K. Explaining the social gradient in coronary heart disease: comparing relative and absolute risk approaches. *J Epidemiol Community Health* 2006 May;60(5):436-41.
- 58 McFadden E, Luben R, Wareham N, Bingham S, Khaw KT. Social class, risk factors, and stroke incidence in men and women: a prospective study in the European prospective investigation into cancer in Norfolk cohort. *Stroke* 2009 Apr;40(4):1070-7.
- 59 Nandi A, Glymour MM, Subramanian SV. Association among socioeconomic status, health behaviors, and all-cause mortality in the United States. *Epidemiology* 2014 Mar;25(2):170-7.
- 60 Whitehead M, Burstrom B, Diderichsen F. Social policies and the pathways to inequalities in health: a comparative analysis of lone mothers in Britain and Sweden. *Soc Sci Med* 2000 Jan;50(2):255-70.
- 61 McFadden E, Luben R, Wareham N, Bingham S, Khaw KT. Occupational social class, educational level, smoking and body mass index, and cause-specific mortality in men and women: a prospective study in the European Prospective Investigation of Cancer and Nutrition in Norfolk (EPIC-Norfolk) cohort. *Eur J Epidemiol* 2008;23(8):511-22.
- 62 Stringhini S, Dugravot A, Shipley M, Goldberg M, Zins M, Kivimaki M, et al. Health behaviours, socioeconomic status, and

- mortality: further analyses of the British Whitehall II and the French GAZEL prospective cohorts. *PLoS Med* 2011 Feb;8(2):e1000419.
- 63 Appleyard M, Hansen AT, Schnohr P, Jensen G, Nyboe J. The Copenhagen city heart study. A book of tables with data from the first examination (1976-1978) and a 5-year follow-up (1981-83). *Scand J Soc Med Suppl* 1989;170(41):1-160.
- 64 Schnohr P, Jensen G, Lange P, Scharling H, Appleyard M. The Copenhagen City Heart Study. *European Heart Journal* 2001;3(suppl H):H1-H83.
- 65 Tjonneland A, Olsen A, Boll K, Stripp C, Christensen J, Engholm G, et al. Study design, exposure variables, and socioeconomic determinants of participation in Diet, Cancer and Health: a population-based prospective cohort study of 57,053 men and women in Denmark. *Scand J Public Health* 2007;35(4):432-41.
- 66 Osler M, Linneberg A, Glümer C, Jørgensen T. The cohorts at the Research Centre for Prevention and Health, formerly The Glostrup Population Studies. *International Journal of Epidemiology* 2011 Jun 1;40(3):602-10.
- 67 Glymour MM, Greenland S. Causal diagrams. In: Rothman KJ, Greenland S, Lash TL, editors. *Modern epidemiology*. 3 ed. Philadelphia: Lippincott Williams & Wilkins; 2008. p. 183-209.
- 68 Greenland S, O'Rourke K. Meta-Analysis. In: Rothman KJ, Greenland S, Lash TL, editors. *Modern epidemiology*. 3 ed. Philadelphia: Lippincott Williams & Wilkins; 2008. p. 652-82.
- 69 Syme SL, Berkman LF. Social class, susceptibility and sickness. *Am J Epidemiol* 1976 Jul;104(1):1-8.
- 70 Andersen I, Diderichsen F, Kornerup H, Prescott E, Rod NH. Major life events and the risk of ischaemic heart disease: does accumulation increase the risk? *Int J Epidemiol* 2011 Aug;40(4):904-13.
- 71 Rod NH, Andersen I, Prescott E. Psychosocial risk factors and heart failure hospitalization: a prospective cohort study. *Am J Epidemiol* 2011 Sep 15;174(6):672-80.
- 72 Lange T, Hansen JV. Direct and indirect effects in a survival context. *Epidemiology* 2011 Jul;22(4):575-81.
- 73 Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol* 2012 Aug 1;176(3):190-5.
- 74 Rod NH, Lange T, Andersen I, Marott JL, Diderichsen F. Additive interaction in survival analysis: use of the additive hazards model. *Epidemiology* 2012 Sep;23(5):733-7.
- 75 Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 1992 Mar;3(2):143-55.
- 76 Pearl J. Direct and indirect effect. In: Kaufmann M, editor. *Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence*. San Francisco CA: 2001. p. 411-20.
- 77 VanderWeele TJ, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface* 2009;2(4):457-68.
- 78 VanderWeele TJ. A three-way decomposition of a total effect into direct, indirect, and interactive effects. *Epidemiology* 2013 Mar;24(2):224-32.
- 79 Nandi A, Glymour MM, Kawachi I, VanderWeele TJ. Using Marginal Structural Models to Estimate the Direct Effect of Adverse Childhood Social Conditions on Onset of Heart Disease, Diabetes, and Stroke. *Epidemiology* 2012 Mar;23(2):223-32.
- 80 Hvidtfeldt UA, Tolstrup JS, Jakobsen MU, Heitmann BL, Gronbaek M, O'Reilly E, et al. Alcohol intake and risk of coronary heart disease in younger, middle-aged, and older adults. *Circulation* 2010 Apr 13;121(14):1589-97.
- 81 Leander K, Hallqvist J, Reuterwall C, Ahlbom A, de FU. Family history of coronary heart disease, a strong risk factor for myocardial infarction interacting with other cardiovascular risk factors: results from the Stockholm Heart Epidemiology Program (SHEEP). *Epidemiology* 2001 Mar;12(2):215-21.
- 82 Stjarne MK, Fritzell J, De Leon AP, Hallqvist J. Neighborhood socioeconomic context, individual income and myocardial infarction. *Epidemiology* 2006 Jan;17(1):14-23.
- 83 Schnall PL, Landsbergis PA, Baker D. Job strain and cardiovascular disease. *Annu Rev Public Health* 1994;15:381-411.
- 84 Steptoe A, Kivimaki M. Stress and Cardiovascular Disease: An Update on Current Knowledge. *Annual Review of Public Health*, Vol 34 2013;34:337-+.
- 85 Koivula RW, Tornberg AB, Franks PW. Exercise and diabetes-related cardiovascular disease: systematic review of published evidence from observational studies and clinical trials. *Curr Diab Rep* 2013 Jun;13(3):372-80.
- 86 Silhol R, Zins M, Chauvin P, Chaix B. Investigating the spatial variability in incidence of coronary heart disease in the Gazel cohort: the impact of area socioeconomic position and mediating role of risk factors. *J Epidemiol Community Health* 2011 Feb;65(2):137-43.
- 87 Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with

- myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 2004 Sep 11;364(9438):937-52.
- 88 Giskes K, Kunst AE, Benach J, Borrell C, Costa G, Dahl E, et al. Trends in smoking behaviour between 1985 and 2000 in nine European countries by education. *J Epidemiol Community Health* 2005 May;59(5):395-401.
- 89 Galobardes B, Lynch J, Smith GD. Measuring socioeconomic position in health research. *Br Med Bull* 2007;81-82:21-37.
- 90 Smith PM, Smith BT, Mustard CA, Lu H, Glazier RH. Estimating the direct and indirect pathways between education and diabetes incidence among Canadian men and women: a mediation analysis. *Ann Epidemiol* 2013 Mar;23(3):143-9.
- 91 Falkstedt D, Hemmingsson T. Educational level and coronary heart disease: a study of potential confounding from factors in childhood and adolescence based on the Swedish 1969 conscription cohort. *Ann Epidemiol* 2011 May;21(5):336-42.
- 92 Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods* 2013 Feb 4.
- 93 Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986 Dec;51(6):1173-82.
- 94 Chaix B, Evans D, Suzuki E. Commentary: socioeconomic status, health behavior, and mortality: old question plus modern methods equals new insights? *Epidemiology* 2014 Mar;25(2):178-81.
- 95 Friedenreich CM. Methods for pooled analyses of epidemiologic studies. *Epidemiology* 1993 Jul;4(4):295-302.
- 96 Ahlbom A. Pooling epidemiologic studies. *Epidemiology* 1993 Jul;4(4):283-4.
- 97 Michiels S, Baujat B, Mahe C, Sargent DJ, Pignon JP. Random effects survival models gave a better understanding of heterogeneity in individual patient data meta-analyses. *Journal of Clinical Epidemiology* 2005 Mar;58(3):238-45.
- 98 Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart LA. Statistical Analysis of Individual Participant Data Meta-Analyses: A Comparison of Methods and Recommendations for Practice. *Plos One* 2012 Oct 3;7(10).
- 99 Ogburn EL, VanderWeele TJ. Analytic results on the bias due to nondifferential misclassification of a binary mediator. *Am J Epidemiol* 2012 Sep 15;176(6):555-61.
- 100 Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004 Jun 26;328(7455):1519.
- 101 Ekholm O, Kjølner M, Davidsen M, Hesse U, Eriksen L, Christensen AI, et al. Sundhed og sygelighed i Danmark 2005 & udviklingen siden 1987 [Danish Health and Morbidity Survey 2005 and trends since 1987]. National Institut of Public Health; 2006.
- 102 VanderWeele TJ, Valeri L, Ogburn EL. The role of measurement error and misclassification in mediation analysis: mediation and measurement error. *Epidemiology* 2012 Jul;23(4):561-4.
- 103 Jensen VM, Rasmussen AW. Danish education registers. *Scandinavian Journal of Public Health* 2011 Jul 1;39(7 suppl):91-4.
- 104 Pallesen PB, Tverborgvik T, Rasmussen HB, Lynge E. Data on education: from population statistics to epidemiological research. *Scand J Public Health* 2010 Mar;38(2):177-83.
- 105 Helweg-Larsen K. The Danish Register of Causes of Death. *Scand J Public Health* 2011 Jul;39(7 Suppl):26-9.
- 106 Jensen HH, Godtfredsen NS, Lange P, Vestbo J. Potential misclassification of causes of death from COPD. *Eur Respir J* 2006 Oct;28(4):781-5.
- 107 Madsen M, Davidsen M, Rasmussen S, Abildstrom SZ, Osler M. The validity of the diagnosis of acute myocardial infarction in routine statistics: a comparison of mortality and hospital discharge data with the Danish MONICA registry. *J Clin Epidemiol* 2003 Feb;56(2):124-30.
- 108 Loucks EB, Buka SL, Rogers ML, Liu T, Kawachi I, Kubzansky LD, et al. Education and coronary heart disease risk associations may be affected by early-life common prior causes: a propensity matching analysis. *Ann Epidemiol* 2012 Apr;22(4):221-32.
- 109 Sondergaard G, Mortensen LH, Andersen AMN, Andersen PK, Dalton SO, Madsen M, et al. Does Shared Family Background Influence the Impact of Educational Differences on Early Mortality? *American Journal of Epidemiology* 2012 Oct 15;176(8):675-83.
- 110 VanderWeele TJ. Unmeasured confounding and hazard scales: sensitivity analysis for total, direct, and indirect effects. *European Journal of Epidemiology* 2013 Feb;28(2):113-7.
- 111 Avin C, Shpitser I, Pearl J. Identifiability of path-specific effects. *Proceedings of the International Joint Conference on Artificial Intelligence*. 2005. p. 357-63.

- 112 Imai K, Yamamoto T. Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments. *Political Analysis* 2013;21(2):141-71.
- 113 Andersen LB, Vestbo J, Juel K, Bjerg AM, Keiding N, Jensen G, et al. A comparison of mortality rates in three prospective studies from Copenhagen with mortality rates in the central part of the city, and the entire country. *Copenhagen Center for Prospective Population Studies. Eur J Epidemiol* 1998 Sep;14(6):579-85.
- 114 Olsen J. Who needs selection bias? *Scand J Work Environ Health* 2014 Jan;40(1):103.
- 115 Nohr EA, Olsen J. Commentary: Epidemiologists have debated representativeness for more than 40 years--has the time come to move on? *Int J Epidemiol* 2013 Aug;42(4):1016-7.
- 116 Chaix B, Evans D, Suzuki E. Commentary: socioeconomic status, health behavior, and mortality: old question plus modern methods equals new insights? *Epidemiology* 2014 Mar;25(2):178-81.
- 117 VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology* 2014 Mar;25(2):300-6.
- 118 Ahmad S, Franz GA. Raising taxes to reduce smoking prevalence in the US: a simulation of the anticipated health and economic impacts. *Public Health* 2008 Jan;122(1):3-10.
- 119 Galan I, Mata N, Estrada C, Diez-Ganan L, Velazquez L, Zorrilla B, et al. Impact of the "Tobacco control law" on exposure to environmental tobacco smoke in Spain. *BMC Public Health* 2007;7:224.