Mechanisms underlying social inequality in post-menopausal breast cancer

Mediation and interaction of behavioral, hormonal and reproductive factors

Ulla Arthur Hvidtfeldt

This review has been accepted as a thesis together with three previously published papers by University of Copenhagen on June 10th, 2014 and defended on June 23rd, 2014

Tutors: Naja Hulvej Rod, Niels Keiding and Anne Tjønneland

Official opponents: Henrik Ravn, Valerie Beral and Jens Peter Ellekilde Bonde

Correspondence: Department of Public Health, Section of Social Medicine, University of Copenhagen, Øster Farimagsgade 5A, P.O.Box 2099, 1014 Copenhagen K, Denmark.

E-mail: ulah@sund.ku.dk

Dan Med J 2014;61(10) B4922

THE THREE ORIGINAL PAPERS ARE:

1. Hvidtfeldt UA, Lange T, Andersen I, Diderichsen F, Keiding N, Prescott E, Sørensen TIA, Tjønneland A, Rod NH. Educational differences in postmenopausal breast cancer – quantifying indirect effects through health behaviors, body mass index and reproductive patterns. *PLoS One*, 2013; 8(10):e78690.
2. Hvidtfeldt UA, Gunter MJ, Lange T, Chlebowski RT, Lane D, Farhat GN, Freiberg MS, Keiding N, Lee JS, Prentice R, Tjonneland A, Vitolins MZ, Wassertheil-Smoller S, Strickler HD, Rod NH. Quantifying mediating effects of endogenous estrogen and insulin in the relation between obesity, alcohol consumption, and breast cancer. *Cancer Epidemiol Biomarkers Prev* 2012; 21(7):1203-12.
3. Hvidtfeldt UA, Tjønneland A, Keiding N, Lange T, Andersen I, Sørensen TIA, Prescott E, Hansen ÅM, Grønbæk M, Bojesen SE, Diderichsen F, Rod NH. Body mass index and alcohol consumption in relation to postmenopausal estrogen-receptor positive and negative breast cancer and serum sex-hormones among users and non-users of hormone therapy. Submitted.

INTRODUCTION

Breast cancer is the most common cancer in adult women worldwide [1], and the main cause of premature death among women in economically developed countries [2]. The incidence is strongly related to age and is predominantly occurring in older ages. In Denmark, the incidence among postmenopausal women (50+ years) has increased considerably during the past decades from 150 to 393 per 100,000 person-years[1] between 1943 and 2010 [3].

The social gradient in cancer is skewed, but whereas cancers of for example the lung and cervix are most prevalent in socially deprived groups, breast cancer is more frequently observed among women of higher socioeconomic position (SEP) [4-8]. However, this tendency appears to be in transition as the increased risk among women of higher SEP attenuates with younger birth cohorts [8,9]. In a broader perspective, the (age-standardized) incidence rates in economically developing countries have also caught up with the high levels observed in the developed part of the world during the last decade [1].

The mechanisms underlying the social inequality in breast cancer incidence are not well described; yet understanding the pathways through which social factors affect the risk of breast cancer is essential to causal inference and thus to effective prevention strategies [10]. The rapid increase in breast cancer incidence in economically developing countries suggests a strong effect of lifestyle and reproductive behaviors, which is also supported in the epidemiologic literature [4,9,11-21]. Previous studies have suggested mediating effects of factors such as age at first birth, parity, hormone therapy (HT) use, alcohol consumption, physical inactivity and obesity on the relation between SEP and breast cancer. For example, Heck et al. reported a relative risk of postmenopausal breast cancer of 2.3 (95% confidence interval (CI): 1.2, 4.3) among women with a high compared to a low level of education [11]. After adjustment for age at first birth, age at menarche and menopause, alcohol consumption, use of HT and oral contraceptives, body mass index (BMI) and height, the relative risk was reduced to 1.5 (95% CI: 0.8, 2.7), which indicates that SEP works partly through these factors.

However, the simplified method of assessing mediating effects by comparing crude versus adjusted models can be biased [22,23]. The main problems discussed are issues of mediator-outcome confounding, exposure-dependent confounding of the mediator-outcome relation, interactions between exposure and mediators as well as interactions between mediators. Another point is that these simple methods do not allow for a decomposition of the total effect into direct and indirect (mediated) pathways [24,25]. In recent years, more advanced methods have been developed to address some of these issues [26,27].

If risk factors interact in synergy, clustering among them will have a stronger impact on the incidence of breast cancer than the sum of their individual effects, and prevention of one factor

[1]Age-standardized according to the world standard population

will reduce the effect of the other [28]. For instance, a person exposed to both alcohol consumption and HT use is at higher risk of breast cancer than would be expected from the sum of their separate effects if these interact [29,30]. Consequently, preventing alcohol consumption will both directly and indirectly decrease the risk of breast cancer by removing part of the effect of HT use. In this regard, interaction is of core public health importance and intervention strategies may be improved through such identifications. Since many of the risk factors for postmenopausal breast cancer cluster among women of higher SEP, it is likely that the social inequality would also be reduced.

This thesis adds to the current knowledge on how SEP affects postmenopausal breast cancer risk by applying new statistical methods in order to quantify mediating effects through lifestyle and reproductive factors and by addressing interaction between the mediators. In addition, assumptions and potential biases involved with such analyses are discussed and investigated through sensitivity analyses. The analyses were based on several well-established prospective cohort studies specifically pooled and linked with register data for the purpose of these mediation analyses. This ensured a large population sample with a broad age range and social distribution as well as a long follow-up time. Additionally, it was possible to investigate detailed hormonal pathways through international collaboration with the Women's Health Initiative. Thus, the work included in this thesis provides insight into the pathways through which SEP may affect the risk of postmenopausal breast cancer, and thereby draws attention to potential paths of intervention.

Aims

The overall objective of this thesis is to determine mechanisms underlying social inequality in postmenopausal breast cancer by addressing mediating effects through HT use, high BMI, lifestyle and reproductive factors. This objective is addressed in three papers and the synopsis which cover different aspects of the pathways from SEP to breast cancer. Specifically, **Paper I** addresses mediation by HT use, alcohol consumption, physical inactivity, high BMI, parity and age at first birth in the relation between educational level and postmenopausal breast cancer. **Paper II** concerns the hormonal pathways from high BMI and alcohol consumption to breast cancer. Finally, **Paper III** addresses interaction between hormone therapy use and BMI and alcohol consumption in relation to postmenopausal breast cancer.

Structure of the synopsis

The synopsis is structured as follows: First, the background section describes a framework for understanding how social factors may affect health in general. This is followed by a conceptual model of the hypothesized pathways from SEP to postmenopausal breast cancer and the research questions forming this thesis. This section also includes an overview of previous papers addressing mediating pathways of the relation between SEP and breast cancer. Second, the data sources and methodology for the papers and additional sensitivity analyses are briefly described. Third, the results of the three studies are summarized and fourth, these results and their potential sources of bias are quantified and discussed. Finally, future perspectives of the findings are discussed.

BACKGROUND

This section provides a brief introduction to the research field of social inequality in health. Also, a conceptual model of the pathways from SEP to postmenopausal breast cancer explored in this thesis is presented, followed by a description of previous studies on mechanisms underlying this relation.

Social inequality in health

The role of SEP in health has been studied for decades, and social inequality in various diseases is widely documented [31,32]. However, much is yet to be learned about these associations and the underlying mechanisms. In the model developed by Diderichsen and colleagues [31], presented in a simplified version below (Figure 1) [33], three mechanisms of social inequality in health are described: I) Social stratification, II) Differential exposure and III) Differential vulnerability. In addition, the model illustrates possible policy entry points for reducing social inequality in health: A) Influencing social stratification, B) decreasing exposures and C) decreasing vulnerability. Social stratification (mechanism I) works at the contextual level encapturing political, cultural, social and environmental elements of society (e.g., legislation, cultural norms, discrimination, access to health care, etc.). The concept of differential exposure (mechanism II) represents the individual exposure to risk factors, which are determined by social position. Risk factors often tend to cluster in certain social groups and their effects may interact with one another thereby causing differential vulnerability (mechanism III). This means that the effect of a specific exposure depends on the presence of other contributing factors [31]. For instance, people of low SEP may be more vulnerable to the effects of smoking due to processes related to childhood environmental circumstances or other risk factors also linked to SEP.

This thesis deals with individual level mechanisms, primarily differential exposure to certain risk factors related to lifestyle and reproduction, but also the aspect of differential vulnerability to these risk factors across socioeconomic groups.

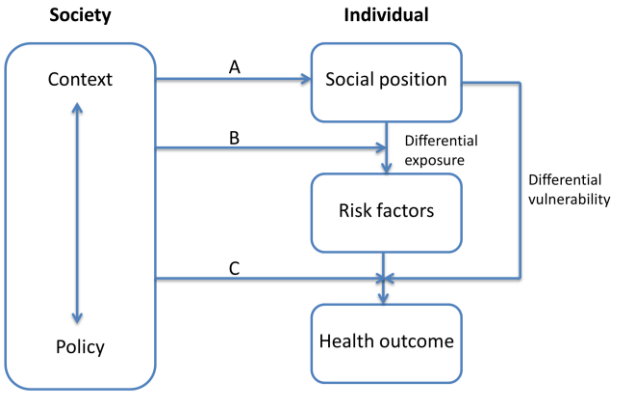

Figure 1
A framework for understanding social inequality in health [33].

How does socioeconomic position affect breast cancer risk?

Women of higher SEP are generally better off concerning nearly all health outcomes, but for breast cancer incidence the social gradient is reverse [4-8]. The conceptual model presented in Figure 2 summarizes the specific pathways (labeled A–H) addressed in this thesis. The model encompasses direct effects of SEP on breast cancer (**path A**) and indirect (mediated) effects through HT use (**path B**), fertility patterns (**path C**), lifestyle factors (**path D**) and high BMI (**path E**). Further, indirect effects of lifestyle factors and high BMI through the hormones estradiol and insulin are explored (**paths F–H**).

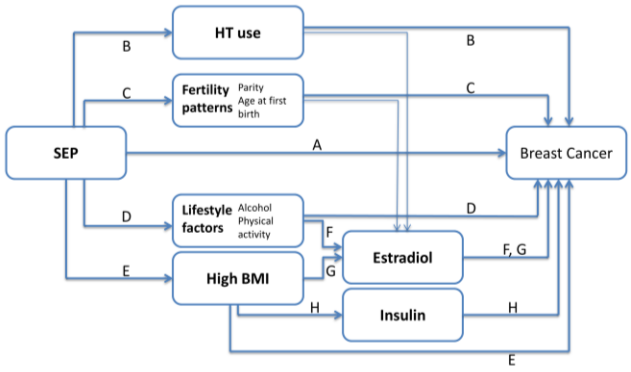

Figure 2
Conceptual model describing the hypothesized direct and indirect pathways from SEP to breast cancer.

In Figure 2, **path A** represents the direct effect of SEP on breast cancer, which in this case represents psychosocial and environmental processes as well as other lifestyle related risk factors not embedded in this thesis (e.g., diet or vitamin use) [18]. Following the framework of Diderichsen et al. presented above, the direct path may represent contextual as well as individual phenomena.

The use of HT is a risk factor of breast cancer, which is more common among women of higher SEP as suggested by **path B** [34-36]. A large study reanalyzing worldwide data found a higher risk of breast cancer in current and recent users of HT compared to never-users as well as a higher risk with increasing duration of use [37]. HT use greatly increases serum estrogen levels [38]. Estrogens stimulate the division of breast epithelial cells, which increases the risk of mutation, and increasing levels of serum estradiol are therefore likely to induce or promote breast cancer [39,40].

The SEP of women also influences family planning behavior (**path C**) as women of higher SEP tend to postpone childbearing and reduce higher-order births [9,21]. Low age at first birth (<30 years) and parity decrease the risk of breast cancer, probably through altered hormonal profiles, differentiation of mammary glands, or changes in the estrogen responsiveness of the gland [41,42].

Lifestyle is closely related to SEP as suggested by **path D**. Women of higher SEP are more likely to drink alcohol and generally consume larger quantities than women of lower SEP [11,20]. The risk of breast cancer has been found to increase with higher levels of alcohol consumption [43,44]. The positive relation between alcohol and breast cancer may be due to increased levels of endogenous estradiol following alcohol consumption **(path F)** [45,46]. Physical activity level is another lifestyle factor found to be higher in women of higher SEP (**path D**) [11]; however, the relation between physical activity and postmenopausal breast cancer is likely inverse [47], and thus physical activity may reduce the overall effect of SEP on breast cancer. Several mechanisms have been proposed to explain this relation. Firstly, physical activity may lead to a reduction in body weight and decrease central adiposity, thereby reducing the aromatization of androgen to estrogen in fat tissue [48]. Further, physical activity has been linked with lowered levels of estrogen (**path F**) in postmenopausal women in both observational and experimental studies, and the association persisted even after adjustment for BMI, suggestive of an independent effect of physical activity [45,48]. Physical activity is also associated with higher levels of circulating concentrations of sex hormone-binding globulin (SHBG), thereby lowering the amounts of free, active hormones in the body [48]. Another potential mechanism is through exercise reduction in insulin [48,49].

High BMI is another risk factor possibly reducing the overall effect of SEP on postmenopausal breast cancer. The prevalence of high BMI/obesity is higher in women with lower SEP (**path E**) [11,13,14], and overweight and obesity have been consistently linked with the risk of postmenopausal breast cancer [50,51]. Body fat directly affects levels of many circulating hormones such as estrogens (**path G**), testosterone and insulin (**path H**) [48]. In the adipose tissue of postmenopausal women, androgens convert into estrogens leading to increased estrogen levels, and several previous observational studies have linked testosterone to breast cancer [51,54]. Hyperinsulinemia has also been suggested as a significant, independent risk factor of breast cancer after adjustment for estradiol and other risk factors [55]. Hyperinsulinemia lowers the levels of SHGB leading to increased levels of bioavailable estradiol and testosterone [48]. Insulin has also been found to stimulate breast cancer cell proliferation [56].

Thus, multiple potential mechanisms in the relation between SEP and breast cancer are at play. Numerous studies have linked SEP with these risk factors of breast cancer separately, but combined effects are not understood in depth. Since these factors all partially take effect through similar hormonal pathways, modification by one factor on the effect of another factor is likely. For instance, physical activity has been found to modify the association between BMI and breast cancer, with inactive women in the upper BMI quartile being at a markedly increased risk compared with their lean and active counterparts [57]. Also, interactions between HT use and BMI have been observed, where the increase in relative risk of breast cancer among users of HT was greater in women with low relative to high weight [37]. Similarly, the effect of alcohol consumption on breast cancer risk may differ according to HT status and BMI [29,30,58].

The focus on these specific pathways in the thesis was guided by the literature reviewed above and on the current knowledge on social inequality in postmenopausal breast cancer presented in Table 1, described in detail below.

Previous studies on mediating effects of social inequality in breast cancer

Table 1 provides an overview of previous prospective studies on mediating effects of social inequality in postmenopausal breast cancer. In general, all of the previous studies have shown a higher risk of breast cancer among women of high versus low SEP as measured by education, income or occupation. In most of the previous studies, reproductive factors such as parity and age at

first birth seem to account for a substantial part of the higher risk of breast cancer in women of high educational level [9,11,14-17,20,21]. Evidence of a significant contribution of HT use and lifestyle factors have also been published previously [11,13,15,17,20]; The large scale Norwegian Women and Cancer Study (NOWAC) found a relative risk of 1.47 among women with a high educational level compared to women with a low educational level [15]. Stepwise adjustment for multiple factors suggested that reproductive factors explained 26% of the increased risk, alcohol consumption accounted for 23% and current HT use and BMI only accounted for approximately 3% of the social inequality. A recent Danish study found a relative risk of 1.2 in the highest versus lowest educated women, which was reduced to 1.06 after adjustment for reproductive factors, HT use and alcohol consumption [13]. Adjustment for BMI did not affect the estimate. However, since most of the previous studies include these factors in the model simultaneously, assessment of the relative contributions of each factor is not possible.

The studies investigating social inequalities in breast cancer as defined according to occupational status generally report a modest decrease of the social inequality after adjustment for reproductive factors and factors such as alcohol consumption and HT use [9,13,18]. The American study by Pudrovska et al. found a relative risk of 1.72 (95% CI: 1.25, 2.36) among women in professional occupations and 1.57 (95% CI: 1.02, 2.42) in women with a managerial occupation compared to housewives. Reproductive factors were found to mediate 23% of the association between professional occupations and breast cancer, but did not affect the higher risk observed for managerial occupation. On the other hand, job authority appeared to mediate 55% of the increased risk among women with a managerial occupation, but did not materially affect the higher risk observed in professionals.

In conclusion, the previous literature supports the hypothesis of mediating pathways from SEP to postmenopausal breast cancer through reproductive patterns, lifestyle factors and HT use. However, the decomposition of effects through each of these factors and detailed analysis on how these mediators may take effect – for example by interactions or through estrogen pathways – is still a rather unexplored area of research.

MATERIALS AND METHODS

The results of Papers I and III were based on data from the Social Inequality in Cancer Cohort Study, and Paper II was based on data from the Women's Health Initiative Observational Study. These data sources and assessments of SEP, mediators and confounders as well as postmenopausal breast cancer are briefly described below followed by a presentation of the applied statistical methods.

Data sources

The Social Inequality in Cancer (SIC) Cohort Study

The aim of establishing the SIC cohort was to elucidate social inequality in different types of cancers and investigate mechanisms behind these inequalities. The database combines data from several large Danish population based cohort studies: The Copenhagen City Heart Study (2nd wave), The Diet, Cancer and Health Study and the Cohorts at the Research Centre for Population and Health (MONICA I–III, the 1936-cohort and INTER99) and register based follow-up. All studies include measurements of lifestyle and biological risk factors for cancer. A cohort profile describing the details of the establishment of the SIC cohort study has been published previously [59], and will be described briefly below.

The Copenhagen City Heart Study (CCHS) was initiated in 1976 where a random sample of citizens in the Copenhagen area aged 20+ years was invited to participate (N=14,223 participants, corresponding to a response rate of ≈74%) [60]. A second wave was completed in 1981–83, which included all previously invited and an additional 500 individuals aged 20–25 years (N=12,698; response rate ≈70%). To date, three subsequent waves have been carried out. At every wave, all participants completed a self-administered questionnaire on health status, lifestyle and reproductive factors and went through physical examinations (including height, weight, blood pressure, etc.). The SIC cohort includes measurements from the second wave in 1981–83.

The Diet, Cancer and Health Study (DCHS) was started in the period of 1993–1997 where almost all men and women aged 50 to 64 years residing in the areas of Copenhagen and Aarhus, who fulfilled the inclusion criteria, were invited [61]. Participants were eligible if they were born in Denmark and free of cancers at the time of inclusion (N=57,053; response rate ≈35%). All participants completed a self-administered questionnaire concerning lifestyle factors. Physical examinations included anthropometric and blood pressure measures as well as samples of blood, urine and fat.

The Cohorts at the Research Centre for Population and Health (CRCPH) include several independent cohort studies of which the three Danish World Health Organization MONICA cohorts, the 1936-cohort (2nd wave) and the INTER99 study were included in the SIC cohort [62]. Participants were drawn as random samples of residents in the greater Copenhagen area and all studies collected baseline information on socioeconomic variables, lifestyle and health by self-administered questionnaires followed by physical examinations (anthropometric measures, blood pressure etc.) and blood samples. The MONICA I–III cohorts included specific birth cohorts of men and women aged 30, 40, 50 and 60 years in 1982–84 (N=3,785; response rate ≈79%), 1987–88 (N=1,504; response rate ≈75%) and 1991–92 (N=2,027; response rate ≈69%). MONICA III also included 70-year olds. The 1936-cohort consisted of men aged 45-years at baseline in 1981–82 (N=992; response rate ≈84%). INTER99 included birth cohorts of men and women in five-year age intervals from 30 to 65 at baseline in 1999–2001 (N=6,784; response rate ≈52%).

The variables of the different cohorts were pooled based on a stepwise harmonization procedure [63], which involved iterative rounds of discussion among members of the SIC steering committee and generation of formal pairing rules to create each variable. In this way, refinement of the harmonized variables was ensured [59]. In total, the pooled SIC cohort included 83,006 men and women aged 20-98 years at baseline. For the analyses in this thesis, all postmenopausal women – defined as women aged 50+ years – who were free of cancer (other than non-melanoma skin cancer) at baseline and who were born after 1920 (due to lack of available information on sociodemographic variables from the registers for women born before 1920) were included. In total, 33,562 women fulfilled these criteria.

Table 1
Overview of prospective studies on mediating effects of the relation between SEP and postmenopausal breast cancer

First author	Population, N (cases)	Study period	Age at baseline	Measure of SEP	Mediators	Breast cancer	Findings, RR (95% CI)
Braaten (2004) [14]	Norway/ Sweden 102,860 (1,090)	1991–1999	30–50 yrs	Education	Parity, Age at first birth, BMI, Height, Age at menarche, Oral contraceptive use, Alcohol consumption	Overall BC	RR=1.51 (1.05–2.16) for postmenopausal BC among highest educated (16+ yrs.) vs. lowest (7–9 yrs). The RR was reduced to 1.09 (0.74–1.61) after adjustment for all mediators. Predominantly due to parity and age at first birth.
Braaten (2005) [15]	Norway 93,638 (3,259)	1991–2001	30–69 yrs	Education	Parity, Age at first birth, Alcohol consumption, BMI, Screening participation, Oral contraceptive use, Current HT use	Overall BC	RR=1.46 (1.19–1.79) for BC among highest educated (16+ yrs.) vs. lowest (7–9 yrs.). Multiple adjusted RR=1.11 (0.89–1.38) where reproductive patterns accounted for 26%, alcohol consumption for 23% and the remaining factors between 3–7% each.
Danø (2004) [9]	Denmark 674,084 (22,884)	1970–1998	20–39 yrs	Education Socioeconomic group (≈occupation)	Age at first birth, Parity	Overall BC	RR=1.38 (1.31–1.45) for BC incidence in women with 12+ yrs. vs. ≤7 yrs. of education. Reduced to RR=1.26 (1.20–1.33) after adjustment for age at first birth and parity. RR=1.09 (0.95–1.26) for BC incidence in academics vs. salaried employees. Reduced to 1.05 (0.91–1.21) after adjustment for age at first birth and parity.
Gadeyne (2012) [16]	Belgium 2.25 mio. (8,224)	1991–1995	35–79 yrs.	Education	Parity, Age at first birth	Overall BC (mortality)	RR=1.16 (1.06–1.28) for postmenopausal BC mortality among highest educated ('tertiary' education) vs. lowest (no or primary) education, which reduced to 1.06 (0.96–1.16) after adjustment for the mediators.
Heck (1997) [11]	USA 6,032 (229)	1971–1993	25–74 yrs.	Education	Age at first birth /nulliparity, Age at menarche, Age at menopause, Oral contraceptive use, HT use, Alcohol consumption, BMI, Height	Overall BC	RR=2.3 (1.2–4.3) for highest (16+ yrs.) vs. lowest (<12 yrs.) educated women. Reduced to RR=1.9 (1.0–3.4) after adjustment for nulliparity/age at first birth and to RR=1.5 (0.8–2.7) after adjustment for all mediators.
Larsen (2011) [13]	Denmark 23,111 (907)	1993–2006	50–64 yrs.	Education Income Occupation	Parity, Age at first birth, HT use, Alcohol consumption, BMI	Overall BC	RR=1.20 (1.01–1.42) for higher vs. basic/high school education. Reduced to RR=1.06 (0.88–1.26) after adjustment. RR=1.46 (1.07–2.00) for self-employed vs. unskilled worker. Reduced to RR=1.36 (0.99–1.86) after adjustment. RR=1.12 (0.89–1.41) for highest income quartile vs. lowest. Reduced to RR=1.07 (0.85–1.34) after adjustment.
Menvielle (2011) [20]	Europe (EPIC) 102,721 (2,389)	1992–1999	50+ yrs.	Education	Parity, Age at first birth, Breast feeding, Age at menarche, Oral contraceptives use, Height, BMI, Alcohol consumption, Physical activity	Overall BC Invasive/ in situ	For invasive breast cancers: RR=1.20 (1.05–1.37) for university or postsecondary vocational vs. primary education. Reduced to RR=1.11 (0.97–1.27) after adjustment for reproductive history and to 1.00 (0.87–1.15) after adjustment for all risk factors. Larger inequalities for in situ cancers which remained after adjustment for all risk factors.
Palmer (2012) [17]	USA 55,895 (1,343)	1995–2009	21–69 yrs.	Education Neighborhood SES	Parity, Age at first birth, Lactation, Age at menarche, Family history of BC, Oral contraceptive use, Age at menopause, HT use, BMI, Alcohol consumption, Physical activity, Geographic region, Mammography use	Overall BC ER status	For *overall* BC: RR=1.17 (0.99–1.37) for highest (17+ yrs.) vs. lowest education (<13 yrs.). Reduced to RR=1.06 (0.90–1.25) after adjustment for parity and age at first birth and to RR=1.02 (0.86–1.21) after further adjustment for the remaining factors. For *ER+* BCs: RR=1.44 (1.14–1.82) for highest (17+ yrs.) vs. lowest education (<13 yrs.). Reduced to RR=1.25 (0.97–1.60) after adjustment for parity and age at first birth and to RR=1.14 (0.88–1.48) after further adjustment for the remaining factors. Similar results for neighborhood SES
Pudrovska (2013) [18]	USA 3,682 (297)	1975–2011	36 yrs.	Occupation	Adiposity,Reproductive history, HT use, Social stress (work hours, job autonomy and authority, job satisfaction etc.)	Overall BC	RR=1.72 (1.25–2.36) for professionals vs. housewives. Reduced to RR=1.59 (1.15–2.20) after adjustment for reproductive history but not affected by adiposity. Social stress also did not affect the RR. RR=1.57 (1.02–2.42) in managerial occupation vs. housewives which was reduced to 1.42 after adjustment for job authority but not affected by reproductive factors.
Strand (2005) [21]	Norway 512,353 (2,052)	1990–2001	35–54 yrs.	Education	Age at first birth, Parity	Overall BC (mortality)	RR=1.25 (1.10–1.41) for BC deaths among women with >12 yrs. of education vs. <10 yrs. Reduced to RR=1.20 (1.06–1.36) after adjustment for parity and to RR=1.08 (0.95–1.23) after adjustment for age at first birth.

Abbreviations: BC, breast cancer; BMI, body mass index; EPIC, European Investigation into Cancer; ER, estrogen receptor; HT, hormone therapy; N, number of participants; PR, progesterone receptor; RR, relative risk; SES, socioeconomic status; vs., versus; yrs., years.
For studies reporting separate findings according to menopausal status, only postmenopausal results are included.

In study III, data on endogenous sex hormone levels were included from a randomly selected subsample of the CCHS (N=1,150). Blood samples were drawn at baseline (1981–83) and stored at -20°C. Duplicate levels of free testosterone and 17β-estradiol (E2) were measured in serum and the means of the two values were applied in the analyses [58].

The Women's Health Initiative (WHI)

In 1991, the American National Institutes of Health established the Women's Health Initiative, which included four clinical trials and an observational study (WHI-OS) [64,65]. The aim was to examine determinants of cardiovascular disease, cancer and other health problems of postmenopausal women. Women were considered eligible based on the following criteria: age between 50–79 years, accessible for follow-up and an estimated survival of at least 3 years. Information on demographic and lifestyle factors, medical history and medication use was collected by a questionnaire and a physical examination at baseline. Blood samples were collected following an overnight fast of at least 12 hours with separated sera stored at -70°C within two hours of collection [66]. In total, the WHI-OS included 93,676 postmenopausal women.

The study population for Paper II included data from two case-cohort ancillary studies of the WHI-OS with measurements of baseline endogenous estradiol (E2) levels [53,55] and fasting insulin determinations [55]. Estradiol was measured in all participants, whereas insulin was only assessed in a subsample of non-diabetics in one of the ancillary studies (N=791) [55]. The procedure of combining the two subsamples of the WHI-OS is described in detail in Paper II [67]. Combining the two studies yielded a total of 601 breast cancer cases and 1,000 subcohort members.

Assessment of socioeconomic position

All participants in the SIC cohort were linked to national registries through a unique personal identification number. Sociodemographic information was available from Statistics Denmark from 1980 and onwards. SEP was defined as highest attained educational level of the woman one year before baseline and categorized as "low" (8–11 years, basic schooling), "medium" (11–14 years, upper secondary or vocational training) and "high" (15+ years) educational level.

In the WHI, educational attainment was assessed by a baseline self-administered questionnaire in 11 specified categories ranging from not attending school at all (<1 year of grade school) to obtaining a higher educational degree (Ph.D., M.D. etc.).

Assessment of lifestyle, BMI, HT use and reproductive factors

Information on lifestyle and reproductive factors was assessed by self-administered questionnaires in all cohorts [59,66]. Alcohol was assessed as consumption of beer, wine and spirits in response categories of "never/almost never", "monthly", "weekly" and "daily" as well as the average number of drinks per week within these categories. In the DCHS, leisure time physical activity was assessed as the average number of hours spent in the past year on various types of activity (e.g., cycling, walking) along with number of hours becoming sweaty or short of breath as a result of these activities. Similarly, the CCHS and the CRCPH assessed the weekly level of physical activity during the past year in four categories ranging from being almost entirely inactive to engaging in vigorous physical activity. The WHI-OS asked about the frequency, duration and intensity of exercise. Metabolic equivalent values (METs) were assigned for the activities and multiplied by the hours exercised at that intensity level per week and summed for all types of activities. Smoking was assessed in categories of never, past and current, and according to daily tobacco use among current smokers in all cohorts. Reproductive factors included self-reported parity and age at first birth. HT use was classified as current HT use (yes versus no). All included cohorts measured baseline weight and height at the physical examination.

Assessment of postmenopausal breast cancer

The *SIC database* was linked with various Danish population-based registers. Time and type of cancer diagnosis was obtained from the Danish Cancer Registry, in which breast cancer is defined according to the International Classification of Diseases (ICD) versions 7 and 10 (ICD7 code 174 and ICD10 code C50) coding schemes [68]. Thus, histologic disease types (ductal, lobular etc.) were considered jointly. The estrogen receptor (ER) determinations of the tumors applied in Paper III were obtained from the Danish Breast Cancer Cooperative Group (DBCG). Since 1977, the DBCG clinical database has covered all breast cancer cases in Denmark with regard to demographic and histopathological variables, therapeutic interventions and follow-up. Cases were classified as positive ER status if immunohistochemical staining revealed >10% ER positivity [69]. Information on emigration and deaths was obtained from the Registry for Population Statistics and Statistics Denmark.

The *WHI* collected information on breast cancer incidence through annual self-administered questionnaires. Subsequently, breast cancer status and clinical and pathological characteristics of the tumors were confirmed through centralized reviews of hospital discharge summaries, operative reports, history and physical examination, radiology reports and oncology consultant reports. Deaths were documented by death certificates and medical records, as well as data linkage to the American National Death Index and the National Center for Health Statistics [70].

In both cohorts, participants were followed from baseline to the date of first breast cancer event, the date of death, emigration or end of follow-up, whichever occurred first.

Identification of confounders

Potential confounders of the relation between SEP and breast cancers were identified through careful consideration of the underlying causal relations based on prior knowledge [71]. The directed acyclic graph (DAG) in the appendix depicts these hypothesized pathways. The model emphasizes underlying causes of SEP and postmenopausal breast cancer in order to evaluate potential confounders of the relation and is therefore not completely exhaustive regarding internal causal relations between variables and regarding intermediate biological processes occurring on the pathway from for example lifestyle factors to postmenopausal breast cancer. This model explicitly states the assumptions underlying the statistical analyses of this thesis, and the inclusion of both measured and unmeasured factors serves as a basis for the discussion of residual confounding.

Statistical methods

The Aalen Additive Hazards Model

The analyses in all three papers were based on the Aalen additive hazards model [72]. This model provides an estimate of the additional number of breast cancer cases associated with a given risk factor (absolute effects) and allows for direct comparison of these numbers across strata of other factors. Like the standard Cox proportional hazards model, the Aalen model with time-constant effects has an unspecified baseline hazard, and the effect of each covariate is modeled by a single parameter. Thus, the two models are equally flexible, but the interpretations of the effect measures are different. For a given exposure, for example SEP, the absolute effect (i.e., rate difference) of high educational level provides an estimate of additional breast cancer cases per 100,000 person-years at risk in the highly educated women compared with women of low educational attainment (adjusted for confounders) [72], whereas the relative effect of a Cox model provides an estimate of how many times greater the hazard is among women with a high versus low educational level (i.e., the hazard ratio).

Assessment of mediation

Mediation analysis and thus the distinction between total, direct and indirect effects are embedded in the counterfactual outcomes framework [24,25]. In this setting, the total individual causal effect (TE) of an exposure, A, on an outcome, Y, is defined as the hypothetical contrast between the outcome that would have been observed under exposure A=a versus A=a^* for the same individual:

$$TE = Y_a - Y_{a^*}$$

Definitions of direct and indirect effects

Traditionally, estimates of mediating effects have been derived from two regression models – one excluding and one including the potential mediator of interest [73]. The results of these two analyses provide the *total* effect of the exposure on the outcome (the unadjusted) and the *controlled direct* effect (CDE) of the exposure on the outcome (the mediator-adjusted), respectively. The term 'controlled' refers to the counterfactual contrast between the two setups in which the exposure is set to A=a and A=a^*, but the mediator is kept fixed ('controlled') at the level M=m [24,25]:

$$CDE = Y_{a,m} - Y_{a^*,m}$$

The term 'controlled' direct effect refers to the effect of the exposure on the outcome when fixing the mediator at some specific level, for example the effect of SEP on breast cancer if an intervention prevented alcohol consumption among all women. Controlled effects do not allow for a straightforward definition of indirect effects [24,25]. This is due to the fact that if the exposure interacts with the mediator to cause the outcome, the controlled direct effect does not equal the total effect, even if there is no effect of the exposure on the mediator. The controlled direct effect will depend on the level at which the mediator is fixed.

The mediation analyses in this thesis were based on the computation of *natural direct* and *indirect* effects as originally proposed by Robins & Greenland [25] and Pearl [24]. The natural direct effect (NDE) differs from the controlled in that the mediator, M, takes the hypothetical value it would have taken under the reference A=a^*:

$$NDE = Y_{a,M(a^*)} - Y_{a^*,M(a^*)}$$

Natural direct effects are in other words defined as the change in outcome that would be observed if the exposure could be changed or fixed (e.g., from *high* educational level to *low*), but leaving the mediators unchanged (corresponding to *high* educational exposure). Thus the natural direct effect encompasses the effect of A on Y through other pathways not involving M. Likewise, natural indirect (i.e., mediated) effects are defined as the change in outcome when the exposure is kept fixed, but the mediator is changed to the value it would take if the exposure was changed:

$$NIE = Y_{a,M(a^*)} - Y_{a,M(a)}$$

The natural indirect effect thus represents the effect of A on Y due to the effect of A on M. The total effect decomposes into the natural direct effect and the natural indirect effect even in situations of nonlinearities and exposure-mediator interaction [24,25].

The total, direct and indirect effects described above are counterfactual measures, and thus not possible to quantify in reality for each person. However, average/population causal effects can be obtained assuming that there are no unmeasured confounding of the exposure-mediator, exposure-outcome and mediator-outcome relation and no confounding of the mediator-outcome relation affected by the exposure (exposure-dependent confounding).

In Paper I and II, natural direct and indirect effects were directly parameterized following the method of Lange, Vansteelandt & Bekaert [27] and the method of Lange & Hansen [26], respectively. The method by Lange, Vansteelandt & Bekaert applied in Paper I, combines effect estimates from two models in three steps: 1) fitting a multinominal logistic regression model of the mediator on exposure and confounders of this relation; 2) constructing weights based on the probabilities of actually obtaining the mediator (from the actual exposure and the auxiliary exposures) and 3) fitting a marginal structural Aalen additive hazards model using these weights to obtain natural direct and indirect effects. The method by Lange & Hansen [26] applied in Paper II combines the Aalen additive hazards model of the direct effect of exposure on outcome (i.e., adjusted for the mediator and potential confounders) with a linear regression model for the exposure-mediator relation. The indirect effect is given by the product of these two parameter estimates. In both approaches, the total effect is derived by summing the direct and the indirect effects. The mediated proportion is given by the indirect effect divided by the total effect. Confidence limits for the direct effect are given in the model output whereas limits for the indirect and total effects as well as mediated proportions (indirect divided by total effect) are computed by parametric bootstrap.

Mediated interactive effects

Assuming no interaction between exposure, SEP, and the mediators on the risk of outcome means that the social inequality in postmenopausal breast cancer is the same across strata of the mediators. In this case, the controlled direct effect equals the natural direct effect. When exposure and mediator interact, the natural direct and indirect effect still sum up to the total effect,

but there are traditionally two ways of accounting for this interactive effect in the mediation framework, depending on which of the involved parameters are ascribed the interaction [74]. Traditionally, the total effect has been decomposed into a *pure direct* effect and a *total indirect* effect, meaning that the interaction is embedded in the indirect effect, or equivalently, a *total direct* effect and a *pure indirect* effect when the direct effect accounts for the interaction [25]. Recently, Vanderweele [74] has suggested a three-way decomposition into a direct, an indirect and an interactive effect, which was applied in this thesis:

$$TE = Y_1 - Y_0 = (Y_{1,m(0)} - Y_{0,m(0)}) + (Y_{0,m(1)} - Y_{0,m(0)}) + (Y_{11} - Y_{10} - Y_{01} + Y_{00})(M_1 - M_0)$$

Thus, in this setup, the total effect decomposes into a pure direct effect, a pure indirect effect and the *mediated interactive* effect given by the product of an additive interaction between the exposure and the mediator on the outcome ($Y_{11} - Y_{10} - Y_{01} + Y_{00}$) and the effect of the exposure on the mediator ($M_1 - M_0$). The mediated interactive effect is present *only* when there is an exposure-mediator interaction and an effect of the exposure on the outcome.

The interpretation of the mediated interactive effect refers back to the differential vulnerability mechanism of social inequality in health outcomes as presented in Figure 2. For instance, the indirect effect of SEP on breast cancer risk through physical activity may vary across SEP strata if other factors such as childhood circumstances affected this relation.

Intertwined pathways

A strong underlying assumption of the applied methods for assessing mediating effects is that all pathways are independent [75-77]. This means that high BMI, reproductive patterns and lifestyle factors for example are assumed to mediate the pathway from SEP to breast cancer independently as depicted in Figure 3. This is a highly unrealistic assumption, since we know that these factors are closely related (cf. the appendix).

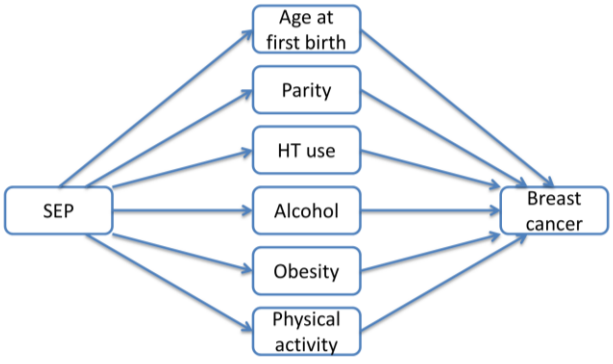

Figure 3
Underlying assumptions of the mediation analysis: Distinct pathways from SEP to breast cancer.

For instance, obesity is probably highly dependent on the level of physical activity and vice versa. Although we do have information on these factors, available methods of estimating the mediating effects, do not allow for the adjustment of other intermediate factors intertwined with this relation [27,75,76]. In a recent paper [76], building on the work by Lange et al [27], the mediation analysis method was extended to include more mediators in the same model, but still assuming independent pathways. The authors suggest a method for investigating whether the mediators are intertwined by a regression analysis of the mediator (M1) on the exposure (E) and the potential intertwined mediator (M2). If the M2-parameter is statistically insignificant (in a reasonably large dataset), non-intertwined pathways can reasonably be assumed [76]. If these pathways prove to be intertwined, the extent may be evaluated by assessing mediation through a variable combined by the intertwined factors [75]. The degree to which the mediating effect of this combined variable differs from the sum of the individual mediating effects provides an estimate of the magnitude of this problem. The results section of this thesis includes estimates of the degree to which the examined pathways are intertwined based on the suggested method described above.

SUMMARY OF RESULTS

This chapter summarizes the main findings from the three papers and the results of the additional sensitivity analyses addressing intertwined mediating pathways. In Paper I, the paths from SEP to breast cancer through HT use, reproductive patterns, lifestyle factors and high BMI (paths A–E in the conceptual model, Figure 2.2) were addressed. Paper II covers the paths from lifestyle factors and high BMI (paths F–H) through estradiol and insulin to breast cancer, and finally, Paper III addresses interaction between the mediators HT use, alcohol consumption and BMI in relation to breast cancer.

Is the incidence of breast cancer socially skewed?

Among women with a medium versus low educational level in the SIC cohort, 70 (95% CI: 29, 112) additional breast cancer cases per 100,000 person-years at risk were observed (Figure 4). Correspondingly, 74 (95% CI: 22, 125) additional cases were observed in women with a high versus low educational level. In relative terms, a medium educational level was associated with a relative risk of 1.19 (95% CI: 1.07, 1.32) compared to a low educational level and similarly, a high educational level was associated with a relative risk of 1.21 (95% CI: 1.05, 1.39).

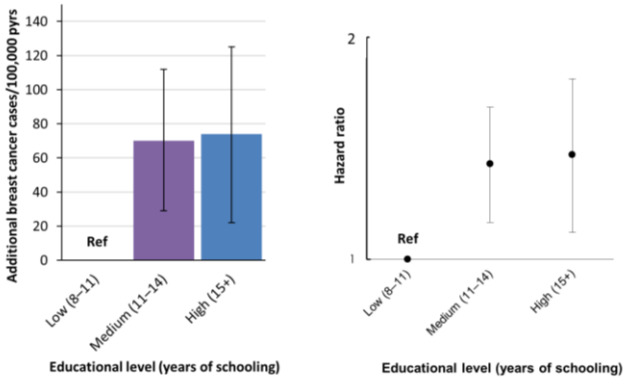

Figure 4
Social inequality in postmenopausal breast cancer as measured by educational attainment in absolute and relative terms in the SIC cohort (adjusted for age and study), N=33,562. Pyrs, person-years.

Are risk factors of breast cancer unevenly distributed across social groups?

In Figure 5, the different risk factors of breast cancer are presented according to educational level in the SIC cohort. The proportion of obese women ranged from 18% in the lowest educated women to 9% in the group of women with the highest education. Likewise, the social gradient in physical inactivity was reversed with 20% inactive women in the group with low education versus 16% in the highly educated group. High alcohol consumption (7+ drinks per week) was highly skewed across the social groups with 29% among the low educated versus 48% percent in the highly educated women. Correspondingly, a positive association was observed for nulliparity (10% and 16% in low and high education, respectively), older age at first birth (12% versus 17%) and to a lesser extent HT use (29% versus 31%).

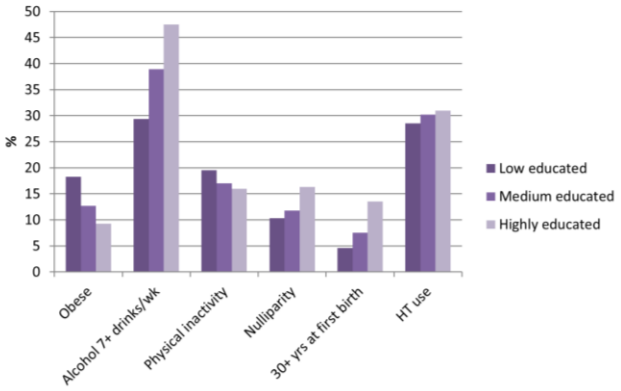

Figure 5

Postmenopausal breast cancer risk factors according to educational level in the SIC cohort, N=33,562. Wk, week; yrs, years.

To what extent do BMI, HT use, lifestyle and reproductive factors mediate the social inequality in breast cancer?

In Paper I, we examined the social inequality in breast cancer among postmenopausal women in the SIC cohort and the mediating effects of BMI, HT use, lifestyle and reproductive factors. The social inequality in postmenopausal breast cancer observed among women in the SIC cohort is given in Figure 4 above. Below, Figure 6 presents the observed additional breast cancer cases according to BMI, lifestyle and reproductive factors and HT use in the SIC cohort.

As expected, the analyses showed an association between alcohol consumption, reproductive factors and HT use and postmenopausal breast cancer. An alcohol consumption of 7+ drinks versus <1 drink per week was associated with 123 (95% CI: 69, 178) additional breast cancer cases per 100,000 person-years, nulliparity versus 3+ children was associated with 180 (95% CI: 108, 251) additional breast cancer cases per 100,000 person-years and 155 (95% CI: 80, 230) additional cases compared to women giving birth before the age of 25 years. Current HT use was associated with 270 (95% CI: 222, 318) additional breast cancer cases per 100,000 person-years compared to women who did not report current HT use. The risk of breast cancer did not seem to vary by BMI or physical activity.

The mediation analyses suggested that alcohol consumption me-

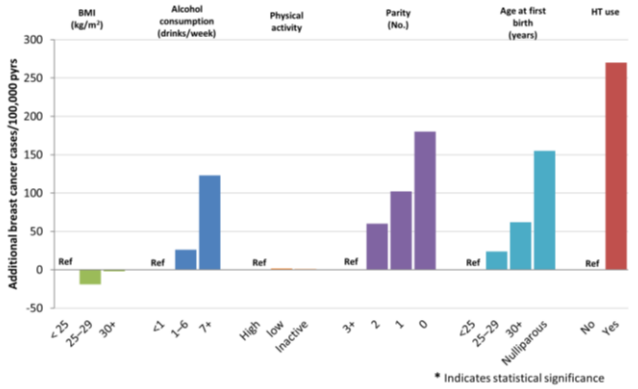

Figure 6

Additional breast cancer cases according to BMI, lifestyle and reproductive factors in the SIC cohort (adjusted for educational level, age and study), N=33,652. Pyrs, person-years.

diated 26% (95% CI: 14%, 69%) of the social inequality in breast cancer. Correspondingly, the mediated proportion of parity was 19% (95% CI: 10%, 45%), age at first birth 32% (95% CI: 10%, 257%) and HT use 10% (95% CI: 6%, 18%) when comparing highly educated to low (Figure 7). High BMI and physical inactivity did not appear to mediate the relation between educational level and postmenopausal breast cancer; however, heterogeneity of effects of educational level was observed across strata of physical activity (*P* for interaction = 0.01). Decomposing this interaction between SEP and physical activity showed a mediated effect through physical activity of 2 (95% CI: -1, 5) additional cases for high compared to low educational level, and a mediated interactive effect of -10 (95% CI: -16, -4). This may mean that women of low educational level are less vulnerable to physical inactivity than women of high educational level (cf. Figure 1), but could on the other hand likely be a chance finding or a result of differential misclassification, as discussed later.

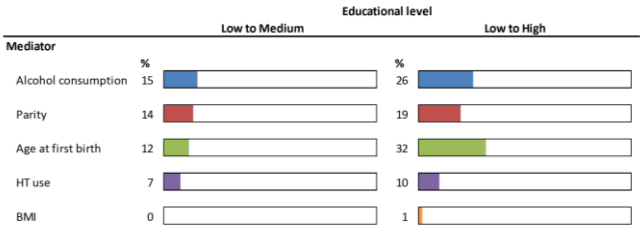

Figure 7

Mediated proportions by each risk factor of the relation between educational level and breast cancer in the SIC cohort (adjusted for age and study), N=33,652.

It must be stressed that the mediated proportions for each risk factor were derived from separate models and thus cannot be added to a total sum of mediating effects due to potentially intertwined pathways (cf. Figure 1, Paper I).

Is the pathway from high BMI and alcohol consumption to breast cancer mediated by estradiol and insulin?

In Paper 2, we addressed the effects of high BMI and alcohol consumption on postmenopausal breast cancer and the indirect

effects through estradiol and insulin (paths F, G and H in the conceptual model Figure 2). A high serum estradiol level is a well-established risk factor for breast cancer, but insulin has also been suggested to be a significant, independent contributor to the relationship between high BMI and breast cancer risk.

In a subsample of women from the WHI-OS, a 5-unit increase in BMI, and to a lesser extent alcohol consumption, were associated with higher levels of estradiol. In the mediation analyses including all breast cancer cases, a 5-unit increase in BMI was associated with 50 (95% CI: 23, 77) additional breast cancer cases per 100,000 person-years, of which 24% (95% CI: 3%, 68%) could be ascribed to higher estradiol levels (Figure 8). Correspondingly, an alcohol intake of 7+ drinks per week compared to abstinence was associated with 178 (95% CI: 60, 297) additional breast cancer cases per 100,000 person-years, however, the mediated effect of estradiol on this relation was minimal (2%; 95% CI: -1%, 11%).

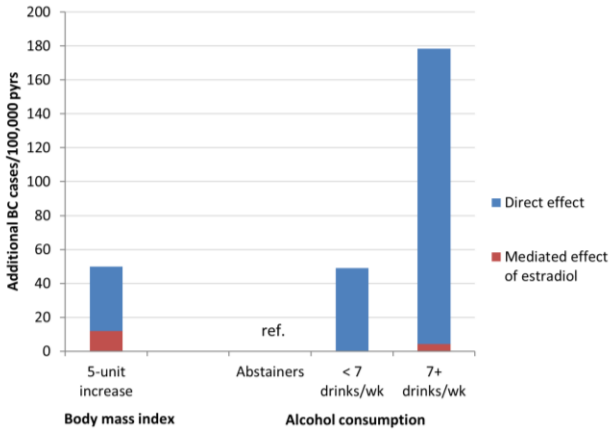

Figure 8

Direct effect of BMI and alcohol consumption and mediated effects through estradiol in the WHI-OS subsample (adjusted for age, ethnicity, education, marital status, physical activity, smoking, age at menarche/menopause, parity, age at 1st birth and 1st degree relative with BC), N=1,601. BC, breast cancer; pyrs, person-years; ref, reference; wk, week.

The potential mediating role of estradiol was further investigated by restricting the analyses to ER positive breast cancers. In these analyses, the contribution of each exposure was similar to the main analysis. The indirect effect of estradiol, however, was remarkably higher for the BMI analysis with 49% (95% CI: 19%, 161%) of the total effect of BMI mediated through estradiol. The associations observed for ER negative breast cancer cases were quite different and statistically insignificant, indicating that the effects on ER-positive breast cancer cases primarily drove the results from the main analysis. However, very few cases (N=126) were ER-negative, and thus, conclusions should be drawn with caution.

The analyses on the subsample of the population with insulin measurements showed that the effect of high BMI on postmenopausal breast cancer risk was partly mediated by estradiol, but to a much higher degree by insulin (Figure 9). The total effect of a 5-unit increase in BMI was 52 (95% CI: 12, 91) additional breast cancer cases per 100,000 person-years at risk. Of this total effect, an indirect effect of 11 (95% CI: -2; 25) additional breast cancer cases per 100,000 person-years was observed for the pathway through estradiol and 34 (95% CI: 9, 59) additional breast cancer cases per 100,000 person-years were observed through the insulin pathway corresponding to 21% (95% CI: -4%, 119%) and 66% (95% CI: 14, 273), respectively. The proportion mediated by estradiol in this analysis corresponded to the analysis without insulin in the model, which suggests that the two factors represent distinct pathways.

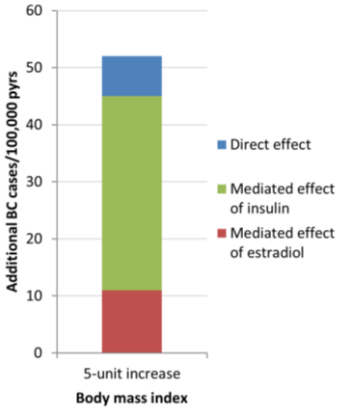

Figure 9

Direct effect of BMI and mediated effects of estradiol and insulin on BC in the WHI-OS subsample (adjusted for age, ethnicity, educational level, marital status, physical activity, smoking, age at menarche/menopause/1st birth, parity and first-degree relative with BC), N=791. BC, breast cancer; pyrs, person-years.

Does hormone therapy use interact with alcohol consumption and BMI according to breast cancer risk?

In Paper III, the objective was to explore the combined effects of HT use and high alcohol consumption as well as high BMI based on the hypothesis that these combinations could increase breast cancer risk beyond the sum of the separate effects.

Evidence of interaction between these factors was observed in this study. In stratified analyses, overweight compared to normalweight was associated with 54 (95% CI: 6, 102) additional breast cancer cases per 100,000 person-years in non-HT users and 121 fewer breast cancer cases (95% CI: -216; -26) per 100,000 person-years in current HT users (*P* for interaction=0.003). A high alcohol consumption (7+ drinks/week) compared to abstinence was associated with 72 (95% CI: 12, 131) additional cases in non-HT users and 180 (95% CI: 42, 319) additional cases in current users per 100,000 person-years at risk (*P* for interaction=0.02).

The combined effects of HT/BMI and HT/alcohol consumption are presented in Figure 10. When combining the effects of HT use with BMI, markedly higher risks of BC were observed in HT users across all BMI groups compared to normalweight non-HT users. For example, 59 (95% CI: -4, 122) additional breast cancer cases per 100,000 person-years were observed among obese non-HT users, and correspondingly 330 (95% CI: 187, 477) additional cases among obese HT-users compared to normalweight non-HT users. For alcohol consumption combined with HT use, a markedly elevated risk of 432 (95% CI: 339, 524) additional breast cancer cases was observed compared to abstinent, non-users of HT.

The analyses according to ER status of the tumor showed, that these effects were largely restricted to ER-positive breast cancer cases. For example, HT use combined with an alcohol consump-

tion of 7+ drinks per week was associated with 360 (95% CI: 285, 436) additional ER-positive cases per 100,000 person-years, 46 (95% CI: 9, 82) ER-negative breast cancer cases and 27 (95% CI: -6, 60) cases of unknown receptor status per 100,000 person-years compared to abstinent non-HT users.

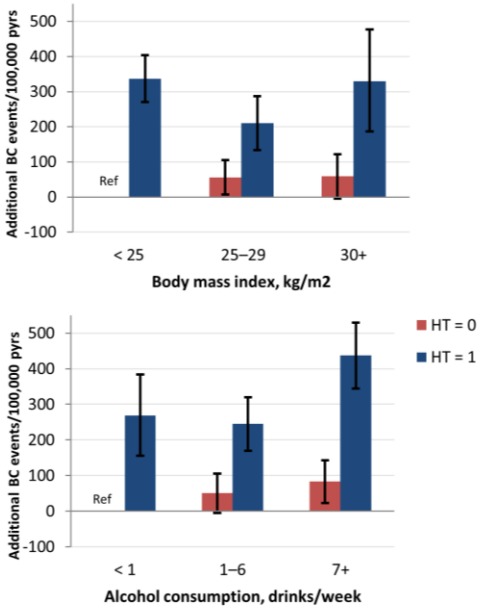

Figure 10

Combined effects of HT use, BMI and alcohol consumption on postmenopausal breast cancer in the SIC cohort (adjusted for age, study, educational level, parity, BMI (analysis of alcohol consumption), alcohol consumption (analysis of BMI), smoking, parity and physical activity). BC, breast cancer; pyrs, person-years; ref, reference.

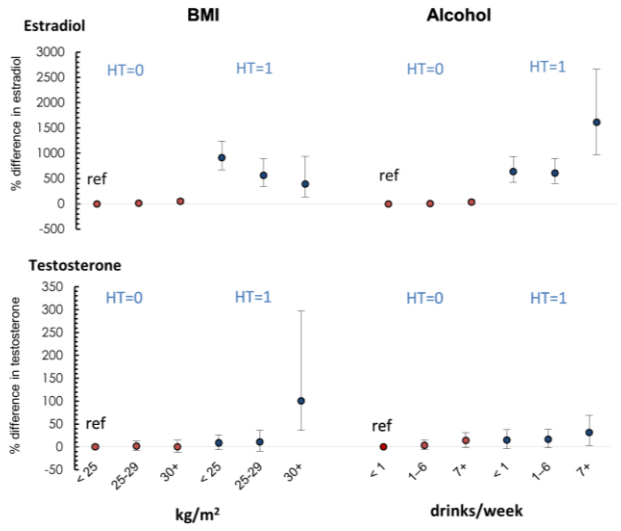

Figure 11

Relative differences in 17β-estradiol and testosterone levels by BMI and alcohol consumption according to HT use in a subsample of the SIC cohort (adjusted for age, educational level, parity, BMI (analysis of alcohol consumption), alcohol consumption (analysis of BMI), smoking, physical activity, parity and time of blood draw). Ref, reference.

Are the investigated pathways intertwined?

As described in the Methods section, intertwined pathways were investigated by a regression analysis of the mediator (M1) on the exposure (E) and the potential intertwined mediator (M2). If the M2-parameter is statistically insignificant, non-intertwined pathways can reasonably be assumed [76]. Table 2 presents the p-values for this analysis in the SIC cohort. According to this, the pathways investigated in Paper I are likely intertwined.

Table 2

P-values for intertwined pathways from a multinominal logistic regression analysis of the mediator (M1) on the exposure (educational level) and the potential intertwined mediator (M2).

M2 \ M1	Physical activity	Alcohol	BMI	HT	Parity	Age at 1st birth
Physical Activity						
Alcohol	0.09					
BMI	<0.001	<0.001				
HT	0.16	<0.001	<0.001			
Parity	0.31	<0.001	<0.001	<0.001		
Age at 1st birth	0.07	<0.001	<0.001	0.01	<0.001	

Adjusted for age and study. Note that the sample sizes for analyses including HT use and age at first birth are reduced due to missing information on these variables (N=2,641 for HT use and N=3,475 for age at first birth).

Table 3 shows the results of the sensitivity analyses combining the potentially intertwined factors for high versus low educational level. The mediated proportion for the variable combined by alcohol and BMI was 27%, which corresponds well to the estimated separate proportions (26% and 1%, respectively). The same applies for most of the other combinations. However, the combination of alcohol with parity gives a mediated proportion of 39%, which is somewhat lower than expected from the separate effects (26% and 19%, respectively). Since the two reproductive factors parity and age at first birth both include the category of nulliparous women, their effects are obviously intertwined, which is also evident from this analysis in which the separate mediated proportions of 19% and 32%, respectively, are reduced to 21% in combination. A mediated proportion of 13% was observed by the combination of BMI and physical activity, which is more than expected from the individual proportions (1% and 3%, respectively). This could be a chance finding or perhaps indicate misclassification of the individual effects of these factors which may be reduced by the combination of the two. Overall, the problem of intertwined pathways does not seem to seriously affect the main conclusions.

Discussion

In this section the findings of the three papers are discussed in relation to previous studies on SEP and postmenopausal breast cancer and in relation to the internal and external validity of the findings. Sensitivity analyses are presented to substantiate the conclusions.

Table 3

Mediated proportions (95% CI) for combinations of mediators (high versus low education) in the SIC cohort.

M1 / M2	Physical activity	Alcohol	BMI	HT	Parity	Age at 1st birth
Physical Activity	3% (-1%,12%)					
Alcohol	33% (13%,107%)	26% (14%,69%)				
BMI	13% (-1%,51%)	27% (10%,109%)	1% (-24%,14%)			
HT	13% (6%,54%)	36% (17%,135%)	6% (-8%,140%)	10% (6%,18%)		
Parity	21% (7%,70%)	39% (19%,130%)	21% (6%,90%)	31% (15%,126%)	19% (10%,45%)	
Age at 1st birth	25% (-23%,148%)	70% (-120%,484%)	32% (-120%,230%)	52% (-145%,339%)	21% (-29%,154%)	32% (10%,257%)

Adjusted for age and study. The analyses including physical activity account for the interaction between SEP and physical activity. Note that the sample sizes for analyses including HT use and age at first birth are reduced due to missing information on these variables (N=2,641 for HT use and N=3,475 for age at first birth).

Main findings

In general, the findings of this thesis support the hypothesis of a social gradient in postmenopausal breast cancer incidence and provide some evidence of mediation through the expected pathways. Below the main results of the three studies are summarized referring to the paths in the conceptual model (Figure 2).

- A higher risk of postmenopausal breast cancer was observed among women with a medium or high educational level compared to women with a low educational level.
- The effect of educational level on postmenopausal breast cancer appeared to be partly mediated through HT use (path B), fertility patterns (path C) and alcohol consumption (path D). The mediating effect of physical activity (path D) was modified by educational level suggesting that women with a high educational level may be more vulnerable to physical inactivity than women of low educational level. BMI did not mediate the education-breast cancer relation (path E).
- A substantial effect of alcohol consumption on postmenopausal breast cancer was observed, but this relation did not, as hypothesized, seem to be mediated by endogenous estradiol levels (path F).
- A higher risk of postmenopausal breast cancer was also observed among women with a higher BMI in the WHI subsample of non-HT users, and the analyses showed mediating effects of both estradiol (path G) and insulin (path H).
- The effect of HT use on breast cancer interacted with alcohol consumption and BMI. Combined with alcohol consumption, HT use increased the number of breast cancer cases to a level markedly higher than would be expected from the sum of their separate effects. For BMI combined with HT use, a modest positive association was observed for non-HT users whereas markedly higher risks were observed across all BMI groups in current HT-users with a tendency towards a U-shaped relation.

Comparison with previous studies

As reviewed in a previous section, several previous prospective studies have observed social inequalities in postmenopausal breast cancer. These studies have consistently reported mediating effects of parity and age at first birth corresponding to the magnitudes observed in Paper I [9,11,14-18,20,21]. Our findings of mediated effects through alcohol consumption and HT use – and lack of effect through BMI – are also consistent with the existing litterature [11,13,15,17,20]. Thus, the findings of this thesis support the conclusions of previous studies and add to the field by quantifying the mediated paths with actual mediation analyses, addressing the issue of potential differential vulnerability across socioeconomic groups (i.e., exposure-mediator interactions), and by investigating the degree to which the mediating pathways are intertwined.

The literature on mediating effects of estradiol and insulin on the relation between high BMI, alcohol consumption and postmenopausal breast cancer is limited. In a large pooled analysis of individual data from eight prospective cohorts, the relation between BMI and postmenopausal breast cancer was found to be largely explained by endogenous estrogen levels – in particular by estradiol levels [51]. Another pooled analysis from the European Investigation into Cancer (EPIC) collaboration confirmed this finding [78]. Studies addressing the impact of insulin as a potential mediator of the relation between BMI and postmenopausal breast cancer are fewer. However, in a recent case-cohort study the role of C-peptide was investigated as a biomarker of insulin secretion in the BMI-breast cancer relation [79]. Adjustment for C-peptide reduced the relative risk of breast cancer among overweight versus normalweight postmenopausal women from 1.62 (95% CI: 1.07, 2.46) to 1.48 (95% CI: 0.76, 2.06). Although the C-peptide concentration cannot directly be translated into insulin secretion levels, the study does lend support to the findings of Paper II, in which insulin levels partially explained the excess breast cancer risk among overweight and obese women.

Several previous studies have addressed the association between alcohol consumption and endogenous estrogens [58,80-82], and generally found higher levels among alcohol consumers. However, a statistically significant relation between alcohol consumption and estradiol is not consistently reported [58,80]. In a controlled feeding study, women receiving a moderate alcohol amount of 15–30 g/day (corresponding to 1–2 drinks/day) had significantly increased levels of estrone sulphate and dehydroepiandrosterone sulphate (DHEAS) – but not estradiol – compared to women in the placebo group [46]. Also, a previous study of 128 healthy postmenopausal women examined the correlation between alcohol consumption and estradiol levels by comparing self-reported alcohol information to prospectively collected food records. The relation between estradiol levels and total weekly alcohol intake was confirmed in analyses of alcohol consumption based on the food record data, but not when considering the self-reported alcohol variable [83]. Thus, the previous literature is somewhat inconclusive regarding the association of alcohol consumption and estradiol. The analysis in Paper III, suggested a modestly higher estradiol level with higher alcohol consumption among non-HT users and an extremely elevated estradiol level in women combining high alcohol intake with HT use. The results of the mediation analyses in Paper II do not lend further support to a mediating role of estradiol in the alcohol-breast cancer relation among non-HT users.

The observed interactions between HT use, high BMI and alcohol consumption are consistent with the previous literature. A recent large meta-analysis of epidemiologic studies on BMI, HT use and breast cancer published from 1980–2012, reported a significantly higher risk of breast cancer in never-users of HT with

a higher BMI [84]. Among ever-HT users, a higher risk of breast cancer was observed in cohort studies but not in the case-control studies. In the EPIC cohort, an analysis of the joint exposure of BMI and HT use showed a more than two-fold higher risk of ER+/PR+ breast cancers among current users of HT across all BMI strata compared to normalweight never-users of HT [85], which is consistent with the findings of Paper III. The markedly higher risk of postmenopausal breast cancer in women combining HT use and alcohol consumption is also in accordance with findings of previous studies [29,30,86-88]. However, the findings of Paper III add to the literature by estimating the additional number of cases associated with the combined effects and by testing the underlying biological hypothesis of a combined effect of these factors on estradiol and testosterone levels.

Strengths

This work on mechanisms underlying social inequality in breast cancer differs from the previous literature by addressing the question through a new analytical approach. The methods developed for the SIC project enabled a decomposition of the pathways of interest within the counterfactual framework of natural direct and indirect effects. The papers were further strengthened by the prospective design, the large sample size of the SIC cohort and linkage to population-based registers on disease, death and emigration. The inclusion of the WHI-data provided a relatively large sample of postmenopausal women with hormone determinations which supported the hypotheses on how mediators may affect the risk of postmenopausal breast cancer.

Limitations

There are several points to consider with regards to the internal validity of the three papers. Below, the internal validity is evaluated focusing on the classical sources of bias: selection, misclassification and confounding as well as the problem of exposure-dependent confounding in mediation analyses. Considerations concerning the methodology and the external validity of the findings conclude this section.

Selection bias

Selection bias of an exposure-outcome estimate arises from conditioning on common effects of the exposure and outcome (or other variables) [89]. In prospective cohort studies, selection bias usually arises from differential follow-up, where for example the drop-out rate could be higher among diseased compared to healthy participants (Figure 12). In both the SIC cohort and in the WHI-OS, loss to follow-up was negligible, which makes selection bias due to differential loss to follow-up unlikely. The choice to participate in a cohort study may depend on the SEP of the individual and at the same time be related to the outcome through a common cause such as family history of breast cancer. In the SIC cohort, the older cohorts (the CCHS and the CRCPH) have a high participation rate of 70% and above, but the low participation rate of 35% in the DCHS has been found to be related to SEP, with higher participation among individuals of a higher educational level [61]. If participants at the same time are more likely to have a relative with breast cancer, and thus be at higher risk of breast cancer, the observed social gradient in breast cancer in the SIC cohort could be partly non-causal. However, a register-based study including all Danish women, have shown a similar relation between educational attainment and postmenopausal breast cancer [9]. The same potential problem applies to the WHI-OS in which spurious associations could be observed if non-participants differed from participants with regard to their alcohol consumption or BMI and with regard to their breast cancer risk. In the overall WHI program, approximately 373,000 women were screened for eligibility. In total, 160,000 women were included in one of the trials or the observational study. Details on excluded women are few, but since alcohol addiction was one of the exclusion criteria, the women are likely to differ according to this factor [64]. Selection also occurred within the group of included women, since the WHI-OS consisted of those who were ineligible or uninterested in participating in the WHI trials. However, comparisons of the distribution of risk factors (e.g., self-rated health, BMI and education) across the different trials and the WHI-OS did not indicate major variations [64], and thus selection bias from conditioning on willingness to participate in the observational part of the WHI was likely small.

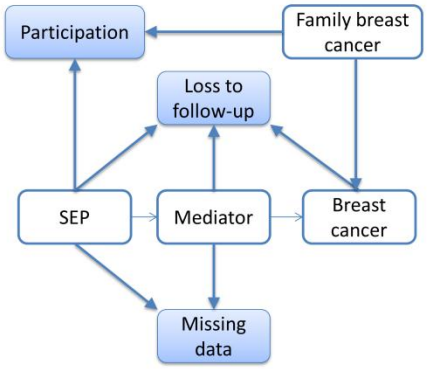

Figure 12
Potential sources of selection bias.

A similar structure may underlie problems of missing data, where a complete case analysis could induce bias if the missing information is a common effect of for example educational level and one of the mediators. In Paper II, multiple imputations were applied to overcome this potential source of bias under the assumption that variables were missing at random conditional on the other predictors and the outcome [90]. However, comparing the complete case analysis to the multiply imputed analysis revealed only negligible differences [67]. In the SIC cohort, approximately 1% of the postmenopausal women were excluded due to missing information. However, the missingness was equally distributed across the SEP groups and thus not likely to have compromised the internal validity of the findings in Papers I and III.

Misclassification of exposure, mediators and outcome

Misclassification occurs when the measured variable is not a good proxy for the actual exposure, confounder, mediator or outcome of interest. This may be due to poor operationalization of the etiological concept of interest (the unobserved construct), problems of measuring the variable correctly or due to crude categorizations of the measurements (Figure 13) [91].

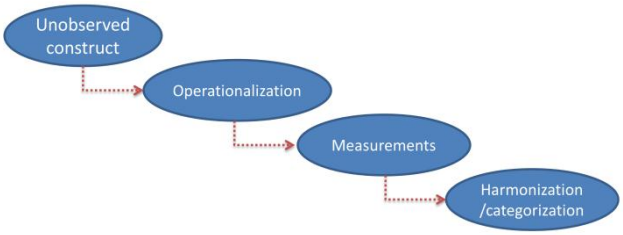

Figure 13
Levels of potential misclassification.

Figure 14 depicts the situation in which the true exposure, SEP, and the true outcome, BC, are measured with error (SEP* and BC*). U_{SEP} and U_{BC} represent measurement errors on exposure and outcome, respectively. The only situation in which the estimate of the association between the measured SEP* and BC* is free of bias, is when there is no true effect of SEP on BC [91].

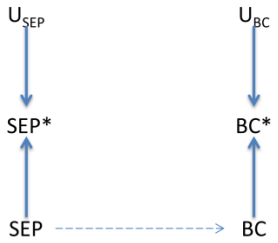

Figure 14
Causal structure of measurement error, U, on exposure, SEP, and outcome, BC.

The impact of measurement error depends on whether the error is *dependent* and/or *differential*. Dependent measurement error arises if some underlying factor influences both the measurement error of the exposure and the disease (e.g., common-method bias in which for example a person's personality affects both the reporting of exposure and outcome in a questionnaire). Differential misclassification of the exposure arises when the measurement error depends on the true value of the outcome (e.g., recall bias in case control studies) or other variables, and likewise the outcome measurement may be differential with respect to the exposure or other variables [91].

The causal diagram above describes the structure of potential misclassification bias; however, evaluating the magnitude or the direction of the bias is somewhat more complex. In the following, potential misclassification of the different measurements with focus on exposure, self-reported mediators and outcome, will be discussed and the consequences evaluated through sensitivity analyses based on plausible scenarios.

Exposure

The concept of SEP encompasses the economic, cultural, political and social resources of individuals or groups determined by their position within the structure of a society (cf. Figure 1) [92]. The classification of the participants' SEP was based on register information on highest attained educational level. Educational level was applied as a measure of SEP because of its wide availability across the relevant data sets, and because it has been found to be a reliable measure of SEP in European countries [93]. Also, compared to income or occupation, educational level is a more constant measure of lifelong social status, i.e., not affected by for example retirement or changes in health status, issues relevant to the aging population of the SIC cohort [8,92]. Inequalities in various health outcomes have been shown to depend on the chosen measure of SEP [94,95]. A recent Danish study of SEP and postmenopausal breast cancer based on the DCHS data, measured SEP by educational attainment, occupation and income. In this study, a higher breast cancer incidence was observed in women of high versus low SEP when SEP was measured by education or occupation (adjusted for education), but a less clear gradient when measuring SEP by individual income (adjusted for education and occupation). The included mediators (reproductive patterns, HT use, alcohol consumption and BMI) seemed to explain the educational differences in breast cancer but not occupational differences [13].

The operationalization of SEP by educational level in this thesis relates to the underlying theory about the mechanisms that link SEP with postmenopausal breast cancer (cf. the appendix). The lifestyle factors, obesity and HT use are hypothesized to be largely driven by health consciousness which probably relates to a construct of knowledge or ability to turn health information into behavior rather than actual wealth/material advantage, labor market position or work-related characteristics (e.g., psychological or physical work environment).

The information on educational level was obtained from a high quality database from Statistics Denmark in which the coverage and validity of educational information is generally considered high [96]. A potential problem with the measurements of education is that the baseline period of the SIC cohort is very broad, and a simple linkage to registers on educational level does not take into account the birth cohort effect of such a measure. For example, the economic implications and credential value of achieving a high school diploma attenuate with time, which may in turn affect the social patterning of breast cancer over time [97]. In Paper I, a sensitivity analysis addressing the potential heterogeneity of the relation between education and breast cancer according to strata of birth cohort did not indicate differential effects. However, the operationalization of SEP as measured by educational level alone may still not accurately reflect the underlying construct of SEP.

Mediators

In this section, misclassification of self-reported mediators is discussed. Since BMI and blood samples were assessed by staff at the baseline physical examination and in similar ways across the pooled cohorts, measurement errors are likely minimized. The self-report of age at first birth and parity are also not likely to be misclassified because such major and well-defined life-events are easier to recall.

In all three papers, lifestyle factors were self-reported and inevitably reported with some degree of error. In Papers I and III we harmonized information across six cohorts, which added further imprecision to the measurements. In situations of mediation by a dichotomous or normally-distributed non-differentially and independently misclassified mediator, the mediating effect will be biased towards the null, and hence the direct effect will be biased away from the null [98,100]; However, in all other scenarios (e.g., where the mismeasured mediator is a non-monotonic

function of the true mediator) the direction of the bias is not straightforward [99-101].

As mentioned previously, in a study comparing self-reported alcohol information to prospectively collected food records, correlations of alcohol consumption and estradiol were only observed when considering the food record data, which may indicate misclassification [83]. However, other studies have found self-reports of alcohol consumption to be generally reliable in observational studies, showing correlations above 0.8 when compared to detailed dietary records [83,102,103]. A study on the Danish MONICA cohort found higher correlations between self-reported alcohol and detailed assessments among men than among women, but did not find evidence of differential reliability according to SEP [102].

Studies on the reliability of self-reported measures of physical activity suggest low to moderate correlations between self-reports and directly observable measures (e.g., accelerometry) [104]. A recent study of older individuals (60+ years) by Sabia et al. compared categories of physical activity based on a self-reported questionnaire and direct measurements, and found that less than 50% of the participants were classified in the same tertile. Approximately 20% who were classified as highly physically active according to the questionnaire had a low physical activity level according to the accelerometer-assessed measure and vice versa [105]. The study also concluded that this discrepancy varied by SEP, with higher correlations observed for high SEP compared to low SEP [105]. Similarly, another study found that women of low educational level tended to exaggerate their physical activity level to a higher degree than women with a higher educational level [106]. Thus, these previous studies indicate that the measurements of physical activity level in this thesis may be seriously misclassified, and the consequences will be further investigated in sensitivity analyses below.

For participants included from the CCHS, mediators were assessed as early as 1981. These baseline measurements were taken as a constant measure of each participant's mediator level throughout the study period of up to 28 years. Obviously, HT use and lifestyle risk factors such as alcohol consumption and physical activity may change over time perhaps due to growing awareness of the health risk it imposes or changes related to aging (e.g., increased sensitivity to alcohol or immobility leading to reduced physical activity). A previous longitudinal study based on all four waves of the CCHS, investigated the latency of alcohol consumption on breast cancer by applying updated information on alcohol consumption. They concluded that baseline measurements of alcohol consumption were more strongly associated with breast cancer probably due to long latency time [107]. This suggests that, at least for alcohol consumption, the application of only baseline measurements may not be a major problem. As discussed in Paper III, HT use may have decreased during follow-up because of the growing awareness of unfavorable health implications. According to one study, the decline was independent of SEP [108]. A sensitivity analysis ending follow-up in 2002 – in which we expect a drop in HT use following the publications of several studies suggesting adverse health effects of HT use – showed a stronger effect of HT use on breast cancer in the restricted analysis. This indicates that misclassification of HT use biased the overall estimates of the HT-breast cancer relation towards the null.

Another point to consider is the fact that the HT measure did not distinguish between duration of use, route of administration or types of HT regimens (estrogen, progestin or combination therapy). The risk of breast cancer has been shown to vary by HT type, with the highest risk observed for combination therapy [109]. Combination therapy is usually prescribed to women with natural menopause whereas therapy of estrogen alone is more common following surgical menopause [109], which has been found to be more common in Danish women of low SEP [110]. The problem of having only a very crude version of the mediator can be considered as a situation of 'multiple versions of the mediator' (Figure 15) [111]. The consequence is that we estimate the indirect effect through the binary HT use variable describing whether or not the women use hormones, while there is another indirect effect through *type* of HT use which we were unable to estimate. This additional indirect effect will be incorporated in the direct effect, and thus, the indirect effect of HT use "yes/no" is underestimated in terms of the 'true' HT use. Thus, considering the potential decrease in HT use over time and the lack of information on HT type, the mediated effect of HT use was likely underestimated in Paper I.

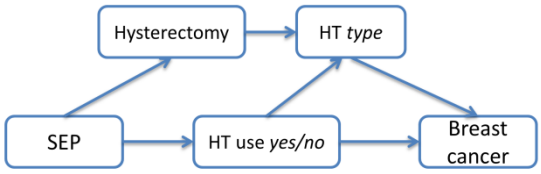

Figure 15
Mediation through multiple versions of the mediator.

Outcome

Breast cancer was assessed prospectively in all included cohorts. The Danish cohorts were linked to the nationwide Danish Cancer Registry [68], and the WHI collected information through self-administered questionnaires, which were further validated by reviews of hospital discharge summaries [70]. The coverage from such sources is considered quite high. However, it does not account for potential case detection variation between subgroups – for instance educational variations, if women of higher educational level were more prone to attend their general practitioner than women of low educational level. In 1991, an organized mammography screening program of women aged 50–69 years was introduced in the Copenhagen area, and an increased detection and overdiagnosis of breast cancer was seen in the following years [112]. If screening attendance is related to educational level or to any of the mediators included in our study, the lack of control for this factor may have biased the results. If women of higher educational level attended the screening visits to a higher degree than women of low educational level, the observed social inequality in breast cancer might be non-causal. A previous study on socioeconomic differences in screening non-attendance in Copenhagen found a U-shaped relation, that is 'never-use' of screening was more frequently observed among women of both low and high education compared to women with a medium educational level [113]. Another study on more recent screening data (2008–2009) from the Region of Central Denmark, from which a large part of the DCHS was sampled, found non-participation to be associated with low SEP [114]. Thus, some degree of differential misclassification due to differences in screening behavior is possible; however, our study included very few cases from the investigated period of the latter study (≈5%) [114].

In Papers II and III, we were able to distinguish breast cancer cases according to ER status, however, the histology of the breast cancer cases was not considered in this thesis. During the 1980s and 1990s, the incidence of lobular tumors increased considerably, whereas the incidence rate of ductal tumors remained rather constant [115,116]. Previous studies have proposed distinct etiologies for these two main types of invasive breast cancers, which suggests different implications for prevention, diagnosis, treatment and prognosis [117,118]. Lobular tumors appear to be more hormonally dependent than ductal [109,119,120], which may indicate that the mediated effects observed in this thesis could be different across histologic subtypes. A recent publication within the Nurse's Health study, examined potential effect heterogeneity of risk factors according to breast cancer histology [121]. Associations of these risk factors were all observed in the expected directions, however, stronger effects of age at menarche, age at first birth and postmenopausal HT use were observed for lobular compared to ductal tumors. The associations observed for BMI, alcohol consumption and family history of breast cancer did not differ according to histologic subtype. Thus, including information on histopathological subtypes of breast cancer would have enhanced the biological understanding of how HT use, lifestyle and reproductive factors affect breast cancer risk.

How does misclassification affect the results?

Misclassifications should be evaluated according to whether they are independent and/or non-differential as mentioned previously. Since SEP, mediators and breast cancer were derived from different sources and collected prospectively at different points in time, dependent measurement errors do not seem obvious. However, the above mentioned study by Sabia et al. suggested differential misclassification of physical activity according to SEP [105]. Figure 16 depicts such differential measurement error in our setting.

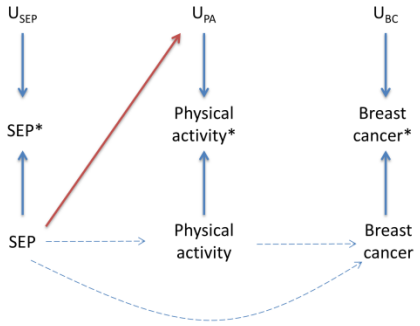

Figure 16
Differential measurement error, U, on the mediator physical activity according to SEP.

The degree to which SEP-dependent misclassification of the physical activity-level would have affected the mediation analysis was evaluated by introducing further misclassification to the observed physical activity variable (Figure 17) [122]. In this scenario, it was assumed that 20% of the women of low educational level over-reported their physical activity level, and that 15% of the women with a medium and 10% with a high educational level over-reported their level of physical activity, respectively. The sensitivity analysis randomly picked the women according to the mentioned percentages and the whole procedure was repeated 1,000 times to explore how the effect estimates changed. The total effect, mediated effect and mediated interactive effect were derived from the median of the replications along with 95% confidence intervals.

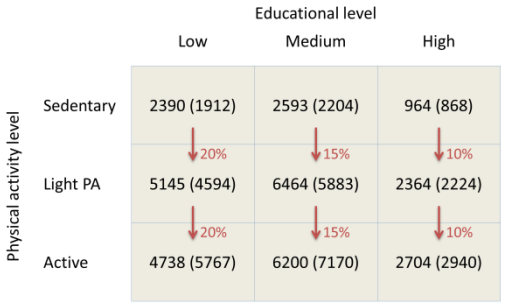

Figure 17
Observed (misclassified) number of participants according to educational level and physical activity.

The total (TE), mediated (ME, and mediated interactive (MIE) effects from the original (observed) data and the results of the simulations are presented in Table 4. In the original data, 70.5 additional breast cancer cases per 100,000 person-years were observed among women with a medium educational level and correspondingly 73.8 additional cases in women with a high educational level compared to women with a low educational level. The mediated effect of physical activity was negligible for both educational groups. A statistically significant mediated interactive effect of 10.1 fewer cases per 100,000 person-years was observed among the highly educated women. The results of the sensitivity analysis suggest that the introduced misclassification of the mediator, as depicted in Figure 17, leads to bias towards the null, as the mediated interactive effect was reduced to -1.3 cases per 100,000 person-years in the misclassified scenario. The second misclassification scenario, in which the contrast in misclassification across educational groups was increased (low educational level 20%, medium educational level 10% and high educational level 5%), led to a slightly positive mediated interaction. Introducing even further misclassification (low 30%, medium 20%, high 10%) led to a clearly positive mediated interactive effect of physical activity. Thus, the mediated interactive effect is very sensitive to differential misclassification of the mediator, which may explain the unexpected interaction observed between SEP and physical activity in Paper I.

Confounding

In the ideal setting, the estimates of breast cancer risks when comparing exposed to unexposed participants would be causally interpretable. However, such conclusions rely on an assumption of exchangeability between exposed and unexposed persons, i.e., that the risk of breast cancer among highly educated women would be equal to that of the low educated women if a hypothetical intervention could change their educational level from high to low [123]. In observational studies, exchangeability is approximated by adjustment for confounding [124]. In the mediation analysis, confounding can occur at several levels as depicted in Figure 18: Between exposure and outcome, exposure and mediator, mediator and outcome – and confounding of the mediator-outcome dependent on the exposure is a further challenge. This will be discussed in a separate section below.

Table 4

Mediated proportions (95% CI) for combinations of mediators (high versus low education) in the SIC cohort.

	Medium education			High education		
	Additional breast cancer cases per 100,000 person-years			Additional breast cancer cases per 100,000 person-years		
	TE	ME	MIE	TE	ME	MIE
Observed						
	70.5	**1.0**	**-0.4**	**73.8**	**2.2**	**-10.1**
95% CI	(29.0; 111.9)	(-1.6; 1.8)	(-2.9; 2.2)	(22.1; 125.6)	(-0.5; 4.9)	(-15.7; -4.5)
Misclassified[1]						
Median	**70.2**	**0.02**	**0.1**	**73.5**	**0.2**	**-1.3**
2.5%–97.5%	(70.1; 70.4)	(-1.0; 1.0)	(-1.2; 1.5)	(73.4; 73.7)	(-0.5; 0.8)	(-2.8; 0.1)
Misclassified[2]						
Median	**70.1**	**-0.8**	**0.6**	**73.3**	**-0.5**	**2.3**
2.5%–97.5%	(69.9; 70.3)	(-2.2; 0.6)	(-1.2; 2.4)	(73.2; 73.6)	(-1.6; 0.5)	(0.7; 3.9)
Misclassified[3]						
Median	**70.0**	**-0.8**	**0.6**	**73.3**	**-1.4**	**6.3**
2.5%–97.5%	(69.8; 70.3)	(-2.5; 0.6)	(-1.3; 2.6)	(73.0; 73.6)	(-3.1; 0.1)	(3.8; 9.0)

TE, total effect; ME, mediated effect; MIE, mediated interactive effect.
Medium and high education defined as 11–14 years (upper secondary or vocational) and 15+ years of education.
[1]Overreport of 20%, 15% and 10% among low, medium and high educational level, respectively.
[2]Overreport of 20%, 10% and 5% among low, medium and high educational level, respectively.
[3]Overreport of 30%, 20% and 10% among low, medium and high educational level, respectively.

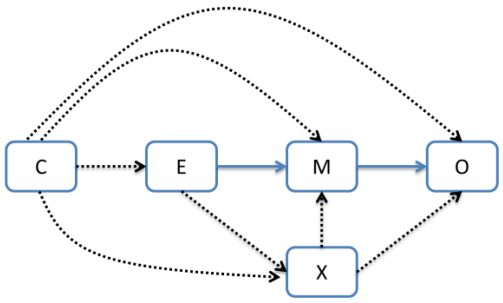

Figure 18
Causal diagram of the relation between exposure, E, and outcome, O, mediated through M and confounded by C and X.

The confounders identified for the analyses in this thesis, were selected on the basis of prior knowledge of causal relations obtained through literature reviews [71]. The DAG presented in the appendix, depicts the assumed causal relation between SEP and postmenopausal breast cancer. As appears from the diagram, most factors are expected to lie on the causal path from SEP to postmenopausal breast cancer. One exception is family history of breast cancer, which was unmeasured in the SIC cohort and thus, not included in Papers I and III. If women with a family history of breast cancer obtained a lower level of education and at the same time were at higher risk of breast cancer, the social inequality in breast cancer observed in Paper I is underestimated due to the lack of control for family history of breast cancer. However, family history of breast cancer is unlikely to have strongly influenced SEP and, as discussed in the paper, the population attributable fraction of family history to breast cancer is modest [125]. Thus, the effect of this particular confounder is likely small.

Early-life SEP is another potential common cause of SEP, the mediators and postmenopausal breast cancer [126]. Childhood SEP may affect breast cancer risk indirectly through adult SEP and the mediators considered in this thesis; however, it is possible that circumstances related to childhood SEP may shape the environment and in turn affect biological processes of later breast cancer independent of adult SEP (cf. a "critical period mechanism") [127]. Factors such as childhood physical activity, diet and chronic inflammation have been put forward as possible risk factors acting independently of adult SEP [128]. Results of a recent analysis addressing the life-course perspective, suggested that the direct effect of childhood SEP on breast cancer was negative, but that the mediated effect through adult SEP was positive [128]. In this way, the additional number of breast cancer cases among women of high educational level may be underestimated in this study due to the lack of control for childhood SEP. In addition, the DAG also suggests potential confounding of the obesity-breast cancer relation through age at menarche. This factor was not measured among women in the SIC cohort [27]. Young age at menarche is a risk factor for both adult obesity [129] as well as postmenopausal breast cancer [130]. We did not observe an overall effect of high BMI on breast cancer risk in Paper I, but lack of control for this factor may have overestimated the effect in Paper III.

In Papers II and III, we considered the association between lifestyle and reproductive factors, BMI, HT use and breast cancer. These relations involved adjustment for several confounding factors. A great concern here is misclassification of confounders, which inevitably leads to residual confounding [91,131,132]. As mentioned in the section above, misclassification is expected for the self-reported lifestyle factors. In Papers II and III, the analyses of the association between high BMI and breast cancer were adjusted for physical activity level. The relation of both BMI and breast cancer with physical activity is expected to be inverse [47,50] and thus, perfect control for physical activity would reduce the BMI-breast cancer association. Since it is likely that physical activity in our population was measured with error, some of the additional number of breast cancer cases observed in women with a higher BMI may be due to physical inactivity rather than obesity. The magnitude of this residual confounding, however, also depends on the extent to which physical activity is correlated with other included confounders [131]. Similar prob-

lems arise in the alcohol-breast cancer analyses (Papers II and III), where adjustment for the measured smoking variable is likely to be insufficient.

Exposure-dependent confounding

A special case of confounding in mediation analyses occurs when a confounder of the mediator-outcome relation is affected by the exposure. In such a situation, the natural direct and indirect effects cannot be identified with current methods irrespective of whether the exposure-dependent confounder is measured or not [75]. The possibility of exposure-dependent confounding may be minimized in situations where the mediator occurs shortly after the exposure, because when the mediator is measured immediately after the occurrence of the exposure, it is less likely that another effect of the exposure would confound the mediator-outcome relation [133]. This is, however, not the case for the majority of mediators considered in this thesis and the inability to take account of potential exposure-dependent confounding is a severe limitation of these methods.

The concept of exposure-dependent confounding is closely related to the problem of intertwined pathways evaluated previously in this thesis. In that scenario, the lifestyle, BMI and reproductive factors were expected to partly take affect through the same pathways. The results of the analyses considering this issue did not, however, suggest the mediators to be strongly intertwined.

When considering the DAG presented in the appendix, other factors which are consequences of SEP, may have confounded the mediator-outcome relations, e.g., 'health consciousness'. According to previous studies, health consciousness may be affected by SEP [134,135], and may in turn affect both obesity and the risk of breast cancer. If health conscious women are less obese and at the same time at lower risk of breast cancer because of a healthier lifestyle, the lack of control for health consciousness overestimates the obesity-breast cancer relation (Figure 19). Because of the inverse relation between SEP and obesity, the indirect effect of SEP on breast cancer through obesity is likely underestimated.

Health consciousness was not measured in the available data. If measured, the problem could be addressed by considering the mediator and the exposure-dependent confounder jointly. However, since the actual interest is in the mediating effect of obesity, this would be unsatisfactory. Other methods have recently been proposed to account for this type of confounding when measured in the data [75,77].

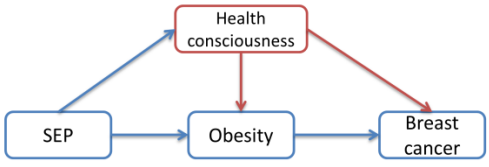

Figure 19
Exposure-dependent confounding of the relation between the mediator obesity and breast cancer through 'health consciousness'.

Other methodological considerations

Aside from the various sources of bias addressed above, other methodological issues should be mentioned. Below, the determination of the study population by the definition of menopausal status and the statistical methods applied are discussed.

Definition of postmenopausality

In this study, we defined postmenopausality according to age at baseline (50+ years), a crude proxy for menopausal status which may have caused some misclassification. Several other definitions have been applied in previous epidemiological studies (e.g., combinations of information according to time since last menstrual period, history of hysterectomy/oophorectomy and/or age) [136,137]. Since the transition from premenopause to postmenopause is complex and takes place over several years, no gold-standard for determining menopausal status has been defined. A recent study comparing various approaches to defining postmenopausality, reported modest overlap between the comprehensive measures taking several aspects into account and the crude measures according to age [136]. Based on the comprehensive definition, 25% of the women aged 50–54 years were classified as premenopausal, whereas 10% of the women aged 40–49 years were classified as postmenopausal. However, the breast cancer incidence did not vary considerably across the definitions. The authors of another study comparing various definitions of postmenopausality, concluded that when lacking information on menstrual history an age-definition of 50 years would be the best proxy for menopausal status [137].

Statistical methods

The methods applied in this thesis were developed in order to quantify mediating effects. As mentioned previously, there are a number of problems with the conventional method of comparing crude and adjusted coefficients in nonlinear models. First of all, the decomposition of the total effect into direct and mediated pathways is not straightforward; and second, estimations are limited by mediator-outcome confounding, exposure-dependent confounding of the mediator-outcome relation, interactions between exposure and mediators as well as interactions between mediators [24,25].

By applying new methods, some – but far from all – of these issues were addressed. The methods allowed for the decomposition of effects, although limited to single pathways under the assumption that all mediators of interest work through distinct, non-intertwined pathways. Sensitivity analyses suggested that this assumption was not severely violated in the present context. The methods also allowed for the three-way decomposition of the direct effect, the pure indirect effect and the mediated interactive effect which is important when considering potential differential vulnerability across SEP groups. We were not, however, able to account for exposure-dependent confounding and interaction between mediators.

Mediation analysis is currently an active area of epidemiologic research and recent discussions have concerned the extent to which the newer methods actually provide qualitatively different results than the conventional mediator-adjustment approach. Most of the previous work on mediating effects of the relation between SEP and breast cancer has been based on the Cox proportional hazards model or logistic regression. As noted previous-

ly, the assumption of collapsibility underlying the comparison of unadjusted to adjusted models of these types is questionable [138]. However, recent work by Vanderweele and colleagues showed that the adjustment approach can be used in situations where the outcome is relatively rare (e.g., <15%) [139]. In situations where the effects of the exposure and mediator do not interact, the controlled direct effect equals the natural direct effect, and the natural indirect effect will equal the total effect minus the controlled direct effect [139,140]. However, when the outcome is common and/or the exposure and mediator interact, actual mediation analyses – like the methods applied in this thesis – are needed [141].

External validity

The participants included in the SIC cohort were all in principle randomly selected from the general population of the urban areas of Copenhagen and Aarhus. The representativeness of the pooled cohort studies is questionable because of their participation rates, which were low – especially in the DCHS (approximately 35%). Studies comparing participants versus non-participants of the cohorts included in the SIC collaboration, have shown a tendency towards a higher participation rate among younger individuals, individuals of a higher educational level and with a slightly healthier risk factor profile [60-62]. These findings suggest, that the diversity of the population according to these factors was limited. Even if the non-response was negligible, the cohorts would not be representative of citizens of more rural areas of Denmark. It is also important to note that the SIC cohort mainly consists of Caucasian women. Since ethnicity is associated with SEP and also breast cancer etiology, there is reason to be cautious about direct extrapolation to ethnic minority women.

The possibility of specific period effects should also be acknowledged. The findings of this thesis is based on cohorts which collected information on risk factors dating back to 1981. These effects are likely dependent on the distribution of various component causes, and their distribution will be specific to the context in which the information was collected.

The same considerations apply for the WHI-OS in Paper II with regards to generalizability to the general American population and the specific time period represented in these data.

CONCLUSION

The results of this thesis suggest that the social inequality in postmenopausal breast cancer is largely mediated by HT use, alcohol consumption and reproductive factors. Alcohol consumption interacted with HT use and their combination was associated with a markedly higher breast cancer risk than expected from their individual effects. The analyses considering the combination of HT use and alcohol consumption in relation to endogenous hormone levels suggested that the higher risk may be driven by noticeable higher estradiol levels and – to some extent – testosterone levels. The importance of estrogen in this relation was also supported by the fact that the findings were largely restricted to ER-positive breast cancer cases.

Overall, high BMI was not a significant mediator of the SEP-breast cancer relation in the SIC cohort. High BMI was more prevalent among women with low compared to high education, but unexpectedly there was no overall association between BMI and breast cancer. However, an interaction between BMI and HT use was identified in further analyses, revealing a tendency towards a higher risk of breast cancer with higher levels of BMI in current non-HT users and a U-shaped relation in current HT users. In the WHI-OS data – which only included current non-HT users – a statistically significant higher risk of breast cancer was observed with higher BMI. The analyses of endogenous hormone levels indicated a role of estradiol, insulin and testosterone in the relation of BMI with postmenopausal breast cancer.

The various sources of bias discussed in this thesis raise concern that the observed relations could be biased. In particular, the impact of HT use and physical inactivity on the relation between SEP and breast cancer may have been affected by misclassification, and the observed mediating effects are likely conservative estimates regarding the impact of each factor.

PERSPECTIVES

The findings of this thesis lend support to the notion that social inequality in breast cancer could be reduced through preventive strategies targeted HT use, alcohol consumption and reproductive factors. However, it is important to keep in mind that findings from epidemiological studies do not generally translate into intervention effects, as the strength of a cause depends on the prevalence of other causes of disease in a given population; and that ethical and economic concerns also apply when planning public health interventions.

According to the model developed by Diderichsen and colleagues presented at the beginning of this thesis (Figure 1), there are three main policy entry points for reducing social inequality in health: A) influencing social stratification, B) decreasing exposures and C) decreasing vulnerability. We are obviously not able to – or interested in – manipulating educational level from high to low (i.e., entry point A), but policies targeted at reducing the effect of SEP on mediators, would be more realistic (i.e., entry point B). For example, policies that would encourage women of higher educational level to give birth to more children at younger ages and reduce HT use and alcohol intake to the levels observed among women of low education have the potential to prevent a large share of the postmenopausal breast cancer cases observed in women of high educational level. The combination of the effects of HT use, alcohol intake and BMI also underpins the potential of decreasing vulnerability, because women who use HT may be particularly vulnerable in terms of postmenopausal breast cancer if they also consume alcohol. Likewise, the risk of breast cancer according to HT use also appeared to vary by BMI. These findings need to be confirmed in other studies, but suggest that changing clinical guidelines for the prescription of HT according to the BMI or the alcohol habits of women may contribute to the prevention of postmenopausal breast cancer.

In a broader perspective, affecting reproductive patterns towards a lower age at first birth and an increase in higher-order births has important implications for public health and society as a whole. The negative consequences of postponement of childbearing include lower fecundity and higher risks of adverse pregnancy outcomes. In addition, the Danish welfare system suffers from a decreasing workforce in proportion to the population of non-working-age (0–15 and 65+ years), and increasing higher-order births would decelerate this development. Thus, the risks and benefits of altering reproductive patterns, alcohol and HT behavior regarding other diseases are also important to

acknowledge, and the apparent protective coronary effect of moderate alcohol consumption is also worth considering.

Another consideration is that the social patterning of reproductive factors, HT use and alcohol consumption may affect the social inequality in one direction, but the increasing and highly skewed prevalence of obesity and physical inactivity may work in the opposite direction. The results of this thesis suggested higher risks of breast cancer among non-HT users. An increase in breast cancer among women of lower SEP due to a rise in obesity would reduce the overall social inequality in breast cancer; but in the interest of public health, it is important not to overlook the future challenge of obesity in relation to postmenopausal breast cancer incidence.

As mentioned previously, it is an ongoing discussion whether the newer counterfactual-based mediation approaches provide better estimates – and thus, whether they add more public health insight relevant for prevention policies – compared to the conventional mediator-adjustment approach. There is no doubt that the development of these methods has elucidated the potential biases that may underlie the reported mediating effects in previous studies. However, recent publications based on the newer methods are also accompanied by increasing numbers of sensitivity analyses, which highlights the need for better data in terms of misclassification and confounding.

Future studies should focus on life-course perspectives of social inequality in breast cancer in order to unravel the relative importance of risk factors and their timing from early childhood to old age. Cohort studies with long follow-up times would enable the identification of sensitive periods, and repeated measurements of mediators would allow for analyses of health behavior trajectories over time. In addition, there is a need for further development of the methodology for the quantification of mediating effects to handle current shortcomings such as exposure-dependent confounding and the potential interactions between mediators.

In conclusion, this thesis adds to the current knowledge on how SEP affects postmenopausal breast cancer risk by quantifying the mediating effects of several modifiable risk factors through the application of new statistical methods. The results were qualitatively in accordance with findings of previous studies, but refined by addressing interactions between SEP and the mediators and between the mediators, by investigating hormonal pathways to support the biological hypotheses and by testing the underlying assumptions and potential biases through sensitivity analyses.

SUMMARY

This thesis is based on studies conducted in the period 2010–2014 at Department of Public Health, University of Copenhagen and at Department of Epidemiology and Population Health, Albert Einstein College of Medicine, New York. The results are presented in three scientific papers and a synopsis.

The main objective of the thesis was to determine mechanisms underlying social inequality (defined by educational level) in postmenopausal breast cancer (BC) by addressing mediating effects through hormone therapy (HT) use, BMI, lifestyle and reproductive factors. The results of previous studies suggest that the higher risk of postmenopausal BC among women of high SEP may be explained by reproductive factors and health behaviors. Women of higher SEP generally have fewer children and give birth at older ages than women of low SEP, and these factors have been found to affect the risk of BC – probably through altered hormone levels. Adverse effects on BC risk have also been documented for modifiable health behaviors that may affect hormone levels, such as alcohol consumption, high body mass index (BMI), physical inactivity, and HT use. Alcohol consumption and HT use are likewise more common among women of higher SEP.

The analyses were based on the Social Inequality in Cancer (SIC) cohort and a subsample of the Women's Health Initiative Observational Study (WHI-OS). The SIC cohort was derived by pooling 6 individual studies from the Copenhagen area including 33,562 women (1,733 BC cases) aged 50–70 years at baseline. The subsample of WHI-OS consisted of two case-cohort studies with measurements of endogenous estradiol (N=1,601) and insulin (N=791). Assessment of mediation often relies on comparing multiplicative models with and without the potential mediator. Such approaches provide potentially biased results, because they do not account for mediator-outcome confounding, exposure-dependent mediator-outcome confounding, exposure-mediator interaction and interactions between mediators. In addition, these simple methods do not allow for a decomposition of the total effect into direct and indirect pathways. The counterfactual-based methods for quantifying mediating effects in this thesis were developed specifically for this project taking into account some of the shortcomings of previous methods.

The results of this thesis showed that a high versus low educational level was associated with a higher risk of postmenopausal BC and that this effect was partly mediated through HT use, fertility patterns and alcohol consumption in the SIC data. Overall BMI did not mediate the education-BC relation. The results from the WHI-OS for the effect of alcohol consumption on BC risk did not – as hypothesized – seem to be mediated by endogenous estradiol levels; however, the observed higher risk of BC with higher levels of alcohol was restricted to estrogen-receptor positive cases, which indicates a role of estrogens in this relation. In the WHI-OS subsample of non-HT users, a higher risk of BC was found with higher levels of BMI; both estradiol and insulin mediated the effect of BMI on BC. The effect of HT use on BC interacted synergistically with alcohol consumption and this combination appeared to be associated with very high serum levels of estradiol in the SIC data. For BMI combined with HT use, a modest positive association was observed for non-HT users whereas markedly higher risks were observed across all BMI groups in current HT-users with a tendency towards a U-shaped relation.

In conclusion, the social inequality in postmenopausal BC seems to be largely mediated by HT use, alcohol consumption and reproductive factors. Various sources of bias – especially misclassification of mediators, but also exposure-dependent confounding – raise some concern about the observed relations. Future studies should focus on life-course perspectives to identify certain windows of susceptibility and collect data on repeated measurements of mediators to enable health behavior trajectories over time. In addition, there is a need for further development of the methodology for the quantification of mediating effects to handle current shortcomings such as exposure-dependent confounding and the potential interactions between mediators.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011;61:69-90.
2. Parkin DM, Bray FI, Devesa SS. Cancer burden in the year 2000. The global picture. *Eur J Cancer* 2001;37:S4-S66.
3. Engholm G, Ferlay J, Christensen N, Johannesen TB, Khan S, Køtlum JE, Milter MC, Ólafsdóttir E, Pukkala E, Storm HH. NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 6.0 (04.12.2013). Association of the Nordic Cancer Registries. Danish Cancer Society. Available from: http://www.ancr.nu, accessed on 19-12-2013.
4. Dalton SO, Schuz J, Engholm G, Johansen C, Kjaer SK, Steding-Jessen M, Storm HH, Olsen JH. 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;44:2074-85.
5. Mackenbach JP, Stirbu I, Roskam AJ, Schaap MM, Menvielle G, Leinsalu M, Kunst AE. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 2008;358:2468-81.
6. Dano H, Andersen O, Ewertz M, Petersen JH, Lynge E. Socioeconomic status and breast cancer in Denmark. *Int J Epidemiol* 2003;32:218-24.
7. Klassen AC, Smith KC. The enduring and evolving relationship between social class and breast cancer burden: a review of the literature. *Cancer Epidemiol* 2011;35:217-34.
8. Strand BH, Kunst A, Huisman M, Menvielle G, Glickman M, Bopp M, Borell C, Borgan JK, Costa G, Deboosere P, Regidor E, Valkonen T, et al. The reversed social gradient: Higher breast cancer mortality in the higher educated compared to lower educated. A comparison of 11 European populations during the 1990s. *Eur J Cancer* 2007;43:1200-7.
9. Dano H, Hansen KD, Jensen P, Petersen JH, Jacobsen R, Ewertz M, Lynge E. Fertility pattern does not explain social gradient in breast cancer in Denmark. *Int J Cancer* 2004;111:451-6.
10. Hafeman DM, Schwartz S. Opening the Black Box: a motivation for the assessment of mediation. *Int J Epidemiol* 2009;38:838-45.
11. Heck KE, Pamuk ER. Explaining the relation between education and postmenopausal breast cancer. *Am J Epidemiol* 1997;145:366-72.
12. Strand BH, Tverdal A, Claussen B, Zahl PH. Birth history the key to highly educated women's higher breast cancer mortality? A follow-up study of 500,000 women aged 35-54. *International Journal of Cancer* 2005;117:1002-6.
13. Larsen SB, Olsen A, Lynch J, Christensen J, Overvad K, Tjonneland A, Johansen C, Dalton SO. Socioeconomic position and lifestyle in relation to breast cancer incidence among postmenopausal women: a prospective cohort study, Denmark, 1993-2006. *Cancer Epidemiol* 2011;35:438-41.
14. Braaten T, Weiderpass E, Kumle M, Adami HO, Lund E. Education and risk of breast cancer in the Norwegian-Swedish women's lifestyle and health cohort study. *Int J Cancer* 2004;110:579-83.
15. Braaten T, Weiderpass E, Kumle M, Lund E. Explaining the socioeconomic variation in cancer risk in the Norwegian women and cancer study. *Cancer Epidemiol Biomarkers Prev* 2005;14:2591-7.
16. Gadeyne S, Deboosere P, Vandenheede H, Neels K. Does birth history account for educational differences in breast cancer mortality? A comparison of premenopausal and postmenopausal women in Belgium. *Int J Cancer* 2012;131:2878-85.
17. Palmer JR, Boggs DA, Wise LA, Adams-Campbell LL, Rosenberg L. Individual and neighborhood socioeconomic status in relation to breast cancer incidence in African-American women. *Am J Epidemiol* 2012;176:1141-6.
18. Pudrovska T, Carr D, McFarland M, Collins C. Higher-status occupations and breast cancer: a life-course stress approach. *Soc Sci Med* 2013;89:53-61.
19. Robert SA, Strombom I, Trentham-Dietz A, Hampton JM, McElroy JA, Newcomb PA, Remington PL. Socioeconomic risk factors for breast cancer: distinguishing individual- and community-level effects. *Epidemiology* 2004;15:442-50.
20. Menvielle G, Kunst AE, van Gils CH, Peeters PH, Boshuizen H, Overvad K, Olsen A, Tjonneland A, Hermann S, Kaaks R, Bergmann MM, Illner AK, et al. The contribution of risk factors to the higher incidence of invasive and in situ breast cancers in women with higher levels of education in the European prospective investigation into cancer and nutrition. *Am J Epidemiol* 2011;173:26-37.
21. Strand BH, Tverdal A, Claussen B, Zahl PH. Is birth history the key to highly educated women's higher breast cancer mortality? A follow-up study of 500,000 women aged 35-54. *Int J Cancer* 2005;117:1002-6.
22. Cole SR, Hernan MA. Fallibility in estimating direct effects. *Int J Epidemiol* 2002;31:163-5.
23. Kaufman JS, Maclehose RF, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. *Epidemiol Perspect Innov* 2004;1:4-16.
24. Pearl J. Direct and indirect effects. *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*. San Fransisco: Morgan Kaufmann, 2001:411-20.
25. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 1992;3:143-55.
26. Lange T, Hansen JV. Direct and indirect effects in a survival context. *Epidemiology* 2011;22:575-81.
27. Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol* 2012;176:190-5.
28. de MR, Jager KJ, Zoccali C, Dekker FW. The effect of joint exposures: examining the presence of interaction. *Kidney Int* 2009;75:677-81.
29. Gapstur SM, Potter JD, Drinkard C, Folsom AR. Synergistic Effect Between Alcohol and Estrogen Replacement Therapy on Risk of Breast-Cancer Differs by Estrogen Progesterone-Receptor Status in the Iowa Womens Health Study. *Cancer Epidem Biomar* 1995;4:313-8.
30. Nielsen NR, Gronbaek M. Interactions between intakes of alcohol and postmenopausal hormones on risk of breast cancer. *Int J Cancer* 2008;122:1109-13.
31. Diderichsen F, Evans T, Whitehead M. The social basis for inequalities in health. In: Evans T, Whitehead M, Diderichsen F, Bhuiya A, Wirth M. *Challenging inequities in health: From ethics to action*. New York: Oxford University Press, 2001:12-23.
32. Elo IT. Social Class Differentials in Health and Mortality: Patterns and Explanations in Comparative Perspective. *Annual Review of Sociology* 2009;35:553-72.

33. 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;50:255-70.
34. Lawlor DA, Smith GD, Ebrahim S. Socioeconomic position and hormone replacement therapy use: explaining the discrepancy in evidence from observational and randomized controlled trials. *Am J Public Health* 2004;94:2149-54.
35. Lokkegaard E, Lidegaard O, Moller LN, Agger C, Andreasen AH, Jorgensen T. Hormone replacement therapy in Denmark, 1995-2004. *Acta Obstet Gynecol Scand* 2007;86:1342-51.
36. Banks E, Beral V, Bull D, Reeves G, Austoker J, Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* 2003;362:419-27.
37. Beral V, Bull D, Doll R, Key T, Peto R, Reeves G, Million Women Study Collaborators. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet* 1997;350:1047-59.
38. Edlefsen KL, Jackson RD, Prentice RL, Janssen I, Rajkovic A, O'Sullivan MJ, Anderson G. The effects of postmenopausal hormone therapy on serum estrogen, progesterone, and sex hormone-binding globulin levels in healthy postmenopausal women. *Menopause* 2010;17:622-9.
39. Key TJ, Appleby PN, Barnes I, Reeves GK, Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: Reanalysis of nine prospective studies. *J Natl Cancer Inst* 2002;94:606-16.
40. Thomas HV, Reeves GK, Key TJA. Endogenous estrogen and postmenopausal breast cancer a quantitative review. *Cancer Causes Control* 1997;8:922-8.
41. Ewertz M, Duffy SW, Adami HO, Kvale G, Lund E, Meirik O, Mellemgaard A, Soini I, Tulinius H. Age at 1St Birth, Parity and Risk of Breast-Cancer - A Metaanalysis of 8 Studies from the Nordic Countries. *International Journal of Cancer* 1990;46:597-603.
42. Britt K, Ashworth A, Smalley M. Pregnancy and the risk of breast cancer. *Endocr Relat Cancer* 2007;14:907-33.
43. Smith-Warner SA, Spiegelman D, Yaun SS, van den Brandt PA, Folsom AR, Goldbohm RA, Graham S, Holmberg L, Howe GR, Marshall JR, Miller AB, Potter JD, et al. Alcohol and breast cancer in women - A pooled analysis of cohort studies. *JAMA* 1998;279:535-40.
44. Hamajima N, Hirose K, Tajima K, Rohan T, Calle EE, Heath CW, Jr., Coates RJ, Liff JM, Talamini R, Chantarakul N, Koetsawang S, Rachawat D, et al. Alcohol, tobacco and breast cancer - collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *Br J Cancer* 2002;87:1234-45.
45. Coyle YM. Lifestyle, genes, and cancer. *Methods Mol Biol* 2009;472:25-56.
46. Dorgan JF, Baer DJ, Albert PS, Judd JT, Brown ED, Corle DK, Campbell WS, Hartman TJ, Tejpar AA, Clevidence BA, Giffen CA, Chandler DW, et al. Serum hormones and the alcohol-breast cancer association in postmenopausal women. *J Natl Cancer Inst* 2001;93:710-5.
47. Monninkhof EM, Elias SG, Vlems FA, van der Tweel I, Schuit AJ, Voskuil DW, van Leeuwen FE. Physical activity and breast cancer - A systematic review. *Epidemiology* 2007;18:137-57.
48. Neilson HK, Friedenreich CM, Brockton NT, Millikan RC. Physical Activity and Postmenopausal Breast Cancer: Proposed Biologic Mechanisms and Areas for Future Research. *Cancer Epidem Biomar* 2009;18:11-27.
49. McTiernan A, Kooperberg C, White E, Wilcox S, Coates R, dams-Campbell LL, Woods N, Ockene J. Recreational physical activity and the risk of breast cancer in postmenopausal women - The Women's Health Initiative cohort study. *JAMA* 2003;290:1331-6.
50. van den Brandt PA, Spiegelman D, Yaun SS, Adami HO, Beeson L, Folsom AR, Fraser G, Goldbohm RA, Graham S, Kushi L, Marshall JR, Miller AB, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol* 2000;152:514-27.
51. Key TJ, Appleby PN, Reeves GK, Roddam A, Dorgan JF, Longcope C, Stanczyk FZ, Stephenson HE, Falk RT, Miller R, Schatzkin A, Allen DS, et al. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. *J Natl Cancer Inst* 2003;95:1218-26.
52. Liedtke S, Schmidt ME, Vrieling A, Lukanova A, Becker S, Kaaks R, Zaineddin AK, Buck K, Benner A, Chang-Claude J, Steindorf K. Postmenopausal Sex Hormones in Relation to Body Fat Distribution. *Obesity (Silver Spring)* 2012;20:1088-95.
53. Farhat GN, Cummings SR, Chlebowski RT, Parimi N, Cauley JA, Rohan TE, Huang AJ, Vitolins M, Hubbell FA, Manson JE, Cochrane BB, Lane DS, et al. Sex hormone levels and risks of estrogen receptor-negative and estrogen receptor-positive breast cancers. *J Natl Cancer Inst* 2011;103:562-70.
54. Cummings SR, Lee JS, Lui LY, Stone K, Ljung BM, Cauleys JA. Sex hormones, risk factors, and risk of estrogen receptor-positive breast cancer in older women: a long-term prospective study. *Cancer Epidemiol Biomarkers Prev* 2005;14:1047-51.
55. Gunter MJ, Hoover DR, Yu H, Wassertheil-Smoller S, Rohan TE, Manson JE, Li J, Ho GY, Xue X, Anderson GL, Kaplan RC, Harris TG, et al. Insulin, insulin-like growth factor-I, and risk of breast cancer in postmenopausal women. *J Natl Cancer Inst* 2009;101:48-60.
56. Gallagher EJ, Leroith D. The proliferating role of insulin and insulin-like growth factors in cancer. *Trends Endocrinol Metab* 2010;21:610-8.
57. Malin A, Matthews CE, Shu XO, Cai H, Dai Q, Jin F, Gao YT, Zheng W. Energy balance and breast cancer risk. *Cancer Epidem Biomar* 2005;14:1496-501.
58. Rod NH, Hansen AM, Nielsen J, Schnohr P, Gronbaek M. Low-risk factor profile, estrogen levels, and breast cancer risk among postmenopausal women. *Int J Cancer* 2009;124:1935-40.
59. Nordahl H, Hvidtfeldt UA, Diderichsen F, Rod NH, Osler M, Frederiksen BL, Prescott E, Tjonneland A, Lange T, Keiding N, Andersen PK, Andersen I. Cohort Profile: The Social Inequality in Cancer (SIC) cohort study. *Int J Epidemiol* 2014;1-9.
60. Schnohr P, Jensen G, Lange P, Scharling H, Appleyard M. The Copenhagen City Heart Study. *European Heart Journal supplements* 2001;3:H1-H83.
61. Tjonneland A, Olsen A, Boll K, Stripp C, Christensen J, Engholm G, Overvad K. 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;35:432-41.
62. Osler M, Linneberg A, Glumer C, Jorgensen T. The cohorts at the Research Centre for Prevention and Health, formerly 'The Glostrup population Studies'. *Int J Epidemiol* 2011;40:602-10.
63. Fortier I, Doiron D, Burton P, Raina P. Invited commentary: consolidating data harmonization--how to obtain quality and applicability? *Am J Epidemiol* 2011;174:261-4.
64. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, Rossouw JE. The Women's Health Initiative recruitment methods and results. *Ann Epidemiol* 2003;13:S18-S77.
65. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials* 1998;19:61-109.
66. Anderson GL, Manson J, Wallace R, Lund B, Hall D, Davis S, Shumaker S, Wang CY, Stein E, Prentice RL. Implementation of the Women's Health Initiative study design. *Ann Epidemiol* 2003;13:S5-17.
67. Hvidtfeldt UA, Gunter MJ, Lange T, Chlebowski RT, Lane D, Farhat GN, Freiberg MS, Keiding N, Lee JS, Prentice R, Tjonneland A, Vitolins MZ, et al. Quantifying mediating effects of endogenous estrogen and insulin in the relation between obesity, alcohol consumption, and breast cancer. *Cancer Epidemiol Biomarkers Prev* 2012;21:1203-12.
68. Gjerstorff ML. The Danish Cancer Registry. *Scand J Public Health* 2011;39:42-5.
69. Moller S, Jensen MB, Ejlertsen B, Bjerre KD, Larsen M, Hansen HB, Christiansen P, Mouridsen HT. The clinical database and the treatment guidelines of the Danish Breast Cancer Cooperative Group (DBCG); its 30-years experience and future promise. *Acta Oncol* 2008;47:506-24.
70. Curb JD, McTiernan A, Heckbert SR, Kooperberg C, Stanford J, Nevitt M, Johnson KC, Proulx-Burns L, Pastore L, Criqui M, Daugherty S. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. *Ann Epidemiol* 2003;13:S122-S128.
71. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37-48.
72. Martinussen T, Scheike TH. Additive Hazards Models. In: Martinussen T, Scheike TH. *Dynamic Regression Models for Survival Data*. New York: Springer-Verlag, 2006:103-73.
73. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;51:1173-82.
74. Vanderweele TJ. A three-way decomposition of a total effect into direct, indirect, and interactive effects. *Epidemiology* 2013;24:224-32.
75. Vanderweele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology* 2014;25:300-6.
76. Lange T, Rasmussen M, Thygesen LC. Assessing natural direct and indirect effects through multiple pathways. *Am J Epidemiol* 2014;179:513-8.
77. Tchetgen Tchetgen EJ, Vanderweele TJ. Identification of natural direct effects when a confounder of the mediator is directly affected by exposure. *Epidemiology* 2014;25:282-91.
78. Rinaldi S, Key TJ, Peeters PH, Lahmann PH, Lukanova A, Dossus L, Biessy C, Vineis P, Sacerdote C, Berrino F, Panico S, Tumino R, et al. Anthropometric measures, endogenous sex steroids and breast cancer risk in postmenopausal women: a study within the EPIC cohort. *Int J Cancer* 2006;118:2832-9.
79. Gaudet MM, Patel AV, Teras LR, Sun J, Campbell PT, Stevens VL, Jacobs EJ, Gapstur SM. Obesity-related markers and breast cancer in CPS-II Nutrition Cohort. *Int J Mol Epidemiol Genet* 2013;4:156-66.
80. Hankinson SE, Willett WC, Manson JE, Hunter DJ, Colditz GA, Stampfer MJ, Longcope C, Speizer FE. Alcohol, height, and adiposity in relation to estrogen and prolactin levels in postmenopausal women. *J Natl Cancer Inst* 1995;87:1297-302.
81. Gavaler JS, Van Thiel DH. The association between moderate alcoholic beverage consumption and serum estradiol and testosterone levels in normal postmenopausal women: relationship to the literature. *Alcohol Clin Exp Res* 1992;16:87-92.
82. Key TJ, Appleby PN, Reeves GK, Roddam AW, Helzlsouer KJ, Alberg AJ, Rollison DE, Dorgan JF, Brinton LA, Overvad K, Kaaks R, Trichopoulou A, et al. Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. *Br J Cancer* 2011;105:709-22.
83. Gavaler JS, Love K. Detection of the relationship between moderate alcoholic beverage consumption and serum levels of estradiol in normal postmenopausal women: effects of alcohol consumption quantitation methods and sample size adequacy. *J Stud Alcohol* 1992;53:389-94.
84. Munsell MF, Sprague BL, Berry DA, Chisholm G, Trentham-Dietz A. Body mass index and breast cancer risk according to postmenopausal estrogen-progestin use and hormone receptor status. *Epidemiol Rev* 2014;36:114-36.
85. Ritte R, Lukanova A, Berrino F, Dossus L, Tjonneland A, Olsen A, Overvad TF, Overvad K, Clavel-Chapelon F, Fournier A, Fagherazzi G, Rohrmann S, et al. Adiposity, hormone replacement therapy use and breast cancer risk by age and hormone receptor status: a large prospective cohort study. *Breast Cancer Res* 2012;14:R76.
86. Suzuki R, Ye W, Rylander-Rudqvist T, Saji S, Colditz GA, Wolk A. Alcohol and postmenopausal breast cancer risk defined by estrogen and progesterone receptor status: a prospective cohort study. *J Natl Cancer Inst* 2005;97:1601-8.
87. Lew JQ, Freedman ND, Leitzmann MF, Brinton LA, Hoover RN, Hollenbeck AR, Schatzkin A, Park Y. Alcohol and risk of breast cancer by histologic type and hormone receptor status in postmenopausal women: the NIH-AARP Diet and Health Study. *Am J Epidemiol* 2009;170:308-17.
88. Gertig DM, Fletcher AS, English DR, Macinnis RJ, Hopper JL, Giles GG. Hormone therapy and breast cancer: what factors modify the association? *Menopause* 2006;13:178-84.
89. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15:615-25.
90. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
91. Hernan MA and Robins JM. Measurement bias. In: Causal inference (chapter 9, pp 109-118). http://www.hsph.harvard.edu/miguel-hernan/files/2013/09/hernanrobins_v1.10.23.pdf, accessed on 24-2-2014.
92. Galobardes B, Lynch J, Smith GD. Measuring socioeconomic position in health research. *Brit Med Bull* 2007;81-82:21-37.

93. Avendano M, Kunst AE, Huisman M, van LF, Bopp M, Borrell C, Valkonen T, Regidor E, Costa G, Donkin A, Borgan JK, Deboosere P, et al. Educational level and stroke mortality: a comparison of 10 European populations during the 1990s. *Stroke* 2004;35:432-7.
94. Geyer S, Hemstrom O, Peter R, Vagero D. Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice. *J Epidemiol Community Health* 2006;60:804-10.
95. Macintyre S, McKay L, Der G, Hiscock R. Socio-economic position and health: what you observe depends on how you measure it. *J Public Health Med* 2003;25:288-94.
96. Jensen VM, Rasmussen AW. Danish Education Registers. *Scand J Public Health* 2011;39:91-4.
97. Chen JT, Beckfield J, Waterman PD, Krieger N. Can Changes in the Distributions of and Associations Between Education and Income Bias Temporal Comparisons of Health Disparities? An Exploration With Causal Graphs and Simulations. *Am J Epidemiol* 2013.
98. Blakely T, McKenzie S, Carter K. Misclassification of the mediator matters when estimating indirect effects. *J Epidemiol Community Health* 2013;67:458-66.
99. Ogburn EL, Vanderweele TJ. Analytic results on the bias due to nondifferential misclassification of a binary mediator. *Am J Epidemiol* 2012;176:555-61.
100. Vanderweele TJ, Valeri L, Ogburn EL. The role of measurement error and misclassification in mediation analysis: mediation and measurement error. *Epidemiology* 2012;23:561-4.
101. le CS, Debeij J, Rosendaal FR, Cannegieter SC, Vandenbroucke JP. Quantification of bias in direct effects estimates due to different types of measurement error in the mediator. *Epidemiology* 2012;23:551-60.
102. Gronbaek M, Heitmann BL. Validity of self-reported intakes of wine, beer and spirits in population studies. *Eur J Clin Nutr* 1996;50:487-90.
103. Giovannucci E, Colditz G, Stampfer MJ, Rimm EB, Litin L, Sampson L, Willett WC. The assessment of alcohol consumption by a simple self-administered questionnaire. *Am J Epidemiol* 1991;133:810-7.
104. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor GS, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act* 2008;5:56.
105. Sabia S, van Hees VT, Shipley MJ, Trenell MI, Hagger-Johnson G, Elbaz A, Kivimaki M, Singh-Manoux A. Association Between Questionnaire- and Accelerometer-Assessed Physical Activity: The Role of Sociodemographic Factors. *Am J Epidemiol* 2014.
106. Lee PH, Yu YY, McDowell I, Leung GM, Lam TH, Stewart SM. Performance of the international physical activity questionnaire (short form) in subgroups of the Hong Kong chinese population. *Int J Behav Nutr Phys Act* 2011;8:81.
107. Thygesen LC, Morch LS, Keiding N, Johansen C, Gronbaek M. Use of baseline and updated information on alcohol intake on risk for breast cancer: importance of latency. *Int J Epidemiol* 2008;37:669-77.
108. Barbaglia G, Macia F, Comas M, Sala M, del M, V, Casamitjana M, Castells X. Trends in hormone therapy use before and after publication of the Women's Health Initiative trial: 10 years of follow-up. *Menopause* 2009;16:1061-4.
109. Tjonneland A, Christensen J, Thomsen BL, Olsen A, Overvad K, Ewertz M, Mellemkjaer L. Hormone replacement therapy in relation to breast carcinoma incidence rate ratios: a prospective Danish cohort study. *Cancer* 2004;100:2328-37.
110. Settnes A, Jorgensen T. Hysterectomy in a Danish cohort. Prevalence, incidence and socio-demographic characteristics. *Acta Obstet Gynecol Scand* 1996;75:274-80.
111. Vanderweele TJ. Mediation analysis with multiple versions of the mediator. *Epidemiology* 2012;23:454-63.
112. Jorgensen KJ, Zahl PH, Gotzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. *BMC Womens Health* 2009;9:36.
113. von Euler-Chelpin M, Olsen AH, Njor S, Jensen A, Vejborg I, Schwartz W, Lynge E. Does educational level determine screening participation? *Eur J Cancer Prev* 2008;17:273-8.
114. Jensen LF, Pedersen AF, Andersen B, Vedsted P. Identifying specific non-attending groups in breast cancer screening--population-based registry study of participation and socio-demography. *BMC Cancer* 2012;12:518.
115. Li CI, Anderson BO, Daling JR, Moe RE. Trends in incidence rates of invasive lobular and ductal breast carcinoma. *JAMA* 2003;289:1421-4.
116. Verkooijen HM, Fioretta G, Vlastos G, Morabia A, Schubert H, Sappino AP, Pelte MF, Schafer P, Kurtz J, Bouchardy C. Important increase of invasive lobular breast cancer incidence in Geneva, Switzerland. *Int J Cancer* 2003;104:778-81.
117. Li CI, Uribe DJ, Daling JR. Clinical characteristics of different histologic types of breast cancer. *Br J Cancer* 2005;93:1046-52.
118. Li CI, Moe RE, Daling JR. Risk of mortality by histologic type of breast cancer among women aged 50 to 79 years. *Arch Intern Med* 2003;163:2149-53.
119. Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol* 2012;13:1141-51.
120. Daling JR, Malone KE, Doody DR, Voigt LF, Bernstein L, Coates RJ, Marchbanks PA, Norman SA, Weiss LK, Ursin G, Berlin JA, Burkman RT, et al. Relation of regimens of combined hormone replacement therapy to lobular, ductal, and other histologic types of breast carcinoma. *Cancer* 2002;95:2455-64.
121. Kotsopoulos J, Chen WY, Gates MA, Tworoger SS, Hankinson SE, Rosner BA. Risk factors for ductal and lobular breast cancer: results from the nurses' health study. *Breast Cancer Res* 2010;12:R106.
122. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. Simulation extrapolation. *Measurement error in nonlinear models: A modern perspective*. Boca Raton: Chapmann & Hall/CRC, 2006:97-128.
123. Hernan MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health* 2004;58:265-71.
124. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol* 2002;155:176-84.
125. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 2001;358:1389-99.

126. Pudrovska T, Anikputa B. The role of early-life socioeconomic status in breast cancer incidence and mortality: unraveling life course mechanisms. *J Aging Health* 2012;24:323-44.
127. Jeffreys M, Warren R, Gunnell D, McCarron P, Smith GD. Life course breast cancer risk factors and adult breast density (United Kingdom). *Cancer Causes Control* 2004;15:947-55.
128. Pudrovska T, Anishkin A, Shen Y. Early-Life Socioeconomic Status and the Prevalence of Breast Cancer in Later Life. *Res Aging* 2012;34:302-20.
129. Pierce MB, Leon DA. Age at menarche and adult BMI in the Aberdeen children of the 1950s cohort study. *Am J Clin Nutr* 2005;82:733-9.
130. Velie EM, Nechuta S, Osuch JR. Lifetime reproductive and anthropometric risk factors for breast cancer in postmenopausal women. *Breast Dis* 2005;24:17-35.
131. Fewell Z, Davey SG, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol* 2007;166:646-55.
132. Ogburn EL, Vanderweele TJ. Bias attenuation results for nondifferentially mismeasured ordinal and coarsened confounders. *Biometrika* 2013;100:241-8.
133. Vanderweele TJ, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface* 2009;2:457-68.
134. Wardle J, Steptoe A. Socioeconomic differences in attitudes and beliefs about healthy lifestyles. *J Epidemiol Community Health* 2003;57:440-3.
135. Iversen AC, Kraft P. Does socio-economic status and health consciousness influence how women respond to health related messages in media? *Health Educ Res* 2006;21:601-10.
136. Phipps AI, Ichikawa L, Bowles EJ, Carney PA, Kerlikowske K, Miglioretti DL, Buist DS. Defining menopausal status in epidemiologic studies: A comparison of multiple approaches and their effects on breast cancer rates. *Maturitas* 2010;67:60-6.
137. Morabia A, Flandre P. Misclassification bias related to definition of menopausal status in case-control studies of breast cancer. *Int J Epidemiol* 1992;21:222-8.
138. Greenland S. Absence of confounding does not correspond to collapsibility of the rate ratio or rate difference. *Epidemiology* 1996;7:498-501.
139. Vanderweele TJ. Causal mediation analysis with survival data. *Epidemiology* 2011;22:582-5.
140. Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *Am J Epidemiol* 2010;172:1339-48.
141. Vanderweele TJ. Unmeasured confounding and hazard scales: sensitivity analysis for total, direct, and indirect effects. *Eur J Epidemiol* 2013;28:113-7.

APPENDIX

Directed acyclic graph (DAG) of the relation between SEP and postmenopausal breast cancer