Breast cancer survivorship after taxane-based chemotherapy

Studies in premenopausal women

PhD Thesis

Cathrine Fonnesbech Hjorth

April 2022

Health

Aarhus University

Department of Clinical Epidemiology, Aarhus University Hospital

Supervisors

Deirdre Cronin Fenton
Associate Professor, PhD (Main supervisor)
Department of Clinical Epidemiology, Department of Clinical Medicine, Aarhus University Hospital and Aarhus University, Aarhus, Denmark

Per Damkier
Professor, MD, PhD
Department of Clinical Pharmacology, Odense University Hospital, Odense, Denmark
Department of Clinical Research, University of Southern Denmark, Odense, Denmark

Timothy Lee Lash
Professor, DSc
Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, USA

Bent Ejlertsen
Professor, MD, PhD, DMSc
Department of Oncology and Danish Breast Cancer Group, Rigshospitalet, Copenhagen University, Denmark

Assessment committee

Bodil Hammer Bech (Chairman),
Associate Professor, MD, PhD
Department of Public Health, Department of Epidemiology
Aarhus University, Denmark

Håvard Søiland
Professor, MD, PhD
Department of Clinical Science,
University of Bergen, Norway

Anders Ekbom
Professor, MD, PhD
Department of Medicine,
Karolinska Institutet, Sweden

Acknowledgements

"If I have seen further, it is by standing on the shoulders of giants" (Isaac Newton, 1675)

The work was carried out at the Department of Clinical Epidemiology, Aarhus University Hospital and Aarhus University, between 2018 and 2022. It has been a privilege to be surrounded by enthusiastic, smart, and dear colleagues and friends every day. Thank you all for the inspiring atmosphere and the helpful spirit.

I wish to express my genuine gratitude to my dedicated main supervisor Dee. I could not have wished for a better supervisor and mentor. You always have one ambitious eye on the target, but also an eye on the persons surrounding you. Thanks for all our professional conversations, but not least the personal ones. Altogether, your guidance has truly supported my development as a researcher. I am extremely grateful to be given the opportunity to continue working with you, Dee.

Thank you to my co-supervisors Tim Lash, Bent Ejlertsen and Per Damkier. Tim, thanks for inspiring discussions at your visits in Aarhus, and thanks for your constructive replies to my e-mails during the time where we were unable to meet in-person. Bent, thank you for sharing your in-depth clinical knowledge and thorough revisions. Per, visiting you and colleagues in Odense has been outstanding. Thanks for being a great host, for your excellent feedback on the projects, and for your skillful linguistic revisions. Also, thanks for introducing me to Tore Stage and Søren Feddersen. Tore, you have no idea how much your great mind helped me during my studies. Søren, thanks for your and Jannie's thorough work in the laboratory, and for meticulously explaining the technical details to me. Also, thanks to Anders Kjærsgaard and Dora Farkas, for your invaluable statistical support and for your patience.

Henrik Bøggild, it was a great pleasure to visit you in Aalborg, and to continue our collaboration. You introduced me to epidemiology and helped me take my first steps into research. Thank you for introducing this world to me.

A special thanks to the outstanding artist who made the title page illustration, my mother Anne Fonnesbech. Thanks to both my parents, Anne and Svend, for always stepping in and supporting me and my family.

Klaus, Edith and Olga, what should I have done without you? Klaus, thanks for your unconditional support and love all the way, for spreading laughs every day in our home, and for being the best and funniest father our girls could ever have wished for. I am so proud of us. Edith and Olga, every day you remind me about the most important things in life: unicorns, cheetahs, and somersaults - and things even bigger than that. It is the greatest honor to be your mom and to kiss you every day.

Cathrine F. Hjorth, April 2022

Grants

Kræftens Bekæmpelse Knæk Cancer [R167-A11045-17-S2]: DKK 1.950.000

Oticon Foundation [18-2137]: DKK 4.500

Kræftens Bekæmpelse [R220-A13165]: DKK 6.927

Dansk Kræftforskningsfond [PLESNER-FAST-Active.FID1839672]: DKK 73.421

Aarhus University: DKK 240.000

Thesis structure

The thesis is based on the following three original studies that are referred to by their Roman numerals (I-III).

Study I: Socioeconomic position and prognosis in premenopausal breast cancer: a population-based cohort study in Denmark.

Hjorth CF, Damkier P, Ejlertsen B, Lash T, Sørensen HT, Cronin-Fenton D. 2021. *BMC Medicine* 19(1). doi: 10.1186/s12916-021-02108-z.

Study II: Single nucleotide polymorphisms and the effectiveness of taxane-based chemotherapy in premenopausal breast cancer: A population-based cohort study in Denmark

Hjorth CF, Damkier P, Stage TB, Feddersen S, Hamilton-Dutoit S, Rørth M, Ejlertsen B, Lash T, Ahern T, Sørensen HT, Cronin-Fenton D. 2021 *[Accepted in Breast Cancer Research and Treatment]*

Study III: The impact of single nucleotide polymorphisms on return-to-work after taxane-based chemotherapy in breast cancer.

Hjorth CF, Damkier P, Stage TB, Feddersen S, Hamilton-Dutoit S, Ejlertsen B, Lash T, Bøggild H, Sørensen HT, Cronin-Fenton D. *[Submitted]*.

In the *Introduction*, I briefly present the research focus of this thesis. This is followed by the *Background*, where I introduce the epidemiology of breast cancer, followed by details on breast cancer treatment and survivorship. I describe concepts of germline variation and socioeconomic position and explore it in the context of response to taxane-based chemotherapy among premenopausal women with breast cancer in literature reviews leading to the identification of the knowledge gaps this thesis aims to fill.

This is followed by a summary of *Methods* and *Results*. I summarize the main findings of this thesis in the *Discussion* and compare them to the existing literature and discuss the validity of the findings. Based on this, I draw the main *Conclusions* and consider the implications and proposed future study aims in the *Outlook* section.

Last, I provide thesis abstracts both as an *English summary* and a *Danish summary*, followed by *References* and *Appendices I-III*, which include the published, accepted, and submitted manuscript drafts and their associated supplemental material.

Abbreviations

ABC: ATP-binding cassette

BCSM: Breast cancer specific mortality

CANTO cohort: Cancer toxicities cohort

CI: Confidence interval

CPR: Central Personal Registration

DBCG: Danish Breast Cancer Group

DFS: Disease-free-survival

DNPR: Danish National Patient Registry

DREAM: The Danish Register for Evaluation of Marginalization

ER: Estrogen receptor

FFPE: Formalin-fixed paraffin-embedded

HER2: Human epidermal growth factor receptor 2

HR: Hazard ratio

IRR: Incidence rate ratio

IQR: Interquartile range

Loess: Locally estimated scatterplot smoothing

OS: Overall survival

PFS: Progression-free survival

PR: Progesterone receptor

ProBe CaRe: Predictors of Breast Cancer Recurrence

RECIST: Response Evaluation Criteria in Solid Tumors

RR: Relative risk

RTW: Return-to-work

SEER: Surveillance, Epidemiology, and End Results

SEP: Socioeconomic position

SLMA: Stable labor market attachment

SNP: Single nucleotide polymorphism

SRR: Summary relative risk

TNM: Tumor node metastasis.

Contents

Introduction

Cancer survivorship is an inclusive term referring to aspects during and beyond cancer diagnosis and treatment. Cancer survivorship relates to a cancer survivor's physical and mental health, but also social and economic issues. Survivorship therefore starts at diagnosis of a primary cancer.[1] This thesis focuses on factors that may influence breast cancer survivorship during and after taxane-based chemotherapy. Survivorship outcomes include the risk of recurrent disease, mortality, and return-to-work. We will address how these outcomes may relate to taxane-induced adverse effects and recovery after treatment.

Premenopausal breast cancer patients undergoing chemotherapy, are generally recommended adjuvant taxane-based chemotherapy. The effectiveness of taxane-based chemotherapy—in terms of cancer recurrence and survival—varies among individuals. Women undergoing taxane-based chemotherapy with seemingly identical clinical and prognostic characteristics at the time of cancer diagnosis vary substantially in the clinical course of their disease. Some women develop recurrent disease and/or die prematurely after completing treatment, others experience disabling toxicities impacting their recovery and quality of life. Clinicians cannot distinguish these women with existing knowledge.[2] It is therefore difficult to personalize treatment and follow-up care to account for this inter-individual variability in survivorship after breast cancer in premenopausal women.

Survivorship during and after breast cancer-directed chemotherapy may be influenced by several factors including socioeconomic position (SEP) and the germline. Worldwide, there is a negative socioeconomic gradient in cancer mortality,[3–6] also after breast cancer.[3,7] Whether such a socioeconomic gradient can influence the inter-individual variability in taxane-effectiveness has not been elucidated.

Survivorship after taxane-based chemotherapy may also be influenced by germline variation. *In vivo* and *in vitro* research suggests that some single variations in the DNA, called single nucleotide polymorphisms (SNPs), influence taxane pharmacokinetics. Yet, findings on the impact of such SNPs on taxane effectiveness in terms of cancer recurrence and mortality is conflicting. Evidence suggests that SNPs are associated with taxane toxicity, particularly neurotoxicity.[8] Adverse effects of chemotherapy can delay the return to normal daily activities, such as returning to work.[9] It is not known if SNPs associated with taxane toxicity influence cancer survivor's resumption of their daily activities and their work life. This may be especially important in premenopausal women, who have a long life-expectancy and who may contribute to the workforce for many years. Identifying women who may not benefit from treatment will help facilitate tailored treatment and follow-up.

To enhance the evidence on factors influencing survivorship after taxane-based chemotherapy in premenopausal women with early breast cancer, we assessed 1) cancer prognosis in terms of recurrence and mortality, focusing on SEP, 2) the association between SNPs and recurrence and mortality, and 3) the association between SNPs and return-to-work and stable labour market attachment.

Background

Breast cancer epidemiology

Female breast cancer is the most frequently diagnosed cancer worldwide, with 2.3 million new cases and 7.8 million 5-year survivors in 2020. Improvements in survival are the result of diagnosis at an earlier stage and increasingly effective treatments.[10] Yet, breast cancer is the cancer with the highest burden in terms of disability-adjusted life years lost.[11] About one-quarter of the women diagnosed with breast cancer are premenopausal at the time of their diagnosis.[12] Risk factors for breast cancer vary among pre- and postmenopausal women. Reproductive factors including low parity, use of hormone replacement therapy, and obesity are risk factors for breast cancer in postmenopausal women, but not in premenopausal women.[13] In contrast, obesity is associated with lower risk in premenopausal women[14] but this may depend on body anthropometrics (body mass index, height, weight and waist/hip ratio).[15,16] Mutations in high penetrance genes, including *BRCA1*, *BRCA2 and PALB2*, are established predictors of breast cancer, especially in premenopausal women.[17–19] For decades, the incidence of breast cancer has been highest among women with higher education,[7] but this gradient seems to have flipped post-millennium.[20] As such, many factors influence breast cancer incidence—these factors differ according to age, menopausal status, and SEP.

In high income countries, the five-year age-standardized relative survival for breast cancer is 90%,[21] with the highest survival for localized stage (99%), and lower survival for regional (86%) and distant stage disease (29%).[22] Mortality due to breast cancer is almost certainly preceded by recurrent disease.[23]

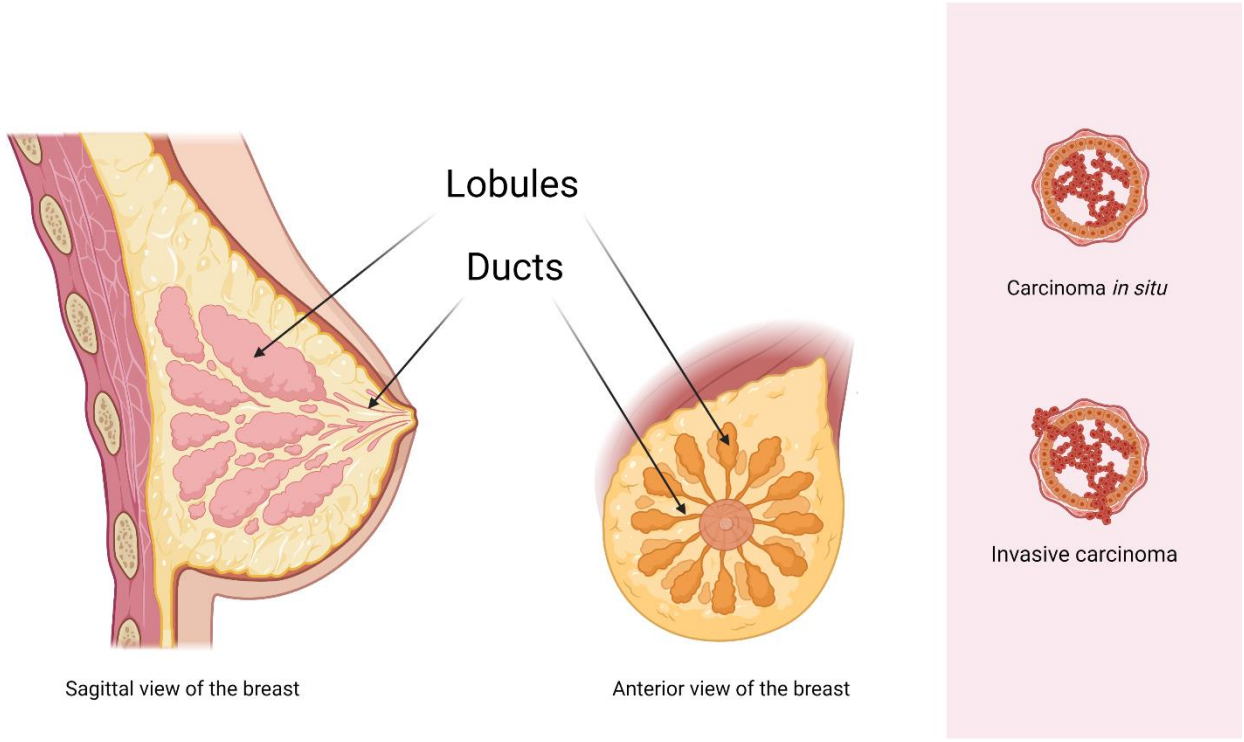
Breast cancer in premenopausal women

Premenopausal women have a higher five-year risk of recurrent breast cancer, compared with postmenopausal women.[24] Different cancer characteristics, family history and genetic risk may contribute to this poorer prognosis.[25] Distributions of breast cancer characteristics differ in pre- and postmenopausal women, but menopausal status in itself may not account for the heterogeneity. Rather, low age is considered a driving force in premenopausal women.[26] Compared with their older counterparts, the youngest premenopausal breast cancer patients (i.e., aged below 40 years at breast cancer diagnosis) have a higher frequency of aggressive tumors characterized by high proliferation, lack of hormone receptor expression, large tumor size and lymphovascular invasion.[27–29] Accordingly, premenopausal women more often undergo axillary lymph node dissection rather than sentinel lymph node biopsy alone, and have a higher frequency of chemotherapy and hormone treatment for estrogen receptor (ER) positive tumors.[30] This may also be influenced by better therapy tolerance in younger patients with higher physical performance.

Breast tumor subtypes

Breast cancer refers to a spectrum of breast tumor subtypes with distinct biological features, treatments, and survivorship. Invasive breast cancers are divided into histological subtypes and malignancy grades according to the World Health Organization. The most common are ductal carcinomas (approximately 75-80%) referring to carcinomas in the lactiferous ducts, and the less common type invasive lobular carcinomas referring to carcinomas in the milk-producing glands called *lobules* (approximately 10-15%).[31] The location of the lobules and the ducts is illustrated in Figure 1. The carcinomas are classified as *invasive* when abnormal cells have invaded surrounding tissue. Abnormal cells located within the ducts or the lobules without spread are referred to as carcinomas *in situ*, which in some cases can become invasive. Malignancy grading of invasive ductal and lobular carcinomas ranges from grade 1 denoting well-differentiated to grade 3 indicating poorly differentiated.[32]

Figure 1. Ductal- and lobular carcinomas.



Anatomy of the ducts and the lobules. In the box, the upper figure is an example of *in situ* carcinoma where abnormal cells are located inside a duct, whereas the lower figure illustrates invasive carcinoma with spread to surrounding tissue.

Breast cancer is also classified according to ER status, progesterone receptor (PR) status and human epidermal growth factor receptor 2 (HER2) status. Around two–thirds of breast tumors are hormone receptor positive, *i.e.*, ER+ and including PR+ (if available), which means that the tumor cells have receptors enabling estrogen

to bind to the cells and facilitate tumor growth. Until 2010, ER and PR were classified as positive if positive nuclear staining was seen in ≥10% of the tumor cells. Since 2010, this threshold has been set to ≥1%.[33] HER2 assessment is more complex, but simplified, HER2 negativity refers to ≤10% expression in tumor cells.[34,35]

Breast cancer stage describes the extent of the cancer including the size of the tumor and how far it has spread to lymph nodes or distant body parts. Stage is usually categorized on a scale ranging from 0 to IV, where stage 0 refers to *in situ* breast tumors with the best prognosis, and stage IV refers to metastatic cancer with spread to distant sites and poorest prognosis. This thesis focuses on stage I-III breast cancers, *i.e.* invasive non-distant metastatic cancers.[36] Sometimes, stage is referred to as localized if the tumor is confined within the breast, regional if the cancer has spread to ipsilateral axillary or subclavicular lymph nodes, and distant if the cancer has spread to other organs.[37]

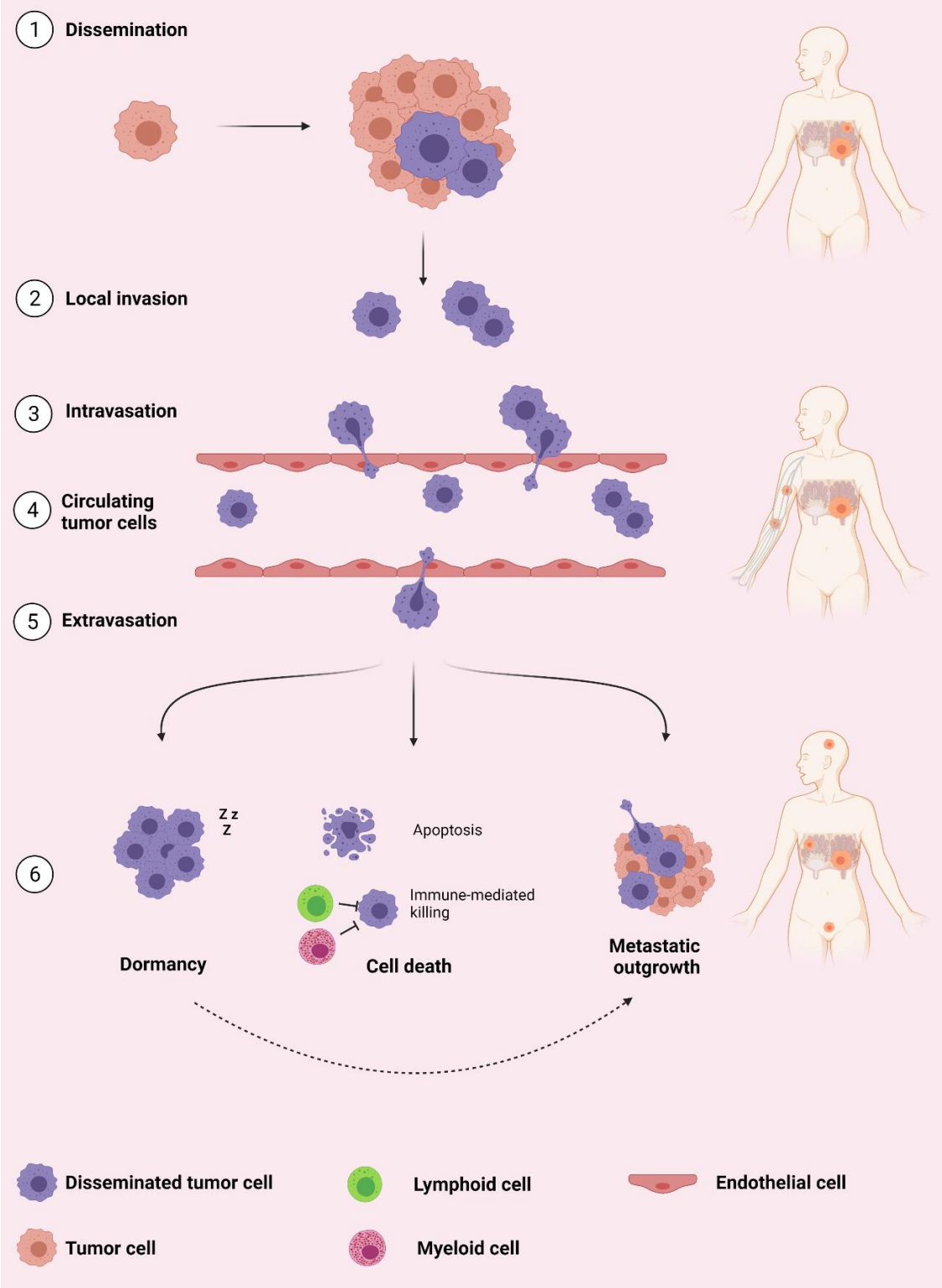
After the study period of interest in this thesis, the following four molecular classifications of breast cancer were introduced; luminal A, luminal B, HER2-enriched and basal-like. Detailed descriptions of these subgroups will not be covered in this thesis.

Breast cancer recurrence

Breast cancer is characterized by a lifelong risk of relapse, which may occur more than 30 years the after primary diagnosis.[38] Though not fully understood, the etiology underlying breast cancer recurrence may involve tumor dormancy, which refers to a latent period where micro metastases are asymptomatic. Tumor cells escape from the primary tumor and disseminate into nearby tissue (Figure 2)—seeding local recurrence. Disseminating tumor cells may also enter the circulation via the circulation or lymphatic system. Most often, the immune system attacks the circulating tumor cells, but some tumor cells evade immune detection, exit the vessels, and seed at a distant site. At this stage, disseminated tumor cells may die by apoptosis or immune-mediated killing. Some cells enter dormancy, which can last for decades. Though poorly understood, changes in the surrounding microenvironment can trigger the regeneration of dormant tumor cells.[39]

Compared to proliferating cancer cells, dormant cancer cells (also called cancer stem cells) are less sensitive to chemotherapy. Accordingly, recurrent tumors can be less treatable than primary cancers. Recurrent tumors also tend to be more aggressive and metastatic, and hence highly fatal.[23,40] Breast cancer recurrence can present as locoregional, distant or contralateral. Most often, the latter may be difficult to distinguish from a new primary breast cancer. Further details on the different recurrence sites are provided in the *Methods* section.

Figure 2. The recurrence cascade.



1) Primary tumors cells escape 2) and invade surrounding tissue. 3) Tumor cells disseminate into the blood or lymphatic system, 4) and circulate. 5) The circulating tumor cells disseminate to distant sites where they 6) either go into dormancy, undergo cell death, or tumor outgrowth. Dormancy may resume tumor growth at a later stage. Illustration created with inspiration from Riggio *et al.*[40] with permission from Springer Nature.[41]

Return-to-work after breast cancer

A return-to-work after cancer diagnosis and treatment may be a marker of recovery and return to everyday life. Return-to-work is reciprocally associated with quality of life.[42,43] As such, delayed or failed return-to-work can have negative effects on health and well-being after cancer.[44]

Up to 80% of breast cancer patients return-to-work after their cancer diagnosis,[45–47] depending on their burden of comorbid disease, SEP, the extent of cancer-directed treatment and adverse effects or complications of treatment.[45] As such, breast cancer survivors have higher unemployment rates compared with their cancer-free counterparts.[45,48] This is particularly important in premenopausal women who, in the absence of cancer, would likely contribute to the workforce for a substantial amount of years. Research in cancer survivors suggests that loss of work has a greater negative impact on quality of life and well-being than loss of health.[43] Thus, researchers have called for early identification of cancer patients at increased risk of work-force detachment, to target supportive care in these patients.[49]

Based on previous studies, Feuerstein *et al.*[50] developed a model conceptualizing barriers to return-to-work (among other work-related outcomes) in cancer survivors. Returning to work after cancer is likely influenced by multiple factors related to the individual (*e.g.*, SEP, age, health, symptoms, physical- and emotional functioning), the job (*e.g.*, type of work, work demands and work environment) and the society (*e.g.*, policies). Literature indicates that both treatment-induced toxicity[51] and chemotherapy[9] may delay the return-to-work or impact stable labour market attachment after breast cancer, though likely influenced by various factors included in Feuerstein's model. Quality of life during chemotherapy is lower than that at the time of breast cancer diagnosis.[52] It is therefore crucial to identify women at risk of delayed return-to-work, to identify those in need of extra support to aid the return to normal daily activities after cancer. Variation in return-to-work may be related to adverse effects, such as neuropathies, as women with such adverse effects may be less able-bodied.

Breast cancer treatments

A detailed description of the many facets of breast cancer treatment is beyond the scope of this thesis. The following section therefore briefly introduces the therapies relevant in this thesis and for premenopausal breast cancer patients. We focus on treatment schedules recommended by the Danish Breast Cancer Group (DBCG), who have guided breast cancer treatment in Denmark since 1977.

Treatment decisions are guided according to breast tumor subgroup, stage, and related prognostic risk groups. The majority of breast cancers diagnosed in premenopausal women are considered high-risk, except for women with tumors that are ≤2 cm, low grade (ductal grade 1 or lobular grade 1-2), ER+, HER– *and* with no lymph nodes involved.[53]

The primary treatment of breast cancer is surgery, often combined with radiotherapy and/or systemic therapy. Randomized trials conducted in the 1980s established breast conserving surgery—also known as lumpectomy—with radiotherapy as the preferred standard of local therapy in patients with early breast cancer.[54] Due to a high risk of a second breast cancer, extensive ductal carcinomas *in situ*, a lack of contour preservation by lumpectomy, or patient preference, some patients receive a total surgical removal of the breast — a mastectomy.[55,56] To reduce the risk of breast cancer recurrence, premenopausal breast cancer patients are recommended systemic therapy. The high-risk group are recommended adjuvant chemotherapy combined with endocrine- and/or HER2-directed therapy according to tumor expression of ER and HER2.

The type of endocrine therapy is determined by menopausal status. The recommended treatment for premenopausal women with ER+ breast cancer is the selective ER modulator, tamoxifen. Tamoxifen and its metabolites block estrogen from binding to the ER, preventing tumor growth and proliferation. The duration of tamoxifen has gradually extended from five years post surgery in 2007,[57] and is now recommended for *up to* 10 years.[58] Women with HER2+ tumors are candidates for immunotherapy with the agent trastuzumab (Herceptin).

In Denmark, high-risk patients have been recommended taxane-based adjuvant chemotherapy since 2007, similar to other countries.[59] Taxanes are a class of chemotherapeutic drugs covering mainly two compounds—docetaxel and paclitaxel. Docetaxel was introduced to routine clinical practice in Denmark in 2007. Taxanes are used in the treatment of various cancers including ovarian, prostate, non-small cell lung, and in breast cancer, where they are combined with other agents as primary chemotherapeutic therapy. Docetaxel-based combination chemotherapy is administered either every three weeks starting with three cycles of epirubicin and cyclophosphamide followed by three cycles of docetaxel, or as sequential docetaxel and cyclophosphamide.[60] Docetaxel dosage is based on body-surface area (60–100 mg/m^{2,} depending on regimen)[60] and administered intravenously. Though still used, docetaxel has since approximately 2013 (after the diagnostic period of interest in this thesis) gradually been replaced by nine weekly cycles of paclitaxel.[57,61]

Taxanes kill both tumor and other non-cancer proliferative cells by promoting microtubule assembly and by blocking disassembly, thereby stabilizing microtubules and prohibiting mitosis.[62] Though this thesis is based on study cohorts diagnosed while docetaxel was the recommended component, we also include knowledge and research regarding paclitaxel, due to their common mechanisms of action, some shared metabolizers and transporters, and similar adverse effects.

Later than the period of interest in this study, neo-adjuvant chemotherapy became more common in Denmark for the treatment of breast cancer in premenopausal women.

Taxane efficacy and effectiveness

The efficacy of taxanes—the result of treatment under ideal, controlled circumstances examined in clinical trials[63]—is high.[59] A meta-analysis by the Early Breast Cancer Trialists' Collaborative Group on 123 randomized clinical trials comparing polychemotherapy regimens for early breast cancer, reported a reduction in breast cancer recurrence and mortality when adding taxane to anthracycline-based treatment. After eight years, the recurrence rates were 30% vs. 35%, breast cancer specific mortality (BCSM) was 21% vs. 24% and overall mortality was 24% vs. 27%.[59] The relative risk (RR) of recurrence was RR=0.84 (95% CI: 0.78–0.91) and both BCSM and overall mortality risks ratios were RR=0.86 (95% CI: 0.79–0.93). However, a proportion of women do not benefit from this treatment. *De novo* resistance occurs in some women, especially those with triple negative breast cancer, but also independently of cancer characteristics.[64,65]

The effectiveness of taxanes—the result of treatment under "real-world" conditions, as examined in observational studies[63]—also varies among individuals. For example, adverse events such as chemotherapy-induced peripheral neuropathies cannot be predicted by age or tumor characteristics.[66] Inter-individual variability in taxane effectiveness is poorly understood, limiting the possibility of an enhanced personalized approach to cancer treatment.[64]

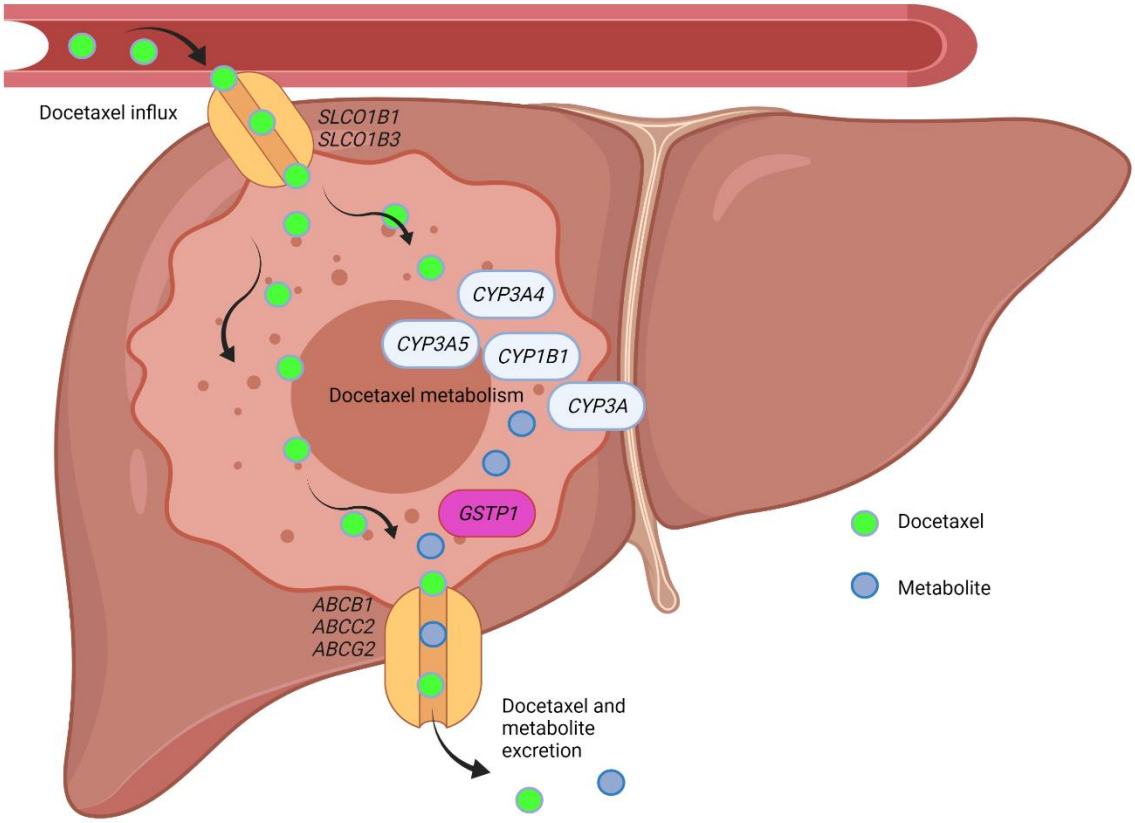
Docetaxel pharmacokinetics

Taxane-effectiveness may be influenced by inherited genetics. Variation at a single position in a DNA sequence—a SNP—is common and most often harmless. Nonetheless, SNPs may alter the function of proteins that catalyze the metabolism or facilitate the transport of these drugs.

Docetaxel is metabolized in the liver (Figure 3). The hepatic uptake is mainly facilitated by the solute carrier organic anion transporters (1B1 and 1B3 encoded by *SLCO1B3* and *SLCO1B1*). Docetaxel is metabolized by the phase I cytochrome-P-450 enzymes, CYP3A4/5 to its main metabolites. The metabolites are in turn conjugated by glutathiones, primarily by the phase II protein glutathione S-transferase pi 1 encoded by *GSTP1*. The ATP-binding cassette (ABC) transporter genes—*ABCB1* encoding the P-glycoprotein 1/multidrug

resistance protein 1, *ABCC2* encoding multidrug resistance protein 2 and *ABCG2* encoding the breast cancer resistance protein—transport the docetaxel and its metabolites into the bile.[67–70] Each of these genes are polymorphic. Docetaxel exposure and effectiveness are linked.[69] Accordingly, genetic variation in these drug metabolizing enzymes and transporters may contribute to the inter-individual effectiveness of docetaxel seen in clinical practice.[69] As such, SNPs may influence cancer survivorship. Whether a SNP has a beneficial or detrimental effect on survivorship is difficult to hypothesize. It can be hypothesized that SNPs in phase 1 metabolizers reducing enzyme activity and thereby limiting metabolism, could lead to increased plasma concentration and enhanced effectiveness. Still such potentially increased drug exposure could lead to toxicities warranting dose-reductions thereby limiting the effectiveness of treatment.

Figure 3. Genes encoding docetaxel metabolizing and transporting proteins.



Docetaxel is administered intravenously and metabolized in the liver. *SLCO1B3* and *SLCO1B1* encode OATP-transporters responsible for hepatic uptake, CYP-genes encode CYP-enzymes metabolizing docetaxel into its metabolites. The metabolites are conjugated by glutathione S-transferase pi 1, encoded by *GSTP1*, rendering the metabolites suitable for excretion. ABC-transporters, encoded by *ABCB1*, *ABCC2* and *ABCG2*, eliminate docetaxel and its metabolites.

Pharmacodynamics of peripheral neuropathy

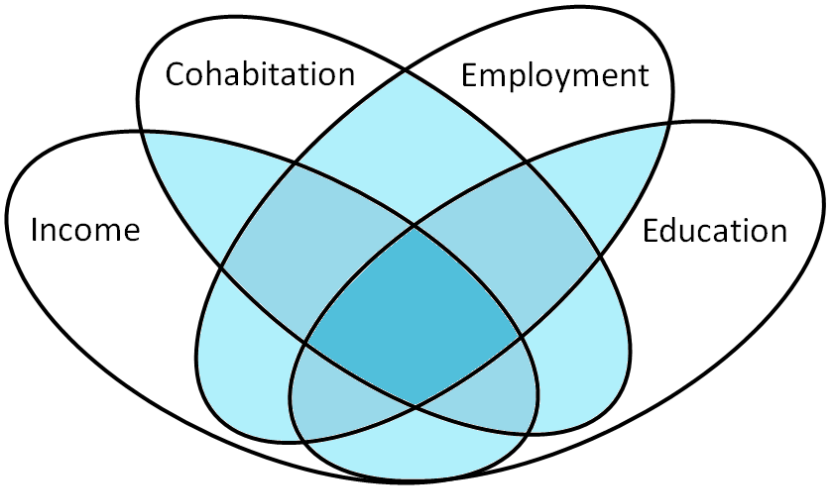
SNPs may also influence the occurrence of chemotherapy-induced peripheral neuropathy.[8] Though the occurrence of adverse effects is not the focus of this thesis, these may influence treatment effectiveness and return-to-work. Docetaxel interferes with neural cells, and can injure both the neural soma and axons.[71] Therefore, peripheral neuropathies can occur after treatment for breast cancer, though the inter-individual variation in the incidence of such neuropathies is poorly understood. Neuropathies associated with taxane treatment usually diminish after completion of treatment.[71,72] However, some patients develop chronic peripheral neuropathy.[72] Chemotherapy-induced peripheral neuropathy has a negative impact on quality of life.[72,73] While the pathogenesis is unclear, chemotherapeutic drugs that are associated with increased risk of neuropathy seem to interfere with microtubule (de-)stabilization. Genes involved in neural processes or DNA-repair, and their associated polymorphisms, have been linked to taxane-induced peripheral neuropathy. These include *EPHA* genes, a group of neural repair genes,[74–76] *ERCC* genes, which encode excision repair cross complementation proteins implicated in nucleotide repair systems,[77] *TRPV1*, which encodes the mechano- and thermal sensor transient receptor potential cation channel subfamily V member 1, which is highly expressed in sensory neurons,[78] and two genes related to Charcot-Marie-Tooth disease: *ARHGEF10* rs9657362 encoding rho guanine nucleotide exchange factor 10 and *FGD4* rs10771973 encoding frabin protein.[74,79–81]

Socioeconomic position

Cancer survivorship may also be influenced by SEP. On a societal level, SEP may influence the extent to which health care systems provide sufficient care for certain groups. At an individual-level, SEP reflects knowledge, skills, and social support, all of which can impact patient compliance with treatment and their recovery.

In line with Krieger *et al.*,[82] the underlying understanding of SEP in this thesis refers to an individual's position within the structure of society. Hence, it is not directly measurable. Instead, we use indicators reflecting different aspects of SEP, which also has a shared *core* dimension,[83] illustrated by the blue areas in Figure 4 in which four examples of SEP indicators are included. The indicators are therefore often correlated, but not interchangeable. According to Krieger *et al.*, SEP is "*an aggregate concept that includes both resource-based and prestige-based measures"*.[82] Examples of resource-based measures are income and education, reflecting material resources as well as social resources. The prestige-based measures cover, among other things, the individuals' knowledge (*e.g.* health literacy) and access to health care, which is linked to their income status and their occupational and educational prestige. SEP is most often approximated by income, education and occupation.[82–84] In the following, I elaborate on the SEP indicators and their distinct features (white areas in Figure 4).

Figure 4. Unique information of SEP by income, cohabitation, employment, and education (white areas), and their shared core dimensions (blue areas).



Income

Income reflects material resources, in addition to more proximal factors like health behavior. In non-tax funded societies, high income may facilitate health care access, thereby directly influencing health and receipt of optimal treatment. In countries with tax-funded health care, income-related differences in health care access should in theory be minimal or even non-existent. Nonetheless, costs for transport and medical out-of-pocket expenses may oppose equality. Mechanisms through which health could be influenced by income may be lifestyle—by having access to food and dwellings of high quality. Moreover, higher income may also influence self-esteem and social participation.[84] Especially in women, household income may integrate a broader picture of income and SEP, compared to individual income, as males are more often the main earner in families. Household income should be corrected for family size to enable comparison across studies.[84]

Education

Education is a well-established indicator of SEP. It is easy to categorize and can approximately be compared across countries. Moreover, it is usually stable in adulthood, is rarely affected by changes in health or other life-changing events and can predict future income and employment. Education may reflect cognitive skills, and be associated with health literacy reflecting an individual's ability to navigate and communicate within the health care system.[84,85]

Employment

Employment, or occupation, reflects health, skills, and social networks, and is naturally correlated with income and education. Detachment from the work force can be due to retirement, unemployment, or social- or health-

related reasons. Workforce detachment may be associated with poor health, lack of material resources, poor self-esteem and a limited social network.[86]

Marital status and cohabitation

Marital status is a social factor related to health and mortality[87] and may reflect social support. Marital status may be quite stable in some individuals, but may also change suddenly with great impact on health and mortality.[87] Cohabitation is related to marital status but can also capture people living together without being married.

Area-based measures of SEP

Area-based SEP can be aggregated from either individual-level data or area data. Area-based SEP can be determined by the proportion of individuals who are unemployed, have higher education, or other single SEP indicators in a particular area. Indices summarizing area-based measures of SEP are also used (*e.g.,* from census databases). Area-based measures are empirically challenged, as they tend to underestimate health associations, compared with individual-level SEP. Area-based measures may be useful in the absence of individual-level data, or when used in combination with individual-level SEP.[86] On the other hand may area-based measures be informative on a person's immediate environment.

Literature review

To identify research gaps related to variability in taxane-effectiveness in premenopausal breast cancer patients, I conducted three literature reviews to retrieve research on the following topics:

1) SEP and breast cancer recurrence and mortality (Study I),
2) SNPs and taxane effectiveness in women with breast cancer (Study II),
3) Return-to-work after breast cancer, including predictors of and barriers to return-to-work (Study III), as no research had been conducted on SNPs and return-to-work (see Table 1).

For each study, I conducted systematic searches using the databases MEDLINE (Pubmed) and EMBASE. The searches were structured by dividing the search fields into blocks, to ensure a systematic search.[88] Within each block I used the Boolean term "OR" to combine synonymous search terms and widen the search field. Then, blocks were combined with "AND" to restrict and focus the search. In MEDLINE, I used Medical Subject Heading (MeSH) terms. In EMBASE, I used the corresponding Emtree terms. The last search date was 10th March 2022.

Selection criteria

I included observational studies, systematic reviews and meta-analyses written in English (we also searched for studies written in Danish, Norwegian or Swedish). In search I (Study I), studies that only focused on race or insurance coverage were excluded. In search II (Study II), I included studies with the endpoints recurrence, disease-free survival (DFS), progression-free survival (PFS), BCSM, all-cause mortality or overall survival (OS). Thus, the search included studies investigating PFS in metastatic breast cancer. I did this to broadly explore potential pharmacokinetic mechanisms that could influence taxane effectiveness, but also to avoid discarding studies looking at DFS, as sometimes PFS was used interchangeably with DFS.[89] In search III (Study III), I included studies examining direct measures of return-to-work or indirect measures (unemployment, sick-leave, disability pension), and excluded studies focusing on interventions to facilitate return-to-work. I included pre-, peri- and postmenopausal breast cancers of all stages. In searches I and II, I limited to publication in years 2002-2022. Due to an extensive yield of literature in search III, I included systematic reviews and meta-analyses published within the past 10 years (2012–2022) and original papers published within the past five years (2017–2022). A summary of the included papers including methods, main findings and limitations are provided in Supplemental Table S1. Summaries of the reviewed literature, including publication information, methods, main findings and limitations, are provided in Supplemental Tables S1–S3.

Table 1. Search strings and search results.

Search 1			
#1 Pubmed	(("breast neoplasms/mortality"[MeSH Terms] OR "Recurrence"[MeSH Terms] OR "neoplasms, second primary"[MeSH Terms] OR "neoplasm recurrence, local"[MeSH Terms] OR "Mortality"[MeSH Terms]) AND "Breast Neoplasms"[MeSH Terms] AND "Socioeconomic Factors"[MeSH Terms]) AND ((fha[Filter]) AND (danish[Filter] OR english[Filter] OR norwegian[Filter] OR swedish[Filter]) AND (2002:2022[pdat]))	545 hits	22 relevant
# 2 Embase	'breast cancer'/exp AND ('breast cancer recurrence'/exp OR 'cancer mortality'/exp OR 'mortality'/exp) AND ('social class'/exp OR 'economic status'/exp OR 'income group'/exp OR 'named groups by marital status'/exp OR 'occupation'/exp) AND ([article]/lim OR [review]/lim) AND ([danish]/lim OR [english]/lim OR [norwegian]/lim OR [swedish]/lim) AND [2002-2022]/py AND [abstracts]/lim AND 'breast cancer'/dm	493 hits	
Search II			
#3 Pubmed	("Taxoids"[MeSH Terms] AND ("polymorphism, genetic"[MeSH Terms] OR "Pharmacogenetics"[MeSH Terms])) AND (english[Filter]) Filters: Danish, English, Norwegian, Swedish, from 2002 - 2022	414 hits	9 relevant
#4 Embase	('pharmacogenetics'/exp OR 'genetic polymorphism'/exp) AND ('docetaxel'/exp OR 'taxoid'/exp) AND 'breast tumor'/exp AND ([danish]/lim OR [english]/lim OR [norwegian]/lim OR [swedish]/lim) AND [2002-2022]/py AND ([article]/lim OR [review]/lim)	485 hits	
Search III			
#5 Pubmed	(("return to work"[MeSH Terms] OR "employment"[MeSH Terms] OR "return to work"[Title/Abstract] OR "labour market attachment"[Title/Abstract] OR "unemployment"[Title])) AND ("Polymorphism, Single Nucleotide"[Mesh])	13 hits	0 relevant
#6 Embase	('pharmacogenetics'/exp OR 'genetic polymorphism'/exp) AND ('docetaxel'/exp OR 'taxoid'/exp) AND ('return to work':ab,ti OR 'labour market attachment':ab,ti OR employment:ti OR unemployment:ti OR 'return to work'/exp)	0 hits	
#7 Pubmed	(("return to work"[MeSH Terms] OR "employment"[MeSH Terms] OR "return to work"[Title/Abstract] OR "labour market attachment"[Title/Abstract] OR "unemployment"[Title]) AND "breast neoplasms"[MeSH Terms]) AND ((danish[Filter] OR english[Filter] OR norwegian[Filter] OR swedish[Filter]) AND (2012:2022[pdat]))	248 hits	17 relevant
#8 Embase	('return to work':ab,ti OR 'labour market attachment':ab,ti OR employment:ti OR unemployment:ti OR 'return to work'/exp) AND 'breast cancer'/exp AND ([danish]/lim OR [english]/lim OR [norwegian]/lim OR [swedish]/lim) AND (2012:py OR 2013:py OR 2014:py OR 2015:py OR 2016:py OR 2017:py OR 2018:py OR 2019:py OR 2020:py OR 2021:py OR 2022:py)	292 hits	

Socioeconomic position and breast cancer recurrence and mortality (Study I).

A summary of the included papers is provided in Table 2. Several studies have focused on SEP and mortality, whereas recurrence was examined in two studies.[90,91] The main SEP indicator investigated was education,[3,13,20,90,92–99] followed by income[3,92–95] and marital status/cohabitation.[3,90,92,97,100–103] Five studies focused on occupation.[3,7,94,104,105] Most studies were registry- and/or population-based studies, with exception of one trial-based study.[97] Nine studies were conducted in the Nordic countries (Denmark, Sweden and Norway), which are characterized by universal, tax-financed health care.

Overall, current evidence agreed that low education was associated with increased mortality after breast cancer.[3,7,20,92–95,98,99,106–108] Trewin *et al.*[20,95] conducted two studies in Norway focused on the association of education with breast cancer risk and mortality. The first described the education gradient and breast cancer incidence and BCSM during 1971 to 2009.[20] They found that higher education was associated with increased breast cancer incidence and BCSM pre-millennium, but the gradient for BCSM attenuated around the year 2000. When they restricted to women aged 35–49 years, the gradient was null during the 80's and 90's. During 2000–2009, a negative education gradient appeared with higher BCSM in women with low or middle education level, trending towards increasing inequality. This was confirmed in both absolute and relative difference estimates. Their later study focusing on stage-specific survival confirmed inequality in prognosis: 5-year stage-specific survival in regional and distant breast cancer improved during 2000–2015, but not among women with low SEP—considering SEP as a composite measure of education and income.[95] One French study did not find an association between education and mortality in breast cancer.[13]

Three studies[92,104,105] found that women with low occupational status (*i.e.*, manual workers, skilled and unskilled workers, farmers,[104,105] and unemployed[104]) had increased all-cause-[104,105] or BCSM,[104] but this was not seen by Dalton *et al.*.[3] One of the studies reported greater inequality among premenopausal women (age <50 years) than in postmenopausal.[105]

Five studies documented higher all-cause and BCSM in unmarried, [100–103] divorced/widowed[97] or women living alone[90] regardless of tumor stage. One study reported that this potential disparity was most pronounced among women with ER+ and/or PR+ and HER2– tumors, though still evident in women with ER–/HER2+ and triple negative breast cancers.[102] One Danish study did not detect an association between cohabitation status and BCSM.[3]

Rasmussen *et al.*[90] used an algorithm to identify recurrence in 27,752 women with breast cancer. They found a cumulative incidence of breast cancer recurrence of approximately 13% after 8 years of follow-up. Using multivariate Cox models, they found a slightly increased rate in women living alone, compared to married/cohabiting, and lower risk in women with medium or long education, compared to those with shorter

education. The latest study in the search (published after Study II) found increased risk of recurrence in women with low (area-based) SEP in women diagnosed with breast cancer before age 40 years.[109]

Given that SEP may be modified by age/menopausal status,[105] it is a limitation that some studies did not distinguish by any of these factors.[3,7,100,102,110] Moreover, metastatic patients, were included in eight studies.[3,20,91,92,98,99,101] The poor prognosis in metastatic breast cancer may overrule any potential role of SEP. Last, the studies generally lacked information on cancer characteristics and comorbidities, which likely influence the associations.

Table 2. Summary of literature on SEP and breast cancer recurrence and mortality (Study I).

Author, journal, year	Design, country, data sources and study period	Population, exposures, relevant outcomes	End of study. median follow-up	Results, limitations
		Reviews and meta-analyses		
Lundqvist *et al.*[111] European Journal of Public Health 2016	Systematic review and meta-analysis including 25 articles. European OECD-countries 2006–2016	European population-based studies including female breast cancer patients, using individual-level data. Exposures: • Education, income, occupation and a SEP index Outcomes: • Breast cancer incidence, case-fatality and mortality	N/A	Decreased case-fatality (SRR: 0.72, 9%% CI: 0.63–0.81) and increased mortality (SRR: 1.25, 95% CI: 1.17–1.32) in women with high SEP. Limitations: • No stratified analysis according to age/menopausal status, ER status or stage.
		Original articles		
Aizer *et al.*[100] Journal of Clinical Oncology 2013	Cohort study USA The National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program 2004–2008	Women aged ≥18 years diagnosed with non-metastatic breast cancer and complete clinical information, and known cause of death (n=183,596). Exposure: • Marital status Outcome: • BCSM	N/A	Married women had decreased risk of BCSM (HR: 0.80, CI: 0.79–0.81), compared to unmarried. Examination of unmarried subgroups (never married, separated, divorced or widowed) showed that all subgroups had higher BCSM than married. Limitations: • No data on chemotherapy • Some unmarried women might be living with a partner • Used area-based SEP indicators for adjustment
Beiki *et al.*[106] Breast Cancer Research 2012	Cohort study Sweden The Swedish Cancer Registry The National Housing Censuses 1961–2007.	Women diagnosed with primary invasive breast cancer who survived one month (n=76,152) Exposure: • Education and ethnicity Outcome: • All-cause mortality and BCSM	31. December 2007	Women with the lowest education level 30% higher risk of dying compared to women with high education. Limitations: • The overall results (not considering ethnicity) lacked information on precision. • Lacked clinical information including hormone receptor status, treatments, stage and comorbidities.
Van Maaren *et al.*[109] Cancer Epidemiology 2022	Cohort study Netherlands Netherlands Cancer Registry 2005	Women aged <40 years, who were operated for stage I–III primary breast cancer (n=525). Exposures: • Area-based SEP Outcomes: • Recurrence and mortality	10 years follow-up	Women with high SEP had lower risk of recurrence, and longer time between surgery and recurrence. Among women with low SEP, 19% died following recurrence, which was more than seen in medium and high SEP women (15% and 13%, respectively). Limitations: • No individual-level indicators of SEP or comorbidities
Di Salvo *et al.*[91] Oncotarget 2017	Cohort study Italy 34 cancer registries within the Italian Cancer Registry Association 2003–2005	Women diagnosed with breast cancer aged ≥15 years regardless of stage (n=3,358). Exposure: • Italian Deprivation Index including education, unemployment, one-parent status, home rental and home overcrowding. Outcomes: • Recurrence, sentinel lymph node biopsy and lumpectomies	31. December 2010	In women aged <65 years, most deprived women had higher cumulative incidence of recurrence than seen in least deprived, especially in women with ER+/PR+ disease. HR showed increased recurrence risk of 1.26 (95% CI: 0.99–1.68) Limitations: • No individual-level indicators of SEP or comorbidities • No information on censoring by death or other relevant events. • Used Kaplan-Meier to examine cumulative incidence (thus, death was not considered as a competing event).

Trewin *et al.*[112] European Journal of Public Health 2017	Population-based cohort study Norway Cancer Registry of Norway, Central Population Registry, National Education Database, Cause of Death Registry 1971–2009	All female inhabitants registered in Norway above 35 years with known education level (n=2,084,143) Exposure: • Education level Outcome: • Breast cancer incidence and BCSM	31. December 2009	Among women aged 35-50 years, there was an education gradient in mortality. Compared with women with low education, women with high education had 28% lower mortality during 2000–2009. Limitations: • No information on tumor characteristics • Included women with metastatic breast cancer • No information on treatments • No information on recurrences
Trewin *et al.*[113] Breast Cancer Research and Treatment 2020	Population-based matched cohort study Norway Cancer Registry of Norway, Central Population Registry, National Education database, Register for Personal Tax Payers 2000–2015	All Norwegian women aged 30–48 years diagnosed with invasive breast cancer (n=7501). Matched comparison cohort included women in Norway matched on age and calendar year Exposures: • Income. Education and combined income and education Outcome: • Stage specific excess mortality	31. December 2017 Median follow-up: Localized: 8.37 years Regional: 7.7 years Distant: 3.4 years Unknown stage: 7.8 years.	During 2000–2015, 5-year survival improved in both regional and distant stage cancers in women with high SEP, but stayed the same (9% and 11%, respectively) in women with low SEP. Increased excess mortality rate in women with low education or income. Limitations: • No stratification on ER status • No information on treatments
Rasmussen *et al.*[90] European Journal of Cancer care 2019	Population-based cohort study Denmark DNPR, Danish Cancer Register and Statistics Denmark 2008–2016	49,208 patients with non-metastatic cancer (7 sites), hereof 27,752 were diagnosed with breast cancer, aged 18 or older. Exposures: • Marital status (cohabitation) and education level Outcomes: • Recurrence and second primary cancer	31. December 2016 Median follow-up: 3.3 years [1.5–5.75]	Increased risk of recurrence in women living alone, and in women with low education. Limitations: • No stratification by ER status • Aggregated treatment information (chemotherapy and radiotherapy) • Marital status collected the completion of treatment
Larsen *et al.*[114] Acta Oncologica 2015	Cohort study Denmark Danish Diet and Health Study, Danish Central Population Register, Danish Cancer Registry, Integrated Database for Labour Market Research, DNPR, National Diabetes Register, DBCG 1993-1997	Postmenopausal women diagnosed with stage I–III breast cancer (n=1,250). Exposure: • Education and income Outcome: • Mortality	31. December 2013 Median follow-up: 9.6 years [IQR: 0.7–12.0]	All-cause mortality was increased in women with low education and in women with medium income, compared to high education and income, respectively. Limitations: • Prone to selection bias, as patients had to accept an invitation to enroll, and only women in Aarhus and Copenhagen were invited. • SEP data collected after breast cancer diagnosis for approx. half of the cohort
Elstad *et al.*[107] European Journal of Public Health 2012	Cohort study Norway Statistics Norway 1971–2002	Women with breast cancer aged 45–74 years (n=2,734) Exposure: • Education Outcome: • Mortality	N/A	Slightly decreased risk in women with higher education Limitation: • No information on stage or ER status

Herndon *et al.* [108] Psychooncology. 2013	The CALGB clinical trials United States CALGB database and questionnaires 1987–1998	Women with early-stage breast cancer (n=5,146) Exposures: • Education, marital status and race Outcomes: • Mortality	November 1997–August 2009 depending on trial	Women with low education had higher mortality and married/single women had lower mortality than divorced or widowed women. Limitations: • Trial patients
Gentil-Brevet *et al.* [104] British Journal of Cancer 2008	Population-based cohort France Côte d'Or breast cancer registry 1995–1997	Women with non-metastatic breast cancer (n=1,011) Exposure: • Occupation (Education if missing) Outcome: • Survival	June 2006	Lower occupation class was associated with increased mortality Limitations: • No stratification by ER status • No information on treatments
Abdoli *et al.* [98] Breast 2017	Population-based matched cohort study Sweden Migration and Health Cohort 2004–2009	Women diagnosed with invasive breast cancer (n= 35,268) Exposures: • Education Outcomes: • Mortality	31. December 2009	Increased all-cause mortality and excess mortality (BCSM) in low educated Limitations: • Included metastatic breast cancer, without presenting stratified results. • Not stratified by age or tumor characteristics
Lagerlund *et al.* [92] Cancer Causes Control 2005	Population-based cohort study Sweden The Swedish Cancer Register, Regional Cancer Registers, the Census databases, the Fertility Register, the Migration Register and the Cause of Death Register	All women diagnosed with breast cancer in 1993 with no prior history of cancer (n=4,645) Exposures: • Socioeconomic index, education, income (individual and household), home ownership and cohabitation Outcomes: • Survival and mortality	31. December 1998	Increased mortality in women with low household SEP, women with low education and women living alone. Limitations: • Included metastatic breast cancer, without presenting stratified results. • Not stratified by age or tumor characteristics • No information on treatments • Short follow-up • Low completeness of SEP indicators.
Dalton *et al.* [115] International Journal of Cancer 2007	Population-based cohort study Denmark DBCG, Integrated Database for Labour Market Research, Statistics Denmark, Danish Psychiatric Case Register, Cause of Death Register 1983–1999	All Danish women diagnosed with breast cancer who underwent protocol-based treatment (n=25,897). Exposures: • Occupation, education, household income, cohabitation, housing tenure, size of dwelling and urbanicity Outcomes: • Survival and BCSM	Survival: 30. September 2005 12.6 years BCSM: 31. December 2001 9.1 years	Increased mortality in women with low education, low income and smaller dwellings after accounting for other prognostic factors Limitations: • Considered tumor characteristics as confounders, rather than mediators. • Included metastatic patients • No analyses examining SEP according to treatments or age.
Hinyard *et al.* [103] Breast 2017	Cohort study United states SEER program 2004–2009	Women aged 25–64 diagnosed with invasive breast cancer (n=166,701) Exposure: • Marital status. Unmarried included domestic partners. Outcomes: • Survival and breast cancer specific survival	31. December 2013	Regardless of stage, unmarried women had increased all-cause and breast cancer specific mortality. Also when adjusted for patient and tumor characteristics. Limitations: • May be confounded by insurance status • Considered tumor characteristics as confounders, rather than mediators (no stratification by ER status).

Bouchardy *et al.* [105] International Journal of Cancer 2006	Population-based cohort study Switzerland Geneva cancer registry 1980–2000	Women diagnosed with breast cancer aged <70 years (n=3,920) Exposure: • Occupation Outcome: • BCSM	N/A	Regardless of patient, tumor and treatment characteristics, women with low SEP had increased 5-year and 10-year BCSM. Stage was less equally distributed in premenopausal women, than in postmenopausal Limitations: • Considered tumor characteristics as confounders, rather than mediators (no stratification by ER status).
Hussain *et al.* [99] International Journal of Cancer 2008	Population-based cohort study Sweden The Swedish Family-Cancer Database, Swedish Cause of Death Register 1990–2004	Women who were cancer-free, alive and Swedish resident 1 July 1990, aged 30-64 (n=1,571,511) Exposure: • Education level Outcomes: • Invasive and *in situ* breast cancer and BCSM	31. December 2004	Risk of dying from invasive breast cancer was highest in women low education, and attenuated by higher education level. Limitations: • Included metastatic breast cancer, without presenting stratified results. • Not information on tumor characteristics • No information on treatments
Rutqvist *et al.* [94] International Journal of Cancer 2006 .	Cohort study Sweden The Swedish Cancer Registry, Stockholm Breast Cancer Database, Swedish Cause-of-Death Registry, Swedish National Censuses, Register of the Total Population and The National Register of Education 1977–1997	Women diagnosed with primary breast cancer during aged below 75 years (n=15,021) in the Stockholm area. Exposures: • Education, income and occupation Outcomes: • OS, stage- and cause-specific survival	8 years, range 1–22	High income, education and occupation correlated with lower clinical stage and vice versa. Focusing on education, no difference in BCSM was observed in stage I and II. In stage III, women with low education had higher BCSM. Limitations: • Not stratified by age, treatments or ER status • 22% of breast cancers were not included (incomplete database) • Primarily urban population, which might give selection of high SEP. • Only presented results from their examination of education, but not income and occupation.
Dasgupta *et al.* [101] Cancer Epidemiology 2016	Population-based matched cohort study Australia Queensland Cancer Registry, Queensland Registrar of Births, Deaths and Marriages and the National Death Index 1996–2012	People diagnosed with one of 10 cancers, aged 20–79 with known marital status. This included 34,217 women with breast cancer. Exposures: • Marital status Outcomes: • Cancer-specific mortality, other and non-cancer mortality	31. December 2012	Women with no partner had increased cause-specific, other cancer and non-cancer mortality. Also when adjusted for age and stage. Limitations: • Not stratified by stage, and stage IV was included (did perform stage adjusted analyses though). • No information on treatments and breast cancer subtypes
Martínez *et al.* [102] Plos One 2007	Population-based cohort study California California Cancer Registry, Census 2000 And the American Community Survey 2005–2012	Women diagnosed with stage I–IV breast cancer (n=145,564) Exposure: • Marital status and neighborhood SEP Outcome: • Mortality	31. December 2013	Unmarried women had increased mortality regardless of tumor subtype, but this was more pronounced in women with ER+//HER2– than in women with triple negative breast cancers. Limitations: • No information on treatments • No age or menopausal specific analyses • No analyses restricted to non-metastatic breast cancer

Menvielle *et al.* [13] British Journal of Cancer 2006	Cohort study France The French National Statistics and French national death registry 1968-1996	Demographic sample, covering 1% of the French population aged 35-74 years. (n= 407,435). Exposure: • Education Outcome: • BCSM, overall and by age of death and birth cohort	An education gradient changed by birth cohorts, suggesting than an increased mortality in higher educated disappeared in the recent cohorts. Limitations: • No data on tumor characteristics or comorbidity.

Abbreviations: BCSM= Breast cancer specific mortality, CALGB= Cancer and Leukemia Group B, CI= Confidence interval, ER= Estrogen receptor, HER2= Human epidermal growth factor receptor 2, HR= Hazard ratio, OECD= Organization for Economic Co-operation and Development, OS= Overall survival, PR= Progesterone receptor, SEER= Surveillance, Epidemiology, and End Results, SEP= Socioeconomic position, SRR=Summary relative risk.

SNPs and breast cancer outcomes (Study II)

A summary of the included papers is provided in Table 3. Eight studies were eligible for inclusion.[89,116–122] One study was cross-sectional,[116] six were cohort studies[89,117–121] and one was a clinical trial.[122] The studies therefore covered both efficacy and effectiveness of taxanes. Seven studies were in Asian populations[89,116–118,120–122] and one study was conducted in the United States and included 89% Caucasian, 8% Asian and 8% African-American.[119]

Three studies[117,118,122] included patients with metastatic breast cancer. One of these studies examined paclitaxel monotherapy in patients previously treated with anthracyclines.[117] Another study included patients treated with docetaxel and capecitabine as either first, second line, or more.[118] The third study in metastatic breast cancer was a clinical trial with patients randomized to either docetaxel and thiotepa, or docetaxel and capecitabine.[122]

Treatment regimens also varied in the studies of patients with non-metastatic breast cancer. Three studies investigated docetaxel-based treatment; one focused on adjuvant docetaxel and epirubicin,[89] one on neoadjuvant docetaxel/thiotepa,[120] and one where patients received both neoadjuvant and adjuvant docetaxel and doxorubicin.[121] Marsh *et al.*[119] included both stage III and metastatic breast cancer patients treated with adjuvant paclitaxel, doxorubicin and cyclophosphamide. Last, the study in non-metastatic triple negative breast cancers included patients who received taxane, adriamycin and cyclophosphamide, with no information on the type of taxane.[116]

Findings by Abdul Aziz *et al.*[116] and Marsh *et al.*[119] suggest that *CYP1B1* rs1056836 variant carriers had decreased risk of recurrence (locoregional or distant) and longer PFS, respectively. Such an association was not detected by Dong *et al.*[123] who reported longer PFS in *CYP1A1* rs1048943 variant carriers. In contrast, Zhou *et al.*[122] found shorter PFS in *CYP1A1* rs1048943 for docetaxel combinations. However, in women treated with docetaxel and thiotepa, variant carriers had longer PFS and OS; when treated with docetaxel and capecitabine the opposite was seen. The authors did not propose any explanation.

Studies on *ABCB1* rs1045642 used several models to examine associations with taxane effectiveness. In a dominant model, Kim *et al.*[121] reported longer OS in wildtypes, compared to variant carriers. Compared to homozygotes, Chang *et al.*[117] reported that heterozygotes had shorter OS than homozygotes. Li *et al.*[89] reported longer PFS in wildtypes. Together, these findings suggest unfavorable impact of variants in *ABCB1* rs1045642. No associations with OS or PFS were detected in *ABCB1* rs2032582[117,119,121] and rs1128503[119,121]. *ABCG2* rs2231142 was examined in two studies,[89,119] but only one study found an association in terms of longer PFS in wildtypes.[89] No studies examined *SLCO1B1* or *SLCO1B3*, or genes encoding DNA-repair pathways.

The studies had methodological limitations. Three studies[117,120,122] fitted multivariate Cox regression models potentially over adjusting the estimates, as patient and cancer characteristics are unlikely to fulfill one of the

criteria as confounders (*i.e.*, the association with the exposure). Several studies may not be directly relevant to the population of interest in the current thesis. First, none of the studies were solely in premenopausal breast cancer patients. Second, the studies were heterogeneous in terms of treatment regimen, and none represented the taxane-combination of interest in the thesis (docetaxel and cyclophosphamide, ±epirubicin). Third, most studies were conducted in Asian populations, which generally have different allele frequencies compared with European populations. Thus, the findings may not be applicable to more heterogeneous, Western populations. Last, the results observed in studies that included metastatic patients may be influenced by the severity of the disease overruling the impact of SNPs. For example, in the study by Dong *et al.*,[123] all patients died during follow-up.

Table 3. Summary of literature on SNPs and taxane effectiveness in women with breast cancer (Study II)

Author, journal, year	Design, setting, DNA sources and study period	Population, exposures, outcome	End of study. median follow-up	Results, limitations
		Original articles		
Abdul Aziz et al.[116] Asian Pacific Journal of Cancer Prevention 2021	Cross-sectional study Malaysia DNA isolated from plasma Period: N/A	Women with non-metastatic triple negative breast cancer (51% premenopausal) treated with taxane (compound not specified), adriamycin and cyclophosphamide (n=76). Exposure: • *CYP1B1* rs1056836 Outcome: • DFS defined as time from first chemotherapy to locoregional or distant recurrence	One year of follow-up	CYP1B1 rs1056836: wildtypes had increased risk of recurrence (HR: 2.5, 95% CI: 1.10–5.66). Limitations: • Blood sample collected during follow-up, risking temporal selection and immortal time bias.[124] • When evaluating time to DFS, follow-up started at first chemotherapy cycle, so the study was likely prone to immortal time bias. • Used Kaplan-Meier to examine cumulative incidence (did not accounted for death as a competing risk). • Underpowered.
Chang et al.[117] Annals of Oncology 2009	Cohort study Korea DNA isolated from plasma 1998–2005	Women with metastatic breast cancer ≥18 years and life expectancy ≥12 months (n=121), receiving paclitaxel. Exposures: • *ABCB1* rs1045642 and rs2032582 Outcomes: • OS, tumor response and toxicities	Median time to progression: 5.7 months (95% CI: 4.6–6.8) OS: 18.1 months (95% CI: 12.5–23.8).	*ABCB1* rs1045642: The CT (*i.e.*, AG) genotype was associated with shorter OS rate, compared to CC (GG) (HR: 3.51, 95% CI: 1.16-3.51) (NB: there appears to be an error in either the estimate or 95% CI) Limitations: • Used multivariate Cox model. • Blood sample collected during follow-up, risking temporal selection and immortal time bias.[124]
Dong et al.[118] Journal of Cancer Research and Clinical Oncology 2012	Single-institute cohort study China DNA isolated from plasma 2007–2010	Females aged ≥18 years with metastatic breast cancer (n=69) treated with docetaxel and capecitabine. Exposures: • 79 SNPs *e.g.*, *CYP1B1* rs1056836, *CYP1A1* rs4646422 and rs1048943, *CYP2D6* rs16947, rs1065852, rs1058164 Outcome: • PFS and OS	5.5 years of follow-up	PFS: Carriers of *CYP1A1* rs1048943 G allele had longer PFS OS: Carriers of *CYP1A1* rs1048943 G allele had longer OS Limitations: • Limited power. • Blood sample collected during follow-up, risking temporal selection and immortal time bias.[124]
Marsh et al.[119] The Pharmacogenomics Journal 2007	Cohort study USA DNA isolated from plasma Period N/A	Women with high-risk or stage III breast cancer who received paclitaxel, doxorubicin and cyclophosphamide either as consolidation therapy following induction adjuvant therapy (stage III, 84%), or as induction therapy (stage IV, 16%) (n=93). Exposures: • *CYP2C8*3* and **4*, *CYP1B1* rs1056836, *CYP3A4* rs2740574, *CYP3A5* rs776746, *ABCB1* rs2032582 and rs1128503, *ABCG2* rs2231142 Outcome: • Paclitaxel clearance and PFS	55 months of follow-up (~4.5 years)	*CYP1B1* rs1056836 GG was associated with shorter PFS than patients with at least one C allele. No associations were found with paclitaxel clearance. Limitations: • No time-to-event analyses. • Did not present estimates of clearance analysis. They report no associations, but did not allow the reader to assess results (based their assessments on p-values).

Li et al.[89] Oncotarget 2017	Cohort study China DNA isolated from carcinoma and para-carcinoma tissue (breast cancer patients) and from plasma (health comparisons). 2010–2016	Women with incident stage II–III breast cancer (n=100, 48% premenopausal), who underwent surgical treatment and treated with adjuvant docetaxel and epirubicin. Exposure: • *ABCB1* s1045642 and *ABCG2* rs2231142 Outcome: • PFS and OS	Maximum 5 years of follow-up	*ABCB1* s1045642 TT vs. CC and *ABCG2* rs2231142 CC had longer DFS (HR: 2.12, 95% CI: 1.04–4.36 and HR: 2.04, 95% CI: 1.13–3.69, respectively). Limitations: • No information on choice of index date and any censoring. • Follow-up maximum 5 years. • Low sample size. • Several mistypes and inconsistencies in the text.
Wang et al.[120] Genetics and Molecular Research 2015	Cohort study China DNA isolated from plasma Period N/A	Women with stage I–II invasive ductal breast cancer (n=262, 59% premenopausal) treated with neo-adjuvant docetaxel and thiopepa. Exposures: • *GSTP1* rs1695, *GSTM1* and *GSTT1*. Outcomes: • Tumor response according to RECIST and OS	N/A	No association with OS. Limitations: • Used multivariate Cox- regression model to examine predictive value of the exposures. • No information on censoring.
Kim et al.[121] Cancer science 2015	Cohort study Korea DNA isolated from plasma 2003–2008	Women with stage II–III breast cancer (n=216, 68% premenopausal) treated with neo-adjuvant docetaxel and doxorubicin, followed by surgery and adjuvant docetaxel and doxorubicin Exposures: • *ABCB1* rs1045642, rs2032582 and rs1128503 and *CYP3A* rs776746 Outcomes: • Response according to RECIST, Recurrence-free survival and OS	Median follow-up: 85.4 months (range 26.2–117.8 months) ~median 7 years of follow-up	75 patients (365%) experienced recurrence, 43 (20%) died. Recurrence-free survival: No associations. OS: *ABCB1* rs1045642 TT carriers had longer OS than CT/CC but not confirmed in multivariate model adjusted for clinical factors: (HR 0.25, 95% CI = 0.06–1.03). Limitations: • Estimated recurrence-free survival using the Kaplan Meier method, without considering death as a competing event. • Underpowered. • No information on censoring. • Phase II clinical trial patients.
Zhou et al.[122]	Prospective study in clinical trial patients China DNA isolated from plasma 2010–2012	Women with metastatic breast cancer (n=130), randomized to receive docetaxel plus thiotepa or docetaxel plus capecitabine. Exposure: • *CYP1A1* rs1048943 Outcomes: • PFS and OS	Not provided, but Kaplan-Meier curves ended after approximately 50 months (~4 years) for OS analyses and 30 months (~2,5 years) for PFS analyses.	Treatment arms combined: No associations with PFS or OS Docetaxel plus thiotepa: Variant carriers (AG/GG) had longer PFS (HR: 1.90, 95% CI 1.04–3.50) and OS (HR: 2.24, 95% CI 1.03–4.80) Docetaxel plus capecitabine arm: Variant carriers had shorter PFS (HR: 0.41, 95% CI 0.23–0.74) and OS (HR: 0.41, 95% CI 0.20–0.83). Limitations: • Used multivariate Cox-regression models. • No information on censoring. • No information on choice of index date.

Abbreviations: CI: Confidence interval, DFS: Disease-free survival, HR: Hazard ratio OS= Overall survival, PFS: Progression-free survival, RECIST: Response Evaluation Criteria in Solid Tumors.

Return-to-work and predictors of returning to work after breast cancer (Study III)

A summary of the included papers is provided in Table 4. We identified one systematic review by Cocchiara *et al.*,[125] which summarized findings from 17 systematic reviews and 8 narrative reviews. Among the review papers we included, three were not included in Cocchiara *et al.'s* review—one published just before[126] and two after.[127,128] We included 11 cohort studies not included in any of the systematic reviews,[42,129–137] where three included patients in a clinical trial.[129,132,137] Return-to-work was self-reported in most studies[42,129–133,135,137–139] and ascertained from registries in one study.[136] As highlighted in the reviews, return-to-work definitions varied widely but were often assessed at annual landmarks where one year after diagnosis/surgery was most common.

Variation in time point of return-to-work assessment complicated condensation of the study findings, but it was evident that the return-to-work after breast cancer varied across countries. For example, for studies conducted in the Netherlands and Germany, 43% and 57%, respectively, had returned to work *within* one year of breast cancer diagnosis, whereas studies from the United States reported a range of 70% to 93%.[9,125,129,135] In two German studies, one study reported that 65% of those eligible to work before diagnosis had returned to work 40 weeks after surgery, of which 13% worked reduced hours[130], and other study reported that 78% of breast cancer survivors had return-to-work 1 year after surgery (56% hours as before diagnosis, 22% reduced working hours). [129] In the latter study, the average time of return-to-work was 4.4 ± 4.5 months after surgery among those who worked the same as they did before diagnosis.[129] Among young German breast cancer survivors aged 18-39 years, 84% returned to work two years after surgery.[138] A fourth German study reported that the cumulative incidence of return-to-work after 10 years was 85%.[131] In Portugal, 70% of women surviving 5 years had returned to work 3 years after breast cancer.[139] One study reported that since surgery, 15% received disability pension during a mean period of 8.3 years.[131] In the same study, the cumulative incidence of working reduced hours was 19% during 10 years.

Predictors of return-to-work

Young age,[9,125,139] higher SEP in terms of education,[9,135,139] income[9] and social support[9,126] promoted return-to-work. Barriers included stage II–III (compared with stage I),[128,130,135] low quality of life,[129] fatique,[129] psychological and cognitive problems[129,136,137,139] and comorbidity.[126,127,129,138] Both qualitative[9,128] and quantitative studies[9,126,133,135,136] highlighted chemotherapy as a major barrier of return-to-work, but also mastectomy, hormone therapy and radiotherapy were mentioned as barriers among other treatment factors.[9,128,139]

Two studies based on the French CANTO (cancer toxicities) cohort included women with stage I–III breast cancer, examined associations between treatments and return-to-work,[137] and family situation and return-to-work.[132] The latter study reported lower odds ratios of return-to-work in women living with a partner and in women older than 50 years.[132] The other CANTO study reported no association between chemotherapy (including taxane-based) and return-to-work. These findings contrast with the other systematic reviews and

the meta-analysis.[9,125,128] Both CANTO studies examined return-to-work two years after breast cancer and short-term effects was therefore not examined.

The studies had considerable methodological issues. Studies based inclusion on survival two-,[132,137,138] three-[139] or five years post-surgery,[129,131] thereby including selected populations of long-term survivors. As such, the studies included women with the best long-term prognosis and were less relevant to newly diagnosed patients. A major limitation of the studies was the lack of analyses incorporating data on the timing of return-to-work. All studies examined landmarks of return-to-work, and thus prevalences, not incidences. This limited the meta-analysis as timing of return-to-work could not be taken into account.[128] As mentioned above, most studies relied on self-reported data, but some of the data were collected retrospectively up to ten years after return-to-work[129,131] and therefore were likely prone to recall bias.[140] Initial response rates ranged from 42%[131] to 88%.[130] Survey responders are most often characterized by higher SEP,[141,142] exemplified in the study by Arfi *et al.*[133] by overrepresentation of wealthiest part of the population. Return-to-work may therefore be differentially misclassified in the questionnaire-based studies and may overestimate the prevalence of return-to-work and bias relative predictor estimates.

Table 4. Summary of literature on return-to-work after breast cancer (Study III)

Author, journal, year	Design, country, data sources and study period	Population, exposures, outcome	End of study, median follow-up	Results, limitations
Reviews and meta-analyses				
Sun *et al.* [126] Supportive Care in Cancer 2017	Literature – review 2004–2014	A total of 25 papers were included. Outcomes: • RTW period, work ability and work Performance	N/A	All studies reported reduced work ability after breast cancer. Time points of RTW were incomparable, but overall RTW most often occurred within one year post-treatment. Barriers to RTW included adverse effects of treatments, work demands and environment, SEP and family support. Limitations: • No flowchart depicting study selection or search history (not systematic).
Islam *et al.* [143] BMC Public Health 2014	Systematic review 2003–2013	Among 12,116 identified papers, 26 were included. Outcomes: • Prevalence of RTW in breast cancer survivors and factors associated with RTW.	N/A	RTW prevalence one year after diagnosis ranged from 43% to 93%, depending on country. Barriers included SEP and treatment related factors, though some papers reported that chemotherapy had no impact on RTW. Limitations: • Studies were heterogeneous, especially according to RTW definitions
Cocchiara *et al.* [125] IOS Press 2018	Systematic review of reviews	Included 26 papers (17 systematic reviews and 9 narrative reviews). Outcomes: • RTW		RTW 1-year prevalence varied from 43% to 93% across countries. Average time to RTW varied from 3 to 11 months. Education, age, marital status, parity, ethnicity and household income influenced RTW. Chemotherapy was reported as one of the major barriers to RTW. Limitations: • Studies were heterogeneous, especially according to RTW definitions. • No quality assessment of primary studies.
Bijker *et al.* [127] Journal of Occupational Rehabilitation 2018	Systematic review 2000–2016	Papers investigating the role of physical impairments on work-related outcomes in breast cancer (n=20 papers) Outcomes: • Work-related outcomes		Functional impairments, *e.g.*, shoulder functioning, hampered RTW. Limited: • Limited comparability across studies regarding definitions of RTW/other work-related outcomes, and different legislations across countries counteracted synthesis of the findings.
Wang *et al.* [128] Journal of Clinical Oncology 2018	Systematic review and meta-analysis	Studies exploring risk factors of unemployment (n=26 studies, 46,927 patients) Outcome: • Unemployment (direct or indirect measures)	Maximum 120 months of follow-up (~10 years)	Predictors of unemployment included access to universal health care, high psychological job demands, childlessness, low SEP, stage ≥II, mastectomy and chemotherapy. Limitations: • Estimates were based on studies with follow-up periods varying from 1–120 months. The estimates may therefore not reflect specific periods after breast cancer.
Original articles				
Monteiro *et al.* [139] The Breast 2019	Cohort study Portugal Self-reported and medical records 2012	Women diagnosed with stage I–III breast cancer proposed for surgery and who survived 5 years (n=462). Comparisons included women with no previous cancer. Outcomes: • Unemployment, early retirement and sick leave	Five years	242 were employed at diagnosis. Among these, 70% had RTW 3 years after breast cancer, which decreased to 67% five years after (mainly due to early- or normal retirement). Higher educated (OR: 0.27, 95% CI: 0.10–0.71) and women who received hormone therapy (OR: 0.38, 95% CI: 0.14–1.07) were less likely to be unemployed. Limitations: • Selected population only including 5-year survivors.

Dumas *et al.* [144] Journal of Clinical Oncology 2020	Cohort study within the CANTO cohort France Medical files and questionnaires	Women with stage I-III breast cancer 18–57 years employed at diagnosis, and who did not experience recurrence or death before end of study (n=1,874) Exposures: • Patient and tumor characteristics, treatments and toxicities Outcome: • Non-RTW two years after ended treatment.	Two years	After two years, 80% had RTW. Toxicities (OR: 1.59, 95% CI: 1.15–2.18) and various psychological symptoms were associated with non-RTW. Limitations: • Excluded women with recurrence, who withdrew consent, were lost to follow-up or died within two years. • Prone to non-responder bias • Did not use "no systemic therapy" as reference group. Hence, it was not possible to judge the association between chemotherapy (yes/no) and RTW. • Used logistic regression, which does not approximate the RR in cohort studies with common events.[145] Therefore, ORs should be interpreted with caution.
Leuteritz *et al*.[138] Journal of Adolescent and Young Adult Oncology 2021	Cohort study Germany Questionnaires 2014–2016	Cancer survivors aged 18–39 years (n=137 in breast) Exposures: • SEP and clinical characteristics Outcome: • Employment status	Two years	Two years after breast cancer 84% of the women employed at diagnosis had RTW. Across all cancer types, having comorbid conditions were associated with non-RTW Limitations: • Prone to non-responder bias • Logistic regression, no rates included.
Arfi *et al.*[133] BMJ Open 2018	Multicentre cohort study Prospective self-reported data in booklet 2015–2016	Women with non-metastatic breast cancer aged ≥18 years working at diagnosis (n=297). Exposures: • Age, marital status, occupation, income, out-of-pocket health expenses and work information. Outcome: • Work during and after breast cancer treatment		Median duration of sick-leave: 155 days (range 5–365). Chemotherapy was associated with longer sick-leave (OR: 3.5, 95% CI: 1.6–7.9). Limitations: • Use logistic regression, no information on timing of RTW. • Selected population including the wealthiest patients. Likely selection by disease, complication and adverse effects severity. • Prone to non-responder bias and recall bias
Heuser *et al.*[130] BMC Health Services Research 2018	Multicenter cohort study Germany Questionnaires 2013	Women with breast cancer who were employed or had the capability of employment (n=577) Outcomes: • RTW 40 weeks after surgery, working time and stress	40 weeks (<1 year).	374 patients (65%) did RTW within 40 weeks, whereas 13% had reduced working time. Women aged 18–44 years and those with higher education more often worked reduced hours. Compared to UICC stage 0/I, stage II and stage III had lower OR of RTW (OR: 0.45, 95% CI: 0.25–0.82 and OR: 0.19, 95% CI: 0.08–0.46, respectively). Limitations: • No analyses on time of RTW. Used logistic regression, which does not approximate RR in cohort studies with common events.[145] • No information on treatments besides surgery.
Arndt *et al.*[131] Acta Oncologica 2019	Multi-regional population-based cohort study Germany Questionnaires and registries 1994–2004	Included both colon cancer- prostate cancer- and breast cancer survivors, who were employed at diagnosis. The latter group included 1070 women. Outcomes: • RTW in former or new job, unemployment, disability pension, early retirement (not-cancer related) and other reasons for leaving job		Among 5-year breast cancer survivors, the cumulative incidences of return to their former job was 64% and of receiving disability pension was 15%. The cumulative incidence of RTW at10 years after breast cancer diagnosis was 85%. 19% had reduced working hours. Limitations: • Prone to non-responder bias, as the study only included those surviving 5 years and the response rate on the questionnaire was 42% • Prone to recall bias–questionnaire emailed 5–10 years after diagnosis.

Caumette *et al.* [132] Current Oncology 2021	Cohort study within the CANTO cohort France Clinical examinations and questionnaires 2012–2017	Women with stage I–III breast cancer <57 years undergoing surgery, who were employed at time of diagnosis and who were recurrence-free and alive 2 years after diagnosis (n=3004) Outcome: • RTW 1- and 2-years post diagnosis		RTW was lower in women living with a partner (OR: 0.63, 95% CI: 0.47–0.86). Married had decreased RTW if aged >50 years (OR: 0.57, 95% CI: 0.34–0.95). Limitations: • Excluded women with recurrence, who withdrew consent, were lost to follow-up, or died within 2 years. • Prone to non-responder bias • Use logistic regression, which does not reflect the RR in cohort studies with common events.[145]
Porro *et al.* [42] Scandinavian Journal of Caring Sciences 2019	Cohort study France Questionnaires and interviews 2014–2015	Women who were 25–60 years old and employed at time of breast cancer diagnosis (n=103) Exposures: • Quality of life at baseline and fatigue (cancer related) Outcome: • RTW	Follow up 3 and 6 months after surgery.	Increasing RTW 3 months and 6 months after breast cancer by higher quality of life at baseline (OR: 1.12, 95% CI: 1.01–1.25 and OR: 1.09, 95% CI: 1.01–1.17, respectively). Limitations: • Prone to selection bias. Initial response rate of 66% and participation may have depended on disease severity and/or complications. • Uses logistic regression, which do not reflect a RR in cohort studies with common events.[145]
Schmidt *et al.* [129] European Journal of Cancer Care 2019	Cohort study including patients from a randomized controlled trial Germany Self-reported 2010–2013	Breast cancer survivors (56% premenopausal) alive 5 years post-diagnosis (n=135). Exposures: • Depressive symptoms, arm morbidity, education, age, cohabitation and fatigue Outcomes: • Work participation and working time, quality of life, physical, cognitive and role function and finances	Follow-up 2016–2017 (about 5 years after diagnosis).	Cessation of work was associated with low quality of life. Fatigue, psychological and cognitive problems, arm morbidity were barriers to RTW. Limitations: • Selected population. • Retrospective self-reported data with risk of recall bias (5 years lookback). • Used logistic regression, which does not approximate the RR in cohort studies with common events.[145] Thus, ORs should be interpreted with caution. • Did not estimate cumulative incidences. • Imprecise estimates due to low sample size.
Rosenberg *et al.* [135] Breast Cancer Research and Treatment 2019	Multi-institutional cohort study USA Surveys at and after breast cancer diagnosis (semiannually for 3 years and annually hereafter) 2006–2016	Women aged ≤ 40 years with breast cancer (n=911). (employed and unemployed) Outcomes: • Employment 1 year post diagnosis • Job satisfaction		70% were employed both before surgery and 1 year after surgery, 7% were employed before and unemployed after surgery. Workforce detachment was associated with stage III tumors (OR: 5.57, 95% CI: 2.63–11.81), chemotherapy (OR: 3.81, 95% CI: 1.5–9.65) and low education (OR: 2.39, 95% CI: 1.30–4.41). Limitations: • No information on adjustment variables used in multivariate analyses. • Used logistic regression, which does not approximate the RR in cohort studies with common events.[145] Therefore, ORs should be interpreted with caution. • Selected population (response rate: 60%). Prone to non-responder bias • Underpowered, CI were wide.

Schärfe *et al.* [136] Scandinavian Journal of Public Health 2017	Population-based cohort study Denmark Danish Breast Cancer Group, the Danish National Prescription Registry, the National Health Service Registry, the DREAM database, the Population's Education Register and the Personal Income Statistics Registry 2000–2012	Self-supporting women diagnosed with breast cancer (n=16,868) Exposure: • History of psychiatric medical usage (2-4 years prior to breast cancer diagnosis). Outcomes: • RTW 1 year	Psychiatric medication use was associated with lower RTW rate during the first year post diagnosis (RR: 0.91, 95% CI: 0.87-0.94). Model was robust when adjusted for demographic-, clinical- and socioeconomic factors. Adjuvant chemotherapy alone (RR adjusted: 0.79, 95% CI: 0.77-0.82) and in combination with endocrine therapy (RR: 0.82, 95% CI: 0.79-0.84) was associated with decreased RTW. Limitations: • No information on breast cancer stage

Abbreviations: CANTO= Cancer toxicities cohort, CI: Confidence interval, OR= Odds ratio, RR= Relative risk, RTW= Return-to-work, SEP= Socioeconomic position.

Knowledge gaps

I identified the following knowledge gaps in the literature.

Study I	SEP and mortality after breast cancer is associated, but research investigating SEP's influence on recurrence is scarce. No studies focused on premenopausal breast cancer patients who underwent guideline therapy, while also incorporating data to address the potential influence of ER status, stage and comorbidities.
Study II	Studies on the associations between SNPs and breast cancer outcomes are underpowered and heterogeneous in terms of patient- and cancer characteristics, and taxane combinations. There were no population-based studies focusing on premenopausal breast cancer patients treated with guideline chemotherapy, and no nationwide studies.
Study III	Return-to-work in breast cancer patients varies across countries, age, SEP and cancer related characteristics. Chemotherapy and adverse treatment effects may be barriers to returning to work. The association between SNPs related to chemotherapy and return-to-work had never been examined.

Aims

The aim of this thesis was to better understand factors influencing survivorship during and after taxane-based chemotherapy in premenopausal women with breast cancer. Literature on the influence of both societal factors and inherited factors is scarce. Therefore, we conducted the following three studies:

Study I

Hypothesis: Women with low SEP may have higher risk of breast cancer recurrence and mortality after taxane-based chemotherapy compared with women with higher SEP.

Aim: To examine the prognosis in terms of recurrence and mortality, focusing on SEP.

Study II

Hypothesis: Women with SNPs in genes related to docetaxel transport, metabolism, DNA-repair, or neural processes, may have altered docetaxel effectiveness, in terms of recurrence and mortality, compared with wildtypes.

Aim: To examine the associations between SNPs and breast cancer recurrence and mortality.

Study III

Hypothesis: Women with SNPs related to docetaxel transport, metabolism, DNA-repair, or neural processes, may have delayed or faster return-to-work and stable labour market attachment, compared with wildtypes.

Aim: To examine the prognosis in terms of return-to-work and stable labour market attachment, focusing on SNPs.

Methods

Ethical aspects

All three studies were approved by the Danish Data Protection Agency record AU 2016-051-000001, serial numbers #994 (Study I) and #808 (Studies II & III), the Regional Ethics Committee (Record no. 1-10-72-4-18), and DBCG (Study I: DBCG-2019-08-20 and Studies II & III: DBCG-2018-01-04). We complied with the World Medical Association's Declaration of Helsinki. The use of registry-based data for scientific studies in Denmark requires no consent from the participants.

Setting

In the following sections I will give a brief introduction to the Danish health care system, and Denmark's strategy to accelerate and optimize cancer diagnostics and treatment. Moreover, I introduce the DBCG, who are responsible for breast cancer clinical guidelines, and follow-up, and the ProBe CaRe cohort, in which our study cohorts were nested.

The Danish health care system

The Danish health care system is tax-financed, with uniform access to all Danish residents. It has three administrative levels. The first is the national level, governed by the Ministry of Health. The Ministry of Health has a steering role and is, among other things, responsible for issuing national guidelines. The second level is the regional level, consisting of five regions responsible for primary health care (*e.g.* general practitioners) and secondary health care (*e.g.* hospitals). Cancer diagnosis and treatments are administered at this level. The third level is the municipalities, counting 98 across Denmark. One of the municipalities roles is social care and some rehabilitation.[146]

Since 2007, all women in Denmark aged 50–69 years have been invited to a national, systematic breast cancer mammography screening program every other year.[147] During 2008, diagnostic pathways (also known as "cancer packages") were integrated in the Danish health care system. When a patient enters a diagnostic pathway, their trajectory through diagnostics and treatments are planned according to predefined time standards. During the study period of this thesis, there was a waiting time guarantee for diagnosis and initiated treatment of one month. If not met, the patients could choose to be treated at a private hospital.[147]

The Danish Breast Cancer Group

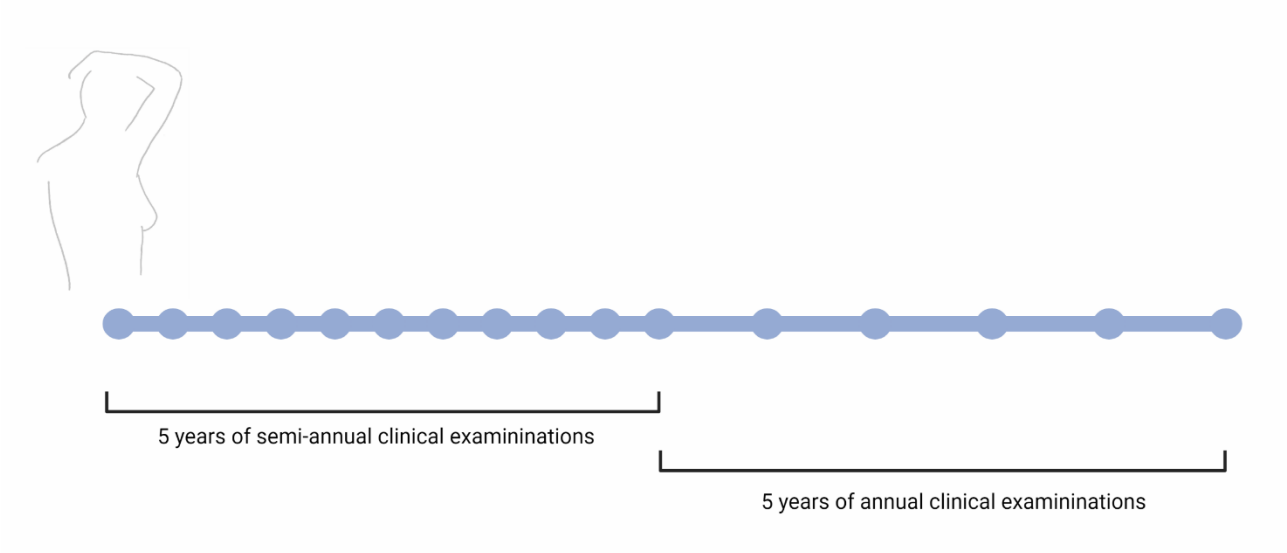
In 1975, the Danish Surgical Society established the Danish Breast Cancer Collaborative Group, later DBCG. At that time, breast cancer prognosis varied geographically with lower mortality in Copenhagen, especially in younger breast cancer patients. In the mid-1970s, treatment options were limited to surgery and radiotherapy,

but systemic therapy was in the offing. The DBCG aimed to consolidate nationwide standardization of treatments to make breast cancer diagnosis, treatment, and prognosis uniform across Denmark. Furthermore, a new database including clinical data was established. The DBCG clinical database is an electronic platform, where all specialties involved in the patient's care at Danish hospitals report patient specific, clinical data. The database collects information from other Danish administrative and medical registries on emigration, vital status, cause of death, pathology reports, other malignancies, comorbidities and hospitalizations.[57]

Clinical follow-up after breast cancer

The recommended follow-up program for women treated for breast cancer has evolved over time. Before 2016, regular outpatient visits were offered semi-annually the first five years after diagnosis, and in the subsequent five years the visits were offered on an annual basis (Figure 5). The examinations included a physical assessment and additional diagnostic work-up when indicated.[148] From 2016, the scheduled visits have been phased out, depending on the region.[149] Women diagnosed with breast cancer before age 50 years are offered annual mammography. The frequency of mammography in women aged 50–70 years, with a history of breast cancer, varies according to breast density.

Figure 5. Scheduled clinical examinations in women diagnosed before 2016.



The Predictors of Breast Cancer Recurrence cohort

Around 2014, the Predictors of Breast Cancer Recurrence (ProBe CaRe) cohort was assembled aiming to investigate inhibition of tamoxifen through studies focusing on pharmacogenomics and breast cancer recurrence and possible interaction with ERβ expression and estrogen-regulating enzymes. The ProBe CaRe cohort included 5,959 premenopausal women diagnosed with early breast cancer between 2002 and 2011 in Denmark, registered in DBCG's clinical database.[150] Women undergoing neo-adjuvant chemotherapy, women

with ER+ disease who were not assigned tamoxifen therapy, women with ER– disease treated with tamoxifen, or missing information on either ER status or tamoxifen were not included.[150]

Data sources

All studies in this thesis are registry-based studies, taking advantage of Denmark's network of administrative and medical registries. In addition, we analyzed biological material from women included in Studies II–III. Individual-level linkage was achieved owing to the unique Central Personal Registration (CPR) number granted every legal resident in Denmark at birth or immigration.[146]

Medical registries

The DBCG clinical database described above, includes virtually all women in Denmark under 70 years diagnosed with breast cancer.[151] DBCG registers patient characteristics: age at diagnosis and dates of diagnosis (surgery); tumor characteristics: hormone receptor status, lymph node status, tumor size and malignancy grade; treatment characteristics: intended and administered chemotherapy, radiotherapy, surgery and endocrine therapy; and follow-up data: dates of other malignancies, last follow-up, death and end of protocol. It is important to note that DBCG stops registrations when a follow-up event occurs. Thus, deaths occurring after recurrence are not registered. The completeness of pathology measures has increased with implementation of biomarker-based therapies. During 2007–2016, proportions of registered HER2 assessment improved from 74% to 99%.[57] Immunohistochemical testing of ER and PR status was initiated in 1990[152] and was implemented into routine diagnostic testing in 1997.[153]

The Danish National Patient Registry (DNPR) holds data on all non-psychiatric hospital-visits since 1977 including all Danish hospitals. For each visit, admission and discharges dates are reported, along with the primary diagnosis at the time of the admission and related diagnoses.[154]

The Danish Registry of Causes of Death includes dates, time, and place of death, along with immediate, underlying and/or contributory causes of death.[155]

The Danish National Pathology Registry records pathology specimens from Danish pathology departments and the practicing pathologists with completeness back to 1997.[156] This includes data on biobanks of routinely stored diagnostic histological and cytological specimens. The biobanks include primary formalin-fixed paraffin-embedded (FFPE) tissue blocks, including tissue from resected breast tumors.

The Danish Medical Birth Registry monitors information on pregnant women and their offspring, including date of birth.[157]

Administrative registries held by Statistics Denmark

The Danish Population Registry receives data from the Danish Civil registration System.[146] The registry includes data on marital transitions, including date of marriage, divorce, or registration of partnership, on cohabitation and residence.

The Income Statistics Registry receives data from smaller registries including information on *e.g.* salaries, taxes and transfer payments.[158] For each individual both personal and household income are reported.

The Population's Education Registry holds data on every person who has attended educational programs authorized by the Danish Ministry of Education and it includes records from immigration, enabling collection of information on education completed outside Denmark. The register has a coverage of 96% in the Danish Population and 85–90% in the immigrant population.[159]

The Danish Register for Evaluation of Marginalization (DREAM) has since 1991 provided weekly information on social security payments to all Danish citizens aged 18–65 years, with high predictive value and completeness.[160]

Biological data (Studies II–III)

For all women included in the ProBe CaRe cohort, FFPE blocks with tumor-infiltrated tissue were identified at Danish hospitals using the women's CPR numbers. The collection was initiated in April 2014 and ended in August 2015. Identification of FFPE blocks at the local pathology departments was done by medical research technicians blinded to any prognostic information of the individual. A total of 5,500 blocks were collected.[150]

Study design

All studies included in this thesis were registry-based cohort studies, nested in the ProBe CaRe cohort. In the following sections I present how the study cohorts were derived and illustrate these in Figure 6.

Study cohorts and follow-up

According to the overall aim to examine survivorship following taxane-based chemotherapy in premenopausal women, we restricted our study cohorts to include women who:

- were diagnosed with non-distant metastatic breast cancer 2007–2011,
- were recommended guideline chemotherapy including docetaxel, and
- were aged 18–55 years at diagnosis, to avoid misclassification of menopausal status.

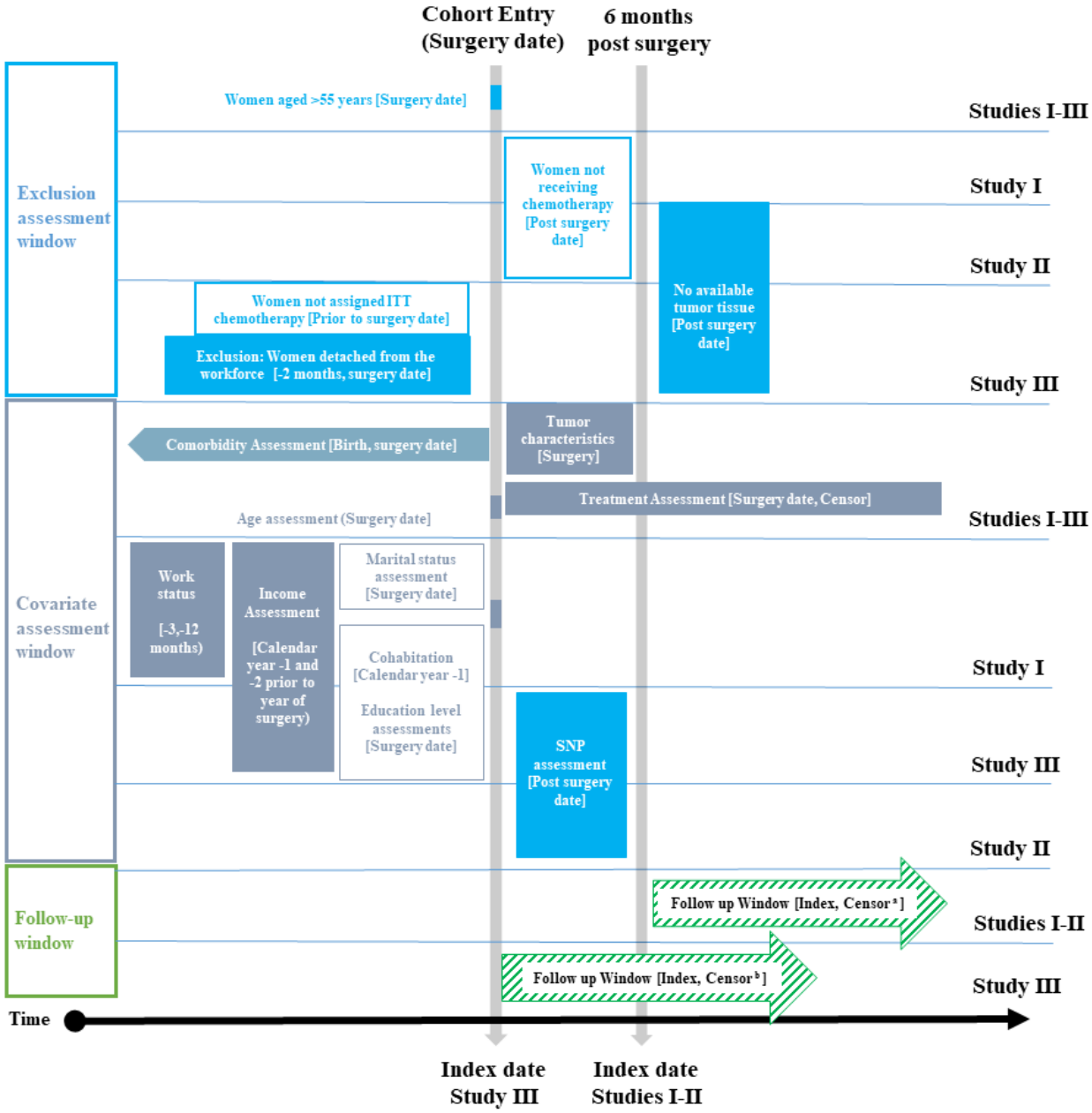
In Studies I–II, we restricted to women who received least one cycle of chemotherapy. Therefore, we began follow-up six months after breast cancer surgery, to approximate the end of adjuvant chemotherapy. We excluded women who experienced recurrence, death or emigrated in the first six months after diagnosis (see

underlying timeline for all studies in Figure 6). We chose this index date to avoid immortal time bias.[161] That is, when observation time in which the outcome cannot occur is included. Had we included the six months in which the women per design could not die (one of the outcomes in Studies I–II), we would have introduced bias as some observation time would have been misclassified. The women who did not receive chemotherapy (and thus were excluded), generally had less severe cancers compared with the overall study cohort (see flowchart in Appendix I).

In Study III, we restricted to women assigned intention-to-treat (ITT) chemotherapy, who were employed prior to breast cancer diagnosis. We anticipated that a considerable proportion of the women were likely to have resumed working during the first six months after diagnosis, so starting follow-up later could lead to selection bias.

In Study I, we ended follow-up on 31th December 2016 in all the analyses, as this was the last available update of the Cause of Death Registry. For Study II, the Cause of Death register was updated until 31st December 2019. We therefore extended our mortality analyses accordingly. Based on the last available update of the DBCG database, follow-up in the recurrence analyses ended on 25th September 2017 in that study. In Study III, we also ended follow-up according to DBCG (*i.e.*, 25th September 2017), as censoring data was derived from that register.

Figure 6. Graphical depiction of the underlying time and the exclusion-, covariate- and follow-up assessment windows.



Abbreviations: ITT= Intention-to-treat, SNP= Single nucleotide polymorphisms. Based on a template provided by Schneeweiss *et al.*.[162]

Exposures

Socioeconomic position (Study I)

In Study I, we considered five SEP indicators as exposures. The SEP indicators included marital status, cohabitation, education level, household income and employment. As illustrated in Figure 6, we collected information on marital status at the date of diagnosis and categorized the women into married including women in registered partnerships, or singles including never married, divorced, separated, or widowed. Statistics Denmark derives cohabitation status from an algorithm collecting information on marital status, parity and living situations, which is updated annually (details can be found in Appendix III)[163] We collected this information in the calendar year before breast cancer and categorized as cohabiting or living alone. Guided by the International Standard Classification of Education,[164] we categorized education into low, intermediate, and high. We determined household income as the average income in the two years preceding the year of breast cancer and categorized it into low, medium and high using sample quartiles (≤Q1, Q2–Q3, ≥Q4) in Study I. This approach was used to avoid misclassification, as income, especially among individuals who are self-employed, may change from year to year. Statistics Denmark corrects the household income for number of persons in the household, to ensure comparability across individuals and studies. Based on a review of current and previous DREAM manuals, we categorized all current and past social benefit codes into three categories; employment, unemployment and health related absenteeism. Detailed categorization can be found in Appendix I. In brief, employed included students, women on maternity leave and part-time work, unemployment included those searching for a job or qualifying for one, health-related absenteeism including women on sick-leave, flexible job, disability pension or equivalent.

Single nucleotide polymorphisms (Studies II–III)

In Studies II–III, we considered SNPs as exposures. We aimed to replicate findings from the pharmacogenetic literature, but also to study other SNPs that theoretically could influence the study outcomes. To identify SNPs, we used the candidate gene approach[165] and selected genes with a biological plausibility association with taxane effectiveness. We considered SNPs as candidates if they were encoding proteins involved in taxane transport or metabolism. We consulted specialists in pharmacology and pharmacogenetics, and performed a comprehensive review of the literature to identify SNPs influencing taxane pharmacokinetics- or dynamics. Finally, we assessed randomized studies, observational studies, reviews, and meta-analyses to identify SNPs with a proposed influence on efficacy, effectiveness, or toxicities. We only included SNPs with a major allele frequency ≥5% according to benchmarks reported for female European non-Finnish cohorts. We examined major allele frequencies in the Genome Aggregation Database.[166] We therefore included 26 SNPs in 20 genes, described in Table 5.

For each SNP, we classified women as wildtypes if they carried two normal alleles or as variant carriers if they carried 1 or 2 variant alleles. We also classified variant carriers as heterozygotes (1 variant allele) or homozygotes (2 variant alleles).

Table 5. Candidate genes and associated variants and relevant function.

Gene	SNP ID	Aliases/proteins encoded	Function (simplified)
ABCB1	rs10248420 rs1045642 rs1128503 rs2032582	Multidrug resistance 1 Transporter P-glycoprotein	Efflux transporters (mainly excretion into bile)
ABCC2	rs12762549	Multidrug resistance protein 2	
ABCG2	rs2231142	Breast cancer resistance protein	
CYP1A1	rs1048943	Cytochrome P450 Family 1 Subfamily A Member 1	Phase 1 metabolizing enzymes
CYP1B1	rs1056836	Cytochrome P450 Family 1 Subfamily B Member 1	
CYP3A	rs10273424	Cytochrome P450 gene cluster	
CYP3A4	rs2740574 rs35599367	Cytochrome P450 Family 3 Subfamily A Member 4	
CYP3A5	rs776746	Cytochrome P450 Family 3 Subfamily A Member 5	
GSTP1	rs1138272	Glutathione S-transferase pi gene	Phase 2 metabolizing enzyme
SLCO1B1	rs2306283 rs4149056	OATP1B1 (Organic anion transporting polypeptide 1B1)	Influx transporters (mainly hepatic uptake)
SLCO1B3	rs11045585	OATP1B3	
EPHA4	rs17348202	Ephrin type-A receptor 4	Neuronal function and repair
EPHA5	rs7349683	Ephrin type-A receptor 5	
EPHA6	rs301927	Ephrin type-A receptor 6	
EPHA8	rs209709	Ephrin type-A receptor 8	
FGD4	rs10771973	Charcot-Marie-Tooth gene	
ARHGEF10	rs9657362	Charcot-Marie-Tooth gene Rho guanine nucleotide exchange factor	
TRPV1	rs879207	Chili-receptor	
ERCC1	rs11615 rs3212986	ERCC Excision Repair 1	DNA-repair
ERCC2	rs13181	ERCC Excision Repair 2	

Outcomes

Breast cancer recurrence (Studies I–II)

We obtained DBCG's classification of recurrence, which includes (1) locoregional recurrence defined as a cancer in the ipsilateral chest wall, skin, or soft tissue, or in the axillary or clavicular lymph nodes, (2) distant recurrence, defined as metastasis located in distant organs, *e.g.*, in the bones or (3) contralateral cancer, defined as cancer in the opposite breast, diagnosed up to 10 years after surgery.[167]

Mortality (Studies I–II)

We ascertained information on all-cause mortality and BCSM from the Cause of Death Registry. In sensitivity analyses, we included BCSM, defined as deaths with breast cancer (ICD-10: C50) as the underlying or contributory cause of death.

Return-to-work and stable labor market attachment (Study III)

Using the DREAM register, we ascertained information on all social benefit payments paid to each woman, or the woman's employer. We categorized women with no benefit payouts as employed, together with women receiving state education grants (payments given to students in Denmark aged 18 years or older) and part-time unemployment benefit (*i.e.*, women working part-time). We defined return-to-work as four consecutive weeks of employment. Using this classification, we ensured that short-term sick leave was not misclassified as work (in Denmark, short-term sick leave was paid by employers for 14–21 days, depending on the calendar period). Some women might suffer from breast cancer sequelae occasionally reducing their workability. Therefore, we also examined stable labor market attachment defined as 12 consecutive weeks of employment.

Covariates

Patient, tumor and treatment characteristics

We described the study cohorts according to relevant patient, tumor and treatment characteristics presented in Table 6. Some of the covariates served as adjustment or stratifying variables. From the DBCG, we collected the following information; age at diagnosis, histological tumor type and malignancy grade (1–3, for ductal and lobular carcinomas only), tumor size, lymph node involvement, receptor status (ER, HER2 and PR), surgery type, ITT adjuvant radiotherapy, tamoxifen therapy and adjuvant chemotherapy (both ITT and administered). Using tumor size and lymph node status we derived stage (I–III) according to the TNM (tumor node metastasis) classification.[36] We classified surgery type as lumpectomy including ITT radiotherapy or mastectomy. All included women with ER+ tumors received endocrine therapy (treatment with tamoxifen). We classified tumors as triple negative if tumors were ER–, HER2– and PR were either negative or missing. We searched the DNPR for somatic comorbid diseases diagnosed prior to the date of surgery. We summarized the

comorbidities using a modified version of the Charlson Comorbidity Index[168] not including breast cancer diagnoses. The algorithm is presented in Appendices I–III.

Socioeconomic indicators

The SEP indicators figuring as exposures in Study I were included as covariates in Studies II–III. In Study III where we conditioned on employment, we changed the categorization not to include women on maternity leave, as they for this reason were not expected to return-to-work during the first year. Moreover, we re-categorized income into two groups determined by the median income in ProBe CaRe to accommodate proper strata size for stratification.

Table 6. Variables applied in the studies.

	Study I	Study II	Study III	Data source
Included as stratifying covariates				
ER status*	✓	✓	✓	DBCG
Stage		✓	✓	DBCG
Cohabitation	✓		✓	The Danish Population Registry
Marital status	✓			The Danish Population Registry
Education	✓		✓	The Population's Education Registry
Employment	✓			DREAM
Income	✓		✓	The Income Statistics Registry
Included as adjustment covariates				
Age	✓			DBCG
Marital status	✓			The Danish Population Registry
Cohabitation	✓			The Danish Population Registry
Education level	✓			The Population's Education Registry
Income	✓			The Income Statistics Registry
Employment status	✓			DREAM
Charlson Comorbidity Index score	✓			DNPR
Included as censoring variables or competing risks				
Emigration	✓	✓	✓	Statistics Denmark
Other malignancies	✓	✓	✓	DBCG
Death	✓	✓	✓	The Danish Registry of Causes of Deaths
Retirement			✓	DREAM
Date of childbirth			✓	The Medical Birth Registry
Date of maternity leave			✓	DREAM
Lost to clinical follow-up	✓	✓		

Abbreviations: DBCG= Danish Breast Cancer Group, DNPR= Danish National Patient Registry, DREAM= The Danish Register for Evaluation of Marginalization.
*ER status included receipt of endocrine therapy.

Genotyping

As described previously,[169] DNA was extracted from the FFPE blocks. Seven of the 26 selected SNPs had already been genotyped (*ABCB1* rs10248420, *ABCB1* rs1045642, *ABCB1* rs1128503, *ABCB1* rs2032582, *CYP1A1* rs1048943, *CYP3A* rs10273424 and *CYP3A5* rs776746). Genotyping performed for the present studies is described in detail in Appendices II-III. In brief, we used commercially available TaqMan assays and a StepOne Plus real-time instrument (Applied Biosystems, Thermo Fisher Scientific, Foster City, California, USA) for polymerase chain reaction. Genotypes were auto-called according to TaqMan VIC/FAM intensity values using QuantStudio Software V1.3. After, the results were inspected, and overridden manually if amplifications curves were irregular. We compared the major allele frequencies with the corresponding benchmarks, and we compared the observed genotype frequencies with the expected genotype frequencies calculated under Hardy-Weinberg Equilibrium and assessed the deviations.

Statistical methods

Patient characteristics (Studies I–III)

We present characteristics of the study cohort in descriptive tables according to patient, tumor and treatment characteristics by presenting frequencies and distributions within each categorical variable. To incorporate time at risk, we included person-years of follow-up across patient, tumor and treatment characteristics in Study I. Moreover, we calculated the median and interquartile range (IQR) of age and follow-up time. In accordance with Danish cell size suppressing rules from Statistics Denmark (suppress cell counts <5), we reported some frequencies (and corresponding proportions) in aggregate, so that re-identification was not possible.

Incidences (Studies I & III)

To examine the incidence of breast cancer recurrence and mortality across indicators of SEP (Study I), we computed incidence rates (IR) per 1,000 person-years (PY) of breast cancer recurrence and of mortality by SEP indicators. We graphically depicted cumulative incidence curves of recurrence and mortality overall, and by SEP indicator (Study I), and the overall and SNP specific cumulative incidences of return-to-work and stable labour market attachment (Study III). For this, we used the Nelson-Aalen estimator.[140] As all survival data, our data were right-censored: that is, some in the cohort did not die during follow-up. Therefore, to keep them in the study, we censored them at end of follow-up. Moreover, we treated death as a competing risk in all analyses where this was not the outcome. In Study III, we also considered early and normal retirement as competing risks. We truncated all presented graphs at maximum 10 years to avoid small strata enabling identification of individuals. Moreover, we used locally estimated scatterplot smoothing (loess) when plotting SNPs to mask sensitive data. When examining recurrence and mortality (Study I), women were censored at emigration, at 10 years or at end of study. In the recurrence model, we also censored at other malignancies. In

Study III, we censored women at recurrence, other malignancy, emigration, maternity leave/childbirth, or end of study.

Regression models (Studies I–III)

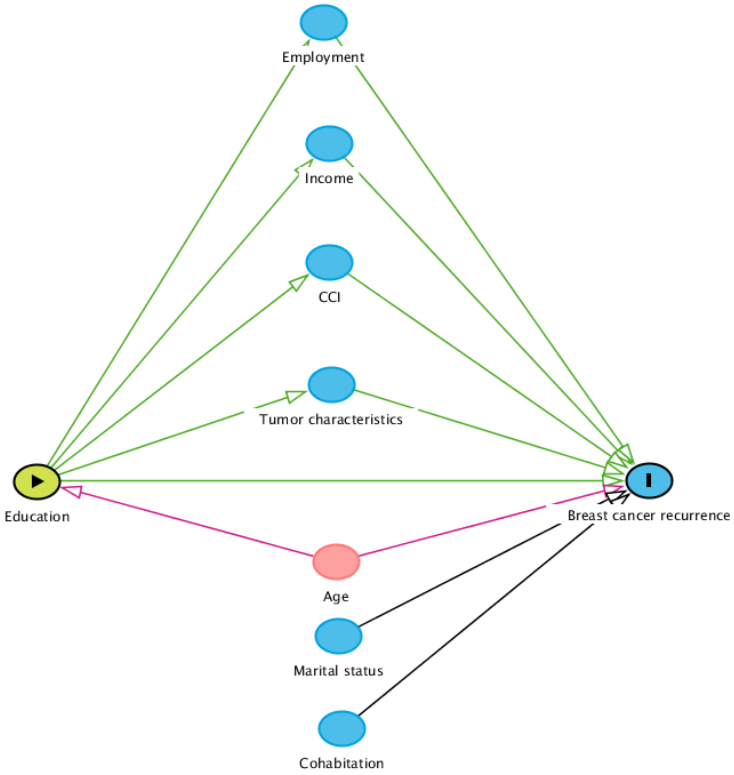
In Study I, we used multivariate Poisson regression to examine incidence rate ratios (IRR) and 95% confidence intervals (95% CI) of 5- and 10-year recurrence and mortality across SEP indicators. We fitted one model for each SEP exposure, assuming sufficient adjustment sets. When using multivariate regression models, equivalent interpretation of the coefficients may not always be appropriate as mistaken interpretations may occur, known as mutual adjustment or Table 2 fallacy.[170] When examining SEP, this leaves out the core-effect of the SEP indicators.[83] Therefore, we conducted separate models for each indicator of SEP, using minimal sufficient adjustment sets evaluated in directed acyclic graphs (DAGS).[171] One of the DAGS is presented in Figure 7. The arrow denotes given associations between all covariates including exposures and outcomes. In the figure, the red path illustrates a backdoor path through which confounding can arise. In this case, the DAG suggests age adjustment only. All the adjustment sets and associated underlying assumptions can be found in Appendix I.

In Studies II & III, we used univariate Cox regression models to compute unadjusted hazard ratios (HR) and associated 95% CIs between SNPs and breast cancer recurrence and mortality (Study II), and return-to-work and stable labor market attachment (Study III). SNPs are inherited, so cannot be affected by other patient, tumor or treatment characteristics and, as such, the other covariates did not fulfill the criteria to be considered confounders.

In Studies I & III, we calculated period-specific incidence IRRs/HRs where each of them started at the index date to avoid bias through selection of healthy survivors, which would be the case if we examined exhaustive time-periods.[172] We conducted several sensitivity analyses to test the robustness of our analyses. Under the assumption that deaths in this young population were likely recurrences, we pooled recurrences and BCSMs to account for potential underreporting of recurrences in DBCG in Studies I & II. We also restricted the mortality analyses only to include BCSM. In addition, we applied several other sensitivity analyses. These are summarized in Table 7 along with a summary of the methods used in Studies I–III. More details can be found in Appendices I–III.

All data management and analyses were generated using SAS software (SAS Institute Inc., Cary, NC, USA).

Figure 7. DAG illustrating associations between SEP exposures, covariates, and the outcomes (Study I).



Simplified version of the DAG provided in Hjorth *et al.*,[173] Appendix I. Education is set as the exposure and breast cancer recurrence as outcome. The model suggests age as the minimum adjustment set. Abbreviations: CCI= Charlson Comorbidity Index.

Table 7 – Summary of materials and methods

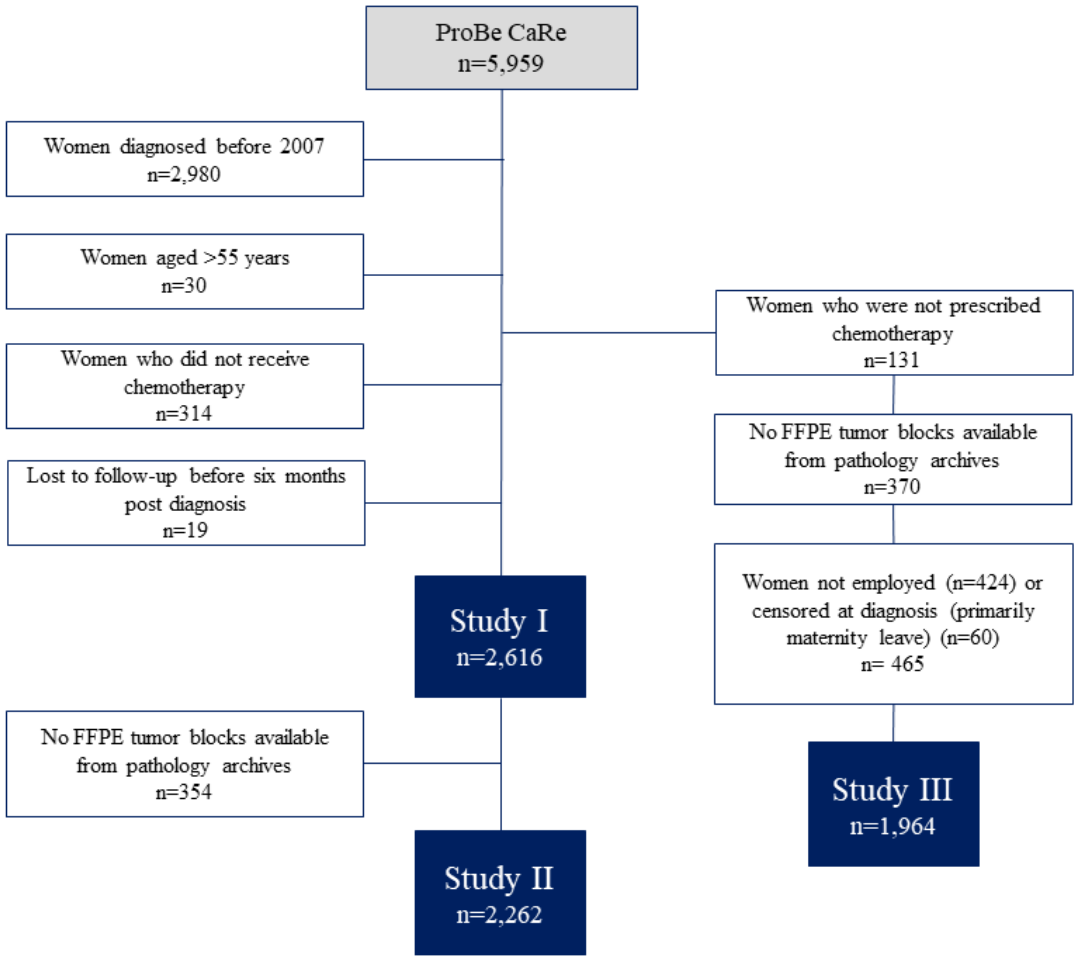
	Study I	Study II	Study III
Objectives	To investigate the influence of SEP on the incidence of breast cancer recurrence and mortality up to 10 years after non-metastatic breast cancer, among premenopausal women who underwent taxane-based combination chemotherapy, and the potential modification by endocrine therapy.	To investigate the influence of SNP on breast cancer recurrence and mortality after taxane-based chemotherapy, and the potential modification through endocrine treatment and stage.	To investigate the influence of SNP on return-to-work during and after taxane-based chemotherapy, and the potential modification through endocrine treatment and SEP.
Setting	Denmark, 2007─2011	Denmark, 2007─2011	Denmark, 2007─2011
Design	Population-based, prospective cohort study	Population-based, prospective cohort study	Population-based, prospective cohort study
Data sources	DBCG, the Danish Population Registry, the Population's Education Registry, the Income Statistics Registry, DREAM, DNPR, Statistics Denmark	DBCG, DREAM, DNPR, Statistics Denmark FFPE archived tumor tissue	DBCG, the Danish Population Registry, the Population's Education Registry, the Income Statistics Registry, the Medical Birth Registry, DREAM, DNPR, Statistics Denmark and FFPE.
Study population	Women in the ProBe CaRe cohort diagnosed with primary invasive non-distant metastatic breast cancer during 2007–2011, who were premenopausal at diagnosis, aged ≤55 years, received adjuvant chemotherapy and who were not lost to follow-up within the first six months after breast cancer diagnosis.	Women in the ProBe CaRe cohort diagnosed with primary invasive non-distant metastatic breast cancer during 2007–2011, who were premenopausal at diagnosis, aged ≤55 years, received adjuvant chemotherapy, had available FFPE archived tumor tissue and who were not lost to follow-up within the first six months after breast cancer diagnosis.	Women in the ProBe CaRe cohort diagnosed with primary invasive non-distant metastatic breast cancer during 2007–2011, who were premenopausal at diagnosis, aged ≤55 years, had available FFPE archived tumor tissue, were employed at breast cancer diagnosis, and were intended to receive chemotherapy.
Exposures	Highest achieved level of education Household income Employment Marital status Cohabitation	26 SNPs in 20 genes: *ABCB1*, *ABCC2*, *ABCG2*, *SLCO1B1*, *SLCO1B3*, *CYP1A1*, *CYP1B1*, *CYP3A*, *CYP3A4*, *CYP3A5*, *GSTP1*, *ERCC1*, *ERCC*, *EPHA4*, *EPHA5*, *EPHA6*, *EPHA8*, *FGD4*, *ARHGEF10*, and *TRVF1*.	26 SNPs in 20 genes: *ABCB1*, *ABCC2*, *ABCG2*, *SLCO1B1*, *SLCO1B3*, *CYP1A1*, *CYP1B1*, *CYP3A*, *CYP3A4*, *CYP3A5*, *GSTP1*, *ERCC1*, *ERCC*, *EPHA4*, *EPHA5*, *EPHA6*, *EPHA8*, *FGD4*, *ARHGEF10*, and *TRVF1*.
Outcomes	Breast cancer recurrence All-cause mortality	Breast cancer recurrence All-cause mortality	Return-to-work Stable labour market attachment
Covariates	Age, comorbidities, ER status*, HER2 status, triple negative breast cancer, lymph node involvement, tumor size, TNM stage, pathological grade, surgery type including ITT chemotherapy and endocrine therapy.	Age, comorbidities, ER status*, HER2 status, triple negative breast cancer, lymph node involvement, tumor size, TNM stage, pathological grade, surgery type including ITT chemotherapy and endocrine therapy. Socioeconomic indicators including education, household income, marital status, cohabitation and employment.	Age, comorbidities, ER status*, HER2 status, triple negative breast cancer, lymph node involvement, tumor size, TNM stage, pathological grade, surgery type including ITT chemotherapy and endocrine therapy. Socioeconomic indicators including education, household income, cohabitation and employment.
Statistical analyses	Incidence rates per 1,000 person-years Cumulative incidences considering death as competing risk when examining recurrence. Crude and adjusted Poisson regression models. Poisson regression models were stratified by ER status.	Cumulative incidences considering death as competing risk when examining recurrence. Crude and adjusted cause-specific Cox-regression models Cox- regression models were stratified by ER status/endocrine treatment and stage	Cumulative incidences considering death and retirement as competing risk when examining recurrence. Crude and adjusted cause-specific Cox-regression models Cox- regression models were stratified by ER status, income, education and cohabitation.
Sensitivity analyses	1) Pooled recurrences with BCSMs 2) Assessed BCSM 3) Narrowed the assessment window of employment status to 1-3 months prior to breast cancer.	1) Pooled recurrences with BCSMs 2) Assessed BCSM	1) Applied an inclusion criterion of least 4 weeks of work prior to breast cancer during 8 weeks 2) Narrowed the employment assessment window to 4 weeks. 3) Included flexible job schedules in the return-to-work assessment

Abbreviations: BCSM= Breast cancer specific mortality, DBCG=Danish Breast Cancer Group, DNPR= Danish National Patient Registry, DREAM= The Danish Register for Evaluation of Marginalization, ER=Estrogen receptor, FFPE=Formalin-fixed paraffin-embedded, HER2=Human epidermal growth factor receptor 2, ITT= Intention-to-treat, ProBe CaRe = Predictors of Breast Cancer Recurrence, SEP= Socioeconomic position and TNM= Tumor node metastasis. *ER status included receipt of endocrine therapy.

Results

The following chapter contains the main findings from Studies I–III. More details are available in Appendices I–III. The flowchart in Figure 8 illustrates study cohort sampling for each study, outlining the number of excluded women, the reasons for exclusion, and the final sample sizes. Patient characteristics for each study can be found in the appendices.

Figure 8. Study cohort overview, outlining the number of excluded women and reasons for exclusion (white boxes) and the final cohort size in each paper (blue boxes)



SEP and breast cancer recurrence and mortality (Study I)

During follow-up (median 6.6 years, IQR: 5.4–7.9), 286 women were diagnosed with breast cancer recurrence, corresponding to a cumulative incidence of 13% (95% CI: 12%–15%). A total of 223 women died during follow-up (median 7.2 years, IQR: 6.0–8.6), corresponding to a CIP of 11% (95% CI: 9%–13%).

For all SEP indicators, women with the lowest SEP (blue lines in Figure 9) had higher cumulative incidence of breast cancer recurrence and mortality, compared with higher SEP (red lines in Figure 9).

Figure 9. Cumulative incidence proportions (%) of breast cancer recurrence (dashed lines) and mortality (solid lines) by marital status, cohabitation, income, education level and employment status.

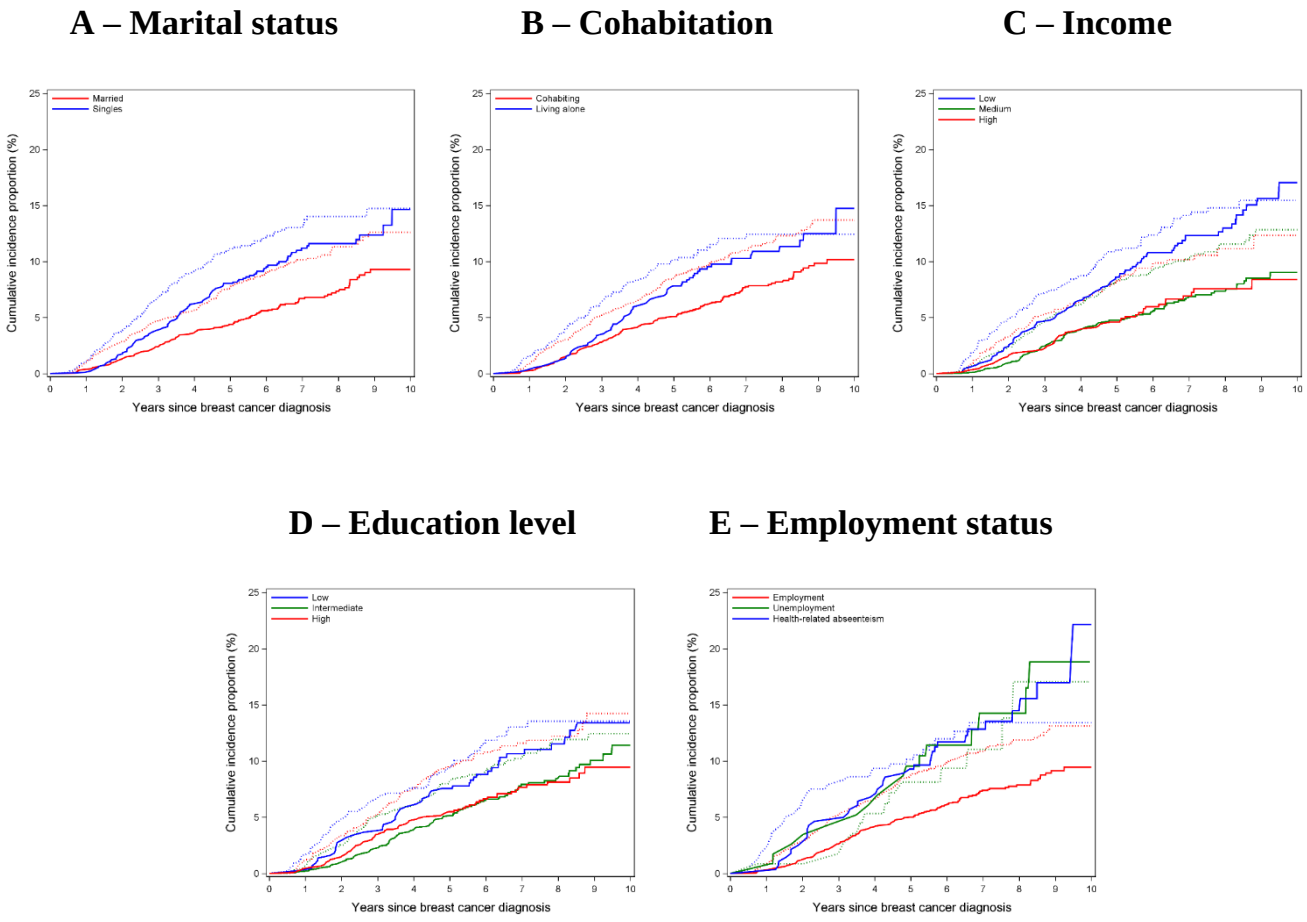


Figure from Hjorth *et al.*[173] Appendix I.

We observed lower incidence rates of breast cancer recurrence and mortality in single women than in married women (recurrence: 27 vs. 18 and mortality: 19 vs. 10 per 1.000 PYs), corresponding to lower 5-year IRRs of

recurrence ($IRR_{adjusted}$: 1.49, 95% CI: 1.11–1.89) and mortality ($IRR_{adjusted}$: 1.83, 95% CI: 1.32–2.52). Women living alone had higher incidence rates of recurrence (24 per 1,000 PY) and mortality (18 per 1,000 PY), than those living with a partner (20 and 12, respectively). Our observed crude IRRs of recurrence at 5 years and mortality at 10 years were elevated (IRR: 1.30, 95% CI: 0.86–1.62 and IRR: 1.41, 95% CI: 1.05–1.88). The estimates attenuated after adjusting for age and marital status. Women with low income had greater incidence rate of recurrence (27 per 1,000 PY) and mortality (20 per 1,000 PY), than those with medium or high income (19 and 11 per 1,000 PY, respectively). The adjusted IRRs of 5-year recurrence and mortality were 1.20 (95% CI: 0.80–1.80) and 1.37 (95% CI: 0.83–2.28) in low compared with high income women, respectively. Women with low and high education had similar 5-year IR of recurrence (23 per 1,000 PYs), but not mortality (18 vs 13 per 1,000 PYs); we found a corresponding elevated IRR of mortality in women with low versus high income ($IRR_{adjusted}$: 1.49, 95% CI: 0.95–2.33). Unemployed women had the lowest 5-year incidence rate of recurrence but the highest incidence rate of mortality (19 and 22 per 1,000 PY, respectively). Accordingly, their 5-year IRR of recurrence was lower than that of employed women ($IRR_{adjusted}$: 0.68, 95% CI: 0.32–1.46), while their IRR of mortality was increased ($IRR_{adjusted}$: 1.61, 95% CI: 0.83–3.13). Compared with employed women, those with health-related absenteeism from work had increased 5-year IRRs of recurrence ($IRR_{adjusted}$: 1.22, 95% CI: 0.80–1.84) and mortality ($IRR_{adjusted}$: 1.80, 95% CI: 1.14–2.82). When stratifying the analyses by ER status, we found that the apparent inequality in the IRRs of recurrence and mortality was most evident in women with ER+ tumors undergoing tamoxifen treatment. All estimates (both crude and adjusted 5-year and 10-year) can be found in Appendix I.

Polymorphisms and breast cancer recurrence and mortality (Study II)

Among the candidate SNPs, 21 were genotyped successfully with call rates ≥95%. Five SNPs were excluded due to call rates <95 % (*ABCB1* rs10248420, *CYP1A1* rs1048943, *TRPV1* rs879207, *ARHGEF10* rs9657362, *EPHA8* rs209709). Details on the excluded SNPs are provided in the *Supplemental materials*.

We found that within 10 years of follow-up, 249 women experienced a recurrence (cumulative incidence: 13%, 95% CI: 12%–15%). During 13 years of follow-up, 259 women died (cumulative incidence: 16%, 95% CI: 12%–20%).

The forest plots presented in Figures 10 and 11 include SNPs where we detected some associations with recurrence and mortality, respectively. Forest plots showing the findings for all investigated SNPS can be found in Appendix II. Two SNPs, *SLCO1B1* rs2306283 and *CYP1B1* rs1056836, showed decreased HR of recurrence in variant carriers (HR: 0.82, 95% CI: 0.64–1.06 and HR: 0.88, 95% CI: 0.68–1.15, respectively). In contrast, *GSTP1* rs1138272 was associated with increased HR of recurrence (HR: 1.16, 95% CI: 0.84–1.62).

Figure 10. Forest plot illustrating HRs and 95% CIs of the association between SNPs and breast cancer recurrence

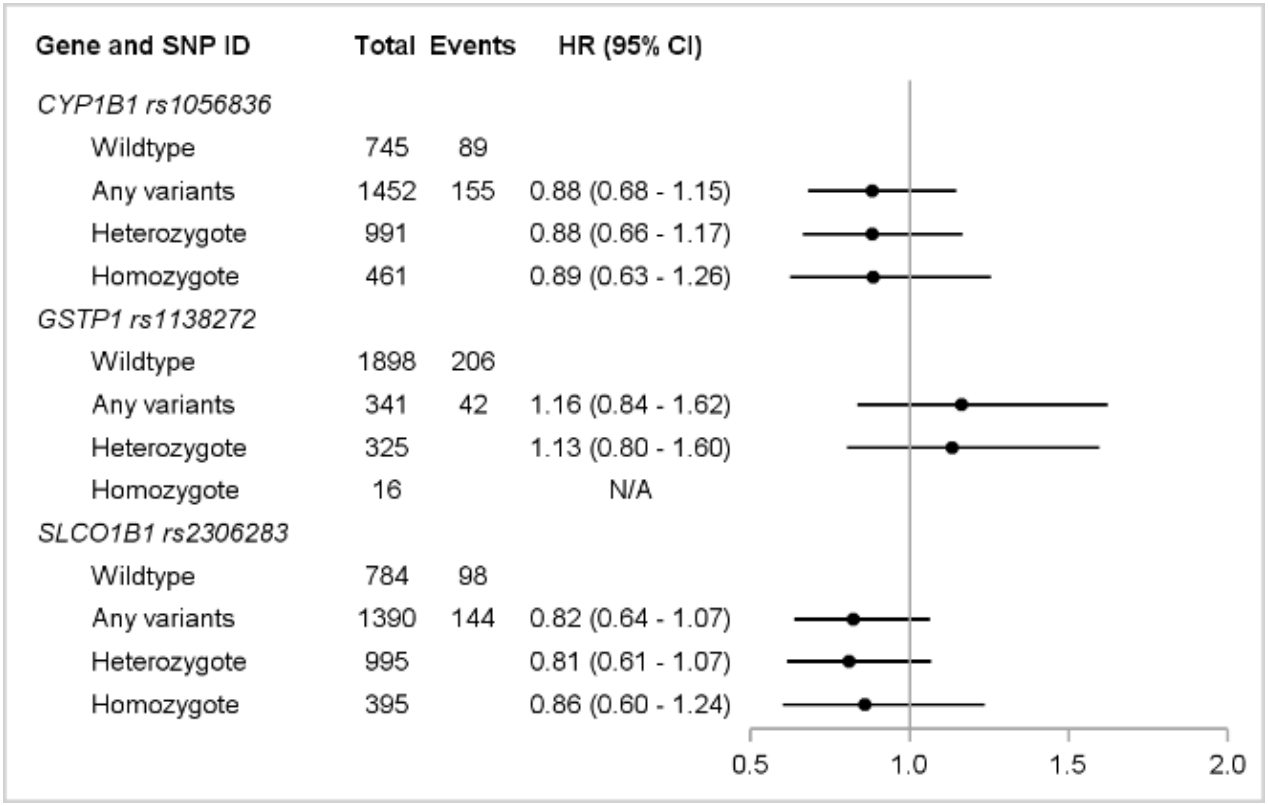


Figure from Hjorth *et al.*, Appendix II.

Figure 11. Forest plot illustrating HRs and 95% CIs of the association between SNPs and mortality.

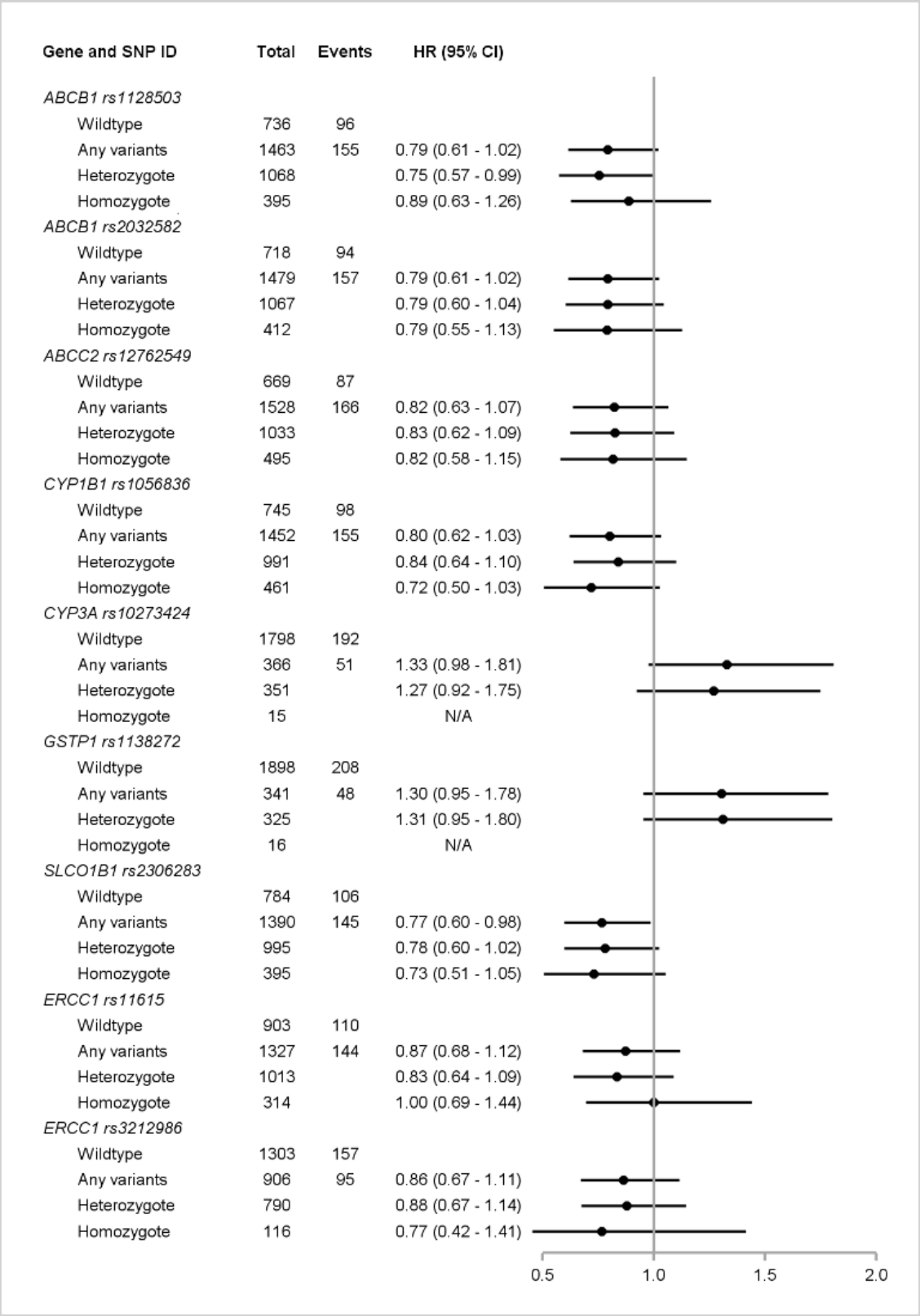


Figure from Hjorth *et al.*, Appendix II.

We identified several SNPs associated with mortality. *SLCO1B1* rs2306283 was associated with decreased HR of mortality (HR: 0.77, 95% CI: 0.60–0.98). *CYP3A* rs10273424 and *GSTP1* rs1138272 were associated with increased HR of mortality (HR: 1.33, 95% CI: 0.98–1.81 and HR: 1.30, 95% CI: 0.95–1.78, respectively), whereas *CYP1B1* rs1056836 was associated with decreased HR of mortality (HR: 0.80, 95% CI: 0.62–1.03) in variant carriers.

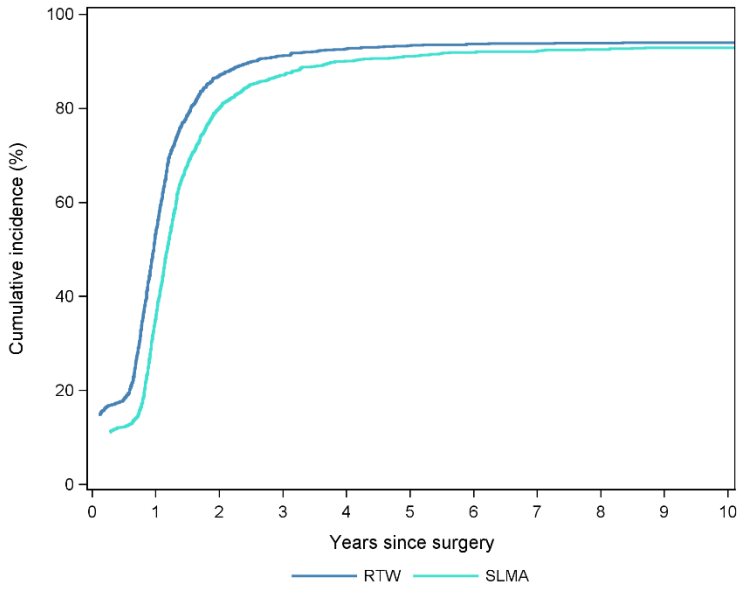
Three SNPs encoding ABC-transporters were associated with lower mortality rates: *ABCB1* rs1128503 (HR: 0.79, 95% CI: 0.61–1.02), *ABCB1* rs2032582 (HR: 0.79, 95% CI: 0.61–1.02) and *ABCC2* rs12762549 (HR: 0.82, 95% CI: 0.63–1.07). Among the SNPs encoding DNA repair genes, *ERCC1* rs11615 (HR: 0.87, 95% CI: 0.68–1.12) and rs3212986 (HR: 0.86, 95% CI: 0.67–1.11) were associated with decreased mortality rates. We observed similar findings considering homozygote variants and heterozygote variant allele carriers (instead of any variant) compared with wildtypes.

Using a composite endpoint of recurrences and BCSM provided similar results to those for recurrence alone, though the HR was elevated in variant carriers of *GSTP1* rs1138272 (HR: 1.23, 95% CI: 0.92–1.64). Analyses for BCSM provided similar results to the all-cause mortality models, though with a declining rate in *SLCO1B1* rs2306283 (HR: 0.71, 95% CI: 0.55–0.93). We did not observe effect measure modification by ER status or stage.

Polymorphisms and return-to-work (Study III)

Compared with the women who were not employed, the study cohort were less likely to be diagnosed with stage III tumors (22% vs. 17%), to receive a mastectomy (44% vs. 38%), to be single (51% vs. 32%), to be living alone (35% vs. 21%) and to have low education (37% vs. 14%).

Figure 12. Cumulative incidence curves of return-to-work and stable labour market attachment.

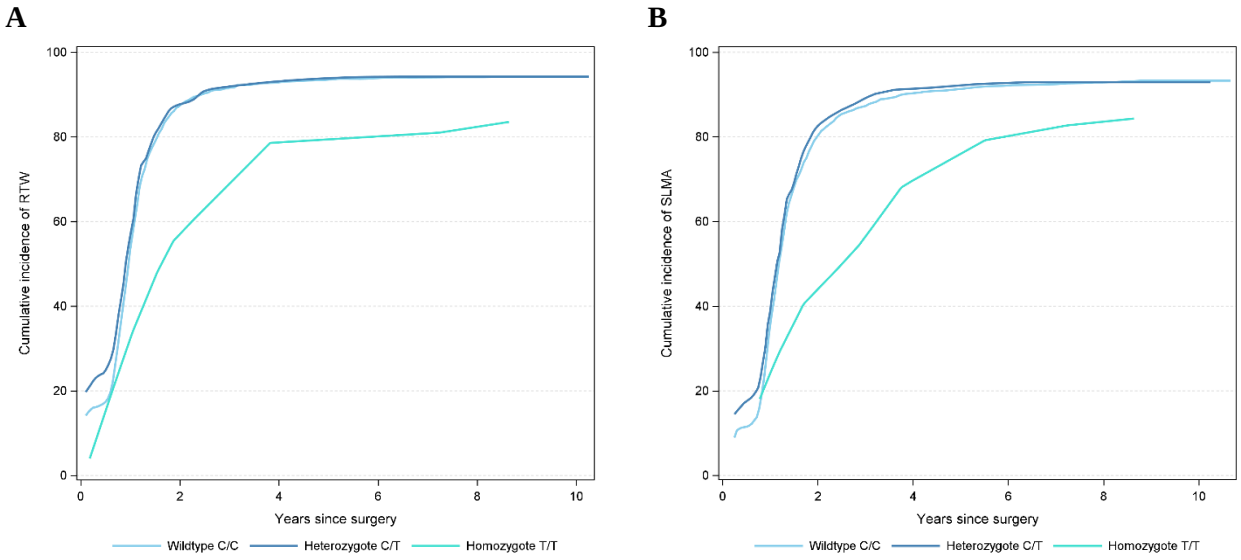


Abbreviations: RTW= Return-to-work, SLMA= stable labour market attachment.

Figure from Hjorth *et al.*, Appendix III (submitted in black and white according to journal recommendations).

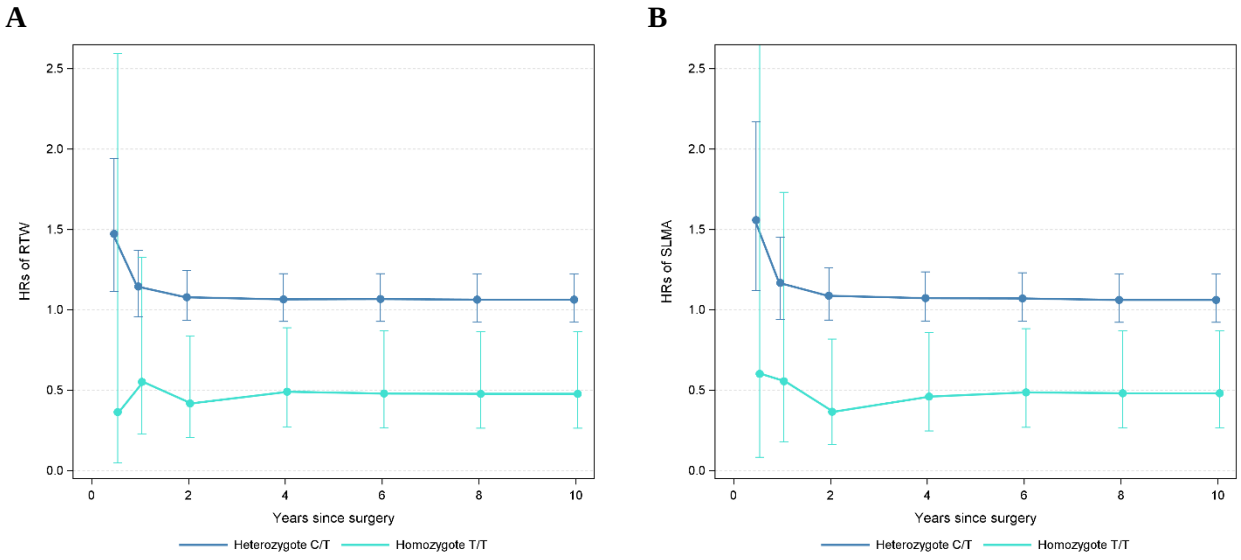
As seen from Figure 12, the cumulative incidence of return-to-work and stable labour market attachment increased steadily during the first three years after diagnosis. At 10 years, the cumulative incidences of return-to-work and stable labour market attachment were similar (94% and 93%, respectively). With regards to *CYP3A5* rs776746, homozygotes had delayed return-to-work (Figure 13, panel A) and stable labour market attachment (Figure 13, panel B) compared to wildtypes and heterozygotes, who had similar incidences.

Figure 13. Cumulative incidence of return-to-work (A) and stable labour market attachment (B) by *CYP3A5* genotype



Abbreviations: RTW= Return-to-work, SLMA= stable labour market attachment. Curves were smoothed using loess function. Figure from Hjorth *et al.*, Appendix III (submitted in black and white according to journal recommendations).

Figure 14. Hazard ratios of return-to-work (A) and stable labour market attachment (B) in *CYP3A5* rs776746 heterozygotes and homozygotes, compared to wildtypes.



Abbreviations: RTW= Return-to-work, SLMA= stable labour market attachment. Following estimates are provided in Appendix III; 0–6 months, 0–1 year, 0–2 years and 0–10 years. Figure from Hjorth *et al.*, Appendix III (submitted in black and white according to journal recommendations).

As shown in Figure 14, return-to-work (panel A) and stable labour market attachment (panel B) were delayed in *CYP3A5* rs776746 homozygotes throughout follow-up with HRs around 0.50 (10-year HRs: 0.48, 95% CI: 0.26–0.86 and 0.49, 95% CI: 0.27–0.88, respectively). The finding was not modified by indicators of SEP and could not be assessed according to ER status because of small strata.

We observed spurious relative associations for other SNPs, but due to imprecise estimates and minor absolute differences observed in cumulative incidence curves, we judged random error to influence these. The estimates can be found in Appendix III.

Discussion

This thesis provides information on factors associated with breast cancer recurrence, mortality and return-to-work in premenopausal women treated for breast cancer. We found that single women had increased risk of recurrence and mortality, while women with low SEP, as indicated by income, education and employment, had increased mortality. The inequality seemed most prominent in women diagnosed with ER+ disease, and thus undergoing endocrine treatment. Using DNA extracted from tumor-infiltrated archival tissue, we found that, in particular, *SLCO1B1* rs2306283 and *GSTP1* rs1138272 were associated with both breast cancer recurrence *and* mortality. Variant carriers of *SLCO1B1* rs2306283 had decreased risk of recurrence and mortality, and variant carriers of *GSTP1* rs1138272 had increased risks. *CYP3A* rs10273424 was associated with increased mortality, and ABC-transporters were associated with decreased mortality. We also observed an association of *CYP3A5* rs776746 with return-to-work—a potential proxy for recovery after breast cancer treatment.

Comparison with existing literature

SEP and breast cancer recurrence and mortality (Study I)

Our findings of increased mortality (all-cause or BCSM) in unemployed, single, and women with low income or low education corroborates findings from studies identified in search I.[3,7,13,90,92–95,98–103,105–108,112] However, none of the published studies investigated women undergoing taxane-based chemotherapy regimens. We did not identify any studies examining the influence of health-related absenteeism from work on breast cancer prognosis.

Our overall incidence of breast cancer recurrence was similar to that reported in a previous Danish study by Rasmussen *et al.*.[90] Like us, Rasmussen *et al.* observed lower rates of recurrence in women with high compared with low education. In general, Rasmussen's study population (including 67,092 cancer survivors, 27,752 of whom were women with breast cancer) had lower education than our study population (30% vs. 15% had low education and 18% vs. 40% had high education, respectively). These differences in education level between our study and the study by Rasmussen are likely due to the inclusion of elderly patients in their study. The studies also differed in the distribution of age/menopausal status. Furthermore, cancer-directed treatment was not examined by Rasmussen *et al.*.

In the study by Rasmussen *et al.*, the adjusted breast cancer recurrence rate was increased in women living alone. We also found an increased but imprecise crude risk of recurrence at five years in women living alone, but this attenuated after adjustment for marital status and age. However, the adjusted models in our study and in the study by Rasmussen *et al.* may be over adjusted. Rasmussen *et al*. incorporated age, marital/cohabitation status, education, comorbidity, calendar period, tumor stage, nodal stage and adjuvant therapy in one

multivariate model rather than considering the minimal sufficient adjustment sets to evaluate each SEP indicator. As an example, cohabitation may not be influenced by calendar period and education. In our cohabitation model, we adjusted for marital status, which, along with other factors, is incorporated in the algorithm used by Statistics Denmark to derive cohabitation.[163] We note, that our crude models suggested increased recurrence and mortality in women living alone.

The underlying mechanism behind poorer prognosis in single women remains unclear but can be hypothesized to reflect the importance of social support. Compared with single women, married women may have a better support system to support symptom recognition, health seeking behavior[174] and to withstand treatment; this may be especially important for those who suffer severe side-effects of treatment.[175] This may be particularly relevant in women also treated with endocrine treatment, which will be discussed below.

Our findings suggest that SEP-related inequalities in cancer prognosis were most prominent in women with ER+ breast cancer. A study by Di Salvo *et al.*[91] included 3,358 women diagnosed with breast cancer during 2003–2005 and evaluated SEP using an area-based deprivation index. They found higher 5-year cumulative incidence of recurrence in women with the highest deprivation levels compared with those with the lowest deprivation levels. Stratification by hormone receptor status showed greater inequality in the ER+/PR+ group. The opposite gradient was seen among women with ER–/PR– tumors. A US-based study reported most pronounced inequality in unmarried versus married women with HR+/HER2– tumors, but inequality was also evident in women with triple negative breast cancers.[102] The authors discussed that the findings could be driven by the poor prognosis of triple negative breast cancer, which may have overruled the effect in the hormone receptor negative group. The inequality seen in women with ER+ tumors may have also been related to patient adherence to endocrine therapy. Discontinuation and non-adherence to endocrine therapy increases the risk of mortality after breast cancer.[176] Adherence to endocrine therapy is particularly low in women with low social support and/or in younger women—women aged below 40 years are 50% less likely to discontinue, and 40% more likely to be non-adherent compared with women aged over 40 years.[177,178] The observed findings are likely a shared effect of these mechanisms.

After the publication of Study I, van Maaren *et al.*[109] published a cohort study examining the impact of SEP on the risk of breast cancer recurrence and mortality in women aged <40 years, diagnosed with stage I–III breast cancer. The study was conducted in the Netherlands, where health care access is uniform, similar to the Danish. Using an area-based SEP indicator (postal code-based algorithm including mean household income, and percentages of people with low income, low education and unemployment), they observed lower 10-year recurrence risk in women with high SEP, compared with women with low SEP. Although they did not use the same indicators of SEP, the inequality reported by van Maaren *et al.* was more pronounced than that seen in our study. The younger age of the study cohort in the study by van Maaren *et al.* may explain these differences, as the influence of SEP may be more pronounced in younger compared with older women.[92,95,112] Another

explanation might be that the area-based indicators do not always provide the same findings as individual-based measures, though area-based measures tend to underperform.[179,180] Nonetheless, taken together, these studies highlight the potentially negative influence of low SEP on breast cancer prognosis.

Polymorphisms and breast cancer recurrence and mortality (Study II)

As Search II indicates, research on SNPs and breast cancer recurrence and mortality is sparse and conflicting.

The most prominent finding in our study was that observed for *SLCO1B1* rs2306283 suggesting decreased risk of recurrence and mortality in variant carriers. In the literature review, I identified no other studies on taxane effectiveness or efficacy examining *SLCO1B1* rs2306283 in breast cancer. One pharmacokinetic assessment suggests that this variant allele is associated with longer paclitaxel clearance in breast cancer patients,[181] whereas one study in 141 white cancer patients (sites not reported) reported no association with docetaxel clearance.[182] Theoretically, *SLCO1B1* rs2306283 reduces OATP1B1 activity, increasing plasma concentrations of docetaxel. This could explain our findings of lower risk of recurrence and mortality in *SLCO1B1* rs2306283 variant carriers.

Our findings of decreased recurrence and mortality risk in *CYP1B1* rs1056836 variant carriers reflects, to some extent, findings in studies by Abdul Aziz *et al.*[116] and Marsh *et al.*.[119] Both studies included breast cancer patients undergoing taxane combination chemotherapy incorporating doxorubicin and cyclophosphamide. Abdul Aziz *et al.* included women with non-metastatic triple negative breast cancer and investigated recurrence and mortality. Marsh *et al.* investigated the risk of PFS in breast cancer patients, 16% of whom had metastatic breast cancer.[119] Neither study was therefore directly comparable to ours due to heterogeneous populations. Moreover, in the study by Marsh *et al.*, taxane-based chemotherapy was used as consolidation therapy, meaning it was given after initial adjuvant treatment.[119] In addition, PFS is not comparable to our assessment of breast cancer recurrence and mortality.[183] PFS implies disease progression indicating that patients were never disease free. In contrast, recurrence implies patients were treated curatively and considered disease-free at some point. However, PFS are poorly standardized, which limits cross study comparisons.[184,185] Nonetheless, our studies suggest a favorable effect of *CYP1B1* rs1056836 variant alleles on prognosis.

In contrast, *CYP3A* rs10273424 was associated with increased mortality, which was unexpected as SNPs in CYPs often decrease enzyme activity.[186,187] Apart from the other included SNPs, *CYP3A* rs10273424 is an intron variant, which may explain why no studies have investigated its influence on the effectiveness of breast cancer treatment. Nonetheless, it has been associated with increased breast cancer risk in premenopausal women.[188] Our findings may reflect gene splicing or linkage disequilibrium with other SNPs. Notably, *CYP3A* rs10273424 was not associated with recurrence.

Eckhoff *et al.*[72] examined the impact of two SNPs in *GSTP1* and three SNPs in *ABCB1* on docetaxel-induced peripheral neuropathy, in a case-control study including 150 Danish trial participants with breast cancer.

GSTP1 rs1138272 was associated with docetaxel-induced peripheral neuropathy. This finding contrasted with findings from other small randomized clinical trials.[119,189] The mechanisms underlying our findings of increased recurrence and mortality are unclear, but concurrent increases in toxicity *and* effectiveness point towards increased drug exposure and reduced drug clearance (via reduced GSTP1 enzyme activity).

Previous studies in breast cancer found no associations between *ABCB1* rs1128503 or *ABCB1* rs2032582 and OS or PFS.[117,119,121] Meta-analyses including multiple cancer sites show conflicting findings.[190,191] One meta-analysis examining *ABCB1* rs1128503 was based on 423 patients from five studies and as such, may have been underpowered.[191] A larger meta-analysis including 3,320 cancer patients (ovarian, breast, gastric, lung and, head and neck) from 15 studies linked *ABCB1* rs1128503 homozygotes with improved OS, consistent with our observed decreased mortality.[190] The study found no association between *ABCB1* rs1045642 and PFS or OS[190]. Similarly, we saw no evidence of an association of *ABCB1* rs1045642 with breast cancer recurrence or mortality in our study. Other studies on breast cancer OS and PFS found an unfavorable effect of *ABCB1* rs1045642 (see literature review for Study II)[89,117,121], but the study samples were small (≤ 216 women) and the findings may have been attributable to chance.

We observed decreased mortality in variant carriers of *ERCC1* rs11615 and rs3212986. In advanced non-small-cell lung cancer treated with docetaxel and cisplatin, *ERCC1* rs11625 variant carriers was associated with improved tumor response.[192] Undoubtedly, a population of advanced non-small cell lung cancer patients is distinct from our cohort of early-stage breast cancer patients. Still, together, these studies may indicate a potential beneficial effect of SNPs in *ERCC1* genes and response to docetaxel-based chemotherapy.

In summary, the findings in *SLCO1B1* rs2306283 and *GSTP1* rs1138272 were most consistent, as these SNPs were related to breast cancer recurrence and mortality.

Polymorphisms and return-to-work (Study III)

This is the first study to investigate the association of SNPs related to taxane metabolism and transport and return-to-work. We focused on the cumulative incidences of return-to-work and stable labour market attachment in our study cohort. Beyond its own significance, return-to-work may also be a marker of recovery, which certainly is related to adverse effects. Yet, we had no information on adverse effects. Instead, I discuss our findings in the context of pharmacogenetic studies focused on toxicities and adverse effects, as they may help to explain our findings of delayed return-to-work in *CYP3A5* rs776746 homozygotes.

Most previous studies reported return-to-work prevalence at one year after breast cancer diagnosis,[9,130] rather than cumulative incidence. Such point prevalences do not accurately reflect time to return-to-work or stable labour market attachment. These studies show substantial variation in 1-year prevalences of return-to-work ranging from 43% to 93%, depending on country. One study suggested a 10-year cumulative incidence of 85% in German breast cancer survivors aged 20–59 years.[131] However, they examined return-to-work in 5-year

survivors. Therefore, the study is likely prone to selection bias as women who died in the first five years after breast cancer diagnosis are likely to have had a poorer clinical course, potentially detaching them from the workforce. The cumulative incidence in the German study was lower than we observed in Study III. We note that the German study included stage IV breast cancer, which has poorer prognosis, and likely contributed to the lower proportion of survivors who returned to work 10 years after diagnosis. The German study was also prone to recall bias, as return-to-work was reported retrospectively by the survivors.

Our analyses showed delayed return-to-work in *CYP3A5* rs776746 homozygotes, suggesting poorer recovery compared with heterozygotes and wildtype carriers. Polymorphisms in *CYP3A5* rs776746 lead to mRNA splicing defects, which impact CYP3A5 enzyme expression. Caucasians who are *CYP3A5* rs776746 wildtype or heterozygote carriers are CYP3A5 non-expressors; *CYP3A5* rs776746 homozygotes are expressors[193] but this variant is rare in Caucasians (minor allele frequency=7%, see Appendix I). The variant allele has been associated with paclitaxel-induced neurotoxicity in Spanish cancer patients (presumably mostly Caucasian),[194] which seems at odds with the potential consequence of the splicing defect. Yet, the study only examined toxicities during treatment, and hence not long-term adverse effects, which could be the mechanism underlying our findings. Eckhoff *et al.*[195] did not find any association between *CYP3A5* rs776746 and docetaxel-induced neuropathy, but they were limited by sample size, as only two homozygotes were included.

Various SNPs included in our study have been associated with toxicities or other prognostic endpoints, but no previous study investigated their association with return-to-work. This may illustrate that adverse effects do not necessarily compromise return-to-work. Still, another explanation may relate to study sizes. Our study was larger than the previous studies examining associations between SNPs and adverse effects and thus better powered to elucidate associations. Some HRs in our study suggested associations between SNPs and return-to-work, but cumulative incidence curves indicate these differences most likely arose due to chance. Further supporting this possibility is that these associations were often observed in SNPs with a low proportion of homozygote variants. Also, many studies were conducted in postmenopausal women, many of whom were beyond the typical working age. Premenopausal women may be more likely to return-to-work owing to younger age and better physical performance enabling them to better cope with adverse effects.

Methodological considerations

As with all epidemiological studies, a critical appraisal of the methodologies is needed before drawing any conclusions. In the following section, I will carefully assess consequences of methodological choices made when we conducted our studies, and their related strengths and limitations. I will discuss the design of the studies, including the data and analyses used. Moreover, I discuss sources of random error and systematic errors (selection and information bias) concerning the internal validity, and the external validity (generalizability) of the studies.

Denmark is reputed as "an entire cohort"[196] owing to the large collection of individual-level, longitudinal data, tracking individuals from early life to death. Data include health and clinical quality data [146] and administrative data on education, marriage, housing, employments, among other data. The individual-level linkage enables life-long follow-up of all citizens, so long as they remain Danish residents. A considerable advantage of data derived from administrative and health registries, is the routine recording and prospective data collection.[106] Nonetheless, important limitations should be considered. The data are registered for administrative purposes only and lack detailed information as outlined below.

Study design

We took advantage of an existing cohort study (ProBe CaRe), which had been assembled for studies of biomarkers and treatment effectiveness in premenopausal breast cancer.[150] We chose a cohort design rather than a case-control design as we aimed to study both absolute and relative risks, and to investigate multiple exposures and outcomes in each of the studies. We could also have used a case-cohort design, but chose a cohort design to optimize the study precision.[140]

In Studies II–III, the proposed associations with SNPs and the outcomes cannot be solely attributed to docetaxel, as docetaxel was recommended in combination with cyclophosphamide and sometimes epirubicin. Clinical trials randomizing docetaxel treatment could elucidate causality between SNPs and docetaxel efficacy but would be unethical given the survival benefit associated with docetaxel. We therefore examined the impact of SNPs on the effectiveness of docetaxel in a "real-world" setting—in women undergoing taxane-based chemotherapy (Study II) or intended for taxane-based chemotherapy (Study III).

Selection of single nucleotide polymorphisms

A limitation of the candidate gene approach is that some potentially important genes may be overseen and not included.[165] We performed a thorough examination of preexisting literature, including genome-wide association studies aimed at uncovering new candidate genes relevant to taxane effectiveness.[74,76] In addition, we consulted specialists and researchers within the research field, to identify less investigated SNPs with plausible impact on the outcomes. We mainly selected SNPs in the coding region, but also one non-coding

variant (*CYP3A* rs10273424), as this has been linked with estrogen levels in premenopausal women and their risk of breast cancer.[188] Non-coding variants can have a functional effect through mRNA- splicing or stability, or gene expression.[165]

Selection of socioeconomic indicators

As defined above, SEP refers to a position within the society,[82] and the SEP indicators we used have distinct features (*i.e.* the white areas in Figure 4), but also reflect a core dimension shared with other indicators of SEP (*i.e.* the blue areas in Figure 4). We were interested in both the unique aspects of the SEP indicators *and* the core SEP. We therefore carefully selected minimal adjustment sets, as minor adjustment could leave residual confounding from the unique aspects of other SEPs, but mutual adjustment would rely on an assumption of no effect of the core SEP.[83] In Study I, the direction of the estimates were similar across the indicators of low SEP, which we believe is a reflection of the core SEP. This reflects the appropriateness of the indicators included.

In Study I, we chose to include both cohabitation and marital status. Focusing solely on marital status gives the opportunity to collect this information on the date of diagnosis, whereas cohabitation in some cases can be registered up to one year before breast cancer diagnosis. Marital status encompasses life-changing events such as divorce and widowing. However, we did not examine the outcomes within these subgroups due to low numbers. Living together without being married, also when having children, is common in Denmark.[87] This is captured in the cohabitation algorithm. We therefore chose to include cohabitation only in Study III.

In the literature,[90,114] SEP is sometimes measured after breast cancer diagnosis, a situation prone to reverse causality as breast cancer diagnosis and treatment can impact certain SEP indicators. Breast cancer possibly motivates marriage among individuals in an ongoing partnership. On the other hand, breast cancer diagnosis is linked with marital stress and divorce in young women.[197] Employment may change, though Danish law protects against firing during disease. Unemployment rates are higher in breast cancer patients, compared with healthy controls.[48] Moreover, some may decide to leave the work force, or may lose their workability due to adverse treatment effects or surgery complications. Naturally, this influences post-diagnostic income if transitioning into social benefits, self-support (or by spouse), or reduced working hours.[44,198]

We assessed employment status including social benefit use. Occupation according to physical and emotional demands could have been informative, as such work demands are important factors for return-to-work.[50] Danish registry-data on occupation rely on industry type. This introduces potential misclassification bias, as job demands cannot be determined by industry. Moreover, type of job and hierarchal position, would be informative for SEP. Statistics Denmark provides such information, but lacks details on those detached from the workforce (type of social benefit),[199] which was a strength of our classification.

Missing values

In general, the positive predictive value, which is the proportion of patients registered with a disease that actually have the disease, is high in the Danish health registries.[146] However, some data may be missing, including some information on recurrences (this will be elaborated in the below section *Misclassification of outcomes*).[200] We evaluated the possibility of performing a probabilistic bias analysis to quantify the potential bias of the underreporting in our analyses with recurrence as outcome, but sufficient bias parameters were not available.[201,202]

In all the studies, we used complete case analysis meaning that some women were left out in a model if one or more covariates were missing. We did not exclude the women from the study cohort but included them in other models if the required covariates were available. Complete case analyses reduce power and precision and introduce bias if there is missing data. In Study I, we lacked some information on income (0.3%), education (1.0%), employment status (0.3%) and cohabitation (0.8%), which may be explained by recent immigration. As such, these data were not missing at random.[203] As reported by Beiki *et al.*,[106] being an immigrant may relate to a poorer breast cancer prognosis. In such a case, this could have biased our findings to the null, though we consider the magnitude of bias to be small considering the low proportions of missing information. In Studies II– III, genotypes were missing for ≤5%, depending on the SNP. We believe the data was missing at random, and therefore unlikely to bias our estimates. Genotype imputation is feasible for example, by using the 1000 Genome Project as reference panel.[204] But considering the low proportion of missing data, we expect that the precision gain would be minor.

Random error and study power

Random error refers to statistical precision,[140] which was expressed using confidence limits in our studies. We did not use the CIs for significance testing to ensure we did not disregard strong associations due to power. Instead, we used the CIs to judge the degree of precision.[205] In several of our analyses, the confidence limits suggested a greater possibility of an association, rather than no association.

We did power calculations prior to the study to ensure sufficient sample size to achieve an informative result. Though sample size were reduced by exclusions and complete case adjustments (Study I), we had sufficient power to provide outcomes with acceptable precision. However, the precision of our estimates was reduced in some of the stratified analyses. For some SNPs, the apportionment ratio[140] departed from 1, where, for example, no homozygotes were present in one stratum, but were present in another stratum, then the ratio would be zero. This prohibited some of the planned stratified analyses.

Selection bias

Selection bias occurs when study cohort members and those *not* included in the cohort, who actually were eligible, differ in their association between the exposure and the outcome.[140]

The ProBe CaRe cohort (*i.e.,* our source population) includes women who were treated according to guideline endocrine treatment. Therefore, only women with ER+ disease who received tamoxifen, and women with ER– who did not receive tamoxifen, were included. Thus, women with unknown ER status and women with ER+ tumors who did not receive endocrine treatments were excluded. The same considerations apply to women who receive neo-adjuvant chemotherapy (which is more frequently used today than during the study period). Accordingly, the ProBe CaRe cohort reflected the guideline treatment regimens during the period of diagnosis. We find it unlikely that SNPs were related to receipt of tamoxifen or assigned ER status, as the SNP status (and thus the potential risk of toxicities) were unknown to the patients and clinicians, and genotyping was not used to determine patient suitability for treatment. Instead, our best explanation is that the excluded women were diagnosed with less severe, stage I breast cancer, and thus only related to the outcomes — just like those who were excluded due to no chemotherapy (Appendix I). According to a US based study, 14% of breast cancer patients refuse chemotherapy.[206] The proportion of patients who refuse treatment in countries with tax-funded health care, such as Denmark, and in younger women is likely to be lower. Refusal could, however, be more frequent in women with low SEP. In this case, the overall social gradient in prognosis would be underestimated. However, the aim of Study I was to examine taxane effectiveness in those treated.

In Studies II–III, genotypes served as exposures. FFPE blocks were archived at the time of primary breast cancer-directed surgery. This mitigated left truncation which would cause bias from temporal selection (*i.e.,* if the availability of archived tumor tissue was subtype dependent). It also avoided bias due to immortal time (*i.e.,* if women died before genotyping), which can occur in studies using DNA collected during follow-up,[124] including three studies identified in Search II.[116–118]

Information bias

Virtually all information is prone to measurement error and misclassification, which can bias estimated effects. Error distribution influences the direction and magnitude of the bias. These should therefore be considered to determine the extent of over- or under-estimation of the effect estimates. If the misclassification is distributed unequally across other variables—for example, the exposure—then it is called *differential* and the direction of the bias can be in either direction. If the misclassification is not related to other variables it is called non-differential and will most often cause underestimation or equalization (depending on the number of covariate categories).[140]

Misclassification of exposures

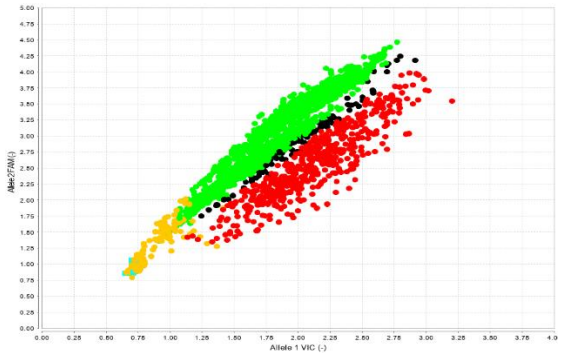
In Study I, the SEP indicators marital status, cohabitation, income, education level and employment status served as exposures. Married women included those in registered partnerships, and singles included never married, divorced, or widowed. The validity of marital status is considered high, but we expect that some women registered as "never married" might have a partner. These are expected to be captured as living with a partner in the cohabitation measure. However, this was only available on an annual basis, and could have changed at the time of breast cancer diagnosis. As such, non-differential misclassification could theoretically bias the associations towards the null.

The validity and completeness of the Danish Population Education Register is high.[159] We collected highest achieved education at breast cancer diagnosis, though some women might be enrolled in higher education. Therefore, some women can have been educated at a higher level since and so could possess a higher SEP. Such non-differential misclassification would typically reduce the strength of associations across the affected categories.

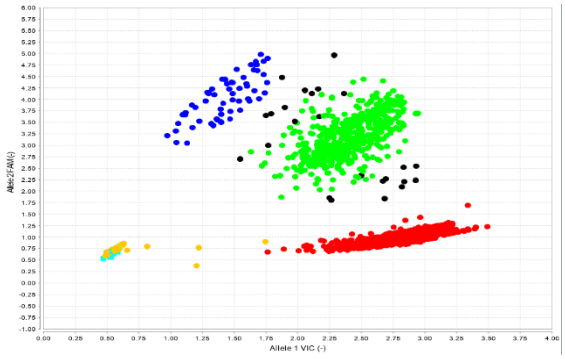
Genotypes can be misclassified due to imperfect fluorescence clustering during genotyping or low quality DNA. When available, DNA for genotyping is preferably extracted from plasma or serum, as this is a superior source of amplifiable DNA.[207] Genotyping of DNA extracted from FFPE can be challenged by fragmented or cross-linked DNA, leading to low call rates or genotype misclassification. However, evidence shows identical genotype classifications between FFPE tissue and plasma samples.[208,209] To ensure high quality genotyping data, we discarded four SNPs (*ABCB1* rs10248420, *CYP1A1* rs1048943, *ARHGEF10* rs9657362 and *EPHA8* rs209709) with poor amplification, as overlapping clusters increased the risk of misclassification, and consequently had call rates <95%. One of the overlapping cluster plots is provided in Figure 15, along with an example of a successful auto call. We also excluded *TRPV1* rs879207, which had well-defined clusters, but a call-rate of 93%. We therefore re-ran this assay in 101 samples (those with sufficient amounts of extracted DNA), but only six samples successfully genotyped. To reach a call-rate of 95%, 38 samples would have been needed, so *TRPV1* rs879207 remained excluded from the analyses.

Figure 15. Auto-called genotyping cluster plot for ARHGEF10 rs9657362 and EPHA6 rs301927

ARHGEF10 rs9657362



EPHA6 rs301927



The color-assigned clusters each represent a genotype. The dots represent samples. The black dots represent samples with no assigned genotype. In the left figure, the clusters overlap, and some samples may be misclassified. In the right figure, the clusters are well differentiated. In this example, the blue cluster represents the G/G genotype (variant homozygote), the green cluster represents the A/G genotype (heterozygote), and the red cluster represents the A/A genotype (wildtypes).

Hardy-Weinberg equilibrium is traditionally assessed using a Chi^2 test, where pre-determined critical values are used to judge equilibrium and eventually discard analyses. The reasoning behind these methods is to detect, stratification bias (according to ethnicity), selection bias or genotyping error, though recent knowledge refutes the latter argument.[210] Our study cohorts may include ethnicities other than Caucasian, but such information was not available. We did have information on immigration, but this is not representative for ethnicity as most immigrants living in Denmark come from European countries.[211] The studies were nationwide studies. As such, selection bias should be avoided. Some FFPE was missing, but we assumed they were missing at random. A further consideration against traditional Hardy-Weinberg equilibrium testing is that it is highly dependent on sample size. We therefore used Hardy-Weinberg equilibrium to calculate expected genotype frequencies and compared these to that observed. Misclassification of genotype may be non-differential, as it may not be related to the exposure or the outcome.

Misclassification of outcomes

In Studies I– II, the outcomes were breast cancer recurrence and mortality. Using medical records as a reference standard, a validation study conducted in a subsample of the ProBe CaRe cohort found that recurrences were registered in 70% (14 out of 20) of those with recurrence registered in their medical record.[200] The positive predictive value was 100%, and the missing records on recurrence may therefore be non-differentially misclassified. Therefore, our relative estimates will not be affected, unless in Study I if the underreporting was differential according to SEP. Our results in Study I indicated either underreporting or underdetection of recurrences. It seems more likely that a recurrence would not be detected, rather than SEP influencing the

reporting of recurrences to DBCG. In contrast, the absolute risks will likely be underestimated by 30% (if the validation cohort reflects our study population). The study did not include contralateral breast cancers, which are included in our study and could increase the completeness. However, it is difficult to distinguish whether a contralateral cancer is a recurrent event or a new cancer. To reduce underreporting, we included contralateral breast cancers as done by others.[149,212]

To examine BCSM, we used the Danish Register of Causes of Death, which in its present form goes back to 1970. From 2007 and onwards, causes of death have been reported using mandatory forms submitted electronically.[155] Autopsies are rarely used in Denmark, so the cause of death is prone to misclassification.[155] According to Statistics Denmark the cause of death statistics do not fully agree with cancer statistics, as the latter includes all deaths in persons with a cancer diagnosis. In the cause of death statistics, cancers are only included if they are registered as the underlying or contributory cause.[213] We therefore examined all-cause mortality to reflect cancer statistics. However, we conducted sensitivity analyses investigating BCSM using the Cause of Death Registry. We expected these statistics to be similar, making the assumption that deaths from causes other than breast cancer were unlikely in our young population.

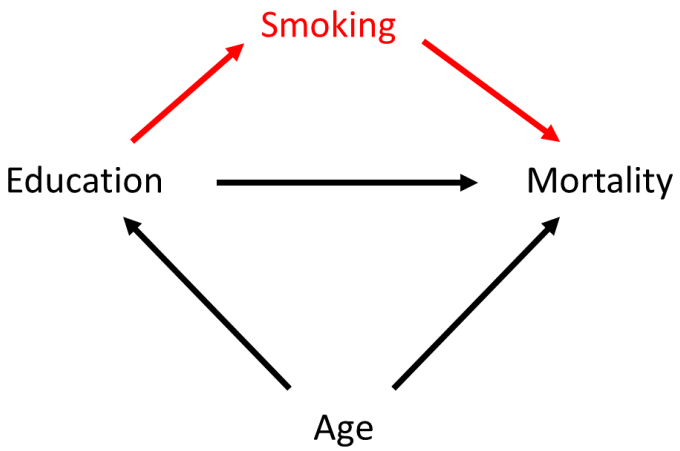
The outcomes in Study III were ascertained from the DREAM database. Misclassification was possible, as the weekly records on social benefits have several limitations. We assumed that no entry in DREAM (*i.e.* no social benefit payout) was equivalent to employment, but this may be inaccurate. First, some work absenteeism may not be recorded. Long-term sick leave is reported in DREAM, but short-term sick leave is not. The cut-off from short-term to long-term sick-leave extended during the study period, being 13 days until 1 January 2008, 14 days until 2 June 2008, and 21 days until 2 January 2012. Since 2012, the cut-off from short-term to long-term sick leave has been 30 days. This means that sick leave will first figure in DREAM between 14 and 30 days after the first day of sick leave. Moreover, vacations paid by the employer are not reported. Second, some employers may not claim reimbursements from the state. We believe reimbursement claims are routine practice at most workplaces and hence the misclassification will be small. Third, some women may choose not to return-to-work or may not claim benefits, due to high-earning spouse or other sources of financial support. An explorative analysis showed that women with high income had faster return-to-work than women with low income. High household income (or a private health insurance) could also be an incentive not to return-to-work or to stop working without claiming social benefits. In that case, the person would figure as employed in our definition. It is therefore possible that SEP will affect return-to-work classification, but non-differentially across SNPs. Thus, the relative estimates may not be influenced.

Confounding and effect-measure modification

Confounding is when one or more factors obscure the effect of an exposure on the outcome.[140] The confusion of the effect occurs when a variable—the confounder—has an effect on the exposure and the outcome, is unevenly distributed across exposure categories and is not on the causal pathway between the exposure and

the outcome. For example, one of the associations of interest in Study I, the influence of education on mortality, could be confounded by age as this is associated with both education and mortality, as illustrated in Figure 16. On the other hand, smoking is more prevalent in individuals with low education,[214] and is related to increased mortality in women with breast cancer.[215] However, we anticipate that smoking lies on the causal pathway, as it seems unlikely that smoking directly or indirectly influences on education. Therefore, smoking does not fulfill the criteria of being a confounder.

Figure 16. Directed acyclic graph for confounding.



In this example, age acts as a confounder, whereas smoking act as a mediator.

Confounding can be addressed by design (randomization, restriction and/or matching) or analytically (by stratification, standardization and/or adjustment). In all the studies, we sought to overcome confounding by place of residence, health care access, among other factors, by conducting population-based studies. In Study I, we also addressed confounding analytically by adjustment—for example, for age, as mentioned in the above example. We used DAGs to identify the minimum set of adjustment variables to eliminate confounding, under the assumed causal structure. Some of the SEP indicators included in the adjustment sets and the Charlson Comorbidity Index scores could vary over time but were only included as baseline adjustment variables due to the risk of reverse causation. The adjustments did not take unmeasured or unknown confounding into account, which observational studies are sensitive to. We considered the potential for unmeasured confounders in all the studies but did not identify such. Still, it is possible that unidentified factors could confound the associations between the exposures and the outcomes in our studies.

Effect-measure modification refers to a situation where the effect/association between the exposure and the outcome changes across strata of another variable.[140] We examined, and handled, effect-measure modification using stratification. In Study I, we supposed that SEP's effect on the outcomes could differ in women with ER− and ER+ tumors, as the latter women were treated with endocrine therapy. Adherence to endocrine therapy may be modified by SEP.[175] Though CIs overlapped, the results indicated that the inequality seen for marital status was particularly evident in women undergoing tamoxifen. We also stratified by ER status in Studies II & III, hypothesizing that the association between SNPs and return-to-work could differ if the SNPs influenced the effect of endocrine therapy. We also stratified by SEP, supposing that recovery could be delayed in women with low SEP, but this was evident in our results.

Immortal time bias

A frequently occurring bias in observational studies is immortal time bias. As briefly mentioned in the *Methods*, immortal time refers to a period during follow-up where the outcome of interest cannot occur. The bias potential arises when this period lies before the exposure of interest.[161] Although not the exposure, the last criterion for cohort entry in Studies I– II was received chemotherapy. Therefore, we could not start follow-up before the end of treatment, as patients had to remain alive to complete their last chemotherapy cycle. However, the date of chemotherapy completion was not available. We therefore began follow-up six months after surgery. The incidence of return-to-work in Study III was high during chemotherapy. We therefore chose to condition on ITT chemotherapy and start follow-up on the date of primary surgery.

Generalizability

Generalizability refers to the extent to which the findings are applicable to other settings and populations. In epidemiological studies, sampling of a study cohort should not focus on overall representativeness if this is incongruous with the research goals.[216] Nonetheless, we restricted the study cohorts to women diagnosed with non-distant metastatic breast cancer from 2007 onwards, as this timeframe approximated the introduction of taxane-based chemotherapy as guideline treatment. Thus, the findings may not be applicable to all patients diagnosed during the study period. Women who were excluded due to a lack of treatment generally had less aggressive breast cancer, and thus better prognosis (see Appendix I). Our study population was nationwide and population-based. We therefore consider the study cohort to be representative of premenopausal women assigned guideline therapy, including taxane-based combination chemotherapy in Denmark and in countries with similar breast cancer treatment and care.

We presented all available information on patient, tumor, and treatment characteristics to the readers, to support their ability to self-assess the transferability. The findings in Study I may be less generalizable to countries that do not have tax-funded health care, where the influence of SEP may be stronger. The findings of studies II–III mainly apply to populations of Caucasian premenopausal women with early breast cancer.

Although docetaxel pharmacokinetics is similar across ethnicities, apart from Japanese patients, there might be biological differences in pharmacogenetics. Allele frequencies vary across ethnicities for some pharmacogenetic variants, so our absolute risks may not necessarily apply to all populations.[165] The biologic consequence of carrying a reduced or enhanced function variant is likely to be the same regardless of race/ethnicity and the prevalence of the reduced function variant in different groups. The generalizability and population impact should be considered in other large population-based studies.

Main conclusions

Based on our studies and a careful evaluation and discussion of previous literature, potential biases, and confounders, we draw the following conclusions.

In Study I, we combined individual-level, population-based registry data related to the women's breast cancer and their clinical course, to information indicating the women's SEP. We found that social inequality in breast cancer prognosis prevailed even in a country with tax-funded health care. Premenopausal women with breast cancer with low SEP—as indicated by being single, having low income, low education, being unemployed or having health-related absenteeism from work—had higher mortality. It is noteworthy that the risk of recurrence only seemed increased in single women. We expected that mortality in these young women was caused by recurrence. This suggests that recurrences are less frequently detected in women with low SEP.

In Studies II and III we combined routinely archived tumor tissue and individual-level registry data in a population-based setting. Study II adds to the evidence of the influence of SNPs on inter-individual variation in taxane effectiveness. *SLCO1B1* rs2306283 and *GSTP1* rs1138272 have previously been investigated for their impact on taxane-induced toxicity, but not in relation to prognostic outcomes in taxane-treated breast cancer. Variant allele carriers of *SLCO1B1* rs2306283 had a better prognosis in terms of lower recurrence and mortality incidence rates and risks, compared with wildtypes. In contrast, *GSTP1* rs1138272 variant carriers had poorer prognosis. In Study III, *CYP3A5* rs776746 homozygotes had delayed return-to-work and stable labour market attachment, indicating a poorer recovery in these women. The underlying mechanisms of our findings in Studies II and III have not yet been elucidated but may rely on altered pharmacokinetics.

In summary, this thesis provides novel insights on inter-individual variation in prognosis in premenopausal women with non-distant metastatic breast cancer after taxane-based combination therapy. Our research suggests that survivorship may be influenced by SEP and inherited SNPs. Our findings emphasize that the inter-individual variation is multifactorial and may not be solved easily.

Outlook

We mostly examined non-modifiable exposures, but this does not exclude the possibility of meaningful interventions improving survivorship after taxane-based chemotherapy. SEP indicators, our exposures in Study I, are frequently non-modifiable, especially core SEP. SNPs are also non-modifiable.

The build-up of the welfare states aimed to reduce socioeconomic inequality in health, but this goal has not been universally achieved.[217] A general misunderstanding is that social inequality emerges from how service users (patients) act within society and in the health care system. Alternatively, social inequality follows how the society and health care system manage to include all groups within the society without distinctions. Therefore, the solution may rely on changes in how the health care system favors certain social classes. During the study period, all the women were enrolled in a 10-year follow-up program. Since 2016, routine follow-up after breast cancer has been altered. Breast cancer survivors are now allocated a physician-, patient- or nurse-led follow-up program. Saltbæk *et al.*[149] studied the mode of recurrence detection in 310 women enrolled in the 10-year clinical follow-up program offered before 2016. The study showed that only one out of six recurrences were detected at scheduled outpatient visits, with the remainder detected through patient or general practitioner requested examinations.[149] In that study, five out of 42 recurrences detected at scheduled outpatient visits were asymptomatic. However, the study did not take SEP into account. Our findings suggest underdetection of recurrences in certain SEP groups. There is a negative socioeconomic gradient in health care seeking behavior and cancer symptom awareness.[218,219] Women with low SEP may therefore benefit from allocation to a physician- or nurse-led follow-up program, rather than patient-led. Moreover, we found greater inequality in women undergoing tamoxifen. Ongoing research in our group is investigating the impact of SEP on tamoxifen adherence. Future research should focus on the underlying mechanisms, and potential interventions that could support tamoxifen adherence. Given our findings, this may be particularly important to improve outcomes in single women.

The overarching aim of pharmacokinetic- and genomic studies on SNPs and taxane outcomes may be to identify biomarkers that can be used to personalize treatment and care, and in this way reduce the inter-individual variability seen in breast cancer outcomes. Our studies of the association of SNPs and breast cancer recurrence and return-to-work do not disentangle whether and how SNPs can be used to personalize taxane-combination therapy. Yet, they suggest that the clinical relevance of some previously reported associations should be questioned, or at least not be generalized to all populations. Nonetheless, we identified two SNPs that warrant more research. In future studies, we will extend our examination of breast cancer recurrence and mortality (Study II) with haplotype- and pathway studies,[169,220] to further elucidate potential mechanisms. Potential genetic testing in guiding chemotherapy management would require clinical trial assessment. In future studies, we also aim to investigate the influence of pharmacokinetic drug interactions by co-medications.[69]

It is essential to examine the mechanisms underlying our findings in Study III to inform potential interventions. The findings *indicate* poorer recovery in women with SNPs in *CYP3A5*. However, whether this is due to physical adverse effects and how this association is impacted by other factors such as dose reductions or drug interactions should be determined. Research shows a considerable willingness in Danish breast cancer patients to reduce their income by taking a flexible job or by working reduced hours.[221,222] If certain groups risking delayed or absent return-to-work can be identified, interventions favoring altered job demands should therefore be investigated.

Summary

Female breast cancer is the most frequent cancer with more than 2.3 million cases diagnosed annually worldwide.[223] Premenopausal breast cancer constitutes one-third of all breast cancer diagnoses.[224] Most of these women are recommended taxane-based chemotherapy. The efficacy of taxanes is high,[59,225] but women with seemingly identical clinical characteristics at diagnosis can vary substantially in their survivorship during and after taxane-based chemotherapy. Some experience breast cancer recurrence and/or premature death. Moreover, some women suffer from disabling side effects, such as neuropathies, delaying their return to everyday activities, such as work, and impeding quality-of-life. This constitutes a major clinical issue, as clinicians are not able to predict who will benefit less from the treatment and consequently have a poor prognosis.

Mechanisms underlying the variation in treatment response may be multifactorial. Socioeconomic position (SEP) influences breast cancer prognosis in general, but whether this applies in premenopausal women undergoing guideline treatment has not been investigated. Moreover, various studies examined the influence of genetic variation in genes related to taxane metabolism, transport, or neuropathies. The findings have been inconclusive, likely due to low sample size and incomparable (selected) populations.

In this thesis, the overarching aim was to investigate both socioeconomic- and genetic factors associated with 10–year breast cancer survivorship after taxane-based chemotherapy, in premenopausal women in the Predictors of Breast Cancer Recurrence (ProBe CaRe) cohort, who were recommended docetaxel-based chemotherapy.

In Study I, we examined the influence of SEP on breast cancer recurrence and mortality. We found that single women had increased risk of recurrence and mortality, especially if they also underwent tamoxifen treatment (and hence had estrogen receptor positive disease). Mortality was increased in women with low income, low education or workforce detachment prior to breast cancer diagnosis.

In Study II, we examined the influence of 26 single nucleotide polymorphisms (SNPs) on breast cancer recurrence and mortality. We reported that women carrying variant alleles in *GSTP1* rs1138272 had increased risk of breast cancer recurrence and mortality. In contrast, variant carriers of *SLCO1B1* rs2306283 had lower rates of recurrence and mortality. These findings were likely explained by altered docetaxel pharmacokinetics.

In Study III, we examined the influence of 26 SNPs on return-to-work. We observed that women who were homozygous for *CYP3A5* rs776746 had delayed return-to-work, compared to wildtypes. These findings were possibly explained by increased toxicities reducing the women's workability.

Dansk resumé

Brystkræft hos kvinder er den hyppigste kræftform med mere end 2,3 millioner tilfælde diagnosticeret årligt på verdensplan.[223] Præmenopausal brystkræft (før overgangsalderen) udgør en tredjedel af alle brystkræftdiagnoser.[224] De fleste af disse kvinder anbefales taxan-baseret kemoterapi. Effektiviteten af taxaner er høj,[59,225] men kvinder med tilsyneladende identiske kliniske karakteristika ved diagnosen kan variere betydeligt i deres sygdomsforløb. Nogle oplever at deres brystkræft vender tilbage (tilbagefald) og/eller dør for tidligt. Desuden lider nogle kvinder af invaliderende bivirkninger, såsom neuropatier. Sådanne bivirkninger kan forsinke deres tilbagevenden til hverdagslivet, såsom at gå på arbejde og dette kan have negativ indflydelse på deres livskvalitet. Dette er et stort problem, da klinikere på nuværende tidspunkt ikke kan forudsige hvem der vil have mindre gavn af behandlingen og som følge heraf har en dårlig prognose.

Variationen i behandlingsrespons er formentlig forårsaget af multiple faktorer. Lavere socioøkonomisk position ses at være associeret med dårligere prognose efter brystkræft og kræft generelt, men hvorvidt dette gælder for præmenopausale kvinder, der gennemgår anbefalet behandling, er ikke undersøgt. Desuden er nuværende forskning af indflydelsen af genetisk variation i gener relateret til taxanmetabolisme, transport eller neuropatier. Undersøgelserne har været tvetydige, sandsynligvis på grund af begrænsninger i den undersøgte populationers størrelser og sammenlignelighed.

I denne afhandling er det overordnede formål at undersøge både socioøkonomiske og genetiske faktorer forbundet med 10-års brystkræftoverlevelse efter taxan-baseret kemoterapi, hos præmenopausale kvinder i Predictors of Breast Cancer Recurrence (ProBe CaRe) kohorten.

I Studie I undersøger vi indflydelsen af socioøkonomisk position på tilbagefald og dødelighed. SAmmenligned med kvinder med en partner, havde enlige kvinder øget risiko for tilbagefald og en højere dødelighed, især hvis de også modtog hormonbehandling med tamoxifen (og dermed havde østrogenreceptor positiv kræft). Dødeligheden var ligeledes øget hos kvinder der havde lav indkomst, lav uddannelse eller som var uden for arbejdsmarkedet før de blev diagnosticeret med brystkræft.

I Studie II undersøgte vi indflydelsen af 26 enkelt-nukleotid polymorfier—på engelsk *single nucleotide polymorphisms* (SNP)—på tilbagefald og dødelighed. Vi fandt at kvinder der bærer variante alleler i *SLCO1B1* rs2306283 havde lavere risiko for tilbagefald og dødelighed. Det omvendte var gældende for *GSTP1* rs1138272. Disse fund kan være forårsaget af ændret farmakokinetik.

I Studie III undersøgte vi de samme SNPs og deres sammenhæng med tilbagevenden til arbejde. Vi observerede, at kvindelige homozygoter for *CYP3A5* rs776746 var senere til at vende tilbage til arbejde, sammenlignet med kvinder med to normale alleler. Dette indikerer at kvindernes arbejdsdygtighed kan være negativt påvirket efter deres kræftforløb, muligvis på grund af bivirkninger.

References

1. Rodriguez MA, Lewis-Patterson P. *Handbook of Cancer Survivorship Care, Chapter 1: Defining Cancer Survivorship*. Springer Publishing Company; 2021. Accessed March 28, 2022. https://connect.springerpub.com/content/book/978-0-8261-3825-5/part/part01/chapter/ch01

2. Noguchi S. Predictive factors for response to docetaxel in human breast cancers. *Cancer Sci*. 2006;97(9):813-820.

3. Dalton SO, Ross L, Düring M, et al. Influence of socioeconomic factors on survival after breast cancer--a nationwide cohort study of women diagnosed with breast cancer in Denmark 1983-1999. *Int J Cancer*. 2007;121(11):2524-2531.

4. Dalton SO, Steding-Jessen M, Gislum M, Frederiksen K, Engholm G, Schüz J. Social inequality and incidence of and survival from cancer in a population-based study in Denmark, 1994-2003: Background, aims, material and methods. *Eur J Cancer Oxf Engl 1990*. 2008;44(14):1938-1949.

5. Singh GK, Jemal A. Socioeconomic and Racial/Ethnic Disparities in Cancer Mortality, Incidence, and Survival in the United States, 1950–2014: Over Six Decades of Changing Patterns and Widening Inequalities. *J Environ Public Health*. 2017;2017:e2819372.

6. Hu K, Ding P, Wu Y, Tian W, Pan T, Zhang S. Global patterns and trends in the breast cancer incidence and mortality according to sociodemographic indices: an observational study based on the global burden of diseases. *BMJ Open*. 2019;9(10):e028461.

7. Lundqvist A, Andersson E, Ahlberg I, Nilbert M, Gerdtham U. Socioeconomic inequalities in breast cancer incidence and mortality in Europe—a systematic review and meta-analysis. *Eur J Public Health*. 2016;26(5):804-813.

8. Frederiks CN, Lam SW, Guchelaar HJ, Boven E. Genetic polymorphisms and paclitaxel- or docetaxel-induced toxicities: A systematic review. *Cancer Treat Rev*. 2015;41(10):935-950.

9. Islam T, Dahlui M, Majid HA, Nahar AM, Mohd Taib NA, Su TT. Factors associated with return to work of breast cancer survivors: a systematic review. *BMC Public Health*. 2014;14(3):S8.

10. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71(3):209-249.

11. World Health Organization. Breast cancer. Accessed January 4, 2022. https://www.who.int/news-room/fact-sheets/detail/breast-cancer

12. Kadakia KC, Henry NL. Adjuvant endocrine therapy in premenopausal women with breast cancer. *Clin Adv Hematol Oncol HO*. 2015;13(10):663-672.

13. Menvielle G, Leclerc A, Chastang JF, Luce D. Social inequalities in breast cancer mortality among French women: disappearing educational disparities from 1968 to 1996. *Br J Cancer*. 2006;94(1):152-155.

14. Premenopausal Breast Cancer Collaborative Group, Schoemaker MJ, Nichols HB, et al. Association of Body Mass Index and Age With Subsequent Breast Cancer Risk in Premenopausal Women. *JAMA Oncol*. 2018;4(11):e181771.

15. Sonnenschein E, Toniolo P, Terry MB, et al. Body fat distribution and obesity in pre- and postmenopausal breast cancer. *Int J Epidemiol*. 1999;28(6):1026-1031.

16. Peacock SL, White E, Daling JR, Voigt LF, Malone KE. Relation between obesity and breast cancer in young women. *Am J Epidemiol*. 1999;149(4):339-346.

17. Breast Cancer Association Consortium, Dorling L, Carvalho S, et al. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med*. 2021;384(5):428-439.

18. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *JAMA*. 2017;317(23):2402-2416.

19. Hu C, Hart SN, Gnanaolivu R, et al. A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N Engl J Med*. 2021;384(5):440-451.

20. Trewin CB, Strand BH, Weedon-Fekjær H, Ursin G. Changing patterns of breast cancer incidence and mortality by education level over four decades in Norway, 1971-2009. *Eur J Public Health*. 2017;27(1):160-166.

21. American Cancer Society. Cancer Facts & Figures 2021. Published online 2021. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2021/cancer-facts-and-figures-2021.pdf

22. National Cancer Institute. Cancer Stat Facts: Female Breast Cancer. SEER. Accessed February 15, 2019. https://seer.cancer.gov/statfacts/html/breast.html

23. Ahmad A. Pathways to breast cancer recurrence. *Int Sch Res Not*. 2013;2013.

24. Lord SJ, Marinovich ML, Patterson JA, et al. Incidence of metastatic breast cancer in an Australian population-based cohort of women with non-metastatic breast cancer at diagnosis. *Med J Aust*. 2012;196(11):688-692.

25. Nelson HD, Zakher B, Cantor A, et al. Risk Factors for Breast Cancer for Women Aged 40 to 49 Years. *Ann Intern Med*. 2012;156(9):635-648.

26. Chollet-Hinton L, Anders CK, Tse CK, et al. Breast cancer biologic and etiologic heterogeneity by young age and menopausal status in the Carolina Breast Cancer Study: a case-control study. *Breast Cancer Res*. 2016;18(1):79.

27. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A. Breast cancer before age 40 years. *Semin Oncol*. 2009;36(3):237-249.

28. Azim HA, Partridge AH. Biology of breast cancer in young women. *Breast Cancer Res BCR*. 2014;16(4):427.

29. Kroman N, Jensen MB, Wohlfahrt J, Mouridsen HT, Andersen PK, Melbye M. Factors influencing the effect of age on prognosis in breast cancer: population based study. *BMJ*. 2000;320(7233):474-478.

30. Ong CT, Thomas SM, Blitzblau RC, et al. Patient Age and Tumor Subtype Predict the Extent of Axillary Surgery Among Breast Cancer Patients Eligible for the American College of Surgeons Oncology Group Trial Z0011. *Ann Surg Oncol*. 2017;24(12):3559-3566.

31. Li CI, Anderson BO, Daling JR, Moe RE. Trends in incidence rates of invasive lobular and ductal breast carcinoma. *JAMA*. 2003;289(11):1421-1424.

32. Møller Talman ML, Jensen MB, Rank F. Invasive lobular breast cancer. Prognostic significance of histological malignancy grading. *Acta Oncol*. 2007;46(6):803-809.

33. Hammond MEH, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Immunohistochemical Testing of Estrogen and Progesterone Receptors in Breast Cancer. *J Clin Oncol*. 2010;28(16):2784-2795.

34. DMCG. Pathology procedures and molecular analyses in breast cancer [In Danish: Patologiprocedurer og molekylærpatologiske analyser ved brystkræft]. Published 2021. Accessed March 7, 2022. https://www.dmcg.dk/Kliniske-retningslinjer/kliniske-retningslinjer-opdelt-paa-dmcg/brystcancer/patologiprocedurer-og-molekylarpatologiske-analyser-ved-brystkraft/

35. Wolff AC, Hammond MEH, Allison KH, et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *Arch Pathol Lab Med*. 2018;142(11):1364-1382.

36. Amin MB, American Joint Committee on Cancer, American Cancer Society, eds. *AJCC Cancer Staging Manual*. Eight edition / editor-in-chief, Mahul B. Amin, MD, FCAP ; editors, Stephen B. Edge, MD, FACS [and 16 others] ; Donna M. Gress, RHIT, CTR-Technical editor ; Laura R. Meyer, CAPM-Managing editor. American Joint Committee on Cancer, Springer; 2017.

37. Ruhl J, Callaghan C, Schussler Ni. Summary stage 2018: Codes and Coding instructions. *National Cancer Institute*. Published online 2021:375.

38. Pedersen RN, Esen BÖ, Mellemkjær L, et al. The Incidence of Breast Cancer Recurrence 10-32 Years after Primary Diagnosis. *J Natl Cancer Inst*. Published online November 8, 2021:djab202.

39. Park SY, Nam JS. The force awakens: metastatic dormant cancer cells. *Exp Mol Med*. 2020;52(4):569-581.

40. Riggio AI, Varley KE, Welm AL. The lingering mysteries of metastatic recurrence in breast cancer. *Br J Cancer*. 2021;124(1):13-26.

41. Rightslink® by Copyright Clearance Center. Accessed March 15, 2022. https://s100.copyright.com/AppDispatchServlet?title=The%20lingering%20mysteries%20of%20metastatic%20recurrence%20in%20breast%20cancer&author=Alessandra%20I.%20Riggio%20et%20al&contentID=10.1038%2Fs41416-020-01161-4&copyright=The%20Author%28s%29&publication=0007-0920&publicationDate=2020-11-26&publisherName=SpringerNature&orderBeanReset=true&oa=CC%20BY

42. Porro B, Michel A, Zinzindohoué C, et al. Quality of life, fatigue and changes therein as predictors of return to work during breast cancer treatment. *Scand J Caring Sci*. 2019;33(2):467-477.

43. van Egmond MP, Duijts SFA, Loyen A, Vermeulen SJ, van der Beek AJ, Anema JR. Barriers and facilitators for return to work in cancer survivors with job loss experience: a focus group study. *Eur J Cancer Care (Engl)*. 2017;26(5).

44. Heinesen E, Kolodziejczyk C. Effects of breast and colorectal cancer on labour market outcomes-average effects and educational gradients. *J Health Econ*. 2013;32(6):1028-1042.

45. Carlsen K, Ewertz M, Dalton SO, Badsberg JH, Osler M. Unemployment among breast cancer survivors. *Scand J Public Health*. 2014;42(3):319-328.

46. Mehnert A. Employment and work-related issues in cancer survivors. *Crit Rev Oncol Hematol*. 2011;77(2):109-130.

47. Hansen JA, Feuerstein M, Calvio LC, Olsen CH. Breast cancer survivors at work. *J Occup Environ Med*. 2008;50(7):777-784.

48. de Boer AGEM, Taskila T, Ojajärvi A, van Dijk FJH, Verbeek JHAM. Cancer survivors and unemployment: a meta-analysis and meta-regression. *JAMA*. 2009;301(7):753-762.

49. Duijts SFA, van Egmond MP, Gits M, van der Beek AJ, Bleiker EM. Cancer survivors' perspectives and experiences regarding behavioral determinants of return to work and continuation of work. *Disabil Rehabil*. 2017;39(21):2164-2172.

50. Feuerstein M, Todd BL, Moskowitz MC, et al. Work in cancer survivors: a model for practice and research. *J Cancer Surviv*. 2010;4(4):415-437.

51. Ganz PA. Survivorship: adult cancer survivors. *Prim Care*. 2009;36(4):721-741.

52. Schwentner L, Harbeck N, Singer S, et al. Short term quality of life with epirubicin-fluorouracil-cyclophosphamid (FEC) and sequential epirubicin/cyclophosphamid-docetaxel (EC-DOC) chemotherapy in patients with primary breast cancer - Results from the prospective multi-center randomized ADEBAR trial. *Breast Edinb Scotl*. 2016;27:69-77.

53. Forero A, Lyons J, Marcom PK, et al. NCCN Guidelines Panel Disclosures. *Breast Cancer*. Published online 2018:209.

54. Ahern TP, Larsson H, Garne JP, Cronin-Fenton DP, Sørensen HT, Lash TL. Trends in breast-conserving surgery in Denmark, 1982–2002. *Eur J Epidemiol*. 2008;23(2):109-114.

55. DMCG. Surgical treatment of breast cancer [In Danish: Kirurgisk behandling af brystkræft]. Published 2021. Accessed February 17, 2022. https://www.dmcg.dk/Kliniske-retningslinjer/kliniske-retningslinjer-opdelt-paa-dmcg/brystcancer/kirurgisk-behandling-af-brystkraft/

56. DMCG. Post-surgical radiotherapy in breast cancer [In Danish: Postoperativ strålebehandling af brystkræft]. Published 2021. Accessed February 17, 2022. https://www.dmcg.dk/Kliniske-retningslinjer/kliniske-retningslinjer-opdelt-paa-dmcg/brystcancer/postoperativ-stralebehandling-af-brystkraft/

57. Jensen MB, Laenkholm AV, Offersen BV, et al. The clinical database and implementation of treatment guidelines by the Danish Breast Cancer Cooperative Group in 2007–2016. *Acta Oncol*. 2018;57(1):13-18.

58. DMCG. Systemic breast cancer therapy - II – (neo)adjuvant systemic therapy of early breast cancer. [In Danish: Systemisk behandling af brystkræft - II – (neo)adjuverende systemisk behandling af tidlig brystkræft]. Published 2021. Accessed February 18, 2022. https://www.dmcg.dk/Kliniske-retningslinjer/kliniske-retningslinjer-opdelt-paa-dmcg/brystcancer/systemisk-behandling-af-brystkraft---ii--neoadjuverende-systemisk-behandling-af-tidlig-brystkraft/

59. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *The Lancet*. 2012;379(9814):432-444.

60. Ejlertsen B, Tuxen MK, Jakobsen EH, et al. Adjuvant Cyclophosphamide and Docetaxel With or Without Epirubicin for Early TOP2A-Normal Breast Cancer: DBCG 07-READ, an Open-Label, Phase III, Randomized Trial. *J Clin Oncol Off J Am Soc Clin Oncol*. 2017;35(23):2639-2646.

61. Danish Breast Cancer Group. Medical treatment (Danish: Medicinsk behandling). Danish Breast Cancer Group. Published 2017. Accessed April 11, 2018. http://www.dbcg.dk/PDF%20Filer/Kap_6_Medicinsk_behandling-07.04.2017.pdf

62. Chen S, Sutiman N, Zhang CZ, et al. Pharmacogenetics of irinotecan, doxorubicin and docetaxel transporters in Asian and Caucasian cancer patients: a comparative review. *Drug Metab Rev*. 2016;48(4):502-540.

63. Fletcher RH, Fletcher SW, Fletcher GS. *Clinical Epidemiology : The Essentials*. 5th ed. Wolters Kluwer Health/Lippincott Williams & Wilkins; 2014.

64. Maloney SM, Hoover CA, Morejon-Lasso LV, Prosperi JR. Mechanisms of Taxane Resistance. *Cancers*. 2020;12(11):3323.

65. Ferguson T, Wilcken N, Vagg R, Ghersi D, Nowak AK. Taxanes for adjuvant treatment of early breast cancer. *Cochrane Database Syst Rev*. 2007;(4):CD004421.

66. Eckhoff L, Nielsen M, Moeller S, Knoop A. TAXTOX – a retrospective study regarding the side effects of docetaxel given as part of the adjuvant treatment to patients with primary breast cancer in Denmark from 2007 to 2009. *Acta Oncol*. 2011;50(7):1075-1082.

67. Di Francia R, Atripaldi L, Di Martino S, et al. Assessment of Pharmacogenomic Panel Assay for Prediction of Taxane Toxicities: Preliminary Results. *Front Pharmacol*. 2017;8:797.

68. Jabir RS, Naidu R, Annuar MABA, Ho GF, Munisamy M, Stanslas J. Pharmacogenetics of taxanes: impact of gene polymorphisms of drug transporters on pharmacokinetics and toxicity. *Pharmacogenomics*. 2012;13(16):1979-1988.

69. Nieuweboer AJM, de Morrée ES, de Graan AJM, Sparreboom A, de Wit R, Mathijssen RHJ. Inter-patient variability in docetaxel pharmacokinetics: A review. *Cancer Treat Rev*. 2015;41(7):605-613.

70. Oshiro C, Marsh S, McLeod H, Carrillo MW, Klein T, Altman R. Taxane pathway. *Pharmacogenet Genomics*. 2009;19(12):979-983.

71. Argyriou AA, Bruna J, Marmiroli P, Cavaletti G. Chemotherapy-induced peripheral neurotoxicity (CIPN): An update. *Crit Rev Oncol Hematol*. 2012;82(1):51-77.

72. Eckhoff L, Knoop A, Jensen MB, Ewertz M. Persistence of docetaxel-induced neuropathy and impact on quality of life among breast cancer survivors. *Eur J Cancer Oxf Engl 1990*. 2015;51(3):292-300.

73. Mols F, Beijers T, Vreugdenhil G, van de Poll-Franse L. Chemotherapy-induced peripheral neuropathy and its association with quality of life: a systematic review. *Support Care Cancer Off J Multinatl Assoc Support Care Cancer*. 2014;22(8):2261-2269.

74. Baldwin RM, Owzar K, Zembutsu H, et al. A genome-wide association study identifies novel loci for paclitaxel-induced sensory peripheral neuropathy in CALGB 40101. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2012;18(18):5099-5109.

75. Abraham JE, Guo Q, Dorling L, et al. Replication of genetic polymorphisms reported to be associated with taxane-related sensory neuropathy in patients with early breast cancer treated with Paclitaxel. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2014;20(9):2466-2475.

76. Leandro-García LJ, Inglada-Pérez L, Pita G, et al. Genome-wide association study identifies ephrin type A receptors implicated in paclitaxel induced peripheral sensory neuropathy. *J Med Genet*. 2013;50(9):599-605.

77. Bosó V, Herrero MJ, Santaballa A, et al. SNPs and taxane toxicity in breast cancer patients. *Pharmacogenomics*. 2014;15(15):1845-1858.

78. da Costa R, Passos GF, Quintão NLM, et al. Taxane-induced neurotoxicity: Pathophysiology and therapeutic perspectives. *Br J Pharmacol*. 2020;177(14):3127-3146.

79. Boora GK, Kulkarni AA, Kanwar R, et al. Association of the Charcot–Marie–Tooth disease gene ARHGEF10 with paclitaxel induced peripheral neuropathy in NCCTG N08CA (Alliance). *J Neurol Sci*. 2015;357(1):35-40.

80. Beutler AS, Kulkarni AA, Kanwar R, et al. Sequencing of Charcot-Marie-Tooth disease genes in a toxic polyneuropathy. *Ann Neurol*. 2014;76(5):727-737.

81. Lam SW, Frederiks CN, van der Straaten T, Honkoop AH, Guchelaar HJ, Boven E. Genotypes of CYP2C8 and FGD4 and their association with peripheral neuropathy or early dose reduction in paclitaxel-treated breast cancer patients. *Br J Cancer*. 2016;115(11):1335-1342.

82. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health*. 1997;18:341-378.

83. Green MJ, Popham F. Interpreting mutual adjustment for multiple indicators of socioeconomic position without committing mutual adjustment fallacies. *BMC Public Health*. 2019;19.

84. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health*. 2006;60(1):7-12.

85. Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health*. 2012;12(1):80.

86. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 2). *J Epidemiol Community Health*. 2006;60(2):95-101.

87. Frisch M, Simonsen J. Marriage, cohabitation and mortality in Denmark: national cohort study of 6.5 million persons followed for up to three decades (1982–2011). *Int J Epidemiol*. 2013;42(2):559-578.

88. Ecker ED, Skelly AC. Conducting a winning literature search. *Evid-Based Spine-Care J*. 2010;1(1):9-14.

89. Li W, Zhang D, Du F, et al. ABCB1 3435TT and ABCG2 421CC genotypes were significantly associated with longer progression-free survival in Chinese breast cancer patients. *Oncotarget*. 2017;8(67).

90. Rasmussen LA, Jensen H, Virgilsen LF, Falborg AZ, Møller H, Vedsted P. Time from incident primary cancer until recurrence or second primary cancer: Risk factors and impact in general practice. *Eur J Cancer Care (Engl)*. 2019;28(5):e13123.

91. Di Salvo F, Caranci N, Spadea T, et al. Socioeconomic deprivation worsens the outcomes of Italian women with hormone receptor-positive breast cancer and decreases the possibility of receiving standard care. *Oncotarget*. 2017;8(40):68402-68414.

92. Lagerlund M, Bellocco R, Karlsson P, Tejler G, Lambe M. Socio-economic factors and breast cancer survival--a population-based cohort study (Sweden). *Cancer Causes Control CCC*. 2005;16(4):419-430.

93. Larsen SB, Kroman N, Ibfelt EH, Christensen J, Tjønneland A, Dalton SO. Influence of metabolic indicators, smoking, alcohol and socioeconomic position on mortality after breast cancer. *Acta Oncol*. 2015;54(5):780-788.

94. Rutqvist LE, Bern A, Stockholm Breast Cancer Study Group. Socioeconomic gradients in clinical stage at presentation and survival among breast cancer patients in the Stockholm area 1977-1997. *Int J Cancer*. 2006;119(6):1433-1439.

95. Trewin CB, Johansson ALV, Hjerkind KV, Strand BH, Kiserud CE, Ursin G. Stage-specific survival has improved for young breast cancer patients since 2000: but not equally. *Breast Cancer Res Treat*. 2020;182(2):477-489.

96. Elstad JI, Torstensrud R, Lyngstad TH, Kravdal O. Trends in educational inequalities in mortality, seven types of cancers, Norway 1971-2002. *Eur J Public Health*. 2012;22(6):771-776.

97. Herndon JE 2nd, Kornblith AB, Holland JC, Paskett ED. Effect of socioeconomic status as measured by education level on survival in breast cancer clinical trials. *Psychooncology*. 2013;22(2):315-323.

98. Abdoli G, Bottai M, Sandelin K, Moradi T. Breast cancer diagnosis and mortality by tumor stage and migration background in a nationwide cohort study in Sweden. *Breast Edinb Scotl*. 2017;31:57-65.

99. Hussain SK, Altieri A, Sundquist J, Hemminki K. Influence of education level on breast cancer risk and survival in Sweden between 1990 and 2004. *Int J Cancer*. 2008;122(1):165-169.

100. Aizer AA, Chen MH, McCarthy EP, et al. Marital Status and Survival in Patients With Cancer. *J Clin Oncol*. 2013;31(31):3869-3876.

101. Dasgupta P, Turrell G, Aitken JF, Baade PD. Partner status and survival after cancer: A competing risks analysis. *Cancer Epidemiol*. 2016;41:16-23.

102. Martínez ME, Unkart JT, Tao L, et al. Prognostic significance of marital status in breast cancer survival: A population-based study. *PLOS ONE*. 2017;12(5):e0175515.

103. Hinyard L, Wirth LS, Clancy JM, Schwartz T. The effect of marital status on breast cancer-related outcomes in women under 65: A SEER database analysis. *Breast Edinb Scotl*. 2017;32:13-17.

104. Gentil-Brevet J, Colonna M, Danzon A, et al. The influence of socio-economic and surveillance characteristics on breast cancer survival: a French population-based study. *Br J Cancer*. 2008;98(1):217-224.

105. Bouchardy C, Verkooijen HM, Fioretta G. Social class is an important and independent prognostic factor of breast cancer mortality. *Int J Cancer*. 2006;119(5):1145-1151.

106. Beiki O, Hall P, Ekbom A, Moradi T. Breast cancer incidence and case fatality among 4.7 million women in relation to social and ethnic background: a population-based cohort study. *Breast Cancer Res BCR*. 2012;14(1):R5.

107. Elstad JI, Torstensrud R, Lyngstad TH, Kravdal O. Trends in educational inequalities in mortality, seven types of cancers, Norway 1971-2002. *Eur J Public Health*. 2012;22(6):771-776.

108. Herndon JE, Kornblith AB, Holland JC, Paskett ED. Effect of socioeconomic status as measured by education level on survival in breast cancer clinical trials. *Psychooncology*. 2013;22(2):315-323.

109. van Maaren MC, Rachet B, Sonke GS, et al. Socioeconomic status and its relation with breast cancer recurrence and survival in young women in the Netherlands. *Cancer Epidemiol*. 2022;77:102118.

110. Carlsen K, Høybye MT, Dalton SO, Tjønneland A. Social inequality and incidence of and survival from breast cancer in a population-based study in Denmark, 1994–2003. *Eur J Cancer*. 2008;44(14):1996-2002.

111. Lundqvist A, Andersson E, Ahlberg I, Nilbert M, Gerdtham U. Socioeconomic inequalities in breast cancer incidence and mortality in Europe-a systematic review and meta-analysis. *Eur J Public Health*. 2016;26(5):804-813.

112. Trewin CB, Strand BH, Weedon-Fekjær H, Ursin G. Changing patterns of breast cancer incidence and mortality by education level over four decades in Norway, 1971-2009. *Eur J Public Health*. 2017;27(1):160-166.

113. Trewin CB, Johansson ALV, Hjerkind KV, Strand BH, Kiserud CE, Ursin G. Stage-specific survival has improved for young breast cancer patients since 2000: but not equally. *Breast Cancer Res Treat*. 2020;182(2):477-489.

114. Larsen SB, Kroman N, Ibfelt EH, Christensen J, Tjønneland A, Dalton SO. Influence of metabolic indicators, smoking, alcohol and socioeconomic position on mortality after breast cancer. *Acta Oncol Stockh Swed*. 2015;54(5):780-788.

115. Dalton SO, Ross L, Düring M, et al. Influence of socioeconomic factors on survival after breast cancer--a nationwide cohort study of women diagnosed with breast cancer in Denmark 1983-1999. *Int J Cancer*. 2007;121(11):2524-2531.

116. Abdul Aziz AA, Md Salleh MS, Yahya MM, Zakaria AD, Ankathil R. Genetic Association of CYP1B1 4326 C>G Polymorphism with Disease-Free Survival in TNBC Patients Undergoing TAC Chemotherapy Regimen. *Asian Pac J Cancer Prev APJCP*. 2021;22(4):1319-1324.

117. Chang H, Rha SY, Jeung HC, et al. Association of the ABCB1 gene polymorphisms 2677G>T/A and 3435C>T with clinical outcomes of paclitaxel monotherapy in metastatic breast cancer patients. *Ann Oncol Off J Eur Soc Med Oncol*. 2009;20(2):272-277.

118. Dong N, Yu J, Wang C, et al. Pharmacogenetic assessment of clinical outcome in patients with metastatic breast cancer treated with docetaxel plus capecitabine. *J Cancer Res Clin Oncol*. 2012;138(7):1197-1203.

119. Marsh S, Somlo G, Li X, et al. Pharmacogenetic analysis of paclitaxel transport and metabolism genes in breast cancer. *Pharmacogenomics J*. 2007;7(5):362-365.

120. Wang J, Wang T, Yin GY, Yang L, Wang ZG, Bu XB. Glutathione S-transferase polymorphisms influence chemotherapy response and treatment outcome in breast cancer. *Genet Mol Res GMR*. 2015;14(3):11126-11132.

121. Kim HJ, Im SA, Keam B, et al. ABCB1 polymorphism as prognostic factor in breast cancer patients treated with docetaxel and doxorubicin neoadjuvant chemotherapy. *Cancer Sci*. 2015;106(1):86-93.

122. Zhou X, Qiao G, Wang X, et al. CYP1A1 genetic polymorphism is a promising predictor to improve chemotherapy effects in patients with metastatic breast cancer treated with docetaxel plus thiotepa vs. docetaxel plus capecitabine. *Cancer Chemother Pharmacol*. 2018;81(2):365-372.

123. Dong N, Yu J, Wang C, et al. Pharmacogenetic assessment of clinical outcome in patients with metastatic breast cancer treated with docetaxel plus capecitabine. *J Cancer Res Clin Oncol*. 2012;138(7):1197-1203.

124. Kehl KL, Schrag D, Hassett MJ, Uno H. Assessment of Temporal Selection Bias in Genomic Testing in a Cohort of Patients With Cancer. *JAMA Netw Open*. 2020;3(6):e206976.

125. Cocchiara RA, Sciarra I, D'Egidio V, et al. Returning to work after breast cancer: A systematic review of reviews. *Work*. 2018;61(3):463-476.

126. Sun Y, Shigaki CL, Armer JM. Return to work among breast cancer survivors: A literature review. *Support Care Cancer Off J Multinatl Assoc Support Care Cancer*. 2017;25(3):709-718.

127. Bijker R, Duijts SFA, Smith SN, de Wildt-Liesveld R, Anema JR, Regeer BJ. Functional Impairments and Work-Related Outcomes in Breast Cancer Survivors: A Systematic Review. *J Occup Rehabil*. 2018;28(3):429-451.

128. Wang L, Hong BY, Kennedy SA, et al. Predictors of Unemployment After Breast Cancer Surgery: A Systematic Review and Meta-Analysis of Observational Studies. *J Clin Oncol Off J Am Soc Clin Oncol*. 2018;36(18):1868-1879.

129. Schmidt ME, Scherer S, Wiskemann J, Steindorf K. Return to work after breast cancer: The role of treatment-related side effects and potential impact on quality of life. *Eur J Cancer Care (Engl)*. 2019;28(4):e13051.

130. Heuser C, Halbach S, Kowalski C, Enders A, Pfaff H, Ernstmann N. Sociodemographic and disease-related determinants of return to work among women with breast cancer: a German longitudinal cohort study. *BMC Health Serv Res*. 2018;18(1):1000.

131. Arndt V, Koch-Gallenkamp L, Bertram H, et al. Return to work after cancer. A multi-regional population-based study from Germany. *Acta Oncol Stockh Swed*. 2019;58(5):811-818.

132. Caumette E, Vaz-Luis I, Pinto S, et al. The Challenge of Return to Work after Breast Cancer: The Role of Family Situation, CANTO Cohort. *Curr Oncol Tor Ont*. 2021;28(5):3866-3875.

133. Arfi A, Baffert S, Soilly AL, et al. Determinants of return at work of breast cancer patients: results from the OPTISOINS01 French prospective study. *BMJ Open*. 2018;8(5):e020276.

134. Lee MK, Kang HS, Lee KS, Lee ES. Three-Year Prospective Cohort Study of Factors Associated with Return to Work After Breast Cancer Diagnosis. *J Occup Rehabil*. 2017;27(4):547-558.

135. Rosenberg SM, Vaz-Luis I, Gong J, et al. Employment trends in young women following a breast cancer diagnosis. *Breast Cancer Res Treat*. 2019;177(1):207-214.

136. Jensen LS, Overgaard C, Garne JP, Bøggild H, Fonager K. The impact of prior psychiatric medical treatment on return to work after a diagnosis of breast cancer: A registry based study. *Scand J Public Health*. Published online August 8, 2017:1403494817722291.

137. Dumas A, Vaz Luis I, Bovagnet T, et al. Impact of Breast Cancer Treatment on Employment: Results of a Multicenter Prospective Cohort Study (CANTO). *J Clin Oncol*. 2020;38(7):734-743.

138. Leuteritz K, Friedrich M, Sender A, et al. Return to Work and Employment Situation of Young Adult Cancer Survivors: Results from the Adolescent and Young Adult-Leipzig Study. *J Adolesc Young Adult Oncol*. 2021;10(2):226-233.

139. Monteiro I, Morais S, Costa AR, et al. Changes in employment status up to 5 years after breast cancer diagnosis: A prospective cohort study. *Breast*. 2019;48((Monteiro I.) Faculdade de Medicina da Universidade Do Porto, Portugal):38-44.

140. Rothman KJ, Lash T, Greenland S. *Modern Epidemiology*. Third edition. LWW; 2012.

141. Ekholm O, Gundgaard J, Rasmussen NKR, Hansen EH. The effect of health, socio-economic position, and mode of data collection on non-response in health interview surveys. *Scand J Public Health*. 2010;38(7):699-706.

142. Lallukka T, Pietiläinen O, Jäppinen S, Laaksonen M, Lahti J, Rahkonen O. Factors associated with health survey response among young employees: a register-based study using online, mailed and telephone interview data collection methods. *BMC Public Health*. 2020;20(1):184.

143. Islam T, Dahlui M, Majid HA, et al. Factors associated with return to work of breast cancer survivors: a systematic review. *BMC Public Health*. 2014;14 Suppl 3:S8.

144. Dumas A, Luis IV, Bovagnet T, et al. Impact of breast cancer treatment on employment: Results of a multicenter prospective cohort study (CANTO). *J Clin Oncol*. 2020;38(7):734-743.

145. Zhang J, Yu KF. What's the Relative Risk?A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes. *JAMA*. 1998;280(19):1690-1691.

146. Schmidt M, Schmidt SAJ, Adelborg K, et al. The Danish health care system and epidemiological research: from health care contacts to database records. *Clin Epidemiol*. 2019;11(11):563-591.

147. World Health Organization. Regional Office for Europe, Policies EO on HS and, Olejaz M, et al. *Denmark: Health System Review*. World Health Organization. Regional Office for Europe; 2012. Accessed February 12, 2022. https://apps.who.int/iris/handle/10665/330321

148. Møller S, Jensen MB, Ejlertsen B, et al. The clinical database and the treatment guidelines of the Danish Breast Cancer Cooperative Group (DBCG); its 30-years experience and future promise. *Acta Oncol Stockh Swed*. 2008;47(4):506-524.

149. Saltbæk L, Horsboel TA, Offersen BV, et al. Patterns in detection of recurrence among patients treated for breast cancer. *Breast Cancer Res Treat*. 2020;184(2):365-373.

150. Collin LJ, Cronin-Fenton DP, Ahern TP, et al. Cohort Profile: the Predictors of Breast Cancer Recurrence (ProBe CaRE) Premenopausal Breast Cancer Cohort Study in Denmark. *BMJ Open*. 2018;8(7).

151. Christiansen P, Ejlertsen B, Jensen MB, Mouridsen H. Danish Breast Cancer Cooperative Group. *Clin Epidemiol*. 2016;8:445-449.

152. Mouridsen HT, Bjerre KD, Christiansen P, Jensen MB, Møller S. Improvement of prognosis in breast cancer in Denmark 1977–2006, based on the nationwide reporting to the DBCG Registry. *Acta Oncol*. 2008;47(4):525-536.

153. Pedersen RN, Öztürk B, Mellemkjær L, et al. Validation of an Algorithm to Ascertain Late Breast Cancer Recurrence Using Danish Medical Registries . *Clin Epidemiol*. 2020;12:1083-1093.

154. Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449-490.

155. Helweg-Larsen K. The Danish Register of Causes of Death. *Scand J Public Health*. 2011;39(7 Suppl):26-29.

156. Erichsen R, Lash TL, Hamilton-Dutoit SJ, Bjerregaard B, Vyberg M, Pedersen L. Existing data sources for clinical epidemiology: the Danish National Pathology Registry and Data Bank. *Clin Epidemiol*. 2010;2:51-56.

157. Bliddal M, Broe A, Pottegård A, Olsen J, Langhoff-Roos J. The Danish Medical Birth Register. *Eur J Epidemiol*. 2018;33(1):27-36.

158. Baadsgaard M, Quitzau J. Danish registers on personal income and transfer payments. *Scand J Public Health*. 2011;39(7_suppl):103-105.

159. Jensen VM, Rasmussen AW. Danish Education Registers. *Scand J Public Health*. 2011;39(7 Suppl):91-94.

160. Hjollund NH, Larsen FB, Andersen JH. Register-based follow-up of social benefits and other transfer payments: accuracy and degree of completeness in a Danish interdepartmental administrative database compared with a population-based survey. *Scand J Public Health*. 2007;35(5):497-502.

161. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008;167(4):492-499.

162. Schneeweiss S, Rassen JA, Brown JS, et al. Graphical Depiction of Longitudinal Study Designs in Health Care Databases. *Ann Intern Med*. 2019;170(6):398-406.

163. Statistics Denamrk. TIMES variabel - FAMILIE_TYPE. Accessed March 10, 2022. https://www.dst.dk/da/Statistik/dokumentation/Times/forebyggelsesregistret/familie-type

164. Unesco. *International Standard Classification of Education (ISCED) 2011*.; 2018. http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

165. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet*. 2002;3(5):391-397.

166. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443.

167. Christiansen P, Al-Suliman N, Bjerre K, Møller S, Danish Breast Cancer Cooperative Group. Recurrence pattern and prognosis in low-risk breast cancer patients--data from the DBCG 89-A programme. *Acta Oncol Stockh Swed*. 2008;47(4):691-703.

168. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383.

169. Ahern TP, Collin LJ, Baurley JW, et al. Metabolic Pathway Analysis and Effectiveness of Tamoxifen in Danish Breast Cancer Patients. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol*. 2020;29(3):582-590.

170. Westreich D, Greenland S. The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients. *Am J Epidemiol*. 2013;177(4):292-298.

171. Textor J, van der Zander B, Gilthorpe MS, Liskiewicz M, Ellison GT. Robust causal inference using directed acyclic graphs: the R package "dagitty." *Int J Epidemiol*. 2016;45(6):1887-1894.

172. Hernán MA. The Hazards of Hazard Ratios. *Epidemiol Camb Mass*. 2010;21(1):13-15.

173. Hjorth CF, Damkier P, Ejlertsen B, Lash T, Sørensen HT, Cronin-Fenton D. Socioeconomic position and prognosis in premenopausal breast cancer: a population-based cohort study in Denmark. *BMC Med*. 2021;19(1).

174. Svendsen RP, Jarbol DE, Larsen PV, Støvring H, Hansen BL, Soendergaard J. Associations between health care seeking and socioeconomic and demographic determinants among people reporting alarm symptoms of cancer: a population-based cross-sectional study. *Fam Pract*. 2013;30(6):655-665.

175. Collin LJ, Cronin-Fenton D, Ahern TP, et al. Early discontinuation of endocrine therapy and recurrence of breast cancer among premenopausal women. *Clin Cancer Res Off J Am Assoc Cancer Res*. Published online December 17, 2020. 4

176. Hershman DL, Shao T, Kushi LH, et al. Early discontinuation and non-adherence to adjuvant hormonal therapy are associated with increased mortality in women with breast cancer. *Breast Cancer Res Treat*. 2011;126(2):529-537.

177. Toivonen KI, Williamson TM, Carlson LE, Walker LM, Campbell TS. Potentially modifiable factors associated with adherence to adjuvant endocrine therapy among breast cancer survivors: a systematic review. *Cancers*. 2020;13(1):107.

178. Wassermann J, Gelber SI, Rosenberg SM, et al. Nonadherent behaviors among young women on adjuvant endocrine therapy for breast cancer. *Cancer*. 2019;125(18):3266-3274.

179. Pardo-Crespo MR, Narla NP, Williams AR, et al. Comparison of individual-level versus area-level socioeconomic measures in assessing health outcomes of children in Olmsted County, Minnesota. *J Epidemiol Community Health*. 2013;67(4):305-310.

180. Moss JL, Johnson NJ, Yu M, Altekruse SF, Cronin KA. Comparisons of individual- and area-level socioeconomic status as proxies for individual-level measures: evidence from the Mortality Disparities in American Communities study. *Popul Health Metr*. 2021;19(1):1.

181. Marcath LA, Kidwell KM, Robinson AC, et al. Patients carrying CYP2C8*3 have shorter systemic paclitaxel exposure. *Pharmacogenomics*. 2019;20(2):95-104.

182. de Graan AJM, Lancaster CS, Obaidat A, et al. Influence of Polymorphic OATP1B-Type Carriers on the Disposition of Docetaxel. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2012;18(16):4433-4440.

183. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer Oxf Engl 1990*. 2009;45(2):228-247.

184. Dancey JE, Dodd LE, Ford R, et al. Recommendations for the assessment of progression in randomised cancer treatment trials. *Eur J Cancer*. 2009;45(2):281-289.

185. Ording AG, Cronin-Fenton D, Ehrenstein V, et al. Challenges in translating endpoints from trials to observational cohort studies in oncology. *Clin Epidemiol*. 2016;8:195-200.

186. Busti AJ. Genetic Polymorphisms of the CYP3A4 Enzyme and Potential Influence on Drug Efficacy and/or Safety. Evidence-based medicine consult. Published 2015. Accessed July 2, 2021. https://www.ebmconsult.com/articles/genetic-polymorphisms-cytochrome-p450-cyp3a4-enzyme

187. Preissner SC, Hoffmann MF, Preissner R, Dunkel M, Gewiess A, Preissner S. Polymorphic Cytochrome P450 Enzymes (CYPs) and Their Role in Personalized Therapy. *PLoS ONE*. 2013;8(12):e82562.

188. Johnson N, Walker K, Gibson LJ, et al. CYP3A variation, premenopausal estrone levels, and breast cancer risk. *J Natl Cancer Inst*. 2012;104(9):657-669.

189. Mir O, Alexandre J, Tran A, et al. Relationship between GSTP1 Ile105Val polymorphism and docetaxel-induced peripheral neuropathy: clinical evidence of a role of oxidative stress in taxane toxicity. *Ann Oncol*. 2009;20(4):736-740.

190. Chen Q, Lin W, Yang J, et al. Prognostic Value of Two Polymorphisms, rs1045642 and rs1128503, in ABCB1 Following Taxane-based Chemotherapy: A Meta-Analysis. *Asian Pac J Cancer Prev APJCP*. 2021;22(1):3-10.

191. Jiang Q, Xu M, Liu Y, et al. Influence of the ABCB1 polymorphisms on the response to Taxane-containing chemotherapy: a systematic review and meta-analysis. *Cancer Chemother Pharmacol*. 2018;81(2):315-323.

192. Isla D, Sarries C, Rosell R, et al. Single nucleotide polymorphisms and outcome in docetaxel-cisplatin-treated advanced non-small-cell lung cancer. *Ann Oncol Off J Eur Soc Med Oncol*. 2004;15(8):1194-1203.

193. Lamba J, Hebert JM, Schuetz EG, Klein TE, Altman RB. PharmGKB summary: very important pharmacogene information for CYP3A5. *Pharmacogenet Genomics*. 2012;22(7):555-558.

194. Leskelä S, Jara C, Leandro-García LJ, et al. Polymorphisms in cytochromes P450 2C8 and 3A5 are associated with paclitaxel neurotoxicity. *Pharmacogenomics J*. 2011;11(2):121-129.

195. Eckhoff L, Feddersen S, Knoop AS, Ewertz M, Bergmann TK. Docetaxel-induced neuropathy: a pharmacogenetic case-control study of 150 women with early-stage breast cancer. *Acta Oncol Stockh Swed*. 2015;54(4):530-537.

196. Frank L. When an Entire Country Is a Cohort. *Science*. 2000;287(5462):2398-2399.

197. Kirchhoff AC, Yi J, Wright J, Warner EL, Smith KR. Marriage and divorce among young adult cancer survivors. *J Cancer Surviv Res Pract*. 2012;6(4):441-450.

198. Jensen L, Overgaard C, Bøggild H, et al. The Long-term financial consequences of breast cancer: a Danish registry-based cohort studyLaura Jensen. *Eur J Public Health*. 2017;27(suppl_3).

199. Statistics Denmark. TIMES variabel - SOCIO13. Accessed February 8, 2022. https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/socio13

200. Cronin-Fenton DP, Kjærsgaard A, Ahern TP, et al. Validity of Danish Breast Cancer Group (DBCG) registry data used in the predictors of breast cancer recurrence (ProBeCaRe) premenopausal breast cancer cohort study. *Acta Oncol*. 2017;56(9):1155-1160.

201. Lash TL, Schmidt M, Jensen AØ, Engebjerg MC. Methods to apply probabilistic bias analysis to summary estimates of association. *Pharmacoepidemiol Drug Saf*. 2010;19(6):638-644.

202. Lash TL, Fink AK. Semi-Automated Sensitivity Analysis to Assess Systematic Errors in Observational Data. *Epidemiology*. 2003;14(4):451.

203. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393.

204. Shi S, Yuan N, Yang M, et al. Comprehensive Assessment of Genotype Imputation Performance. *Hum Hered*. 2018;83(3):107-116.

205. Rothman KJ. Six Persistent Research Misconceptions. *J Gen Intern Med*. 2014;29(7):1060-1064.

206. Moya JJ, Moazzez A, Ozao-Choy JJ, Dauphine C. Patients With Invasive Breast Cancer Who Refuse Treatment: An Analysis of Associated Factors and Impact on Survival. *Am Surg*. 2021;87(10):1627-1632.

207. Baak-Pablo R, Dezentje V, Guchelaar HJ, van der Straaten T. Genotyping of DNA Samples Isolated from Formalin-Fixed Paraffin-Embedded Tissues Using Preamplification. *J Mol Diagn JMD*. 2010;12(6):746-749.

208. Hertz DL, Kidwell KM, Thibert JN, et al. Genotyping concordance in DNA extracted from formalin-fixed paraffin embedded (FFPE) breast tumor and whole blood for pharmacogenetic analyses. *Mol Oncol*. 2015;9(9):1868-1876.

209. Sjöholm MIL, Hoffmann G, Lindgren S, Dillner J, Carlson J. Comparison of archival plasma and formalin-fixed paraffin-embedded tissue for genotyping in hepatocellular carcinoma. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol*. 2005;14(1):251-255.

210. Waples RS. Testing for Hardy-Weinberg proportions: have we lost the plot? *J Hered*. 2015;106(1):1-19.

211. Statistics Denmark. Immigrants and their descendants. Accessed April 4, 2022. https://www.dst.dk/en/Statistik/emner/borgere/befolkning/indvandrere-og-efterkommere

212. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *The Lancet*. 2005;365(9472):1687-1717.

213. Statens Serum Institut. Registry declaration for the Cause of Death Registry [In Danish: Registerdeklaration for Dødsårsagsregisteret]. Published online 2013. Accessed March 4, 2022. https://sundhedsdatastyrelsen.dk/-/media/sds/filer/registre-og-services/nationale-sundhedsregistre/doedsaarsager-og-biologisk-materiale/doedaarsagsregisteret/registerdeklaration-doedsaarsagsregisteret.pdf?la=da

214. Huisman M, Kunst AE, Mackenbach JP. Inequalities in the prevalence of smoking in the European Union: comparing education and income. *Prev Med*. 2005;40(6):756-764.

215. Sollie M, Bille C. Smoking and mortality in women diagnosed with breast cancer—a systematic review with meta-analysis based on 400,944 breast cancer cases. *Gland Surg*. 2017;6(4):385-393.

216. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42(4):1012-1014.

217. Mackenbach JP. *Health Inequalities : Persistence and Change in Modern Welfare States*. Oxford University Press; 2019.

218. Hvidberg L, Wulff CN, Pedersen AF, Vedsted P. Barriers to healthcare seeking, beliefs about cancer and the role of socio-economic position. A Danish population-based study. *Prev Med*. 2015;71:107-113.

219. Hvidberg L, Pedersen AF, Wulff CN, Vedsted P. Cancer awareness and socio-economic position: results from a population-based study in Denmark. *BMC Cancer*. 2014;14:581.

220. Baurley JW, Kjærsgaard A, Zwick ME, et al. Bayesian Pathway Analysis for Complex Interactions. *Am J Epidemiol*. 2020;189(12):1610-1622.

221. Kollerup A, Ladenburg J. Willingness to pay for accommodating job attributes when returning to work after cancer treatment: A discrete choice experiment with Danish breast cancer survivors. *LABOUR*. 2021;35(3):378-411.

222. Kollerup A, Ladenburg J, Heinesen E, Kolodziejczyk C. The importance of workplace accommodation for cancer survivors – The role of flexible work schedules and psychological help in returning to work. *Econ Hum Biol*. 2021;43:101057.

223. International Agency for Research on Cancer. *Latest Global Cancer Data: Cancer Burden Rises to 19.3 Million New Cases*. World Health Organization; 2020. https://www.iarc.who.int/wp-content/uploads/2020/12/pr292_E.pdf

224. Heer E, Harper A, Escandor N, Sung H, McCormack V, Fidler-Benaoudia MM. Global burden and trends in premenopausal and postmenopausal breast cancer: a population-based study. *Lancet Glob Health*. 2020;8(8):e1027-e1037.

225. Ejlertsen B. Adjuvant chemotherapy in early breast cancer. *Dan Med J*. 2016;63(5).

Appendices

Appendix I — Paper I

Appendix II — Paper II

Appendix III — Paper III

The papers have been removed from the file due to copyright issues