From the Department of Medical Epidemiology and Biostatistics
Karolinska Institutet, Stockholm, Sweden

# SURVIVORSHIP IN HODGKIN LYMPHOMA: CHILDBEARING AND TREATMENT-RELATED DISEASE

Caroline E. Weibull


Karolinska Institutet

Stockholm 2018

All previously published papers were reproduced with permission from the publisher.
Cover illustration by Caroline E. Weibull.
Published by Karolinska Institutet.
Printed by E-Print AB.

Typeset by the author using LaTeX 2$_\varepsilon$

ISBN 978-91-7831-154-5

Institutionen för medicinsk epidemiologi och biostatistik

# Survivorship in Hodgkin lymphoma: Childbearing and treatment-related disease

## AKADEMISK AVHANDLING

som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i hörsal Samuelssonsalen, Tomtebodavägen 6, Karolinska Institutet, Solna

**Fredagen den 9 november 2018, kl 09.00**

av
**Caroline E. Weibull**

*Huvudhandledare:*
Professor Paul W. Dickman
Karolinska Institutet
Institutionen för medicinsk epidemiologi och biostatistik

*Bihandledare:*
Doktor Sandra Eloranta
Karolinska Institutet
Institutionen för medicin, Solna

Professor Paul C. Lambert
University of Leicester, UK
Department of Health Sciences
Biostatistics Research Group
samt
Karolinska Institutet
Institutionen för medicinsk epidemiologi och biostatistik

Professor Magnus Björkholm
Karolinska Institutet
Institutionen för medicin, Solna

Docent Ingrid Glimelius
Uppsala Universitet
Institutionen för immunologi, genetik och patologi

*Fakultetsopponent:*
Professor James R. Cerhan
Mayo Clinic, USA
Division of Epidemiology
Department of Health Sciences Research

*Betygsnämnd:*
Professor Joanna Tyrcha
Stockholms Universitet
Institutionen för matematik

Docent Michael Fored
Karolinska Institutet
Institutionen för medicin, Solna

Docent Thomas Relander
Lunds Universitet
Avdelningen för onkologi och patologi
Medicinska fakulteten

**Stockholm 2018**

# Abstract

Cancer is often considered a disease with, historically, poor survival that affects middle-aged and elderly individuals. Hodgkin lymphoma (HL) is a lymphatic malignancy that affects both young and old individuals, with the age-specific incidence curve having its first peak at ages 20-30 years. As survival has improved substantially over the last decades, there is an increasing number of survivors – some still at a young age. The primary purpose of this thesis was to address issues related to childbearing and treatment-associated disease among HL survivors. As a means to investigate these issues, novel statistical methods were developed and applied.

**Childbearing in relation to HL**
Some HL survivors will be at the beginning, or in the midst, of their childbearing years. Both pregnancy and HL are associated with changes to the immune system, making it plausible that pregnancy could affect the progression of the disease. **Study I** in this thesis aimed to answer if pregnancy affects the risk of relapse among female patients in remission from HL.

Ever since the introduction of radio- and chemotherapy with the possibility to cure HL, the negative effects of therapy on fertility have been a concern. Contemporary treatments are believed to be less gonadotoxic than those previously used, but few studies have compared childbearing potential between the main treatment regimens administered today (ABVD and BEACOPP) in a real-world setting. In **Study II**, temporal trends in childbearing among female HL patients (of childbearing ages) were investigated, within groups of treatments, and in comparison with the general population.

Both studies utilized a cohort of women diagnosed between 1992 and 2009, at ages 18-40 years, for whom detailed information on relapse as well as patient and disease characteristics was available. For the purpose of comparing childbirth rates with the general population, HL patients were individually matched to HL-free comparators. Childbirth rates were studied separately within two time windows during follow-up: 0-3 years and 3-7 years, and cumulative probability of childbirth was calculated in the presence of the competing risks of death and relapse.

No evidence to support the hypothesis of pregnancy-associated relapse was found. However, since the absolute risk of relapse is at its highest levels during the first 2-3 years after diagnosis, female HL patients could, if possible, be advised to delay childbearing to avoid co-occurence. Childbearing potential improved over calendar time, reflecting reduced toxicity and changes in counseling. Women treated during recent years had childbirth potential in line with that of matched comparators three years after diagnosis. Even women treated with BEACOPP, the most gonadotoxic chemotherapy, had an increasing cumulative probability of having children after HL. Importantly, no women had children after a relapse within the first seven years after diagnosis, which motivates a need for fertility advice and counseling at time of HL diagnosis.

**Treatment-related disease**

Late effects from cancer therapy are becoming increasingly important to quantify as the number of cancer survivors grows. Chemo- and radiotherapy used to treat HL increases the risk of cardio- and cerebrovascular diseases, and secondary malignancies (SMs).

In **Study III**, excess incidence of diseases of the circulatory system (DCS) among HL patients was investigated to describe temporal trends in DCS morbidity attributable to HL and its treatment. Data on patients diagnosed with HL between 1985 and 2013 at ages 18-80 years, for whom information on inpatient DCS records was available, was used. Relative survival methods were applied to estimate excess incidence rates indirectly from the observed and expected rates of DCS. Cumulative excess incidence of DCS was calculated in the presence of competing risks.

The treatment-related incidence of DCS declined between the mid-1980s and mid-1990s, after which no substantial improvements were observed. The risk of a treatment-related DCS persists for up to 10 years among patients who completed their treatment in the new millennium.

When studying late effects it is important to attempt to capture the additional disease incidence associated with cancer treatment. Additionally, to gain insight in real-life risks, it is of interest to study not only time to first event, but continue to follow patients as they experience different types of late effects before reaching an absorbing state. Doing both of these simultaneously requires estimating excess transition rates to transient states in a multi-state model framework, for which no methods have existed. **Study IV** suggested a way to achieve this, using a recently developed simulation strategy to predict transition probabilities. As an illustrative example, data from Study III on HL patients and DCS incidence was used. Combining methods from relative survival with the multi-state framework enables investigation of complex patient pathways and can be useful for several applications related to survivorship among cancer patients.

# List of publications

I. Weibull CE, Eloranta S, Smedby KE, Björkholm M, Kristinsson SY, Johansson ALV, Dickman PW, Glimelius I. **Pregnancy and the Risk of Relapse in Patients Diagnosed With Hodgkin Lymphoma.** *Journal of Clinical Oncology* 2016 Feb 1;34(4):337-44.

II. Weibull CE, Johansson ALV, Eloranta S, Smedby KE, Björkholm M, Lambert PC, Dickman PW, Glimelius I. **Contemporarily treated Hodgkin lymphoma patients have childbearing potential in line with matched comparators.** *Journal of Clinical Oncology* 2018 Sep 10;36(26):2718-2725.

III. Weibull CE, Björkholm M, Glimelius I, Lambert PC, Andersson TML, Smedby KE, Dickman PW, Eloranta S. **Temporal trends in treatment-related incidence of diseases of the circulatory system among Hodgkin lymphoma patients.** *Submitted.*

IV. Weibull CE, Lambert PC, Eloranta S, Andersson TML, Dickman PW, Crowther MJ. **A multistate model incorporating estimation of excess hazards and different time scales.** *In manuscript.*

The articles are referred to by their roman numerals throughout, and are reproduced in full at the end of this thesis.

# Related publication

- Björkholm M, Weibull CE, Eloranta S, Smedby KE, Glimelius I, Dickman PW. **Greater attention should be paid to developing therapies for elderly patients with Hodgkin lymphoma  a population-based study from Sweden.** *European Journal of Heamatology* 2018 Jul;101(1):106-114.

# Contents

# List of abbreviations

| | | | |
|---|---|---|---|
| AER | Absolute excess risk | IPS | International prognostic score |
| CABG | Coronary artery bypass grafting | MBR | Medical birth register |
| CAD | Coronary artery disease | MGR | Multi-generation register |
| CDR | Cause of death register | MI | Myocardial infarction |
| CHF | Congestive heart failure | NCI | National Cancer Institute |
| CHL | Classical Hodgkin lymphoma | NHL | Non-Hodgkin lymphoma |
| CI | Confidence interval | PD-1 | Programmed cell death protein 1 |
| CIF | Cumulative incidence function | PET | Positron emission tomography |
| CT | Chemotherapy | PH | Proportional hazards |
| CVD | Cardiovascular disease | PIN | Personal identification number |
| DCS | Diseases of the circulatory system | PROM | Premature rupture of membranes |
| df | Degrees of freedom | RCC | Regional cancer centers |
| EBV | Epstein-Barr virus | RR | Relative risk |
| EIRR | Excess incidence rate ratio | RT | Radiotherapy |
| ESR | Erythrocyte sedimentation rate | RTP | Register for total population |
| HIV | Human immunodeficiency virus | SCR | Swedish cancer register |
| HL | Hodgkin lymphoma | SIR | Standardized incidence ratio |
| HR | Hazard ratio | SLR | Swedish lymphoma register |
| HR-S | Hodgkin Reed-Sternberg | SM | Secondary malignancies |
| ICD | International classification of disease | SMR | Standardized mortality ratio |
| IPR | Inpatient register | WHO | World health organisation |

| | |
|---|---|
| ABVD | doxorubicin, bleomycin, vinblastine, and dacarbazine |
| BEACOPP | bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone |
| ChlVPP | chlorambucil, vinblastine, procarbazine, and prednisone |
| CHOP | cyclophosphamide, doxorubicin, vincristine, and prednisone |
| LVPP | chlorambucil, vinblastine, procarbazine, and prednisone |
| MOPP | mechlorethamine, vincristine, procarbazine, and prednisone |
| OEPA | vincristine, etoposide, prednisone, and doxorubicin |

# 1. Introduction

Cancer survivorship is conceptually different to cancer patient survival. While the latter is the probability of surviving up to a certain point after a cancer diagnosis, survivorship incorporates more than just that probability. Being diagnosed with cancer, undergoing therapy over a potentially long period, and adjusting oneself to living with a history of cancer, has implications on a physiological as well as a psychological level. Although every cancer survivor has a different experience of life thereafter, some feelings are shared by most people, such as an increasing concern over their own and their immediate family's health, as well as a higher appreciation of how precious life is.

According to `www.cancer.net` "cancer survivorship has at least two common meanings: (1) Having no signs of cancer after finishing treatment and (2) Living with, through, and beyond cancer." The second point is the focus of this thesis; studying life beyond cancer among patients in remission from Hodgkin lymphoma (HL).

Until the middle of the twentieth century, patients diagnosed with HL had very limited chances of surviving. Irradiation and single agent chemotherapy were the only treatment options, which meant that patients with disseminated disease were difficult to treat with curative intent. Since then, the development of efficient therapies and more accurate staging procedures has resulted in large reductions in HL mortality.

The age-specific incidence of HL has its first peak around ages 20-25 years. With more and more HL patients being cured from their disease, there is now a large group of relatively young cancer survivors with many more years left to live. Some of these are yet to start families, and the experience of cancer may give rise to worries related to childbearing. Will I be able to have children? Can a future pregnancy result in my cancer coming back? These are two issues that are addressed in this thesis.

In a life-span perspective, HL therapy has been shown to increase the risk of a number of late complications, including cardiovascular disease and secondary malignancies (SMs). While treatments are continuously being developed to minimize the risk of late effects without jeopardizing the chances of curing the HL, there is a balance in terms of maximizing the life expectancy among the patients. In some situations, it might be worth risking late effects of treatment if cure can be reached. It is vital that real-world (as opposed to clinical trial) evidence on morbidity and mortality from treatment-related diseases is continuously gathered to support decision making for the treating physicians. It is also important to partition the risk of such diseases into the component that can be expected in the absence of HL and the component which is related to the HL and treatment thereof. For this to be possible, sophisticated statistical methods have been, and are being developed. In this thesis, such methods are extended and applied to study temporal trends in the excess incidence of circulatory system disease among HL patients.

Cancer patient survivors are not only at risk of death due to cancer, or treatment-related cardiac disease, or SMs. They are at risk of all of these things, all at once. Therefore, these risks should be studied simultaneously. The last study included in this thesis presents a solution to this analytical problem, by incorporating estimation of excess morbidity and mortality into a multi-state model.

# 2. Aims of this thesis

As a result of the vast improvements in survival among HL patients, clinical focus and research has partly shifted towards improving quality of life after curative treatment and reducing adverse events potentially caused by the treatment. The studies included in this thesis all concern life after diagnosis among survivors of HL. The overarching purpose is however twofold, covering both statistical methods development and clinical application.

The primary aim was to investigate clinically relevant hypotheses concerning childbirth potential after HL, pregnancy-associated relapse, and incidence of treatment-related DCS. As a notable proportion of patients are diagnosed at a young age when they have not yet finished their reproductive life, issues related to fertility and childbearing are of great concern to the survivors and their families. Likewise, treatment-induced complications are a major problem, especially among younger patients with many years still to live.

The second aim was to apply and extend novel statistical methods within relative survival and multi-state modeling to facilitate proper and thorough investigation of the above-mentioned clinically relevant questions.

More specifically, the aims were to:

- Investigate if pregnancy triggers relapse among women in remission from HL [Study I].
- Describe temporal trends in childbirth patterns among female HL survivors, by clinical characteristics and compared to the general population [Study II].
- Study treatment-related morbidity due to DCS among HL patients [Studies III].
- Incorporate estimation of excess incidence rates and the use of multiple time scales into a multi-state model framework [Study IV].

# 3. Background

## 3.1 Hodgkin lymphoma

Lymphoma is a malignant condition where the tumor cells originate from *lymphocytes*, a type of white blood cell that are part of our immune system. Lymphocytes are the main type of cell found in the lymphatic system but they are also found in the blood. There are three different types of lymphocytes: T (thymus-derived) cells, NK (natural killer) cells, and B (bone marrow derived) cells. Lymphomas can arise in either T/NK or B cells, and classification into sub-types is based on the origin of the malignant cell. According to the World Health Organization (WHO) classification system, there are three main categories of lymphoma: B cell lymphomas (make up around 80%), T/NK cell lymphomas (10%), and HL (10%) which also originates from B cells. The two first sub-types together make up the non-Hodgkin lymphomas (NHL). HL is named after the pathologist Thomas Hodgkin, who in 1832 was the first to describe the disease post-mortem in seven cases. In 1898, Carl Sternberg claimed that diagnosis should be based on a histological investigation, and a couple of years later Dorothy Reed described the malignant cells that are the microscopic hallmark of HL as "giant cells" and further noted that "eight or ten nuclei have been found in a single cell" [1]. These cells, illustrated in Figure 3.1, are referred to as Hodgkin Reed-Sternberg (HR-S) cells and are today necessary for a diagnosis of HL. Although HR-S cells can be low ($\sim$1%) in number, they have a larger mass than other B-cells, and produce factors that attract inflammatory cells, which in turn makes up the majority of the tumor burden.

There are two main types of HL; Classical HL (CHL), that affects around 95% of the patients, and nodular lymphocyte predominant HL (the remaining 5%). Classical HL can be further divided into:

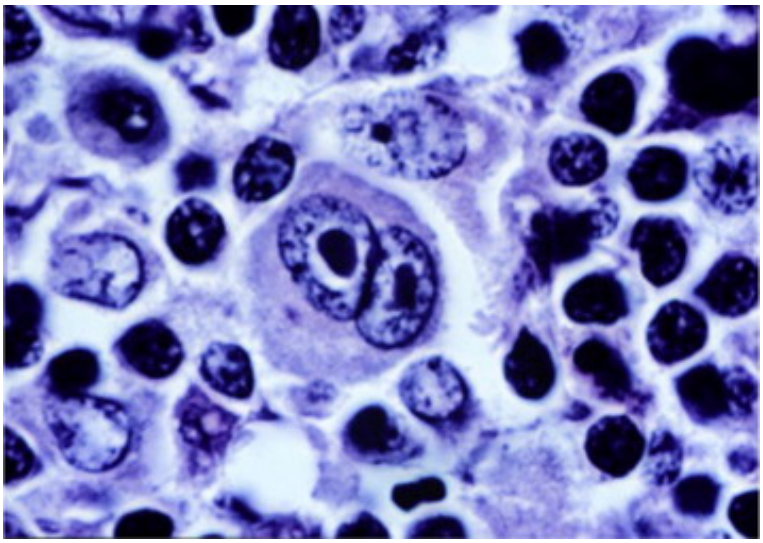
- Nodular sclerosis CHL – accounts for 60-80% and is typically seen in young adults.
- Mixed cellularity CHL – accounts for 15-30% and is most commonly seen in elderly patients.
- Lymphocyte-rich CHL – accounts for around 5%.
- Lymphocyte-depleted CHL – rare (less than 1%) and is most commonly seen in advanced stages and among elderly patients.

For the purpose of this thesis, no separation was done between any of these sub-types.

### 3.1.1 Signs and symptoms, staging, and prognosis

HL is usually detected due to enlargement of one or more lymph nodes – often in the neck, under the arm, or in the groin. These are typically described as having a rubbery feeling and are not tender or

Figure 3.1: Characteristic microscopic picture of a HR-S cell. Source: Terezakis, S. 2015 [2].



painful for the patient. Alternatively, the patient seeks medical care due to heavy night sweats, fever or unexpected weight loss. This group of symptoms are called *B symptoms* and are often a sign of more advanced disease.

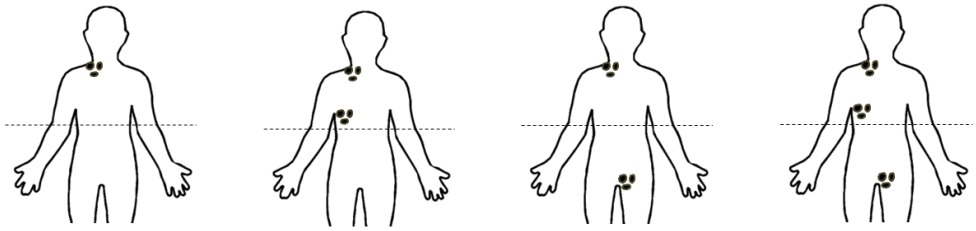
Staging of HL is done according to the modified Ann Arbor staging classification[3], which is based on:

1. Number of lymph node regions involved.
2. Location of the affected lymph nodes.
3. Other organ involvement.
4. Whether or not the patient is experiencing B symptoms.

Historically, laparotomy with splenectomy and lymphangiography were used as staging procedures, but today positron emission tomography (PET) and computed tomography (CT) scans together with biopsies are primarily used. Figure 3.2 illustrates how stage I-IV disease is defined. Presence of B symptoms is further denoted with the suffix letter “B” and “A” in the absence of symptoms. Stage IA-IIA is typically considered *limited* stage, while IIB-IVB is *advanced* stage.

Stage of HL is supplemented by an assessment of prognosis. For limited stage patients, the following factors are associated with an unfavorable prognosis: bulky disease, extra nodal disease, more than three affected lymph node regions, presence of B symptoms, and an erythrocyte sedimentation rate (ESR; the rate at which red blood cells sediment) $\geq$50 mm/hour. For advanced stage patients, prognosis is determined by seven verified prognostic factors - male sex, age at diagnosis above 45, presenting with stage IV, hemoglobin (Hb) concentration less than 105 g/l, serum albumin less than 40 g/l, lymphopenia (lymphs $< 600 \cdot 109$/l or lymphocyte count <8% of white blood cell count), and leukocytosis (white blood cell count $\geq 15 \cdot 10^9$/l). These prognostic factors are summarized in the International Prognostic Score (IPS) score, where presence of each risk factor scores one. After therapy initiation, the most important negative prognostic factor is not having a PET scan showing complete metabolic remission after two cycles of treatment.

Figure 3.2: Classification of staging in HL.



**Stage I**
One lymph node region involved

**Stage II**
Two or more lymph node regions involved on the same side of the diaphragm

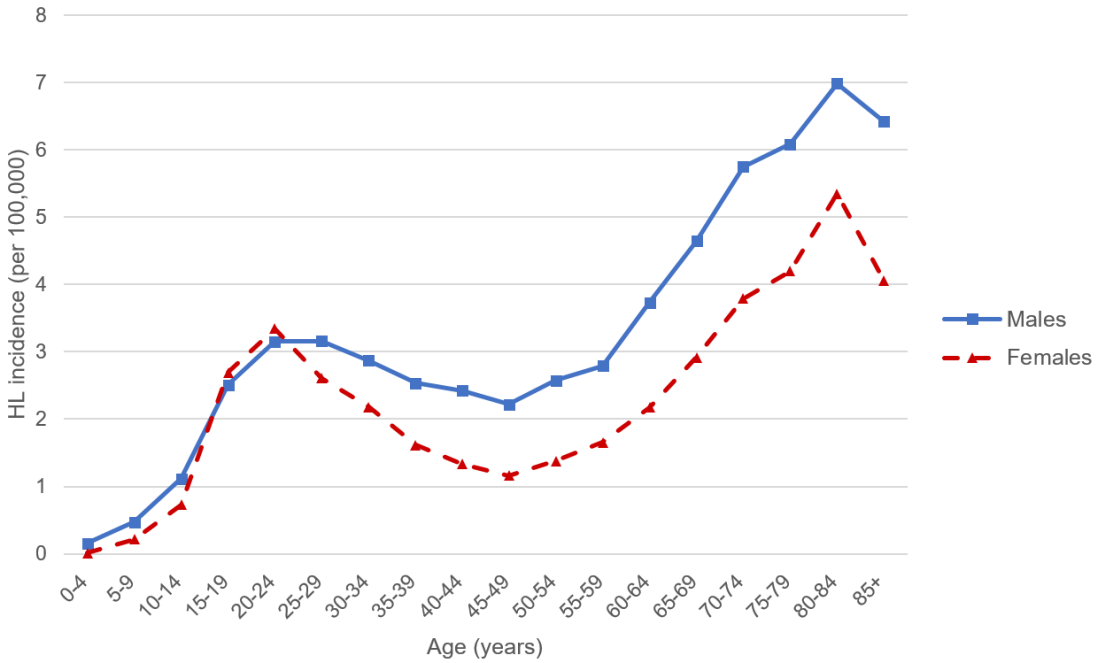
**Stage III**
Two or more lymph node regions involved, on both sides of the diaphragm

**Stage IV**
Widespread disease outside of the lymphatic regions

### 3.1.2 Epidemiology

HL is a rare malignancy, around 200 new cases are diagnosed in Sweden each year constituting approximately 0.3% of all cancer cases. However, it is one of the most frequently diagnosed malignancies in young adults. The incidence in high income countries has a bimodal shape by age, as shown in Figure 3.3. Very few children and adolescents get diagnosed with HL; the first peak in incidence appears in ages 20-25, and after a decrease among middle-aged people it increases again from age 45-50.

Figure 3.3: Age-specific incidence of HL in Sweden 1970-2016. Source: National board of Health and Welfare.



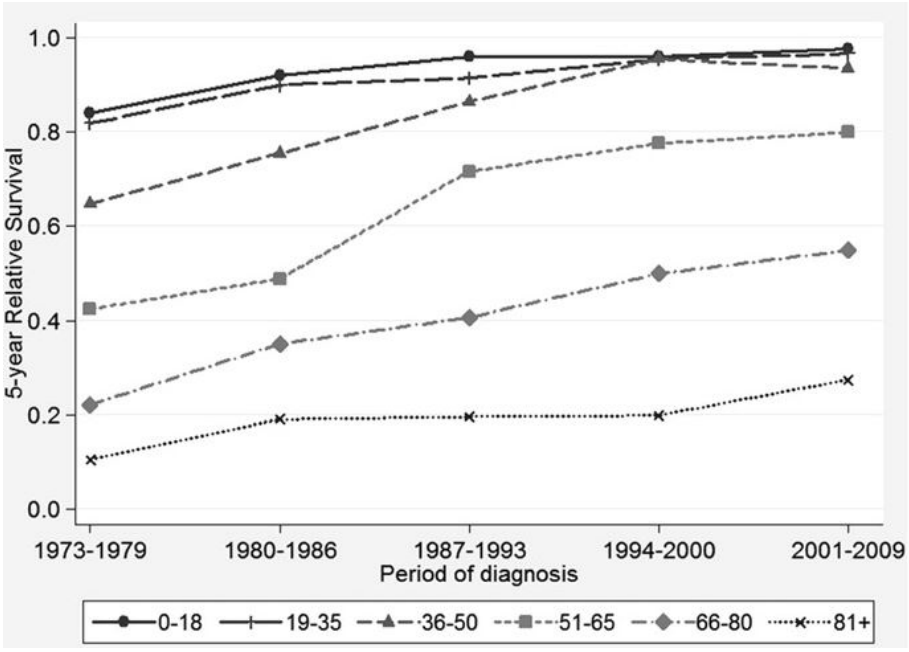

Knowledge on established risk factors for HL is scarce. Besides high socioeconomic status and male sex, familial aggregation has been observed. Individuals with a first degree family member diagnosed

with HL, or some other type of lymphoma, have a 3- to 7-fold increase in risk [4–6]. Immunodeficiency increases the risk of HL. Especially, having an infectious mononucleosis ("mono") caused by the Epstein-Barr virus (EBV) has been shown to increase the risk of the EBV-positive subtype of HL among young adults [7]. However, given the high prevalence of EBV in the general public, the etiology is still not fully understood. HIV (human immunodeficiency virus) infected individuals are also at greater risk of developing HL.

### 3.1.3 Treatment principles

Untreated, survival among HL patients is generally short (less than a year) and half a century ago, HL was associated with a very unfavorable prognosis. Today, advances in chemo- and radiation therapy have made it possible to cure the majority of both limited and advanced stage patients (even if first-line treatment fails). Figure 3.4 illustrates the improvements in 5-year relative survival between 1973 and 2009, for patients of different ages.

Figure 3.4: Temporal trends in five-year relative survival for HL patients in Sweden, by age at diagnosis. Source: Sjöberg, J. *et al.* 2012 [8].




Before the introduction of effective chemotherapy (CT), HL patients were treated with extended field radiotherapy (RT), e.g., mantle field (torso) or inverted Y-field (pelvis). For patients with limited stage, remission rates were fairly high, but relapse was common. Patients with advanced stage invariably had a fatal outcome. The first attempts to treat HL with CT consisted of single agent drugs – mechlorethamine, chlorambucil, or vincristine. In 1964, DeVita and colleagues successfully combined four CT drugs (mechlorethamine, oncovin/vincristine, procarbazine, and prednisone) and could report on the first cures of advanced stage HL [9]. This therapy is known as the MOPP regimen, and has been extensively used since. Besides enabling cure of advanced disease patients, this opened up the field of combined modality treatment (i.e., treatments with a combination of CT and RT) for patients with limited-stage disease, which in turn further reduced relapse rates among this group of patients. Unfortunately, cure came at a cost in terms of increased risk of infertility and SMs, especially leukemia and breast cancer.

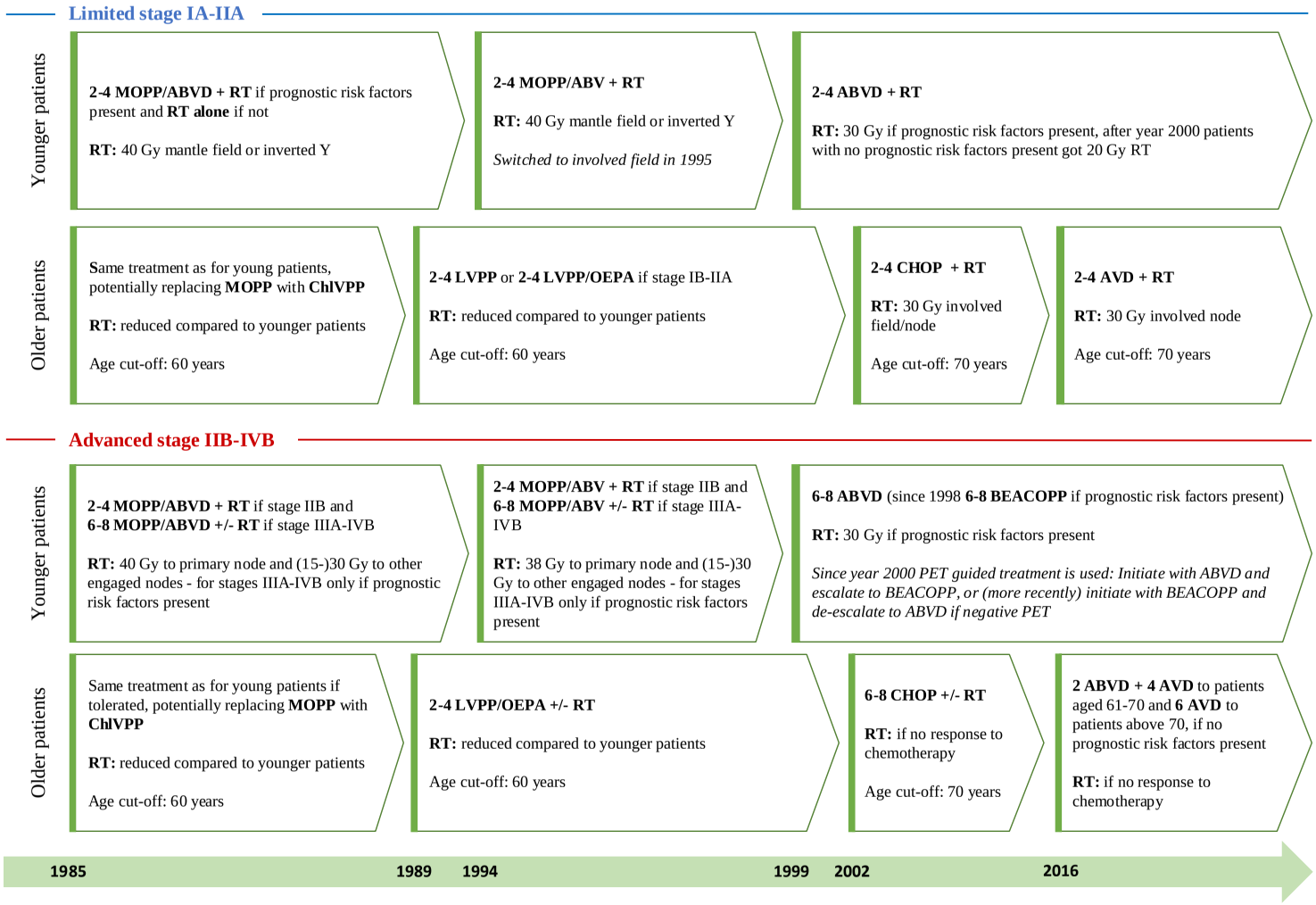
Over time, alterations of the MOPP regimen have been suggested, with the primary aim to reduce associated toxic effects. In 1975, ABVD (Adriamycin/doxorubicin, bleomycin, vinblastine, and dacarbazine) was introduced to patients failing MOPP. It was later adopted as first-line treatment since the cure rates were higher and the side-effects fewer. Combinations and other regimens have also been used over the years, such as MOPP/ABV (for young patients), and CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone) administered to older patients.

In the 1990s, The German Hodgkin Study Group developed the BEACOPP (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone) CT regimen to address the fact that approximately 30% of advanced-stage patients treated with ABVD were still not cured [10]. BEACOPP has a more severe toxicity profile than ABVD, and in Sweden this therapy has primarily been used for patients with advanced stage and unfavorable risk factors according to IPS. The trend today is towards giving BEACOPP initially and de-escalate to ABVD or AVD if the patient has PET negative disease after two cycles of chemotherapy.

Parallel to advances in multi-agent CT, RT has gone from extended field (mantle/para-aortic/inverted Y-field/total nodal irradiation) to more targeted (involved field/node). There has also been a shift towards smaller doses of irradiation (from 40 Gy to 20-30 Gy for patients with limited stage). The development of standards for HL treatment in Sweden, as recommended by the Swedish Lymphoma Group, are illustrated in Figure 3.5.

The advances in chemo- and radiotherapy, together with better staging techniques, have made HL a curable disease with a 10-year relative survival exceeding 90% for patients aged under 65 at diagnosis [8]. However, still today approximately 2% of patients with classical HL are refractory to first line therapy and around 13% relapse [11]. These patients are currently treated with high-dose therapy or salvage chemotherapy and autologous stem cell transplantation. As this therapy is not tolerated by all patients, and for some not sufficient for cure, alternative treatments are emerging (primarily aimed at refractory patients). One example is the drug *brentuximab vedotin* that targets CD30, a molecule found on tumor cells, and by doing so delivers chemotherapy only to those cells. In a study on pre-treated HL patients with relapse/refractory disease, 75% responded and 34% achieved complete remission [12]. Another recent study on advanced stage (III+IV) patients treated with ABVD versus A+AVD (brentuximab vedotin instead of bleomycin), reported a lower risk of disease progression and death at two years after diagnosis among those treated with A+AVD [13]. Other novel therapies include PD-1 (programmed cell death protein 1) inhibitors - in a recent phase II trial of *nivolumab* among patients with relapsed or progressive disease who had failed other lines of treatment, 66% of patients responded to treatment and 9% went into complete remission [14].

Figure 3.5: HL treatments development in Sweden between 1985 and 2016. LVPP/OEPA is chlorambucil, vinblastine, procarbazine, and prednisone alternating with vincristine, etoposide, prednisone, and doxorubicin. ChlVPP is chlorambucil, vinblastine, procarbazine, and prednisone.



## 3.2 Childbearing in relation to HL

The first peak in HL incidence precedes and covers the prime childbearing years. This raises several interesting research topics on pregnancy and childbearing in relation to HL.

### 3.2.1 Pregnancy as a risk factor for HL

As HL incidence starts to diverge between males and females around the childbearing ages, reproductive factors (such as parity and age at first birth) have been suggested to have an effect on HL risk. Although higher parity has been observed to increase the risk for HL [15], more evidence points to a reduced [16, 17] or no effect on risk at all [18]. Today, the latter is the general belief, and the National Cancer Institute (NCI) even states that "Pregnancy is not a risk factor for Hodgkin lymphoma".

### 3.2.2 Pregnancy and prognosis of HL

A recent childbirth prior to diagnosis does not appear to be an adverse prognostic factor in HL, as opposed to findings in melanoma and breast cancer [19]. Nor do women diagnosed during pregnancy have inferior cause-specific survival compared to women diagnosed outside of pregnancy [20, 21].

Nevertheless, a cancer diagnosis during pregnancy is both traumatic for the patient and poses problems in terms of clinical management. The biggest challenge in the management of HL during pregnancy is timing of delivery versus treatment initiation. The main goal is to continue the pregnancy to full term without jeopardizing the safety of the woman. Phase of pregnancy (i.e., trimester), stage of HL, and whether the disease is slow-growing or not, are the main deciding factors. If HL is diagnosed during the first trimester and immediate treatment is needed, termination of the pregnancy is recommended [22]. If HL is diagnosed during the second or third trimester, treatment is either deferred until after delivery, or antenatal chemotherapy is given. Irrespective of treatment alternative, the most common complications are induction of labor, PROM (premature rupture of membranes) and Cesarean section [23]. Several studies have concluded that antenatal CT with standard regimen during the second or third trimester is safe with regards to both the fetus and the mother's prospects of remission [22–24].

Data on how a pregnancy among women in remission from HL influences the risk of relapse has been sparse. A French study from 1988 comprising 12 patients who were pregnant during treatment for HL or shortly thereafter, found no evidence that pregnancy influences the course of HL [25]. However, as the study lacked a comparison group the results should be interpreted with caution.

### 3.2.3 Fertility and childbearing after HL

Since the introduction of curative HL therapy, its negative effect on fertility has been a major concern and one motivation for the development of new less gonadotoxic treatments.

RT to the pelvic region is particularly toxic among men. The testis are among the most radio-sensitive tissues, and thus even low doses of pelvic irradiation can cause loss in gonadal function [26]. For women, the effect is dependent on both dose and age (as the oocyte reserve decreases with increasing age), but normally small doses (less than 4 Gy) are considered safe. Today, for limited-stage patients with

subdiaphragmal HL, RT is often replaced with full CT and no RT. However, CT can also have a negative effect on fertility. In particular, alkylating agents (such as mechlorethamine, included in MOPP, and procarbazine, included in MOPP and BEACOPP) have been found to induce oligospermia/azoospermia (low or no sperm count in the semen) in men and amenorrhea (absence of menstruation) in women, temporary or permanently [26–29]. Men are more likely to experience loss in gonadal function due to CT compared to women [27].

For adult men, semen cryopreservation (i.e., sperm banking) is easily achieved and therefore standard procedure. For women, there are less options available in situations when immediate treatment is needed. If time allows, preservation of in vitro fertilized embryos can be done in adult women. Cryopreservation of oocytes and ovarian tissue can also be performed. Another option for women is adjuvant gonadotropin-releasing hormone analogue treatment. This has been shown to be effective in retaining ovarian function in some studies [30, 31] but not in others [32], and is not standard in HL patients today. In the rare cases of adolescents who have not entered puberty when HL therapy is initiated, fertility preservation becomes more complicated. Cryopreservation of immature testicular tissue among boys is not yet possible, while for girls, ovarian tissue cryopreservation can be done to fully guarantee fertility preservation. To maximize the chances of retaining fertility among very young HL patients, treatments currently used tries to limit the gonadotoxic effects as much as possible.

Whether the possibility of having children differs between HL survivors and people in general has been somewhat unclear. A small Canadian study showed encouraging results on limited-stage, 3-year relapse-free HL survivors treated with ABVD, concluding that pregnancy rates among female HL patients who attempted to become pregnant did not differ from those of matched comparators [33]. A similar pattern was observed for female survivors of early-stage unfavorable HL in the German Hodgkin Study Group HD14 trial [30]. A Norwegian study from 2011 reported lower childbirth rates among female but not male HL patients, compared to controls [34]. Similarly, a slight but significant lower number of children was seen in female HL survivors treated 1964 to 2004, who had children since before, compared to controls [35]. A Swedish study from 2013 stated, although not specific to HL survivors, that “cancer survivors are less likely to give birth compared with the background population” [36].

## 3.3 Treatment-related mortality and morbidity

HL patients treated with CT and/or RT are at risk of developing a range of treatment-related complications. These may present shortly after treatment, or many years later. The two most common late complications, with risk of fatal outcome, are DCS and SMs. Due to its bimodal incidence and vast improvements in survival, HL is often considered a “model disease” for gaining knowledge about late effects from CT and RT also applicable to other cancer types. As a consequence, the literature in this field is extensive, and several excellent reviews have been published that summarize the current knowledge on late effects [37–40].

### 3.3.1 Diseases of the circulatory system (DCS)

Both RT and CT may induce damage to the circulatory system. In cases of RT to the chest the risk of several cardiac sequelae, such as valvular heart disease and heart failure, has been observed to increase [41–44]. Moreover, a linear relationship between radiation dose (number of Gy) and risk of coronary

artery disease (CAD) has been reported [45]. In cases where the head and neck have been exposed to irradiation, damage to the carotid arteries may occur which in turn can lead to stroke [46–48]. As the mean RT doses have decreased over time, the excess risk of stroke has also been reported to decline [49].

Table 3.1: Late effects among HL survivors. Relative risk (RR) refers to standardized mortality ratio (SMR) or corresponding measure of risk in relation to the general population. Absolute excess risk (AER) is defined as the observed number of deaths minus the number expected, divided by person-time at risk in the patient cohort.

| **Author (year)** | **N** | **Diagnosed** | **Outcome** | **RR** | **AER**[1] |
|---|---|---|---|---|---|
| Aleman (2003) | 1,261 | 1965-1987 | CVD death | 6.3 *(sign.)* | 17.8 |
| Hull (2003) | 415 | 1962-1998 | CABG | 2.4 (1.1-3.7) | - |
| | | | Valve surgery | 8.4 (3.2-13.6) | - |
| Aleman (2007) | 1,474 | 1965-1995 | CAD | 3.2 (2.7-3.7) | 61.7 |
| | | | MI | 3.6 (2.9-4.4) | 35.7 |
| | | | CHF | 4.9 (3.6-6.4) | 25.6 |
| Swerdlow (2007) | 7,033 | 1967-2000 | MI death | 2.5 (2.1-2.9) | 125.8 |
| Myrehaug (2008) | 615 | 1988-2000 | CVD | 1.9 (1.2-3.0) | 35.6 |
| | | | MI | 1.9 (1.0-3.6) | 18.2 |
| Andersson (2009) | 6,946 | 1965-1995 | CAD | 1.6[2] (1.4-1.8) | - |
| | | | CHF | 1.5[2] (1.2-1.9) | - |
| Kiserud (2010) | 557 | 1971-1991 | CVD death | 4.9 (3.1-7.9) | - |
| Castellino (2011) | 2,633 | 1970-1986 | Heart disease death | 12.7 (9.8-16.2) | 13.1 |
| Galper (2011) | 1,279 | 1969-1998 | CABG | 3.2 (2.8-3.5) | 18.2 |
| | | | Valve surgery | 9.2 (8.1-10.3) | 14.1 |
| van Nimwegen (2015) | 2,524 | 1965-1995 | CAD | 3.2 (3.0-3.5) | 70.0 |
| | | | CHF | 6.8 (5.9-7.6) | 58.0 |

*Abbreviations: N, number of included HL patients; CVD, cardiovascular disease; CABG, coronary artery bypass grafting; CAD, coronary artery disease; MI, myocardial infarction; CHF, congestive heart failure.*
[1] Per 10,000 person-years
[2] 10-19 years after treatment for HL

The cardiac toxicity of chemotherapy is mainly driven by anthracyclines (such as doxorubicin included in the HL therapy regimen ABVD). However, other anti-cancer drugs (such as alkylating agents) are also recognized to lead to long-term cardiac sequelae. The clinical manifestation of anthracycline-induced cardiac toxicity is often arrhythmias or cardiomyopathy, which in turn can lead to congestive heart failure (sometimes presenting many years after treatment).

While reductions in RT dose has most probably reduced the risk for RT-associated DCS, contemporary CT treatment still involves high cumulative doxorubicin doses. Most scientific literature addressing the cardiotoxicity of HL therapy is based on patients treated with a combination of CT and RT. Among the few studies addressing the toxicity of CT alone, a Mexican study from 2005 showed that patients treated with epirubicin rather than doxorubicin in ABVD, had better overall survival [57]. Moreover, a Dutch study on patients diagnosed 1965-1995 reported that, irrespective of RT exposure, the rate of heart failure increased 3-fold in patients treated with anthracycline CT [58].

Key findings on cardiac morbidity and mortality among HL patients are summarized in Table 3.1. Most studies are based on patients treated with outdated therapy. As there are many HL survivors still alive

today who were treated with, e.g., extended field RT and remain at risk of these complications, the knowledge gained from these studies is highly relevant. However, there is a lack of studies based on patients treated with contemporary HL therapy investigating whether changes in treatment regimes (i.e., the shift towards less RT and more CT, and especially the anthracycline-containing ABVD) has influenced the incidence of circulatory system diseases. In a Swedish study on patients diagnosed between 1973 and 2006, it was concluded that treatment-related DCS mortality among HL patients have declined since the mid-1980s and "when accounting for competing causes [...], excess DCS mortality constitutes a small proportion of the overall mortality among HL patients" [59]. However, a reduction in treatment-related DCS mortality does not necessarily imply a reduction in treatment-related incidence of DCS, as improved treatment of DCS could explain the mortality reduction.

### 3.3.2 Secondary malignancy (SM)

Although not the focus of this thesis, SMs constitute the other group of serious late effects of HL therapy, and should therefore be covered, albeit in less detail.

SMs in HL patients are traditionally grouped into leukemia, non-Hodgkin lymphoma (NHL), and solid tumors. The risk of a SM has been shown to peak at five to ten years after HL therapy in cases of treatment with CT alone [60]. Irrespective of treatment regimen, risks have been observed to remain increased for more than 25 years [60, 61].

The largest standardized incidence ratios (SIRs) have been observed for leukemia, especially acute leukemia (SIR=20-30)[62–64]. The increased risk is especially seen among patients treated with alkylating agents such as mechlorethamine and procarbazine [61, 65].

HL patients have a 14-16 times higher risk of NHL than the general population [62, 64]. This effect is of the same magnitude irrespective of treatment and remains approximately the same throughout life [63]. Due to the close relationship between HL and other malignant lymphomas, it is unlikely that the HL therapy alone causes the increased risk, other factors such as the underlying biology of lymphoma and shared risk factors, are probably involved as well.

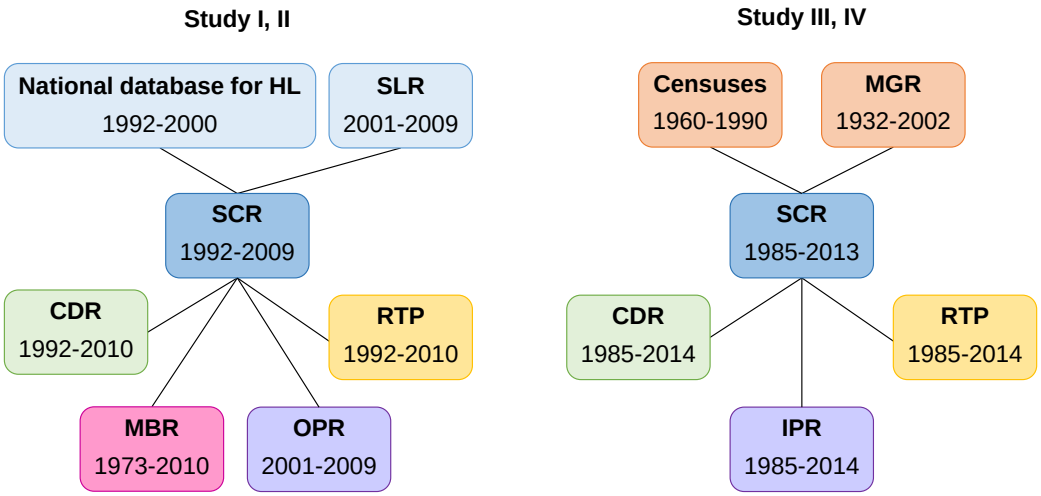
Solid tumors account for 70% to 80% of all SMs in HL patients [62, 64, 66]. The most common sites are breast-, lung-, and gastrointestinal (GI) tract cancers [33, 63]. Considering all SMs together, HL patients experience a 2-3 times higher risk compared to the general population, and the occurrence is related primarily to RT [61]. In a study on breast cancer risk among female HL survivors it was shown that reducing the radiation volume (to limit the irradiation to the breast tissue) significantly lowers the risk, and that gonadotoxic treatments, by inducing early menopause, further reduce the breast cancer risk (among young patients) [48]. In 2015, Shaapveld *et al.* reported a 30-year cumulative incidence of 28.5% for all solid cancers combined; breast cancer 16.6%, lower respiratory 7.1%, and GI tract 7.0% [61].

# 4. Data material

## 4.1 National registers

The studies in this thesis utilized data from several national health and population registers (see Figure 4.1). The Swedish national registers are held at either the National Board of Health and Welfare (Socialstyrelsen) or Statistics Sweden (Statistiska Centralbyrån), except for the Swedish lymphoma register which is managed jointly by six different Regional Cancer Centers (RCCs). The personal identification number (PIN, personnummer) unique to all Swedish residents [67, 68] enabled linkage between different registers.

Figure 4.1: Illustration of the national registers used, and for which years, in Study I-II and III-IV.



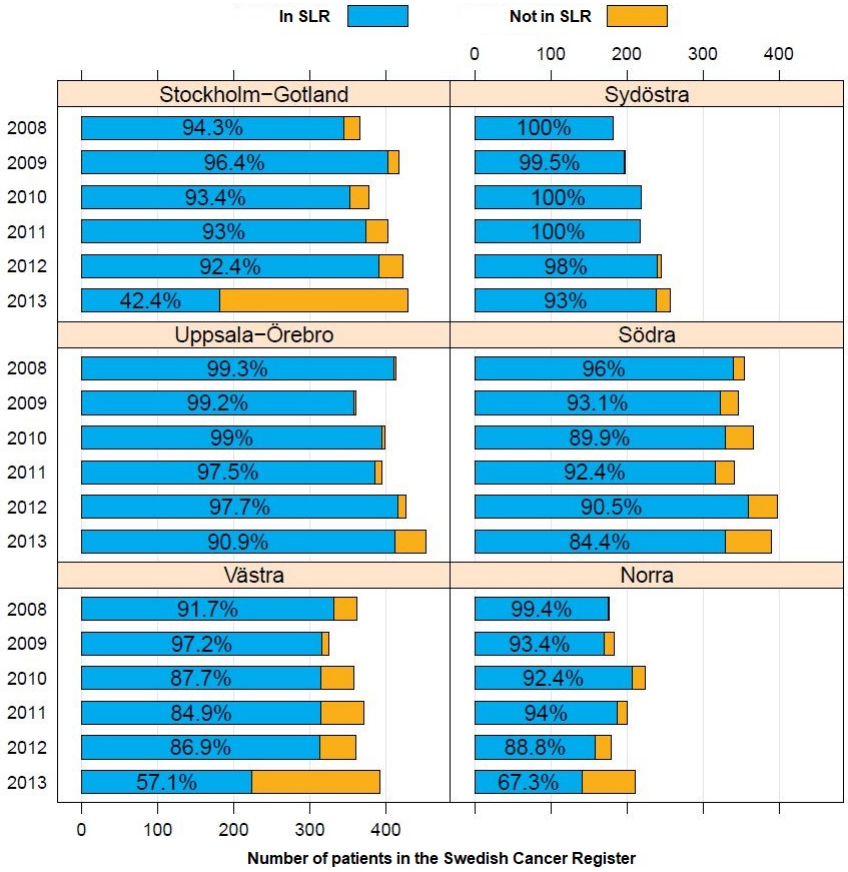

### The Swedish Cancer Register (SCR)

In all four studies, the Swedish Cancer Register (SCR) was used to identify incident cases of HL. The SCR was established in 1958 and contains all newly diagnosed primary cancers in Sweden. Reporting to the register is done by clinicians, pathologists and cytologists and is mandatory by law. The register re-codes all diagnoses to the $7^{th}$ revision of the International Classification of Diseases (ICD). ICD-7 code 201 was used to identify HL. The completeness for HL has been shown to be very high (>95%) [69].

## The Swedish Lymphoma Register (SLR)

The Swedish Lymphoma Group was formed in 1979 with the main purpose to optimize the care of lymphoma patients in Sweden. As part of this, the National Database of Hodgkin lymphoma was initiated in 1992, recording more detailed clinical information (such as stage and treatment) than what could be found in the SCR. Between 1992 and 1998, all Swedish healthcare regions except Stockholm reported to the database, and 1999-2000 all six regions were included. The nationwide Swedish Lymphoma Register (SLR) was established in year 2001, and covers around 95% of all lymphoma cases found in the SCR [70]. Figure 4.2 shows the region-specific coverage between 2008 and 2013. The SLR was utilized to link information on clinical characteristics, especially relapse information, to HL patients in Study I-II.

Figure 4.2: Coverage of the Swedish Lymphoma Register (SLR) between 2008 and 2013 for all types of lymphoma found in the Swedish Cancer Register (SCR). Source: Svenska lymfomregistret, nationell kvalitetsrapport för diagnosår 2013 [70].




## The Cause of Death Register (CDR)

The Swedish Cause of Death Register (CDR) includes data on all deceased Swedish residents, who die in Sweden or abroad, since 1961. Both underlying and contributing causes of death are recorded, as well as date of death. For Study I-II, dates of death were used, and for Study III-IV both dates and the underlying cause of death were extracted.

### The Medical Birth Register (MBR)

Since 1973, all pregnancies resulting in a live birth in Sweden, and all stillbirths delivered after 28 full gestational weeks (January 1973 - June 2008) or 22 gestational weeks (from July 2008), are recorded in the Swedish Medical Birth Register (MBR). The MBR contains information on maternal, pregnancy, and offspring characteristics. Both actual date of delivery and estimated delivery date based on ultra sound is recorded. For Study I, the latter was used to estimate date of conception (est. delivery date - 280 days). In Study II, the estimated delivery date was used as the time point of the outcome.

### The In- and Outpatient Registers (IPR and OPR)

Six of the Swedish counties started recording inpatient visits in 1964, initiating the Inpatient Register (IPR). Successively, remaining counties have followed and the register reached national coverage in 1987. However, already in the early 1980s the coverage was around 95%[1]. For the purpose of this thesis, the IPR was used to extract hospitalizations for DCS (Study III-IV) from 1985 and onward. At that time only the counties Kronoberg and Bohus were not reporting, constituting 2% and 3% of the total population size in 1985, respectively. Since 2001, outpatient visits, including day surgery and psychiatric care, are recorded in the Outpatient Register (OPR). Information on primary care visits is not included in any national patient register. The OPR was used in Study II to retrieve infertility- and fertility preservation diagnoses.

### The Multi-Generation Register (MGR)

The Multi-Generation Register (MGR) contains all individuals born in 1932 or later, who have resided in Sweden at some point since 1961, and their parents. The MGR is useful for creating family structures by linking index persons to their relatives. However, as the register includes essentially everyone residing in Sweden, it can also be used to represent the Swedish general population. For Study III-IV, the MGR was (together with censuses, see below) used for this purpose.

### Censuses

To further ensure that the cohort based on the MGR could serve as a representation of the Swedish general population, individuals included in censuses were added. Between 1860 and 1990, Sweden conducted population and housing censuses to collect information on population size, educational level, and occupation, to name a few.

### The Register for Total Population (RTP)

Information on emigration from Sweden was extracted from the Register for Total Population (RTP) initiated in 1968. The RTP holds data on, e.g., number of childbirths and deaths, marriages and divorces, and is the basis for many official statistics regarding the Swedish population.

[1] Based on population size data from Statistics Sweden and information on reporting from the National Board of Health and Welfare.

## 4.2 The Social Mobility database

The Social Mobility database includes data from the IPR (1964-2010), the CDR (1961-2010), and the RTP (1968-2010) for index persons found in either the MGR (1932-2002) or in any of the Swedish censuses 1960-1990. This database was used in Study III-IV to calculate expected incidence rates of DCS in a cohort representing the Swedish population.

## 4.3 Ethical considerations

Using the Swedish national health and population registers for scientific research is governed by law, and approval from the regional ethical review board is necessary. Except for quality registers, such as the SLR, individuals included in national registers have no possibility to opt out, and might not even be aware that they are included in a range of registers and by that several research projects. It is important to weigh the potential harm (both physical and psychological) versus the benefit of each study.

Although the research included in this thesis has the potential to benefit future HL patients, it does not necessarily benefit previously diagnosed patients. Thus, we need to make certain that no individuals included are harmed, that their personal integrity is not breached, and that the research is sufficiently important to motivate the use of the data. To ensure that data is handled as safe and respectful as possible, registers are held at Governmental institutions (such as Statistics Sweden and the National Board of Health and Welfare), and only de-identified data is accessed at the researchers end (i.e., the personal identification numbers are removed), and small (generally less than 5) cell counts are not presented in the published reports.

# 5. Statistical methods

## 5.1 Survival Analysis

### 5.1.1 Key concepts and measures

Survival analysis is a field in statistics dealing with *time to event* data. Examples include studying time to discharge after admittance to hospital, or time to death following a diagnosis of cancer. Individuals are followed over time, for which an underlying time scale (such as time since diagnosis or attained age) is predefined. One key feature of survival analysis is the presence of incomplete observations in terms of outcome status. This means that for one reason or another, individuals are not followed long enough for them all to experience the outcome of interest. This is referred to as (right) *censoring*, and requires special statistical concepts and methods. There are other types of censoring, such as left- and interval-, but they are beyond the scope of this thesis. In case of right censoring, it is only observed that the individual has not experienced the outcome up until some point in time. It is further assumed that, conditional on adjusting variables, the individuals who remain under study represent those who were censored.

Individuals do not necessarily have to be in the study from time zero. This is known as *delayed entry*. For example, when attained age is the time scale, and study participants enter at some time point after birth, this will be the case for the full cohort.

If we denote by $T_E$ the time until an event of interest, and $T_C$ the censored survival time, we will only observe $T = \min(T_E, T_C)$. For each study participant $i$, data is represented in pairs $(t_i, \delta_i)$ where $t_i$ is the observed value of $T$ for that individual and $\delta_i$ an indicator of whether the outcome occurred ($\delta_i = 1$) or not ($\delta_i = 0$).

The density function, $f(t)$, and cumulative distribution function, $F(t)$, can be used to characterize the distribution of $T$. $F(t)$ is simply the integrated $f(t)$ and assuming that the censoring is non-informative[2], it represents the probability of experiencing the outcome before time $t$. In epidemiological applications this is sometimes referred to as the *cumulative incidence function* (CIF) [71, 72]. Two other functions are considered the corner stones of survival analysis – the survival function, $S(t)$, and the hazard function, $h(t)$. The survival function is a monotonic decreasing function with $S(0) = 1$ and $\lim_{t\to\infty} S(t) = 0$. It describes the *probability of being event free at time t*, and has a one-to-one relationship with the

[2]This assumption will be further covered in Section 5.2.

cumulative distribution function (and density function):

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) = 1 - \int_0^t f(u)\,\mathrm{du}. \tag{5.1}$$

The hazard function is the instantaneous rate of failure, and is defined as:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \tag{5.2}$$

The term "hazard" is generic and is typically exchanged with incidence or mortality in medical applications, depending on the outcome. By integrating the hazard function we get the *cumulative hazard function* $H(t)$, which is related to the survival function as:

$$H(t) = -\log S(t). \tag{5.3}$$

The cumulative hazard function describes the *accumulated risk up until time t*, and although hard to interpret in itself, it is useful for modeling purposes.

The most commonly used non-parametric estimator of the survival function in medical applications is the Kaplan-Meier estimator [73]. For testing purposes, the log-rank test (which compares survival curves) can be applied. Similarly, the Nelson-Aalen estimator of the cumulative hazard function (described, e.g., in [74]) is widely used. The studies in this thesis are all observational studies. As opposed to randomized trials, such studies are typically not balanced with regards to confounding factors. Thus, a modeling approach is preferred to non-parametric methods.

### 5.1.2 The Cox proportional hazards model

The Cox model [75] is the best known of all survival models, and is defined as:

$$h(t; \mathbf{x}) = h_0(t) \cdot \exp(\mathbf{x}\boldsymbol{\beta}) \tag{5.4}$$

where $h_0(t)$ is the baseline hazard function, and $\mathbf{x}\boldsymbol{\beta}$ the linear predictor (the covariates and a vector of coefficients). The model gained popularity due to the fact that no parametric shape needs to be assumed for the baseline hazard function, avoiding misspecifications. However, this implies that $h_0(t)$ is not estimated, only hazard ratios (HRs) as measures of relative effects. One key assumption of the standard Cox model defined in Equation 5.4 is the proportional hazards (PH) assumption, meaning that HRs are assumed constant over follow-up time. However, this assumption is not unique to the Cox model, but common to many survival models. The PH assumption can be relaxed in several ways, such as by including interaction terms between covariates and follow-up time.

The Cox proportional hazards model was used in Study I to estimate HRs comparing relapse rates between women exposed to pregnancy and women unexposed to pregnancy. In Study II, the Cox model was used to estimate HRs of childbirth between women with different covariate patterns, separately for two distinct time periods during follow-up (0-3 years and 3-7 years), assuming PH within each period but allowing the HRs to differ between the two time windows.

During recent years parametric models are becoming increasingly used, as they offer an estimate of the rates themselves (as opposed to just relative rates). In 1994, even Sir David Cox himself said [76]:

"In the light of further results one knows since, I think I would normally want to tackle the problem parametrically. [. . .] I'm not keen on non-parametric formulations normally."

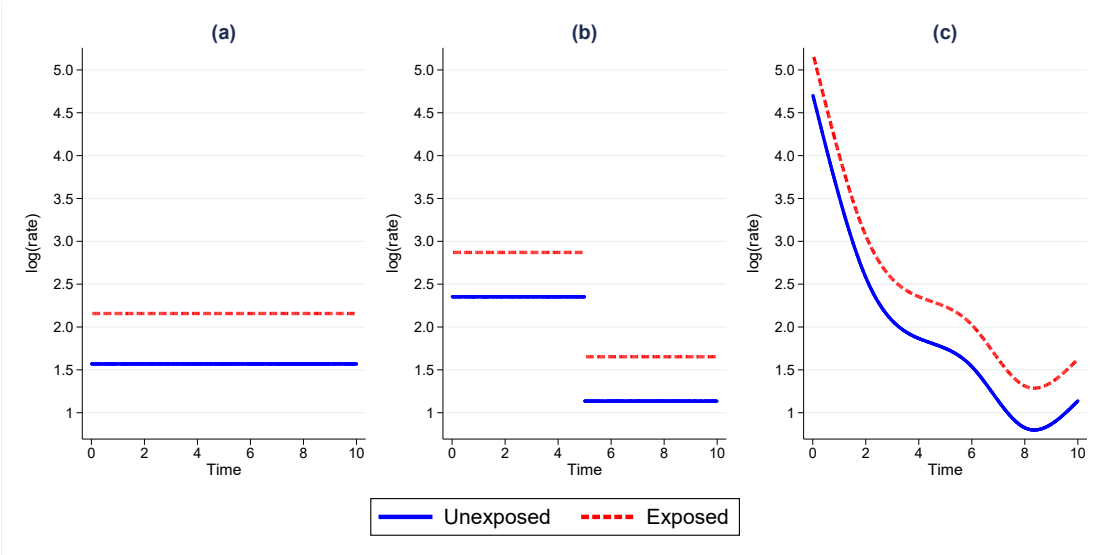
### 5.1.3 The Poisson model

Parametric models enable comparisons between individuals with different covariate pattern on the absolute scale as well as the relative scale. One commonly used model is the Poisson model, which was used in Study IV to estimate expected DCS rates using a population incidence file grouped on calendar year, age, and sex. The Poisson model is defined (on the log scale) as:

$$\log[h(t;\mathbf{x})] = f(t) + \mathbf{x}\boldsymbol{\beta}. \tag{5.5}$$

The baseline rate is modeled with $f(t)$, which can take a range of functional forms. Figure 5.1 shows some common examples of Poisson models (for illustrative purposes, all assume PH).

Figure 5.1: Examples of three commonly used functional forms for the baseline hazard rate in a Poisson PH model setting in case of two exposure levels. Model **(a)** assumes a constant rate across follow-up time, i.e., $f(t) = \lambda$. Model **(b)** is a piece-wise constant rate model, allowing the baseline rate to differ between the two intervals, but assuming constant rate within each interval. In model **(c)** the baseline rate is modeled using restricted cubic splines with three internal knots.




Models where the baseline rate is modeled using a data driven flexible function of time, similar to model **(c)** in Figure 5.1, are becoming more common. On such model, applied in all studies in this thesis, is the flexible parametric survival model.

### 5.1.4 The flexible parametric survival model

The flexible parametric survival model was first introduced by Royston and Parmar in 2002 [77] and has been further developed since [78]. Modeling is done on the log cumulative hazard scale with the baseline rate modeled using *restricted cubic splines*. Splines are mathematical functions defined by piece-wise polynomials which in turn are based on basis functions of increasing degree (up to order 3 for

cubic splines). A restricted cubic spline as a function of $x$, with $K$ *knots* (i.e., joining points), can be written as:

$$s(x;\boldsymbol{\gamma}) = \gamma_0 + \gamma_1 \cdot \nu_1(x) + \gamma_2 \cdot \nu_2(x) + ... + \gamma_{K-1} \cdot \nu_{K-1}(x) \tag{5.6}$$

where $\nu_i$ are the *basis functions*:

$$\nu_i(x) = \begin{cases} x & \text{for } i = 1 \\ (x-k_i)_+^3 - \lambda_i(x-k_1)_+^3 - (1-\lambda_i)(x-k_K)_+^3 & \text{for } i = 2, ..., K-1 \end{cases}$$

with $k_i$ being the i$^{th}$ knot, $(x-k_i)_+ = \max(0, x-k_i)$ and $\lambda_i = \frac{k_K - k_i}{k_K - k_1}$. Before the first ($k_1$) and beyond the last ($k_K$) knot, the spline function reduces to a linear function, which is the key characteristic of a *restricted* spline, providing stability in the tails where data tends to be sparse. The spline function is furthermore forced to have continuous first and second derivatives at the knots, ensuring a smooth function.

A proportional hazards model can be written on the log cumulative hazard scale as:

$$\log(H(t;\mathbf{x})) = \log(H_0(t)) + \mathbf{x}\boldsymbol{\beta} \tag{5.7}$$

where $H_0(t)$ is the baseline log cumulative hazard function. Expressing this function in terms of a restricted cubic spline as a function of log time gives the flexible parametric model:

$$\log(H(t;\mathbf{x})) = s(\log(t);\boldsymbol{\gamma}) + \mathbf{x}\boldsymbol{\beta} \tag{5.8}$$

As the flexible parametric survival model is a proportional hazards model, the effects of covariates are interpreted in the same manner as for other PH models. Extending to allow for time-dependent effects, i.e., relaxing the PH assumption, is easily incorporated using interaction terms in the model.

The main advantage of the flexible parametric survival model, as compared to other parametric models, is the possibility to predict hazard rates and survival functions without pre-specifying a parametric shape of the baseline rate. In cases where the underlying rates have a more complex shape, choosing the appropriate distribution can be difficult. And although it is possible to predict hazard rate from a Cox model, e.g., by using the method of Kalbfleisch and Prentice [79], the baseline rate is treated as a nuisance parameter with those methods, and is thus highly erratic.

For Study I-II, the flexible parametric survival model was used to predict relapse and childbirth rates, respectively. It was the main model underlying all results in Study III, and in Study IV it was used to model all transitions except the expected DCS rate.

## 5.2 Statistical methods for competing risks

So far, the term "survival" has been used without further specification. In studies of cancer patient survival, the entity of interest is often *the probability that the patients will die from their cancer*. Thus, deaths due to other causes can be considered *competing risks*; events that prevents or alters the probability of the event of interest. Depending on the research question at hand, these deaths are accounted for in different ways.

If interest lies in studying etiology or making comparisons between different groups of patients, the

measure of interest is *net* survival[3] – the survival in a hypothetical world where the cancer of interest is the only possible cause of death. For example, net survival is used in studies comparing cancer patient survival between countries. As other cause mortality might differ between countries, this will bias the comparison. Therefore individuals who die due to causes other than cancer are censored from the analysis, and the survival function is interpreted as the proportion of patients still alive at time $t$, given that you can only die due to cancer. Implicitly, this requires the assumption that the time to censoring (including the deaths due to other causes) is independent of the time to death due to cancer, conditional on the covariates included in the analysis. This is typically referred to as non-informative censoring, or the independence assumption.

However, if the study objective is related to risks in a real-world setting where every different cause of death is present, net survival is not a relevant entity. Examples of this can be studies assessing resource allocation, or where risk communication is the endpoint. In such situations, competing causes of death need to be accounted for so that the estimated measures accurately represent the actual survival among the patients, i.e., in the presence of competing risks. Interpretation of the survival function, $S(t)$, becomes difficult in such analyses so traditionally the *cause-specific cumulative incidence functions* (cause-specific CIFs) are presented. These are also known as the cause-specific cumulative probabilities of an event.

Figure 5.2: Illustration of competing causes of death among cancer patients.

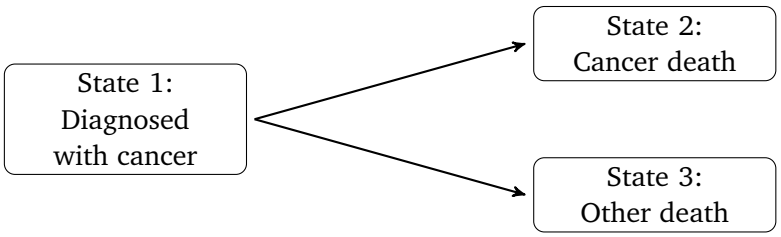



Figure 5.2 illustrates a typical situation with competing causes of death. Individuals start out alive, diagnosed with cancer in State 1. They can then either die from their cancer (State 2) or die from something else (State 3). As you can only die once, the two causes of death prevent each other from occurring. In a study estimating net survival, patients dying from non-cancer would be censored, and the survival proportion would typically be presented. In a competing risks study, two cause-specific CIFs would be presented (the cumulative probability of dying from cancer and non-cancer, respectively).

Similarly, and more relevant for the studies in this thesis, the states in Figure 5.2 do not necessarily have to be different causes of deaths, but can represent incident cases of some illness as well.

The cumulative probability of dying from cause $k$ in the presence of competing risks can be expressed as:

$$\text{CIF}_k(t) = \int_0^t S(u) \cdot h_k(u)\, \text{du} \qquad k \in (1, ..., K) \tag{5.9}$$

where $h_k(t)$ is the cause-specific hazard function and $S(t)$ the probability of not having experienced any of the $K$ events up until time $t$, defined as [80]:

$$S(t) = \exp\Big(-\sum_{k=1}^{K} H_k(t)\Big). \tag{5.10}$$

The cause-specific CIFs can be estimated non-parametrically using the Aalen-Johansen formula [81], or

[3] In the competing risks literature, net survival is sometimes called “marginal survival” [80].

from modeling. If a modeling approach is taken, one of the methods described next is commonly used to obtain estimates of the cause-specific CIFs.

One alternative is to model cause-specific[4] hazard functions, and from these predict the cause-specific CIFs. The standard modeling approach is using a Cox regression model for each cause $k$:

$$h_k(t;\mathbf{x}) = h_{0k}(t) \cdot \exp(\mathbf{x}_k \boldsymbol{\beta}_k) \tag{5.11}$$

where $h_{0k}(t)$ is the underlying cause-specific baseline rate and $\mathbf{x}_k \boldsymbol{\beta}_k$ the linear predictor for cause $k$. As the baseline rate function is not explicitly estimated, it is difficult to obtain a smooth estimate of the cause-specific CIFs. Thus, a parametric model is often preferred. Except for cases where very simplistic models (i.e., the exponential) have been used, the cause-specific CIFs are not analytically intractable, so numerical integration [82, 83] or simulation [84] is typically used instead.

Another option is to model the *subdistribution* hazards using a Fine & Gray model [85]. The difference here lies in the definition of the risk set. When estimating cause-specific hazard rates, individuals who experience **either** of the events are removed from the risk set. However, for estimation of the subdistribution hazard rates, individuals who experience the competing event **remain** in the risk set. The subdistribution hazards (and associated hazard ratios) can be difficult to interpret, but have the advantage that there is a one-to-one relationship with the cause-specific CIFs.

Recently developed methods also suggest modeling the cause-specific CIFs directly, using flexible parametric survival models [86].

Cause-specific CIFs can also be estimated in a relative survival framework, where they are often referred to as "crude probabilities of death" [87–89]. This will be covered in Section 5.3.2.

## 5.3 Relative survival

In population-based cancer patient survival, relative survival is the framework of choice for estimating net survival and the mortality associated with a diagnosis of cancer. The relative survival ratio, $R(t)$, is defined as [90]:

$$R(t) = \frac{S(t)}{S^*(t)} \tag{5.12}$$

where $S(t)$ is the observed all-cause survival in the cancer patient group and $S^*(t)$ is the expected all-cause survival from a comparable group assumed free from the cancer in question. Its analogue on the hazard scale, the excess mortality rate, denoted $\lambda(t)$, is

$$\lambda(t) = h(t) - h^*(t) \tag{5.13}$$

where $h(t)$ is the all-cause mortality rate among the cancer patients and $h^*(t)$ is the expected all-cause mortality rate in a comparable group assumed free from the studied cancer. When the excess mortality equals zero, the patients are no longer experiencing a mortality above that expected in the absence of a cancer diagnosis. The expected mortality is typically retrieved from publicly available population mortality tables ("life tables") stratified by calendar year, sex and age, which can be found in online databases such as `mortality.org`, or via the National Bureau of statistics.

[4] Sometime referred to as component- or transition-specific.

Relative survival has become popular for estimating the mortality associated with cancer as it is does not require information on cause of death. Thus, besides being insensitive to accurate classifications of death, it also captures both direct and indirect (e.g., treatment-related) mortality attributable to the cancer and its treatment.

Several assumptions are made when applying relative survival. One is that the expected mortality, calculated using life tables, is free from variation (thus treated as a fixed covariate). Moreover, cancer patients are assumed to be exchangeable with the cancer-free comparison group, conditional on the stratification variables in the life tables, with respect to non-cancer mortality. There are scenarios where this assumption does not hold. For example, lung cancer patients differ from the general population in terms of lifestyle habits (such as smoking), which in turn increase their non-cancer mortality. To apply relative survival on lung cancer patients, the life tables should be stratified on smoking (and potentially other lifestyle factors as well). In practice however, due to the high lung cancer mortality, the bias is negligible even when not doing so [91]. In situations where interest lies in the effect of a specific lifestyle factor (e.g., smoking) adjusting the life tables becomes more important [92].

Relative survival can be estimated non-parametrically or be modeled using a range of models [89, 93]. For Study III-IV in this thesis, the flexible parametric relative survival model was used to estimate excess incidence rates of DCS.

### 5.3.1 Flexible parametric relative survival models

The flexible parametric survival model has been extended to estimate excess mortality in a relative survival framework, by incorporating the expected hazard into the likelihood [94]. Modeling is done on the cumulative excess hazards scale. In accordance with Equation 5.13, the cumulative hazard can be written as a function of the cumulative expected and excess hazard:

$$H(t) = H^*(t) + \Lambda(t). \tag{5.14}$$

The cumulative excess hazard is expressed similarly to the model in Equation 5.8, and the model is defined as:

$$\log(\Lambda(t, \mathbf{x})) = \log(\Lambda_0(t)) + \mathbf{x}\boldsymbol{\beta} \tag{5.15}$$

with

$$\log(\Lambda_0(t)) = s(\log(t), \boldsymbol{\gamma_0}). \tag{5.16}$$

### 5.3.2 Estimating cause-specific CIFs in relative survival

In the relative survival framework, Cronin *et al.* showed how cause-specific CIFs can be calculated non-parametrically using life-tables for population-based cancer patient survival [95]. Their work has been further extended to modeling by Lambert and colleagues [87]. In 2012, a method for partitioning the excess mortality rates and corresponding cause-specific (cancer and non-cancer death) CIFs into components was proposed [88]. The CIF for cancer-specific death in a relative survival framework can be expressed as:

$$\text{CIF}_{\text{cancer}}(t) = \int_0^t S^*(u) \cdot R(u) \cdot \lambda(u)\, \text{du} \tag{5.17}$$

where $S(t)$ in Equation 5.9 has been replaced with $S^*(u) \cdot R(u)$ and $h(t)$ with its excess counterpart, $\lambda(u)$. For deaths due to other causes, the CIF is given by:

$$\text{CIF}_{\text{non-cancer}}(t) = \int_0^t S^*(u) \cdot R(u) \cdot h^*(u)\, \text{du}. \tag{5.18}$$

Given partitioning of excess mortality rates, $\text{CIF}_{\text{cancer}}$ can be divided further into mortality due to, e.g., cardiovascular disease (CVD) and remaining excess mortality due to cancer. The expression in Equation 5.17 for cause $k$ becomes:

$$\text{CIF}_{\text{cancer, k}}(t) = \int_0^t S^*(u) \cdot R(u) \cdot \lambda_k(u)\, \text{du}, \qquad k \in (\text{CVD, remaining}) \tag{5.19}$$

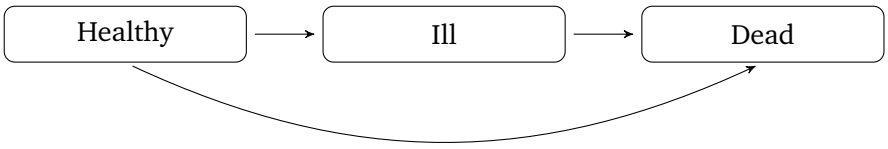
with $\lambda_k(t)$ being the excess mortality rate due to cause $k$. In the example of CVD and remaining excess cancer mortality, three CIFs would be predicted: $\text{CIF}_{\text{non-cancer}}$, $\text{CIF}_{\text{cancer, CVD}}$, and $\text{CIF}_{\text{cancer, remaining}}$.

For the purpose of Study III in this thesis, the CIF of DCS was predicted using the expression in Equation 5.17.

## 5.4 Multi-state models

It is natural to imagine how adding more states to the competing risks situation illustrated in Figure 5.2 would be useful to gain a better understanding of patient trajectories. Typically, we consider multi-state models as a generalization of competing risks models where intermediate states have been introduced. An example of a multi-state model in its simplest form is the illness-death model depicted in Figure 5.3.

Figure 5.3: Illustration of an illness-death model. The initial "healthy" state is generic in the sense that it can represent either healthy individuals or patients with a certain disease (such as cancer patients).




However, also the very simple situation of following cancer patients over time to death can be viewed as a multi-state model with two states (namely "cancer" and "dead"). Essentially, multi-state models can be very simple or extremely complex, depending on the structure and number of states. Before describing estimation in a multi-state model framework, some notation and definitions need to be introduced.

Let $\{Y(t), t \geq 0\}$ be a stochastic process taking on values in the finite state space $\mathscr{S} = \{1, ..., \text{S}\}$. States are (in broad) classified as either *transient* (states that you can both enter and exit) or *absorbing* (states that you cannot leave, e.g., dead). Transient states that can be re-entered are called *recurrent* states. The *history of the process* until time $s$, $\mathscr{H}_s = \{Y(u); 0 \leq u \leq s\}$, consists of all previous observations of the process. The probability of being in state $b$ at time $t$, given that the process was in state $a$ at time $s$ and its history until time $s$ ($\mathscr{H}_{s-}$) is defined as:

$$P(Y(t) = b | Y(s) = a, \mathscr{H}_{s-}) \tag{5.20}$$

where $(a, b) \in \mathscr{S}$. This is the *transition probability*. The expression can be simplified further assuming that future transitions only depend on the current state of the process (and not its history):

$$P(Y(t) = b|Y(s) = a, \mathscr{H}_{s-}) = P(Y(t) = b|Y(s) = a). \tag{5.21}$$

This equality defines the **Markov property** and is a central part in multi-state modeling. In the special setting when everyone starts in state $a$ at time $s = 0$, the transition probability simplifies further into the *state (occupation) probability*; the probability of being in state $b$ at time $t$. The hazard rate of going from one state to the next is known as the *transition intensity*, and for a Markovian process it is defined as:

$$h_{ab}(t) = \lim_{\Delta t \to 0} \frac{P(Y(t + \Delta t) = b|Y(t) = a)}{\Delta t}. \tag{5.22}$$

The transition intensity essentially has the equivalent interpretation as for survival models in general; the instantaneous probability of going from state $a$ to state $b$, given that you were in state $a$ at time $t$. The collection of transition intensities, for all possible transitions, governs the multi-state process.

In situations where the Markov assumption is unrealistic, it can be relaxed so that the future of the process depends not only on the current state, but also on the time point at which the current state was entered. Such models are known as *semi-Markov* models, and the transition probability is given by:

$$P(Y(t) = b|Y(s) = a, \mathscr{H}_{s-}) = P(Y(t) = b|Y(s) = a, T_a) \tag{5.23}$$

where $t$ is the time since starting state and $T_a$ is the time at which state $a$ was entered. Modeling the transition rate from state $a$ to $b$ is done in a delayed entry model setting, adjusting for $T_a$ as a fixed covariate. In practice, it is common to not only take into account the time point at which the current state was entered, but have time since entry to current state as the time scale. The resulting process is a *Markov renewal*, or *clock-reset*, process with $t - T_a$ as the underlying time scale. This simplification is often suitable in situations where state $a$ represent a more severe event, such as recurrence of cancer or events associated with a high initial mortality. As an example, imagine an illness-death model following cancer patients from diagnosis ("Healthy") until they suffer from CVD ("Ill") and/or death ("Dead"). The mortality rate will depend on not only which state the process is currently in (cancer or CVD), but also on when the current state was entered. More importantly, if time since CVD diagnosis is believed to be of greater importance than time since cancer, a clock-reset approach might be appropriate for estimating the transition from CVD to dead.

More complex alternatives of the semi-Markov model includes modeling one or more transitions with multiple time scales. In the example above, time since CVD and cancer might both be important for the mortality rate after CVD. If so, a second time scale can be added to that transition model. Including more than one underlying time scale is by no means exclusive to multi-state modeling. In case of a single outcome, modeling with multiple time scales, e.g., adding attained age as a second time scale to time since diagnosis, is not uncommon.

Although multi-state models can be extremely complex and usually carry specific assumptions related to the transitions (such as the Markov property), estimation of transition intensities is straight-forward, as every transition can be viewed as a survival model. This means that the same methods as for competing risks modeling can be applied. However, it is more complicated to predict transition probabilities.

For some simplistic parametric models, such as a constant or piecewise constant rate model, the transition probabilities can be calculated analytically using maximum likelihood methods [96]. However, for

more complex models, this is not possible. Several alternative approaches have been suggested, including numerical integration [97, 98] and ordinary differential equations [99]. In more recent papers, a simulation-based approach together with parametric transition models has been proposed [84, 100]. Simulation methods can be seen as superior to the other alternatives, as they are less computer-intensive compared to numerical integration in case of complex multi-state models.

Until this point, no implemented methods have existed for combining multi-state modeling with relative survival to estimate excess transition rates to transient states. Existing methods are somewhat limited as they only allow for estimating excess mortality rates (i.e., methods for relative survival are applied on absorbing states alone) [101–103]. The method proposed in Study IV in this thesis incorporates estimation of excess incidence rates into a multi-state model. Additionally, different time scales can be applied to different transitions, and it is possible for states to share transition model.

# 6. Results

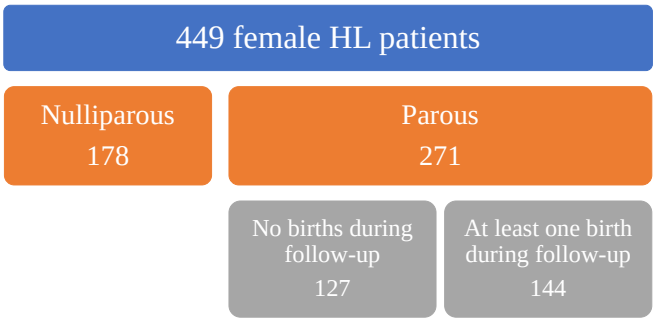
In summary, the main findings were:

- Among women in remission from HL, pregnancy does not increase the risk of relapse [Study I].
- The prospect of having children after being diagnosed with, and treated for, HL has improved among female survivors over calendar time, and is comparable with that of the general population three years after diagnosis [Study II].
- The incidence of treatment-related DCS among Swedish HL patients declined between the mid-1980s and mid-1990s, but since then small to no improvements have been seen [Study III].
- Incorporating estimation of excess incidence in multi-state models offers a possibility to study treatment-related morbidity and mortality among cancer survivors simultaneously, which can enhance the understanding about underlying mechanisms and total burden of disease [Study IV].

## 6.1 Study I

This was the first scientific study to address if a post-diagnosis childbirth could have an impact on HL relapse risk. For descriptive purposes, women were classified as either nulliparous (no childbirths before diagnosis or during follow-up), parous without births during follow-up (meaning that they were parous already at HL diagnosis), or parous with at least one birth during follow-up. The parity distribution among the 449 women in the cohort, aged 18-40 at diagnosis, is illustrated in Figure 6.1.

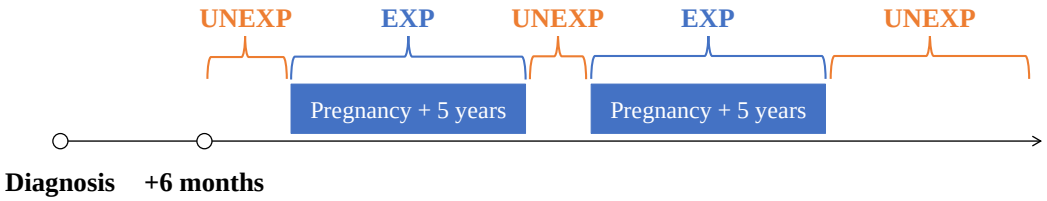
Figure 6.1: Parity status at the end of follow-up.




For the main analysis, pregnancy was considered a time-varying exposure. Follow-up started six months after HL diagnosis. All 144 women with a childbirth during follow-up contributed with exposed person-time during their pregnancies, and for another five years after delivery. Relapses occurring during that

time window were noted as “pregnancy-associated”. In studies of pregnancy-associated cancer, a time window of 1-2 years after delivery is commonly used [104]. However, here it was extended to include long-term effects of the pregnancy and to allow for time to register the relapse. Figure 6.2 gives an example of exposed and unexposed time windows for a woman with two childbirths during follow-up.

Figure 6.2: Example of pregnancy as a time-varying exposure. Abbreviations: UNEXP, unexposed; EXP, exposed.




In total, 47 women had a relapse during follow-up. However, only one woman had her relapse within the predefined pregnancy exposure window. Under the assumption that women exposed to pregnancy would experience the same relapse rate as women not exposed to pregnancy, the expected number of relapses was 3.76. Taken together, this suggests that in relation to risk of relapse, it is safe for female HL survivors to have children after completed treatment. However, as the absolute risk of relapse is at its highest level during the first 2-3 years after diagnosis, delaying childbirth is recommended when possible.

## 6.2 Study II

This study described trends in childbearing patterns over time among female HL survivors. Childbirth rates were compared between patients with different clinical characteristics, and between patients and HL-free women from the general population (“comparators”). Moreover, cumulative probabilities of a childbirth (i.e., CIFs) were calculated non-parametrically for patients and comparators in the presence of the competing events of death and/or relapse.

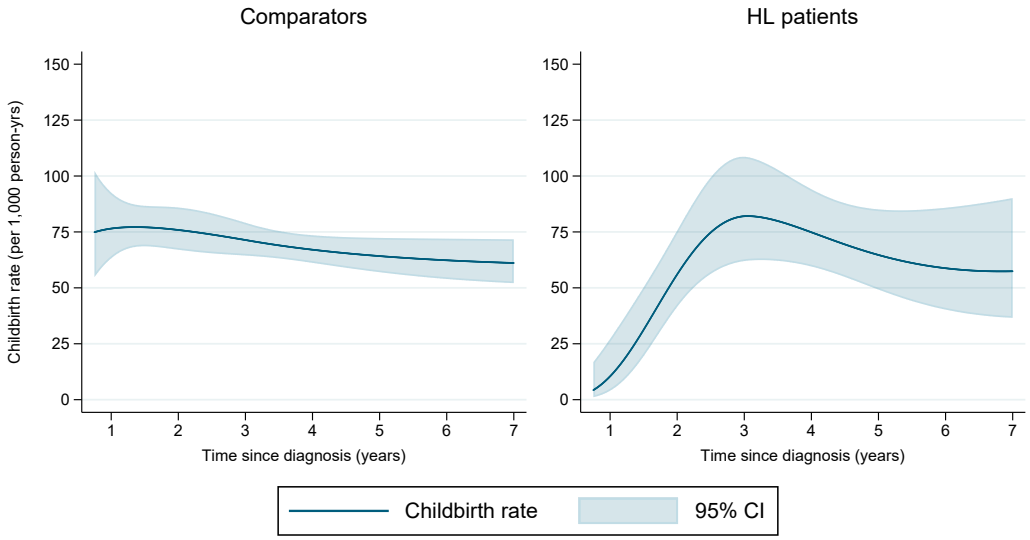
As shown in Figure 6.3, childbirth rates among HL patients varied considerably during follow-up. Due to the long treatment period (six months to one year), childbirth rates are expected to be low during the first years following diagnosis. As such, it is not reasonable to assume a constant HR (childbirth rate among HL patients divided by the rate among comparators) across the entire follow-up. HR’s were therefore estimated separately for two distinct time periods (0-3 years and 3-7 years) during follow-up.

Childbirth rates and the cumulative probability of having children after a diagnosis of HL increased over calendar time in this cohort of Swedish female HL survivors. Although rates were generally lower in relation to matched comparators the first years after completed treatment, no differences were observed three years or more after HL diagnosis, irrespective of clinical characteristics at diagnosis and treatment.

One objective of this study was to gain insight into the real-world possibilities of having children among ABVD- versus BEACOPP-treated patients, in the presence of competing risks. Reassuringly, also women treated with BEACOPP had an increasing cumulative probability of childbirth, albeit at a lower level than comparators.

During 2001-2009, when information on fertility preservation (from specialist visits in the OPR) could be linked to the HL patients, the proportion of women with a fertility referral or preservation around

Figure 6.3: Childbirth rates among comparators and HL patients. Rates were predicted from a flexible parametric survival model, where the exposure was having a diagnosis of HL, and the effect of being an HL patient was assumed non-proportional over follow-up.




the time of diagnosis, was only 6.5%. No woman in the cohort had a childbirth after relapse during the first seven years after diagnosis, stressing the importance of curative first-line treatment, and that relapse is an important factor to incorporate in studies of childbirth potential after HL.

## 6.3 Study III

The purpose of this study was to investigate temporal trends in treatment-related incidence of DCS among HL patients. Importantly, the risk of DCS was separated into “excess” (related to the HL and its treatment) and “expected” (in the absence of HL), complementing reports that show high risks of DCS without disentangling the risk patients would have faced also in the absence of HL.

Methods in relative survival were used to obtain estimates of excess incidence. No publicly available population incidence file for DCS exist, so this had to be constructed from individually level data. As expected rates are treated as fixed parameters (implying without variance), they have to be based on a large number of individuals, preferably the whole population. Therefore, a very large cohort (n=10,020,379), representing the Swedish population, was followed over time for incident cases of DCS. Before estimating the rates, data was grouped on calendar year, age, and sex. Rather than incorporating the empirical rate (essentially the number of events divided with the amount of person-time for each combination of the year, sex, and age) in the relative survival model a modeling approach was taken. Three different Poisson models for the expected rate were evaluated:

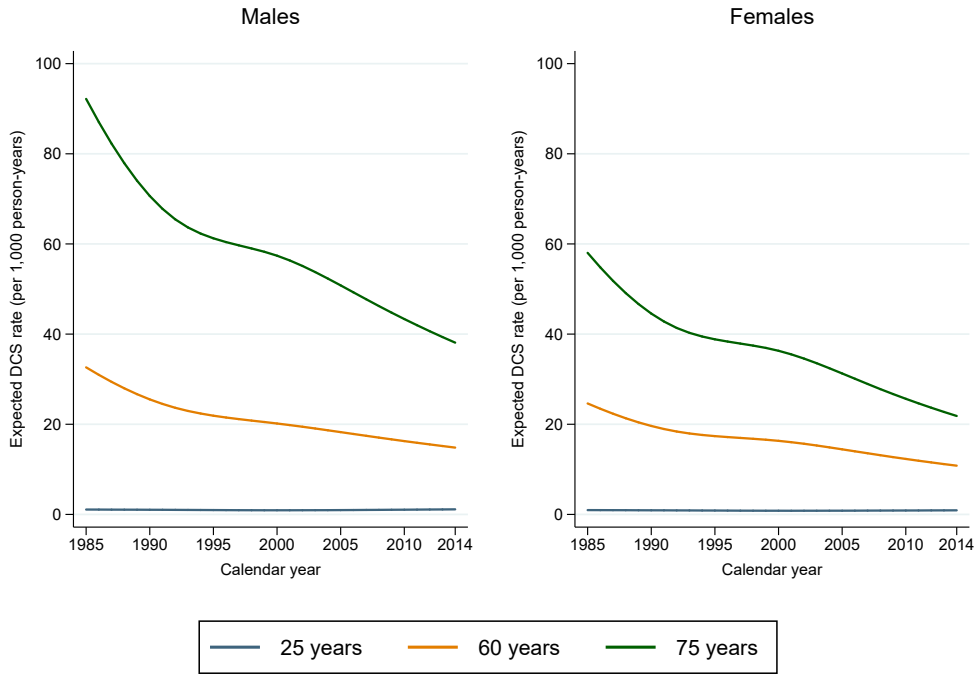
1. A main effects model where age and calendar year were the time scales (parameterized using restricted cubic splines).

2. A two-way interactions model, allowing for interactions between age and sex, age and calendar year, and sex and calendar year.

3. A three-way interactions model, allowing for all two-way interactions plus one interaction between all three covariates.

For the final analyses, the second model was chosen. Figure 6.4 shows the temporal trends of expected rates for the three age groups of HL patients presented in the main results in the study. Expected rates of DCS declined over calendar time, which is in accordance with official statistics on DCS incidence in Sweden[5].

Figure 6.4: Expected rates of DCS over calendar time for males (right panel) and females (left panel) aged 25, 40 and 75 years, predicted from a Poisson model with 2-way interaction terms.




Excess incidence rates, interpreted as treatment-related, of DCS were estimated indirectly from observed and expected rates using a flexible parametric relative survival model. Treatment-related incidence rates of DCS declined for patients aged 25 and 60 years at diagnosis between the beginning of the study period and the mid-1990s. After that point, no further reductions were observed. For elderly patients (aged 75 years at diagnosis), no improvements were seen. Moreover, the risk of a treatment-related DCS was seen to persist for up to 10 years among HL patients who completed their treatment in the new millennium.

## 6.4 Study IV

Relative survival methods can be used to estimate excess incidence of some disease among cancer patients, interpreted as the incidence above and beyond that expected in the absence of cancer. Under certain assumptions, the excess incidence can be interpreted as treatment-related incidence, a measure that is difficult to capture using other methods. In Study III, treatment-related DCS among HL patients was investigated. However, the incident DCS state, which in practice is a transient state, was treated

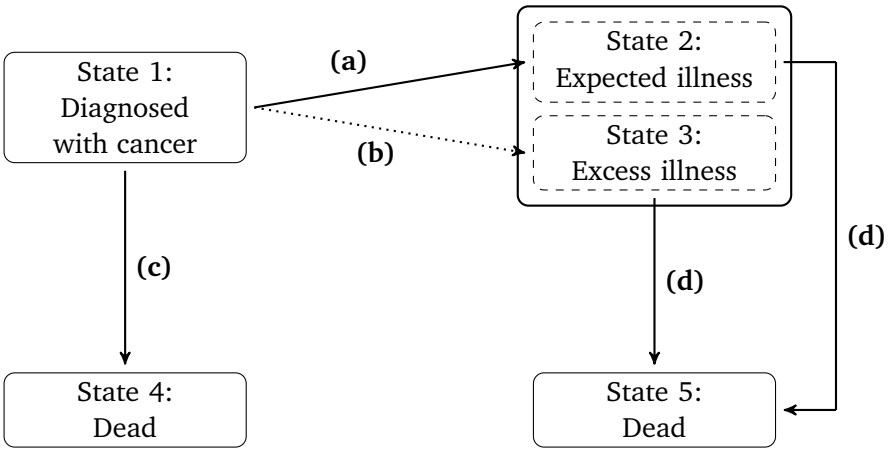
[5]Based on data from http://www.socialstyrelsen.se/statistik/statistikdatabas

as an absorbing state. At the time of that study, no methods existed to estimate excess incidence in a multi-state model framework. Thus, the aim of Study IV was to combine methods in relative survival with those in multi-state modeling to enable prediction of excess transition probabilities.

Using recently developed simulation-based methods in multi-state modeling, a method incorporating estimation of excess incidence rates in a multi-state model was suggested. The proposed approach extended existing methods by (with reference to Figure 6.5 below):

- Including a model with multiple time scales in the multi-state model. (Transition **(a)**).
- Indirectly, via the observed and the expected transition rates, incorporate estimation of the excess transition rate. (Transition **(b)**).
- Let two distinct states share the same transition model. This is necessary, since the transient disease state contains two theoretical states, each with its own transitions into the state (the expected and the excess). However, once the patients have entered the state, these two theoretical states are indistinguishable. (Transition **(d)**).
- Allowing for different transitions in the multi-state model to have different time scales. (Transition **(a)** has calendar time and attained age as time scales, transition **(b)** and **(c)** have time since cancer, and transitions **(d)** has time since illness).

Figure 6.5: Illustration of a multi-state model with a transient state, partitioned into expected and excess illness.




Once the transition models are in place, several measures can be predicted, such as transition probabilities (and function thereof) and length of stay in a certain state.

The example in this study was, for illustrative purposes, modeled with possibly over-simplified transition models. Thus, interpretation of the predicted transition probabilities presented should be avoided.

All modeling was done in Stata, using the `multistate` and `merlin` packages.

# 7. Discussion and conclusions

## 7.1 Childbearing in relation to HL

There are several reasons why reproductive factors are interesting to study in relation to HL. By observing the pattern in Figure 3.3 it is striking to see how the incidence curves for males and females follow each other closely up until the childbearing ages, after which they start diverging. Additionally, males have worse prognosis than females (although this is true for most types of cancer [105]). There is also a plausible link related to the immune system. Immunodeficiency is related to HL risk, and pregnant women alternate between having a strong and a weak immune system, depending on phase. For example, during the first trimester, the immune system is lowered to avoid rejection of the fetus.

So what is known about the relationship between childbearing and HL in women? High parity does not appear to influence the risk of HL, nor does a recent pregnancy affect prognosis. Being diagnosed with HL during pregnancy can be a complication in terms of treatment strategy, but does not otherwise affect survival. Given these null findings, there has been a common consensus that a post-diagnosis pregnancy does not trigger relapse among female HL survivors. However, no scientific evidence had existed prior to Study I in this thesis. The findings therein confirmed the common impression that in relation to relapse risk, it is safe to become pregnant after HL. In the latest version of the Swedish clinical guidelines for HL, information based on these results has been added, providing data for clinicians to rely on when communicating with their patients.

HL therapy can result in temporary or permanent infertility. The negative effects of HL therapy on fertility has been one of the motivations for developing new, less toxic, treatments, and contemporary treatments are considered less harmful than those used historically. Study II showed that the childbearing potential among female HL survivors was in line with, or close to, that of matched comparators.

One could argue that with this in mind, fertility counseling is both a waste of resources and that it might even cause unnecessary concern for the patients. However, individuals who have just received a cancer diagnosis, should get all the facts related to the treatment they are about to undergo. Moreover, fertility counseling does not necessarily imply referral to a fertility clinic, it can simply be information given by the oncologist or a well-informed nurse at the oncology clinic. It is also important to stress that, at time of diagnosis, it is not possible to determine who will relapse and not. Since treatment for relapse carries such a high risk of infertility, the need for fertility advice remains.

Studying *fertility*, the capacity to produce offspring, in a register-based setting is a very difficult task. What can be captured in population registers and compared between groups of women is parity, the number of children born. Note that, in demography, the definitions are different. In demography, fertility refers to offspring actually produced whereas capability to produce is termed fecundity. In this thesis, the

more generally understood definition of fertility is used, i.e., the capacity to produce offspring. For Study II, differences in childbirth rates between female HL patients does not necessarily reflect differences in fertility; only conclusions related to actual childbearing should be drawn. It is also worth noting that studying parity is not without complication. Nulliparous women are a heterogeneous group. Some are not in a "steady" relationship, some are but do not want to have children, and some have a wish for children but have not (yet) succeeded.

Information on terminated pregnancies was due to privacy protection of sensitive data not available for research purposes until recently (October of 2016). This is relevant especially to Study I, where it is possible that women in early stages of pregnancy who relapse, choose to terminate their pregnancy.

To conclude, pregnancy-associated relapse does not need to be taken into account when counseling patients on future reproductive plans. If possible, patients could be advised to wait until the risk of a relapse is lower. The probability of having children after finished HL therapy is in general good. However, patients should receive information on fertility preservation options and the effects of second line treatment on fertility.

## 7.2 Treatment-related morbidity and mortality

Study III in this thesis was not the first attempt to address late effects of HL therapy. Many previous studies have, based on data with detailed treatment information, showed how the risk of, e.g., CAD increases with increasing amount of irradiation to the chest. However, there is a gap in knowledge related to the difference between the observed risk and the risk related to treatment. Imagine that you follow HL survivors for fifty years – they are more or less bound (due to aging) to experience either circulatory system disease, a secondary cancer, and/or death. The key question is: how much of the observed risk can be attributable to the HL therapy? Of course, this is a very difficult outcome to capture. It is not noted in the patient records whether or not a heart attack is associated with previous cancer therapy, since it is not possible to separate between different types of heart attacks in that way.

Relative survival is the gold standard for studying population-based cancer patient survival. Recent methods have enabled estimation of cause-specific CIFs as well. It is natural to extend this methodology to study excess incidence, as a measure of cancer-related incidence. Study III used sophisticated relative survival models to estimate excess incidence rate ratios. A user-written Stata package was used to predict the cumulative probabilities (i.e., CIFs) of treatment-related ("excess") DCS, and DCS expected also in the absence of HL ("expected").

Interpreting excess risk as treatment-related risk should be done with caution. There are two layers of assumptions that need to be fulfilled. Firstly, when interpreting the excess risk as "the risk above and beyond that expected in the absence of cancer", the cancer patients are assumed to be exchangeable with the general population (on which the expected rates are calculated), conditional on the life table stratification variables. Secondly, interpreting the excess risk as treatment-related assumes that the cancer itself is not affecting the risk.

The fact that treatment-related DCS incidence did not decline after the mid-1990s might reflect that treatment strategies have been more or less unchanged after that point and until the end of the study period. It could also be a result of a more aggressive strategy along the lines "cure the HL and take care of the heart disease later". Taken together, this stresses the importance of continuous follow-up of

patients treated for HL and primary prevention (stress tests, healthy lifestyle, etc.). Investigating late effects of HL therapy is an ongoing effort; as new treatments develop, new scientific real-world evidence of safety is needed.

Related to this is the methods development done in Study IV in this thesis. Development and implementation of statistical methods in medical research is crucial. As patients are now surviving their cancer to a greater extent, and are at risk of treatment-related morbidity and mortality, new tools are needed to study survival and survivorship. The possibility to estimate excess transition rates in a multi-state model enables studying late effects in the bigger picture where patients are at risk of multiple events. Importantly, there is one reservation to using the suggested approach. Any estimation of excess hazards requires an estimate of the expected hazard. To achieve this, an appropriate life table (with information on incidence or mortality) for a comparable cancer-free group of people is needed. When the outcome is excess mortality among cancer patients who are still in the starting state of cancer, these are publicly available. However, if the outcome is excess incidence of a specific disease, they might need to be constructed using individual level data which can be difficult to access. The same goes for studying excess mortality among patients diagnosed with, e.g., DCS. The appropriate life table would in such cases need to be based on a cohort of all individuals in the country with a diagnosis of DCS, including information on deaths.

In conclusion, recent improvements in relation to the risk of treatment-related DCS among HL survivors are absent, calling for continued efforts towards less toxic treatments and primary prevention strategies. Methods for studying patient trajectories are useful for answering questions related to long-term risks of treatment-related morbidity and mortality.

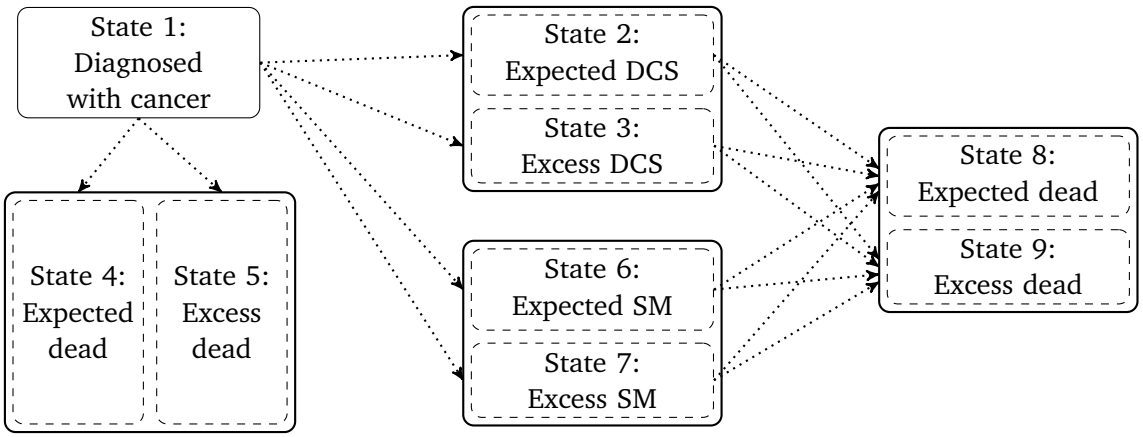
# 8. Future perspective

A lot remains to be done in relation to the questions touched upon in this thesis.

From the findings in Study II, we observe that none of the women who went through second line treatment after a relapse had children afterwards. Relapse in HL is fairly common, in our cohort 10% relapsed. It is possible that the follow-up (restricted to the first seven years after diagnosis) was not long enough to record children born to these women later, at a point where their fertility had recovered. To re-visit the issue of childbearing among relapsed women could reveal more information on these women's possibility to have children. It would also be interesting to take a closer look at the women who were treated with BEACOPP, potentially with longer follow-up, and more detailed information on fertility preservation.

I remain intrigued by the difference in incidence between males and females. The effect of reproductive factors (parity and age at first birth) on HL incidence has been studied before, and the consensus is that there is no relationship. The most recent study on Swedish data was published twenty years ago, in 1998 [18]. It would be interesting to reproduce those analyses on a larger cohort, possibly including subtype of HL. Additionally, by analyzing and presenting two characteristics of exposure jointly [106, 107], comparison of specific covariate patterns (e.g., women with 2 childbirths aged 25 at first birth versus nulliparous women) is made, which could add to the knowledge on reproductive factors in relation to HL risk.

Figure 8.1: A multi-state model including transitions to excess and expected: mortality (state ), diseases of the circulatory system (DCS), secondary malignancies (SM), and mortality after either DCS or SM.




However, the future work closest to my heart is applying the multi-state model developed in Study IV to the application that motivated it: late effects among HL patients. As mentioned, HL survivors are

at risk of a multitude of different complications, all at the same time. Studying these in isolation does not give a complete picture. Combining different events into composite outcomes (such as incident plus fatal cases of DCS, or DCS plus SM) is not the solution. Multi-state models offers the possibility to follow patients through life, via all kinds of different transient states. With the extension suggested in Study IV in this thesis, such a multi-state model can now capture treatment-related morbidity and mortality. An additional extension to our proposed model includes partitioning the absorbing state of death into excess and expected mortality. In the bigger picture, several late effects would be included, such as SM, all in competition with each other. Combining all of these extensions could result in the multi-state model illustrated in Figure 8.1. It looks complex, but given the appropriate population life tables, this is just a collection of survival models.

## Final words

While improving cancer patient survival is the first and most important step, improving survivorship is the natural next step. In the studies included in this thesis, we have shown that, in relation to relapse risk, it is safe to have children after a diagnosis of HL. And encouragingly, we have seen that the possibility of a future childbirth among HL survivors has improved over time. The risk of treatment-related disease persists, but it remains to be studied in a more complex setting to be better understood. Studying survivorship is a huge and very diverse task, addressing many different measures of quality of life. Real-world evidence of risks and complications among cancer survivors are vital to find areas that need improvements. It is my hope that the four studies included in this thesis in some way have contributed to the knowledge on life after cancer.

# 9. Sammanfattning på svenska

Det övergripande syftet med denna avhandling var att undersöka och ämna besvara frågor av relevans för människor som har överlevt Hodgkins lymfom (HL), en typ av blodcancer som uppkommer i lymfkörtlarna. Till skillnad från många andra cancertyper, är detta en sjukdom som inte bara drabbar äldre personer, utan även unga män och kvinnor. Runt 200 nya fall diagnosticeras varje år i Sverige, och även om HL är en ovanlig cancerform i stort, så är den i topp tre av cancer bland unga.

Historiskt sett har överlevnaden i lymfom varit låg. Lyckligtvis har stora framsteg gjorts vad gäller både stråk- och cytostatikabehandling under de senaste femtio åren, och idag botas över 90% av patienter under 65 år vid diagnos. Detta har resulterat i ett ökat antal unga människor som diagnosticerats med, behandlats för, och överlevt HL. Därför är frågor relaterade till livet efter cancer, såsom barnafödande och behandlingsrelaterad sjukdom, mer och mer aktuella.

**Studie I** syftade till att besvara huruvida en graviditet bland kvinnor som friskförklarats från HL påverkar risken för recidiv (återfall). Bland de 449 kvinnor som ingick i studien, födde 144 stycken barn under uppföljningstiden. Totalt sett recidiverade 47 kvinnor, dock var det bara en av dessa kvinnor som fick sitt recidiv kort efter en graviditet. Resultaten från denna studie pekar därför mot att en graviditet efter HL ej påverkar risken för recidiv.

I **Studie II** studerades trender i barnafödande efter behandling för HL, då både cytostatika- och strålbehandling kan påverka ens fertilitet. Eftersom det inte alltid är möjligt att genomföra fertilitetsbevarande åtgärder för kvinnor, är det viktigt att undersöka om chanserna att få barn förbättras i takt med att nya typer av behandlingar introduceras. Denna studie kunde visa att tre år efter avslutad behandling har kvinnliga HL-patienter samma takt på barnafödande som kvinnor i allmänhet. Även patienter som behandlats med en mer toxisk typ av cytostatika hade chans att få barn efter HL.

**Studie III** undersökte om risken att drabbas av behandlingsrelaterad hjärt-kärlsjukdom minskat över tid bland svenska HL-patienter. Då stråk- och cytostatikabehandling kan öka risken att drabbas av exempelvis hjärtattack och stroke, har utvecklingen av nya behandlingsmetoder till stor del fokuserat på att minska allvarliga biverkningar. För att fånga upp hur stor del av sjukdomsrisken som kan antas vara till följd av behandling för HL, tillämpades avancerade statistiska metoder. Denna studie visar att riskerna minskade från mitten av 1980-talet och fram till mitten av 1990-talet, men sedan dess har inga större förbättringar skett.

Canceröverlevare är inte bara vid risk för en, utan flera typer av behandlingsrelaterad sjukdom, samt död, samtidigt. I **Studie IV** utvecklades metoder för att studera sannolikheten att drabbas av olika typer av behandlingsrelaterad sjukdom i en så kallad flertillståndsmodell, där man tar hänsyn till risken för att dö.

# 10. Acknowledgements

The last couple of years have truly been fantastic, and I am extremely grateful for the opportunity to pursue a PhD at MEB. There are many people who have, directly or indirectly, contributed to the fact that you are now holding this thesis in your hands. This is my attempt to thank them.

First of all, a big thank you to **all the people who donate money to Cancerfonden**, without whom none of the research projects in this thesis would be possible.

To my main supervisor **Paul Dickman**. Thank you for making my PhD education better than I could have ever imagined. Your door has always been open and you have said "yes" to me going to courses and conferences whenever and wherever I wanted. I have really enjoyed our discussions (not always ending up the way you want) and I have learnt so much from you, in all aspects of research. I am so thankful for your trust in me, and for allowing me to steer this ship along my own path.

To one of the most caring people I know, my co-supervisor **Sandra Eloranta**. Ever since our first student-supervisor experience during my MSc project in 2008, you have not only been a fantastic role model and teacher but also a very good friend. You have the extraordinary talent of mixing work with pleasure, and some of the best times during my PhD have been in your company. I look forward to future **#teamactivity** successes!

To my co-supervisor **Paul Lambert**. Throughout my PhD, you have continuously told me to slow down and take one (or several) step(s) back. And you have always been right. I'm very thankful for this. Also, thank you for always making time for me whenever I had "a quick question" (that never ended up being quick) during your Stockholm weeks.

To **Magnus Björkholm**, my co-supervisor and Hodgkin lymphoma mentor. You humble me with your knowledge and years of experience. Thank you for telling me that you were going to be one of my supervisors, and for teaching me so much about lymphoma and things like where the spleen is located. It is a pleasure to collaborate with you!

Thank you to **Ingrid Glimelius**, my co-supervisor for the latter part of my PhD journey. I have truly enjoyed our collaboration, and I believe one feature we share is that we never back down from correcting, checking, or appending the smallest of things in every project. I am also very thankful for my visit at Uppsala Akademiska Sjukhuset. The patients I met touched me, and the insight into your every day life in the clinic was an eye-opener.

To **Sigurdur Kristinsson**, my co-supervisor for the first part of my PhD. Thank you for all the support during the first years of my PhD, I hope we get to collaborate on the projects that fell out of this thesis some day in the future. Hopefully on Iceland!

**Hasse Walum** – my scientific mentor and friend. Without you there would not be a MSc in mathematical statistics and there would definitely not be a PhD thesis. Thank you for believing in me at a time in my life when I didn't myself.

To **Anna Johansson** – my mentor, research collaborator, and good friend; how can I even begin to thank you for everything you have done for me during the last 10 years? We might have cried a lot but we have mostly laughed. You are the best office neighbor ever.

**Therese Andersson** – Whatever problem I have encountered in any of my PhD projects, you have had to listen to it - I'm sorry (but thanks!). You are such a talented biostatistician and researcher, and I'm lucky to not only get to collaborate with you but also call you my friend.

To **#teamactivity** mate **Michael Crowther**. You made my idea for the fourth study reality! Thanks for not making me feel (too) stupid when we co-program in Stata, and for illustrating how lines can turn to dots. But more importantly, thanks for epic text messages, tweets, and for fun times in Norway.

Thank you to **Karin Ekström Smedby** for sharing your vast knowledge, and for your thoughtful and thorough input on my PhD studies.

To **Marcus Aly** and **Roxanna Hellgren** – thank you for welcoming me to visit the Urology clinic at KS, and Bröstcentrum at SöS, respectively.

The one and only **Hans H. Storm** – thank you not only for fantastic dancing and interesting conversations at the ANCR/IACR meetings, but for taking the time to review my research at my half-time control.

**Jingcheng Zhao** and **Sara Ekberg**, thank you for volunteering (?) to be on my pre-dissertation committee.

Thank you to current and previous **Heads of Department** for shaping MEB into such a great place to do research. And to **Pär Sparén** and **Ninoa Malki** - thank you for all your work with the SocMob database, and for your generosity to let me use it.

**Camilla Ahlqvist**, without you it would be so much more difficult to get a PhD. It has been a luxury having your help, and always with a smile. The same goes for **Gunilla Sonnebring** – always smiling, always helping everyone out. Thank you for proof reading this text.

To the brilliant **Hannah Bower** for being such a good friend and for making this journey so much fun with everything from Italian summer school to gambling in Vegas! And for being on my pre-dissertation committee. Thanks also to **Mark Rutherford**, **Betty Syriopoulou**, **Sarwar Islam**, and other past and present members of the MEB + Leicester Survival Group. You are all ridiculously talented.

Thank you to past and present **PhD students in the Biostat Group** for lunches on the roof, support, and for making the last couple of years simply the best: **Gabriel Isheden**, **Xingrong Liu**, **Bénédicte Delcoigne**, **Andreas Karlsson**, **Thorgerdur Palsdottir**, **Elisabeth Dahlqwist**, **Zheng Ning**, **Johan Zetterquist**, **Daniela Mariosa**, **Henrik Olsson**, **Peter Ström**, **Linda Abrahamsson**, **Wenjiang Deng**, **Rickard Strandberg**, **Robert Szulkin**, and **Shuang Hao**.

To the rest of the fantastic **Biostat Group at MEB** – every single one of you inspire me and makes going to work a pleasure! Especially thanks to **Juni Palmgren** for supervising my MSc project and opening the doors to MEB for me. You are such an inspiration and I'm forever grateful to you. **Marie Jansson**, you are the kernel of the MEB Biostat group, and I don't know what I would have done without you. Thank you for taking care of me, I will miss you the most. **Cecilia Lundholm** – thank you for being

a generous and simply amazing head of the Applied Biostat Group. And a special thank you to **Arvid Sjölander** for chairing my dissertation.

Thank you to past and present **PhD students at MEB** including, but not excluded to, **Cecilia Radkiewicz**, **Frida Lundberg**, **Kathleen Bokenberger**, **Tor-Arne Hegvik**, **Isabell Brikell**, **Malin Ericsson**, **Johanna Holm**, **Favelle Lamb**, **Anna Plym**, **Johanna Sieurin**, **Anne Örtqvist**, **Louise Eriksson**, **Carolyn Cesta**, **Nelson Ndegwa Gichora**, **Andreas Jangmo**, and **Emilio Ugalde Morales**.

To **Hatef Darabi**, **Edoardo Colzani**, **Davide Valentini**, and **Gaetano Marrone** for great memories at the Swedish place, and for becoming my first friends at MEB 10 (!) years ago.

Sometimes you need a break from science – thank you to **Anthony Sullivan** and everyone at **Evolve** for giving me two hours each week when my only thoughts have been to survive.

**Erik** and **Marianne Walum** – thank you for Sunday dinners, for encouraging me when I started studying, and for all the love you have shown me.

To my best friends since over 20 years: **Emma**, **Michaela**, and **Miriam**. Your support and interest in my research means so much. Many of my decisions at age 16 were poor, but the choice of best friends is my best ever.

Thank you to **Anders Weibull** and **Karin Thomsson** for being fantastic parents-in-law, and for always showing an interest in my work.

To my mother **Gittan S:t Cyr** and sister **Andrea Bernström** – for being so extremely talented at something completely different than what I do and know. You are my biggest inspiration of successful independent women. To my father, **Holger Dietrich**. Wish you were here.

**Peter Weibull** – Thank you for the endless support and encouragement. You are my stability and my best friend. Jag älskar dig.

To my sons, **Wilhelm** and **Alexander**, without whom nothing would be worth anything. ❤

# References


[1] M. Zwitter, J. R. Cohen, A. Barrett, and E. D. Robinton. Dorothy Reed and Hodgkin's disease: a reflection after a century. *International Journal of Radiation Oncology, Biology, Physics*, 53(2):366–75, 2002.

[2] S. Terezakis. Dorothy Reed Mendenhall: expressions of a pioneer in Hodgkin disease. *International Journal of Radiation Oncology, Biology, Physics*, 92(1):8–10, 2015.

[3] P. P. Carbone, H. S. Kaplan, K. Musshoff, D. W. Smithers, and M. Tubiana. Report of the committee on Hodgkin's disease staging classification. *Cancer Research*, 31(11):1860–1, 1971.

[4] E. T. Chang, K. E. Smedby, H. Hjalgrim, A. Porwit-MacDonald, G. Roos, B. Glimelius, and H. O. Adami. Family history of hematopoietic malignancy and risk of lymphoma. *Journal of the National Cancer Institute*, 97(19):1466–74, 2005.

[5] C. Crump, K. Sundquist, W. Sieh, M. A. Winkleby, and J. Sundquist. Perinatal and family risk factors for Hodgkin lymphoma in childhood through young adulthood. *American Journal of Epidemiology*, 176(12):1147–58, 2012.

[6] J. R. Cerhan and S. L. Slager. Familial predisposition and genetic risk factors for lymphoma. *Blood*, 126(20):2265–73, 2015.

[7] H. Hjalgrim, J. Askling, K. Rostgaard, S. Hamilton-Dutoit, M. Frisch, J. S. Zhang, M. Madsen, N. Rosdahl, H. B. Konradsen, H. H. Storm, and M. Melbye. Characteristics of Hodgkin's lymphoma after infectious mononucleosis. *New England Journal of Medicine*, 349:1324–1332, October 2003.

[8] J. Sjöberg, C. Halthur, S. Y. Kristinsson, O. Landgren, U. A. Nygell, P. W. Dickman, and M. Björkholm. Progress in Hodgkin lymphoma: A population-based study on patients diagnosed in Sweden from 1973-2009. *Blood*, 119(4):990–6, 2012.

[9] V. T. Devita, A. A. Serpick, and P. P. Carbone. Combination chemotherapy in the treatment of advanced Hodgkin's disease. *Annals of Internal Medicine*, 73(6):881–895, 1970.

[10] V. Diehl, J. Franklin, D. Hasenclever, H. Tesch, M. Pfreundschuh, B. Lathan, U. Paulus, M. Sieber, J. U. Rueffer, M. Sextro, A. Engert, J. Wolf, R. Hermann, L. Holmer, U. Stappert-Jahn, E. Winnerlein-Trump, G. Wulf, S. Krause, A. Glunz, K. von Kalle, H. Bischoff, C. Haedicke, E. Duehmke, A. Georgii, and M. Loeffler. BEACOPP, a new dose-escalated and accelerated regimen, is at least as effective as COPP/ABVD in patients with advanced-stage Hodgkin's lymphoma: Interim report from a trial of the German Hodgkin's lymphoma Study Group. *Journal of Clinical Oncology*, 16(12):3810–21, 1998.

[11] I. Glimelius and A. Diepstra. Novel treatment concepts in Hodgkin lymphoma. *Journal of Internal Medicine*, 281(3):247–260, 2017.

[12] A. Younes, A. K. Gopal, S. E. Smith, S. M. Ansell, J. D. Rosenblatt, K. J. Savage, R. Ramchandren, N. L. Bartlett, B. D. Cheson, S. de Vos, A. Forero-Torres, C. H. Moskowitz, J. M. Connors, A. Engert, E. K. Larsen, D. A. Kennedy, E. L. Sievers, and R. Chen. Results of a pivotal phase II study of brentuximab vedotin for patients with relapsed or refractory Hodgkin's lymphoma. *Journal of Clinical Oncology*, 30(18):2183–9, 2012.

[13] J. M. Connors, W. Jurczak, D. J. Straus, S. M. Ansell, W. S. Kim, A. Gallamini, A. Younes, S. Alekseev, Á. Illés, M. Picardi, E. Lech-Maranda, Y. Oki, T. Feldman, P. Smolewski, K. J. Savage, N. L. Bartlett, J. Walewski, R. Chen, R. Ramchandren, P. L. Zinzani, D. Cunningham, A. Rosta, N. C. Josephson, E. Song, J. Sachs, R. Liu, H. A. Jolin, D. Huebner, and J. Radford. Brentuximab vedotin with chemotherapy for stage III or IV Hodgkin's lymphoma. *New England Journal of Medicine*, 378:331–344, January 2018.

[14] A. Younes, A. Santoro, M. Shipp, P. L. Zinzani, J. M. Timmerman, S. Ansell, P. Armand, M. Fanale, V. Ratanatharathorn, J. Kuruvilla, J. B. Cohen, G. Collins, K. J. Savage, M. Trneny, K. Kato, B. Farsaci, S. M. Parker, S. Rodig, M. G. Roemer, A. H. Ligon, and A. Engert. Nivolumab for classical Hodgkin's lymphoma after failure of both autologous stem-cell transplantation and brentuximab vedotin: a multicentre, multicohort, single-arm phase 2 trial. *Lancet Oncology*, 17(9):1283–94, 2016.

[15] A. B. Miller, T. H. Barclay, N. W. Choi, M. G. Grace, C. Wall, M. Plante, G. R. Howe, B. Cinader, and F. G. Davis. A study of cancer, parity and age at first pregnancy. *Journal of Chronic Diseases*, 33(10):595–605, 1980.

[16] O. Kravdal and S. Hansen. Hodgkin's disease: the protective effect of childbearing. *International Journal of Cancer*, 55(6):909–14, 1993.

[17] O. Kravdal and S. Hansen. The importance of childbearing for Hodgkin's disease: new evidence from incidence and mortality models. *International Journal of Epidemiology*, 25(4):737–43, 1996.

[18] M. Lambe, C. Hsieh, S. Tsaih, J. Adami, B. Glimelius, and H. O. Adami. Childbearing and the risk of Hodgkin's disease. *Cancer Epidemiology, Biomarkers & Prevention*, 7(9):831–834, 1998.

[19] H. Moller, A. Purushotham, K. M. Linklater, H. Garmo, L. Holmberg, M. Lambe, D. Yallop, and S. Devereux. Recent childbirth is an adverse prognostic factor in breast cancer and melanoma, but not in Hodgkin lymphoma. *European Journal of Cancer*, 49(17):3686–93, 2013.

[20] M. Lishner, D. Zemlickis, P. Degendorfer, T. Panzarella, S. B. Sutcliffe, and G. Koren. Maternal and foetal outcome following Hodgkin's disease in pregnancy. *British Journal of Cancer*, 65(1):114–7, 1992.

[21] H. Stensheim, B. Moller, T. van Dijk, and S. D. Fosså. Cause-specific survival for women diagnosed with cancer during pregnancy or lactation: a registry-based cohort study. *Journal of Clinical Oncology*, 27(1):45–51, 2009.

[22] B. Brenner, I. Avivi, and M. Lishner. Haematological cancers in pregnancy. *The Lancet*, 379(9815):580–7, 2012.

[23] A. M. Evens, R. Advani, O. W. Press, I. S. Lossos, J. M. Vose, F. J. Hernandez-Ilizaliturri, B. K. Robinson, S. Otis, L. Nadav Dagan, R. Abdallah, A. Kroll-Desrosiers, J. L. Yarber, J. Sandoval, K. Foyil, L. M. Parker, L. I. Gordon, K. A. Blum, C. R. Flowers, J. P. Leonard, T. M. Habermann, and N. L. Bartlett. Lymphoma occurring during pregnancy: antenatal therapy, complications, and maternal survival in a multicenter analysis. *Journal of Clinical Oncology*, 31(32):4132–9, 2013.

[24] D. Pereg, G. Koren, and M. Lishner. The treatment of Hodgkin's and non-Hodgkin's lymphoma in pregnancy. *Haematologica*, 92(9):1230–7, 2007.

[25] J. P. Jouet, B. Buchet-Bouverne, P. Fenaux, J. P. Pollet, M. Simon, M. P. Walter, J. L. Leroy, F. Puech, M. Delecour, and F. Bauters. [Influence of pregnancy on the development of Hodgkin's disease]. *Presse Medicale*, 17(9):423–7, 1988.

[26] S. Howell and S. Shalet. Gonadal damage from chemotherapy and radiotherapy. *Endocrinology and Metabolism Clinics of North America*, 27(4):927–43, 1998.

[27] Z. Blumenfeld, E. Dann, I. Avivi, R. Epelbaum, and J. M. Rowe. Fertility after treatment for Hodgkin's disease. *Annals of Oncology*, 13 Suppl 1:138–47, 2002.

[28] K. Behringer, K. Breuer, T. Reineke, M. May, L. Nogova, B. Klimm, T. Schmitz, L. Wildt, V. Diehl, and A. Engert. Secondary amenorrhea after Hodgkin's lymphoma is influenced by age at treatment, stage of disease, chemotherapy regimen, and the use of oral contraceptives during therapy: A report from the German Hodgkin's lymphoma Study Group. *Journal of Clinical Oncology*, 23(30):7555–64, 2005.

[29] K. Behringer, H. Mueller, H. Goergen, I. Thielen, A. D. Eibl, V. Stumpf, C. Wessels, M. Wiehlputz, J. Rosenbrock, T. Halbsguth, K. S. Reiners, T. Schober, J. H. Renno, M. von Wolff, K. van der Ven, M. Kuehr, M. Fuchs, V. Diehl, A. Engert, and P. Borchmann. Gonadal function and fertility in survivors after Hodgkin lymphoma treatment within the German Hodgkin Study Group HD13 to HD15 trials. *Journal of Clinical Oncology*, 31(2):231–9, 2013.

[30] K. Behringer, I. Thielen, H. Mueller, H. Goergen, A. D. Eibl, J. Rosenbrock, T. Halbsguth, D. A. Eichenauer, M. Fuchs, K. S. Reiners, J. H. Renno, K. van der Ven, M. Kuehr, M. von Wolff, V. Diehl, A. Engert, and P. Borchmann. Fertility and gonadal function in female survivors after treatment of early unfavorable Hodgkin lymphoma (hl) within the German Hodgkin Study Group HD14 trial. *Annals of Oncology*, 23(7):1818–25, 2012.

[31] M. Huser, L. Smardova, P. Janku, I. Crha, J. Zakova, P. Stourac, J. Jarkovsky, J. Mayer, and P. Ventruba. Fertility status of Hodgkin lymphoma patients treated with chemotherapy and adjuvant gonadotropin-releasing hormone analogues. *Journal of Assisted Reproduction and Genetics*, 32(8):1187–93, 2015.

[32] K. Behringer, L. Wildt, H. Mueller, V. Mattle, P. Ganitis, B. van den Hoonaard, H. W. Ott, S. Hofer, A. Pluetschow, V. Diehl, A. Engert, and P. Borchmann. No protection of the ovarian follicle pool with the use of GnRH-analogues or oral contraceptives in young women treated with escalated BEACOPP for advanced-stage Hodgkin lymphoma. final results of a phase ii trial from the German Hodgkin Study Group. *Annals of Oncology*, 21(10):2052–60, 2010.

[33] D. C. Hodgson, M. Pintilie, L. Gitterman, B. Dewitt, C. A. Buckley, S. Ahmed, K. Smith, A. Schwartz, R. W. Tsang, M. Crump, W. Wells, A. Sun, and M. K. Gospodarowicz. Fertility among female hodgkin lymphoma survivors attempting pregnancy following ABVD chemotherapy. *Hematological Oncology*, 25(1):11–5, 2007.

[34] H. Stensheim, M. Cvancarova, B. Moller, and S. D. Fosså. Pregnancy after adolescent and adult cancer: a population-based matched cohort study. *International Journal of Cancer*, 129(5):1225–36, 2011.

[35] M. A. van der Kaaij, N. Heutte, P. Meijnders, E. Abeilard-Lemoisson, M. Spina, L. C. Moser, A. Allgeier, B. Meulemans, B. Dubois, A. H. Simons, P. J. Lugtenburg, B. M. Aleman, E. M. Noordijk, C. Ferme, J. Thomas, A. Stamatoullas, C. Fruchart, P. Brice, I. Gaillard, J. K. Doorduijn, C. Sebban, W. G. Smit, S. Bologna, J. M. Roesink, F. Ong, M. P. Andre, J. M. Raemaekers, M. Henry-Amar, and H. C. Kluin-Nelemans. Parenthood in survivors of Hodgkin lymphoma: An EORTC-GELA general population case-control study. *Journal of Clinical Oncology*, 30(31):3854–63, 2012.

[36] M. Hartman, J. Liu, K. Czene, H. Miao, K. S. Chia, A. Salim, and H. M. Verkooijen. Birth rates among female cancer survivors: a population-based cohort study in Sweden. *Cancer*, 119(10):1892–9, 2013.

[37] A. K. Ng. Review of the cardiac long-term effects of therapy for Hodgkin lymphoma. *British Journal of Haematology*, 154(1):23–31, 2011.

[38] A. K. Ng. Current survivorship recommendations for patients with Hodgkin lymphoma: focus on late effects. *Hematology ASH Education Program*, 2014(1):488–94, 2014.

[39] F. E. van Leeuwen and A. K. Ng. Long-term risk of second malignancy and cardiovascular disease after Hodgkin lymphoma treatment. *Hematology ASH Education Program*, 2016(1):323–330, 2016.

[40] F. E. van Leeuwen and A. K. Ng. Late sequelae in Hodgkin lymphoma survivors. *Hematological Oncology*, 35 Suppl 1:60–66, 2017.

[41] M. C. Hull, C. G. Morris, C. J. Pepine, and N. P. Mendenhall. Valvular dysfunction and carotid, subclavian, and coronary artery disease in survivors of hodgkin lymphoma treated with radiation therapy. *JAMA*, 290(21):2831–7, 2003.

[42] M. J. Adams, S. R. Lipsitz, S. D. Colan, N. J. Tarbell, S. T. Treves, L. Diller, N. Greenbaum, P. Mauch, and S. E. Lipshultz. Cardiovascular status in long-term survivors of Hodgkin's disease treated with chest radiotherapy. *Journal of Clinical Oncology*, 22(15):3139–48, 2004.

[43] F. A. van Nimwegen, M. Schaapveld, C. P. Janus, A. D. Krol, E. J. Petersen, J. M. Raemaekers, W. E. Kok, B. M. Aleman, and F. E. van Leeuwen. Cardiovascular disease after Hodgkin lymphoma treatment: 40-year disease risk. *JAMA Intern Med*, 175(6):1007–17, 2015.

[44] D. J. Cutter, M. Schaapveld, S. C. Darby, M. Hauptmann, F. A. van Nimwegen, A. D. Krol, C. P. Janus, F. E. van Leeuwen, and B. M. Aleman. Risk of valvular heart disease after treatment for Hodgkin lymphoma. *Journal of the National Cancer Institute*, 107(4), 2015.

[45] F. A. van Nimwegen, M. Schaapveld, D. J. Cutter, C. P. Janus, A. D. Krol, M. Hauptmann, K. Kooijman, J. Roesink, R. van der Maazen, S. C. Darby, B. M. Aleman, and F. E. van Leeuwen. Radiation

dose-response relationship for risk of coronary heart disease in survivors of Hodgkin lymphoma. *Journal of Clinical Oncology*, 34(3):235–43, 2016.

[46] L. D. Dorresteijn, A. C. Kappelle, W. Boogerd, W. J. Klokman, A. J. Balm, R. B. Keus, F. E. van Leeuwen, and H. Bartelink. Increased risk of ischemic stroke after radiotherapy on the neck in patients younger than 60 years. *Journal of Clinical Oncology*, 20(1):282–8, 2002.

[47] D. C. Bowers, D. E. McNeil, Y. Liu, Y. Yasui, M. Stovall, J. G. Gurney, M. M. Hudson, S. S. Donaldson, R. J. Packer, P. A. Mitby, C. E. Kasper, L. L. Robison, and K. C. Oeffinger. Stroke as a late treatment effect of Hodgkin's disease: A report from the childhood cancer survivor study. *Journal of Clinical Oncology*, 23(27):6508–15, 2005.

[48] M. L. De Bruin, L. D. Dorresteijn, M. B. van't Veer, A. D. Krol, H. J. van der Pal, A. C. Kappelle, W. Boogerd, B. M. Aleman, and F. E. van Leeuwen. Increased risk of stroke and transient ischemic attack in 5-year survivors of Hodgkin lymphoma. *Journal of the National Cancer Institute*, 101(13):928–37, 2009.

[49] M. V. Maraldo, P. Brodin, M. C. Aznar, I. R. Vogelius, P. Munck af Rosenschold, P. M. Petersen, and L. Specht. Doses to carotid arteries after modern radiation therapy for Hodgkin lymphoma: is stroke still a late effect of treatment? *International Journal of Radiation Oncology, Biology, Physics*, 87(2):297–303, 2013.

[50] B. M. Aleman, A. W. van den Belt-Dusebout, W. J. Klokman, M. B. Van't Veer, H. Bartelink, and F. E. van Leeuwen. Long-term cause-specific mortality of patients treated for Hodgkin's disease. *Journal of Clinical Oncology*, 21(18):3431–9, 2003.

[51] A. J. Swerdlow, C. D. Higgins, P. Smith, D. Cunningham, B. W. Hancock, A. Horwich, P. J. Hoskin, A. Lister, J. A. Radford, A. Z. Rohatiner, and D. C. Linch. Myocardial infarction mortality risk after treatment for Hodgkin disease: A collaborative British cohort study. *Journal of the National Cancer Institute*, 99(3):206–14, 2007.

[52] S. Myrehaug, M. Pintilie, R. Tsang, R. Mackenzie, M. Crump, Z. Chen, A. Sun, and D. C. Hodgson. Cardiac morbidity following modern treatment for Hodgkin lymphoma: supra-additive cardiotoxicity of doxorubicin and radiation therapy. *Leukemia and Lymphoma*, 49(8):1486–93, 2008.

[53] A. Andersson, U. Naslund, B. Tavelin, G. Enblad, A. Gustavsson, and B. Malmer. Long-term risk of cardiovascular disease in Hodgkin lymphoma survivors–retrospective cohort analyses and a concept for prospective intervention. *International Journal of Cancer*, 124(8):1914–7, 2009.

[54] C. E. Kiserud, J. H. Loge, A. Fosså, H. Holte, M. Cvancarova, and S. D. Fosså. Mortality is persistently increased in Hodgkin's lymphoma survivors. *European Journal of Cancer*, 46(9):1632–9, 2010.

[55] S. M. Castellino, A. M. Geiger, A. C. Mertens, W. M. Leisenring, J. A. Tooze, P. Goodman, M. Stovall, L. L. Robison, and M. M. Hudson. Morbidity and mortality in long-term survivors of Hodgkin lymphoma: A report from the childhood cancer survivor study. *Blood*, 117(6):1806–16, 2011.

[56] S. L. Galper, J. B. Yu, P. M. Mauch, J. F. Strasser, B. Silver, A. Lacasce, K. J. Marcus, M. A. Stevenson, M. H. Chen, and A. K. Ng. Clinically significant cardiac disease in patients with Hodgkin lymphoma treated with mediastinal irradiation. *Blood*, 117(2):412–8, 2011.

[57] A. Avilés, N. Neri, J. M. Nambo, J. Huerta-Guzman, A. Talavera, and S. Cleto. Late cardiac toxicity secondary to treatment in Hodgkin's disease. A study comparing doxorubicin, epirubicin and mitoxantrone in combined therapy. *Leukemia and Lymphoma*, 46:1023–1028, July 2005.

[58] F. A. van Nimwegen, G. Ntentas, S. C. Darby, M. Schaapveld, M. Hauptmann, P. J. Lugtenburg, C. P. M. Janus, L. Daniels, F. E. van Leeuwen, D. J. Cutter, and B. M. P. Aleman. Risk of heart failure in survivors of Hodgkin lymphoma: effects of cardiac exposure to radiation and anthracyclines. *Blood*, 129(16):2257–2265, 2017.

[59] S. Eloranta, P. C. Lambert, J. Sjöberg, T. M. L. Andersson, M. Bjorkholm, and P. W. Dickman. Temporal trends in mortality from diseases of the circulatory system after treatment for Hodgkin lymphoma: a population-based cohort study in Sweden (1973 to 2006). *Journal of Clinical Oncology*, 31(11):1435–41, 2013.

[60] A. J. Swerdlow, C. D. Higgins, P. Smith, D. Cunningham, B. W. Hancock, A. Horwich, P. J. Hoskin, A. T. Lister, J. A. Radford, A. Z. S. Rohatiner, and D. C. Linch. Second cancer risk after chemotherapy for Hodgkin's lymphoma: a collaborative British cohort study. *Journal of Clinical Oncology*, 29:4096–4104, November 2011.

[61] M. Schaapveld, B. M. P. Aleman, A. M. van Eggermond, C. P. M. Janus, A. D. G. Krol, R. W. M. van der Maazen, J. Roesink, J. M. M. Raemaekers, J. P. de Boer, J. M. Zijlstra, G. W. van Imhoff, E. J. Petersen, P. M. P. Poortmans, M. Beijert, M. L. Lybeert, I. Mulder, O. Visser, M. W. J. Louwman, I. M. Krul, P. J. Lugtenburg, and F. E. van Leeuwen. Second cancer risk up to 40 years after treatment for Hodgkin's lymphoma. *New England Journal of Medicine*, 373:2499–2511, December 2015.

[62] A. J. Swerdlow, J. A. Barber, G. V. Hudson, D. Cunningham, R. K. Gupta, B. W. Hancock, A. Horwich, T. A. Lister, and D. C. Linch. Risk of second malignancy after Hodgkin's disease in a collaborative British cohort: The relation to age at treatment. *Journal of Clinical Oncology*, 18:498–509, February 2000.

[63] G. M. Dores, C. Metayer, R. E. Curtis, C. F. Lynch, E. A. Clarke, B. Glimelius, H. H. Storm, E. Pukkala, F. E. van Leeuwen, E. J. Holowaty, M. Andersson, T. Wiklund, T. Joensuu, M. B. van't Veer, M. Stovall, M. Gospodarowicz, and L. B. Travis. Second malignant neoplasms among long-term survivors of hodgkin's disease: A population-based evaluation over 25 years. *Journal of Clinical Oncology*, 20:3484–3494, August 2002.

[64] A. K. Ng, M. P. Bernardo, E. Weller, K. H. Backstrand, B. Silver, K. C. Marcus, N. J. Tarbell, J. Friedberg, G. P. Canellos, and P. M. Mauch. Long-term survival and competing causes of death in patients with early-stage Hodgkin's disease treated at age 50 or younger. *Journal of Clinical Oncology*, 20(8):2101–8, 2002.

[65] F. E. van Leeuwen, A. M. Chorus, A. W. van den Belt-Dusebout, A. Hagenbeek, R. Noyon, E. H. van Kerkhoff, H. M. Pinedo, and R. Somers. Leukemia risk following Hodgkin's disease: relation to cumulative dose of alkylating agents, treatment with teniposide combinations, number of episodes of chemotherapy, and bone marrow damage. *Journal of Clinical Oncology*, 12:1063–1073, May 1994.

[66] F. E. van Leeuwen, W. J. Klokman, M. B. Veer, A. Hagenbeek, A. D. Krol, U. A. Vetter, M. Schaapveld, P. van Heerde, J. M. Burgers, R. Somers, and B. M. Aleman. Long-term risk of second malignancy in survivors of Hodgkin's disease treated during adolescence or young adulthood. *Journal of Clinical Oncology*, 18:487–497, February 2000.

[67] J. F. Ludvigsson, P. Otterblad-Olausson, B. U. Pettersson, and A. Ekbom. The Swedish personal identity number: Possibilities and pitfalls in healthcare and medical research. *European Journal of Epidemiology*, 24(11):659–667, 2009.

[68] J. F. Ludvigsson, C. Almqvist, A. K. Bonamy, R. Ljung, K. Michaelsson, M. Neovius, O. Stephansson, and W. Ye. Registers of the Swedish total population and their use in medical research. *European Journal of Epidemiology*, 31(2):125–36, 2016.

[69] I. Turesson, M. S. Linet, M. Bjorkholm, S. Y. Kristinsson, L. R. Goldin, N. E. Caporaso, and O. Landgren. Ascertainment and diagnostic accuracy for hematopoietic lymphoproliferative malignancies in Sweden 1964-2003. *International Journal of Cancer*, 121(10):2260–6, 2007.

[70] The Swedish Lymphoma Group. Svenska lymfomregistret, nationell kvalitetsrapport för diagnosår 2013. Technical report, 2013.

[71] K. J. Rothman. *Epidemiology: An Introduction*. New York: Oxford University Press, 2002.

[72] Per Kragh Andersen, Ronald B Geskus, Theo de Witte, and Hein Putter. Competing risks in epidemiology: possibilities and pitfalls. *International journal of epidemiology*, 41:861–870, June 2012.

[73] E. L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

[74] O. O. Aalen, O. Borgan, and H. K. Gjessing. *Survival and event history analysis: A process point of view*. Springer-Verlag New York, 1st edition, 2008.

[75] David R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society*, pages 187–220, 1972.

[76] N. Reid. A conversation with Sir David Cox. *Statistical Science*, (9):439–455, 1994.

[77] P. Royston and M. K. B. Parmar. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine*, 21:2175–2197, August 2002.

[78] P. C. Lambert and P. Royston. Further development of flexible parametric models for survival analysis. *Stata Journal*, 9(2):265–290, 2009.

[79] J. D. Kalbfleisch and R. L. Prentice. *The Statistical Analysis of Failure Time Data*. The Statistical Analysis of Failure Time Data. John Wiley & Sons, Inc., second edition edition, 2002.

[80] H. Putter, M. Fiocco, and R. B. Geskus. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine*, 26:2389–2430, May 2007.

[81] O. O. Aalen and S. Johansen. An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scandinavian Journal of Statistics*, pages 141–150, 1978.

[82] S. R. Hinchliffe and P. C. Lambert. Flexible parametric modelling of cause-specific hazards to estimate cumulative incidence functions. *BMC Medical Research Methodology*, 13, 2013.

[83] S. R. Hinchliffe and P. C. Lambert. Extending the flexible parametric survival model for competing risks. *Stata Journal*, 13(2):344–355, 2013.

[84] M. J. Crowther and P. C. Lambert. Parametric multistate survival models: flexible modelling allowing transition-specific distributions with application to estimating clinically useful measures of effect differences. *Statistics in Medicine*, 36:4719–4742, December 2017.

[85] J. P. Fine and R. J. Gray. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446):496–509, 1999.

[86] S. I. Mozumder, M. Rutherford, and P. C. Lambert. Direct likelihood inference on the cause-specific cumulative incidence function: A flexible parametric regression modelling approach. *Statistics in Medicine*, 37(1):82–97, 2018.

[87] P. C. Lambert, P. W. Dickman, C. P. Nelson, and P. Royston. Estimating the crude probability of death due to cancer and other causes using relative survival models. *Statistics in Medicine*, 29(7-8):885–895, 2010.

[88] S. Eloranta, P. C. Lambert, T. M. L. Andersson, K. Czene, P. Hall, M. Björkholm, and P. W. Dickman. Partitioning of excess mortality in population-based cancer patient survival studies using flexible parametric survival models. *BMC Medical Research Methodology*, 12:86, June 2012.

[89] P. W. Dickman, E. Coviello, and M. Hills. Estimating and modelling relative survival. *Stata Journal*, 2015.

[90] F. Ederer, L. M. Axtell, and S. J. Cutler. The relative survival rate: A statistical methodology. *NCI Monograph*, 6:101–121, 1961.

[91] S R Hinchliffe, M J Rutherford, M J Crowther, C P Nelson, and P C Lambert. Should relative survival be used with lung cancer data? *British journal of cancer*, 106:1854–1859, May 2012.

[92] T. Blakely, M. Soeberg, K. Carter, R. Costilla, J. Atkinson, and D. Sarfati. Bias in relative survival methods when using incorrect life-tables: Lung and bladder cancer by smoking status and ethnicity in New Zealand. *International Journal of Cancer*, 131:E974–E982, September 2012.

[93] P. W. Dickman, A. Sloggett, M. Hills, and T. Hakulinen. Regression models for relative survival. *Statistics in Medicine*, 23:51–64, January 2004.

[94] C. P. Nelson, P. C. Lambert, I. B. Squire, and D. R. Jones. Flexible parametric models for relative survival, with application in coronary heart disease. *Statistics in Medicine*, 26:5486–5498, December 2007.

[95] K. A. Cronin and E. J. Feuer. Cumulative cause-specific mortality for cancer patients in the presence of other causes: a crude analogue of relative survival. *Statistics in Medicine*, 19(13):1729–1740, 2000.

[96] P. K. Andersen and N. Keiding. Multi-state models for event history analysis. *Statistical Methods in Medical Research*, 11:91–115, 2002.

[97] H. J. Hsieh, T. H. H. Chen, and S. H. Chang. Assessing chronic disease progression using non-homogeneous exponential regression Markov models: an illustration using a selective breast cancer screening in Taiwan. *Statistics in Medicine*, 21(22):3369–3382, 2002.

[98] S. R. Hinchliffe, D. A. Scott, and P. C. Lambert. Flexible parametric illness-death models. *Stata Journal*, 13(4):759–775, 2013.

[99] A. C. Titman. Flexible nonhomogeneous Markov models for panel observed data. *Biometrics*, 67(3):780–787, 2011.

[100] M. Fiocco, H. Putter, and H. C. van Houwelingen. Reduced-rank proportional hazards regression and simulation-based prediction for multi-state models. *Statistics in Medicine*, 27:4340–4358, September 2008.

[101] E. Huszti, M. Abrahamowicz, A. Alioum, C. Binquet, and C. Quantin. Relative survival multistate Markov model. *Statistics in Medicine*, 31:269–286, February 2012.

[102] S. Gilard-Pioc, M. Abrahamowicz, A. Mahboubi, A. M. Bouvier, O. Dejardin, E. Huszti, C. Binquet, and C. Quantin. Multi-state relative survival modelling of colorectal cancer progression and mortality. *Cancer Epidemiology*, 39:447–455, June 2015.

[103] F. Gillaizeau, E. Dantan, M. Giral, and Y. Foucher. A multistate additive relative survival semi-Markov model. *Statistical Methods in Medical Research*, 26:1700–1711, August 2017.

[104] A. L. V. Johansson, T. M.-L. Andersson, C.-C. Hsieh, S. Cnattingius, and M. Lambe. Increased mortality in women with breast cancer detected during pregnancy and different periods postpartum. *Cancer Epidemiology, Biomarkers & Prevention*, 20:1865–1872, 2011.

[105] C. Radkiewicz, A. L. V. Johansson, P. W. Dickman, M. Lambe, and G. Edgren. Sex differences in cancer risk and survival: A Swedish cohort study. *European Journal of Cancer*, 84:130–140, October 2017.

[106] B. McKnight, L. S. Cook, and N. S. Weiss. Logistic regression analysis for more than one characteristic of exposure. *American Journal of Epidemiology*, 149:984–992, June 1999.

[107] C. E. Weibull, S. Eloranta, D. Altman, A. L. V. Johansson, and M. Lambe. Childbearing and the risk of bladder cancer: A nationwide population-based cohort study. *European Urology*, 63:733–738, April 2013.