August 2020

A Multistate Survival Analysis for Sequential Follicular Lymphoma Treatment Leveraging National Lymphocare Study

Aashay Mahesh Mahesh Mehta
Clemson University, aashaymehta246@gmail.com

A MULTISTATE SURVIVAL ANALYSIS FOR SEQUENTIAL FOLLICULAR LYMPHOMA TREATMENT LEVERAGING NATIONAL LYMPHOCARE STUDY

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Industrial Engineering

---

by
Aashay Mahesh Mehta
August 2020

Accepted by:
Dr. Çağlar Çağlayan, Committee Chair
Dr. Scott Mason
Dr. William Ferell

Abstract

Follicular lymphoma (FL), caused due to the abnormal growth of B-cells, is the most common form of indolent Non-Hodgkin's Lymphoma (NHL) in the United States. Majority of the patients diagnosed with FL are subjected to a series of treatment during their clinical course, thereby providing researchers an opportunity to evaluate sequential effect of different treatment regimens and identify significant risk factors affecting the survival of patients.

National Lymphocare Study (NLCS) is the largest (n = 2,740) prospective study conducted in the United States that collected information about different prognostic factors and outcomes required to conduct a comprehensive multi-state survival analysis. Using the data set, we analyzed the effect of 11 stationary and 5 dynamic variables on the death-specific transition and compared the results obtained from two multi-state models. We identified significant clinical factors impacting all-cause and FL-specific death using Cox proportional hazard model and predicted the course of disease using Aalen-Johansen estimator.

The risk of all-cause death was 16.37% following 5-years from diagnosis. At the same time point, patients initially kept under watchful waiting (WW) were at a lower risk of FL-specific death, 4.52%, compared to death due to other-causes, 8.25%. Similarly, patients receiving an induction treatment without WW had an reduced risk of FL-specific death compared to death due to other cause at the 5-years time point.

We further identified that the risk of FL-specific death increase after first-, second-, and third-line treatment with an increase in age of diagnosed patients. Dynamic variables, such as low albumin and elevated lactate dehydrogenase, were associated with poor outcomes after first-, second-, and third-line treatment. The presence of B-symptoms at diagnosis was not associated with an increased risk of FL death. On the other hand, being female reduced the risk (hazard ratio: 0.46 [0.3 - 0.8]) of FL death following treatment 2.

Using a multi-state survival analysis framework allowed to quantify the effect of prognostic factors on cause-specific death. We identified that the risk of death due to other-cause is higher compared to the risk of death due to FL.

Dedication

I would like to dedicate my thesis to my beloved grandfather Dr. Kantilal Mehta. Your wisdom puts my mind and heart at ease. I miss you.

Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Dr. Çağlar Çağlayan, for continuous support throughout my thesis research, for his patience, motivation, and immense knowledge. His detailed and constructive comments were vital to the development of this thesis. I could not have imagined having a better advisor and mentor.

Second, I would like to acknowledge all the doctors and clinicians who were the pioneers of the National Lymphocare Study, the staff member who supported the study vision and patients for being a part of the clinical trial. Particularly, I would like to thank Dr. Christopher Flowers for his guidance and help.

Third, I want to thank the members of my master thesis committee, Dr. Scott Mason and Dr. William Ferrell. Finally, my profound gratitude goes to my family and friends for their continuous and unparalleled love, help and support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Follicular Lymphoma (FL) is a type of cancer that is caused by the abnormal growth of B-Lymphocytes, a type of white blood cells, found in the lymph nodes and other parts of the body (Figure 1.1). The family of lymphomas is divided into two categories: Hodgkin's lymphoma and Non-Hodgkin's lymphoma (NHL), where FL belongs to the class of the NHLs. In general, FL has a slow-growing (i.e., indolent) prognosis after diagnosis rather than aggressive, and is the most prevalent form of the slow-growing NHLs in the United States (U.S.), accounting for nearly 75% of slow-growing NHL cases [1].

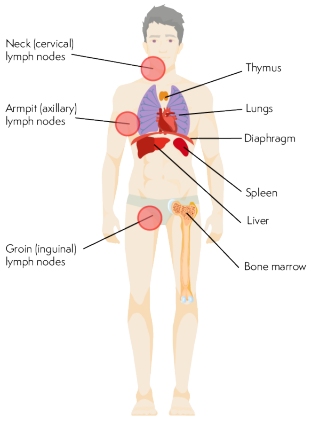

Figure 1.1: Human Anatomy with Lymph Nodes [2]

In 2019, approximately 15,000 - 20,000 new patients are projected to be diagnosed with FL, accounting for 20% to 25% of NHL patients [3–5]. FL is not known to be an inherited disease and is more common in people aged 50 years and above, where the mean age at diagnosis is 65 years [6]. A considerable proportion of FLs relapse (i.e., grows again) after the initial treatment [7, 8] and patients diagnosed with FL are usually subjected to a series of therapies during the course of their treatment [9].

Currently, there is no standard care available for the initial and sequential treatments of FL in the U.S. [10–13]. A variety of disease management options are available and implemented in the clinical practice, including watchful waiting, single-agent therapy, chemotherapy with immunotherapy, radiotherapy and bone marrow transplant [14–16]. At present, there is no treatment strategy that has been shown to yield superior FL outcomes compared to all other alternative strategies [10, 12, 17, 18]. The selection of the corresponding therapy is often guided by the personal and local expertise of the clinicians and depends on patient characteristics and various clinical factors such as age and stage of the disease at diagnosis.

Limited data are available regarding the outcomes associated with the sequential FL treatments. Some of the main challenges preventing new comprehensive randomized controlled trials (RCTs) from being performed are high financial costs, the lengthy follow-up durations, and the requirement for relatively large number of patients to show a statistically significant difference between treatment groups [19]. Still, there are several existing randomized clinical trials and observational studies, conducted to assess the performance of FL treatments over time. In particular, National LymphoCare Study (NLCS) is the largest prospective cohort study conducted in the U.S. for FL [8]. Recruiting more than 2,700 FL patients between 2004 and 2007, NLCS was conducted to collect information about clinical outcomes of newly diagnosed FL patients over a period of 10 years. Further, majority of the NLCS

participants received multiple lines of treatments after diagnosis. Accordingly, NLCS provides researchers a unique opportunity to conduct a comprehensive analysis to examine the effectiveness of sequential FL treatments over time.

Prior analytical approaches predominantly focused on the effects of a single line of therapy on a single outcome of interest such as progression free survival (PFS) or overall survival (OS). Yet, given a fair portion of the FL cases relapse and most patients require more than a single therapy, analyzing sequential FL treatments remains critical. In addition, while analyzing multiple lines of FL treatments, it is important to study both FL-associated deaths and deaths associated with other causes, rather than solely focusing on a single outcome such as OS, in order to accurately identify risk factors affecting cause-specific mortality. Such an analysis would require an analytical model that accounts for competing risks and separately captures multiple treatment stages with cause-specific deaths within a single framework.

Attempting to make an initial step in closing this gap, we develop a comprehensive continuous-time multi-state model that thoroughly captures the clinical course of FL and conduct a multi-state survival analysis study by utilizing the largest prospective cohort study dataset that is currently available for FL. Our main objectives are to (i) investigate the clinical outcomes associated with the multiple lines of FL treatment, (ii) identify the significant socio-demographic and clinical risk factors affecting FL treatment and cause-specific mortality, and (iii) quantify their impact on FL outcomes at each line of therapy by leveraging the NLCS dataset.

The main findings of this study are as follows: First, we identified that the risk of death due to FL is relatively less compared to risk of death due to other-cause at the 5- and 8-year time point for all the patients in the data set. In fact, for WW patients the risk of other-cause death is 1.5 times the risk of FL related death. Second, we identified that old age, elevated lactate dehydrogenase, and low albumin

level were associated with poor outcomes, i.e. increased the risk of FL-specific death, after first-, second-, and third-line treatment. Third, being female reduced the risk of FL-specific death following treatment 2. On the other end, cancer stage of patient, the presence of B-symptoms, number of nodal and extra nodal sites involved, and low hemoglobin level did not have statistical significant effect on FL-specific death following each treatment stage.

The rest of this thesis is organized as follows. In Chapter 2, we provide a detailed descriptive statistic of various treatment regimens, and clinical and socio-demographic factors in the NLCS data set. In Chapter 3, we review the literature about different methodologies used to analyze time-to-event data sets and shed light on important findings from previous FL studies. In Chapter 4, we give a detailed illustration of the methods and models utilized in the identification of key prognostic. In Chapter 5, we present the results obtained from our analysis. In Chapter 6, we discuss the strengths and limitation of our analysis and provide scope for future research.

Chapter 2

Patients and Data Input

We identified a total number of 2,740 patients from the NLCS dataset for our multistate survival analysis study. The NLCS is the largest observational study and database in FL in the U.S., conducted to examine FL treatments and clinical outcomes for newly diagnosed patients [16]. All the eligible patients in the NLCS were diagnosed with FL within six months of their enrollment and none of them had a prior history of lymphoma. The study was multicentered and was conducted by biotechnological companies Genentech, located in San Francisco, CA, and Biogen Idec, located in Cambridge, MA. The study collected information about treatment regimens, patient outcomes such as response to treatment and cause of death, patient demographics, and clinical features. We briefly discuss the parameters and variables in the data set and report the proportion of each category for each factor in the following subsection. Conceptual and operational details about NLCS have been published elsewhere [16].

2.1 Data Input

Out of a total of 2,740 patients in the dataset, two were excluded from our analysis due to missing information about treatment strategies. In addition, we excluded thirty patients due to conflicting information about the cause of death. The remaining 2,708 patients were included in the current analysis.

The median age of the 2,708 adult patients included in our study were 61.5 years. We categorized patients into six age groups as follows: age < 50 (22.5%), age between 50-60 (24.8%), 60-70 (25.3%), 70-80 (19.7%), 80-90 (7.2%), and age $>$ 90 (0.7%). This categorization is in line with the published medical literature [20]. Other patient characteristics that were available at diagnosis were sex, Ann Arbor stage, blood hemoglobin level, the presence of B-symptoms, the number of nodal and extra nodal site involvements, histologic FL grade, and the presence of bone marrow (BM) involvement. The Ann Arbor stage of FL at diagnosis was classified into two categories: stage I & II (early), and stage III & IV (advanced). Hemoglobin levels less than and above 12 g/dL were respectively classified as low and normal. Providing critical information about the area and the growth of FL, the number of nodal sites involved were divided into two categories, < 5, and ≥ 5, and the number of extra nodal sites were divided into three categories, 0, 1, and ≥ 2 sites involved. Histological FL grade were classified as grade I (low) and grade II or III (high). Some of these variables were not available for several patients.

Two other constant variables that were measured at the time of diagnosis and included in the NLCS dataset were Eastern Cooperative Oncology Group (ECOG) performance scale [21] and Follicular Lymphoma International Prognostic Index (FLIPI) score [22]. ECOG is utilized to assess and quantify the impact of the disease on a patient's functional status and daily abilities. Based on their performance

status, patients were categorized into 3 classes in our dataset: normal (0), restricted but capable (1), requires assistance or disabled (2).[1] FLIPI is a risk assessment and classification tool, developed by *Solal-Céligny et al.* [22] in 2004. It uses five adverse prognostic factors (namely, age, FL stage, number of nodal areas, hemoglobin level, and lactic acid dehydrogenase (LDH) level) to classify patients into three risk groups: low/good risk, intermediate risk, and high/poor risk.

There were several dynamic variables that were collected at the time of diagnosis and prior to the initialization of each treatment regimen. These variables were serum LDH level, platelet level, white blood cell (WBC) level, and albumin level. These dynamic variables are categorized into three categories; low/high, normal and not available. Albumin level < 3 g/dL, WBC level $<$ 3,000/mm3, and platelet level $<$ 100,000 /mm3 are labelled as low or high-risk factor as they are associated with adverse outcomes. Whereas, Elevated LDH level associated with adverse outcome is labelled as high. On the other hand, low-risk factors for each of these four dynamic variable are labelled as normal and are considered as the baseline or reference category for future analysis.

Time and presence of cause-specific death were the primary prediction outcomes of our study. Cause-specific deaths were classified into two categories: death due to FL, and death from other causes. For the patients who died during the clinical trial (i.e., NLCS) or before the last follow-up, the time of deaths were documented in the dataset. Yet, since the NLCS data is right-censored, it is not known whether any patient experienced death after the time of the last follow-up (i.e. 03/2014). We report the proportions of patient characteristics and clinical factors together with the percentages of cause-specific deaths in Table 2.1.

[1]ECOG classes 0, 1 and 2 respectively correspond to 90-100, 70-80, and $\leq$60 in Karnofsky performance status (KPS) [21].

Table 2.1: Patient Characteristics and Clinical Factors

Patient Characteristics and Clinical Features	**Categories for Prognostic Factors**	**Proportion of Patients (n = 2,708)**	**Percentage of Death**	
			Death FL (n = 261)	**Death Other (n = 360)**
Age	≤ 50	22.5%	10.0%	6.1%
	51-60	24.8%	15.7%	11.7%
	61-70	25.3%	24.5%	23.3%
	71-80	19.7%	32.6%	36.9%
	81-89	7.2%	15.3%	19.7%
	≥ 90	0.7%	1.9%	2.2%
Sex	Male	48.4%	47.1%	51.7%
	Female	51.6%	52.9%	48.3%
Follicular Lymphoma International Prognostic Index (FLIPI)	Good	28.2%	7.3%	11.1%
	Intermediate	24.2%	21.5%	20.3%
	Poor	27.3%	52.5%	43.1%
	Not Available	20.3%	18.8%	23.1%
Stage of Disease	I & II	32.5%	17.6%	27.2%
	III & IV	66.5%	81.2%	71.1%
	Not Available	1.0%	1.1%	1.7%
Hemoglobin Level	≥ 12 g/dL	73.8%	57.5%	60.8%
	< 12 g/dL	19.5%	36.0%	30.6%
	NA	6.6%	6.5%	8.6%
Histology (Cancer Type)	Grade I	42.1%	37.5%	39.2%
	Grade II & III	46.3%	47.9%	47.2%
	Not Available	11.6%	14.6%	13.6%
B-Symptoms	No	74.7%	64.0%	69.4%
	Yes	25.3%	36.0%	30.6%

Table 2.1: Patient Characteristics and Clinical Factors (continued)

Patient Characteristics and Clinical Features	Categories for Prognostic Factors	Proportion of Patients (n = 2,708)	Percentage of Death	
			Death FL (n = 261)	Death Other (n = 360)
Extra Nodal Site Involved	None	45.2%	36.0%	38.6%
	1	32.8%	33.3%	40.8%
	≥ 2	18.4%	27.2%	16.4%
	Not Available	3.5%	3.4%	4.2%
Nodal Site Involved	< 5	63.4%	57.9%	60.3%
	≥ 5	32.4%	38.7%	34.4%
	Not Available	4.2%	3.4%	5.3%
Bone Marrow Involvement	No	46.6%	34.9%	36.9%
	Yes	28.5%	34.1%	29.2%
	Not Available	24.8%	31.0%	33.9%
Eastern Cooperative Oncology Group (ECOG)	0	46.8%	32.6%	30.0%
	1	18.9%	23.0%	26.1%
	2	3.3%	6.5%	6.9%
	Not Available	31.0%	37.9%	36.9%
Albumin Level at Diagnosis	≥ 3.0 g/dL	83.5%	74.3%	82.5%
	< 3.0 g/dL	4.5%	10.3%	6.1%
	Not Available	12.1%	15.3%	11.4%
Lectate Dehydrogenase (LDH)	Normal	59.9%	45.2%	53.3%
	Elevated	16.2%	29.5%	18.1%
	Not Available	24.0%	25.3%	28.6%
White Blood Cell (WBC) Count	$\geq 3{,}000/\text{mm}^3$	94.1%	95.0%	93.3%
	$< 3{,}000/\text{mm}^3$	3.4%	3.1%	3.9%
	Not Available	2.6%	1.9%	2.8%

Table 2.1: Patient Characteristics and Clinical Factors (continued)

Patient Characteristics and Clinical Features	Categories for Prognostic Factors	Proportion of Patients (n = 2,708)	Percentage of Death	
			Death FL (n = 261)	Death Other (n = 360)
Platelet Level at Diagnosis	$\geq$ 100,000/mm^3	93.8%	93.5%	90.8%
	$<$ 100,000/mm^3	3.5%	3.8%	6.1%
	Not Available	2.8%	2.7%	3.1%

Therapies received by the patients over the course of their treatment process were categorized into 7 classes in the dataset: chemotherapy, rituximab (R)-chemotherapy, R-monotherapy, investigational therapy, non-investigational therapy, radioimmunotherapy, radiotherapy, and bone marrow transplant (Table 2.2).

Table 2.2: The Distribution of Therapies Received at Each Treatment Stage

Treatment (TX)	TX 1 (n = 2475)	TX 2 (n = 1026)	TX 3 (n = 472)	TX 4 (n = 248)	TX 5 (n = 144)	TX 6 (n = 77)
Chemotherapy	4.2%	9.5%	15.5%	15.7%	23.6%	18.2%
Rituximab (R) Chemotherapy	61.5%	40.3%	38.6%	35.5%	31.3%	37.7%
R-Monotherapy	18.7%	25.7%	17.4%	14.9%	10.4%	9.1%
Investigational Therapy	7.5%	4.5%	7.0%	7.7%	7.6%	5.2%
Non-Investigational Therapy	0.7%	1.6%	3.0%	3.2%	6.9%	5.2%
Radioimmunotherapy	0.1%	4.3%	6.1%	4.8%	3.5%	16.9%
Radiotherapy	7.2%	13.4%	8.7%	9.7%	8.3%	7.8%
Bone Marrow Transplant	0.0%	0.9%	3.8%	8.5%	8.3%	0.0%

Chapter 3

Literature Review and The Key Differences of Our Study

In this section we provide an overview on two streams of research regarding lymphoma treatment literature. First, we discuss different analytical or computational methods used in the previously published literature for the management of lymphoma, such as survival analysis, statistics, and machine learning. Second, we discuss the medical literature relevant to FL treatment. Finally, we confer about how our current study contributes to the existing literature.

3.1 Theoretical Background and Methodology Review

The occurrence and time of disease progression, death, and other clinical events and outcomes are among the key measures in most medical and epidemiological studies. Yet, most clinical data used in these studies are *censored*, where information on time-to-event is usually not completely available for all patients participating in

the study. As a result, the direct use of standard statistical methods such as logistic regression and analysis of variance has a limited application for clinical datasets. Instead, survival analysis, a branch of statistics that can cope with *censored* data, is commonly utilized to analyze time-to-event type data in clinical research.

Especially within the last three decades, survival analysis methods have been increasingly utilized to study different chronic and infectious diseases such as breast cancer, lung cancer, HIV/AIDS, and tuberculosis [23–26]. As a result, survival analysis-based approaches have led the identification of numerous factors affecting the time and rate of the events of interest, such as death or disease progression, for various diseases and different patient cohorts [23, 24, 27]. In addition, the use of survival analysis techniques have also gained popularity in other areas including non-medical real-life problems such as time until stock market crash, time until next earthquake, time until divorce, and time until equipment failure. [28].

In general, the survival analysis techniques used for analyzing censored data can be broadly classified into non-parametric (e.g., Kaplan-Meier product limit), parametric (e.g., Weibull distribution) and semi-parametric (Cox proportional hazards regression) methods. The most common uses of these techniques are (i) to identify significant factors affecting the time of the event of interest and (ii) to estimate of survival probability as a function of a risk factor over time. In particular, Kaplan-Meier estimator is used to project the survival probability at each point in time, Cox proportional hazards models are used to quantify the impact of risk factors on survival, and parametric models are used for both.

The most common and simplest framework for survival analysis methods is two-state models such as alive-and-death, treatment-and-death, and diagnosis-and-progression. As the name implies, a two-state model consists of two model states and a single transition between these states. Commonly, one of the states is transient

(e.g., alive) and the other one is absorbing (e.g., death), where the rate and time of the transition to the absorbing state and the factors affecting this transition are of the main interests.

Despite its popularity in clinical literature, two-state models have a couple of pitfalls. First, previous studies have utilized survival curves, calculated by Kaplan-Meier (KM) estimator [29] over two-state models, to make comparison between subgroup of patients using log-rank test [30]. However, application of simple log-rank test might lead to misleading results because of the presence of heterogeneous samples in cancer studies [31] and instead, more robust regression methods, such as Cox regression [32], should be applied to evaluate the effect of multiple explanatory variables on the outcome variable [31]. Second, two-state models lead to biased estimates and potentially misleading conclusions in the presence of multiple interdependent clinical events, such as progression of disease and death, and competing risk for multiple transitions such as cause-specific mortality [33, 34]. In such cases, these clinical events and the transitions between them should be studied with a more comprehensive multi-state model rather than each transition being studied separately by a two-state model in isolation [35].

Following the availability of multi-event datasets and increased awareness about the pitfalls of two-state models, multi-state models, an extension of competing risk models with multiple transient and absorbing states, have found increased application in medical research. One of the earliest use of multi-state model was proposed by *Fix et al.* in the year 1951 while studying relapse and recovery of cancer patients [36]. More recently, *Çağlayan et al.* developed a multi-state model to study the combined effect of sequential therapies and other explanatory factors on the treatment-specific survival of older cohort of FL patients using Cox PH regression and Aalen-Johansen estimator [9, 37]. Another study conducted by *Çağlayan*

et al. evaluated competing risk of death amongst patients diagnosed with Diffuse Large B-Cell Lymphoma (DLBCL), an advanced form of follicular lymphoma, within a multi-state modelling framework [38]. The study identified significant factors by fitting Cox PH model between all the model states in the multi-state model. Similarly, studies involving the understanding of complex biological process such as, cancer, diabetics, leukemia, etc., have seen widespread application of multi-state models in recent history [39–41].

Aalen-Johansen (AJ) estimator, a generalization of Kaplan-Meier estimator, is the canonical non-parametric estimator for multi-state survival analysis models [37]. It is used to calculate the empirical transition matrices and to project the course of the underlying stochastic process over time. The AJ estimator can cope with censored observations in the clinical datasets, makes no assumption on the probability distribution generating the stochastic process of interest, and is shown to produce reliable estimates both for Markov and non-Markov models [38]. There are also recent variations of the AJ estimator attempting to generate more consistent estimates particularly for non-Markov multi-state models [42, 43].

To quantify the effect of the covariates on the transitions between model states, parametric or semi-parametric regression models are utilized in the literature. The difference between these two modeling approaches lies in the way in which researchers make assumption about the underlying probability distribution for the hazard function corresponding to the transition of interest. For parametric models, it is assumed that the baseline hazard function follows a known distribution such as Exponential or Weibull distribution. This allows researchers to make prediction about the survival of patients beyond the n^{th} observed survival time. On the other hand, semi-parametric models, such as Cox PH, makes no assumption about the underlying distribution for the baseline hazard function. It allows to quantify the increased risk/hazard of

a factor compared to the reference group/level of that factor but does not specify the baseline function of the underlying probability distribution. Cox PH is the most widely used semi-parametric regression model in the clinical literate and the studies analyzing FL patients have made extensive use of Cox PH model [9, 38, 44–47]. In particular, *Provencio et al.* utilized univariate and multivariate Cox PH model to identify significant risk factors associated with long-term survival of FL patients in Spain [44] whereas *Swenson et al.* compared the survival of FL patients in three different decades and showed significant improvement in overall survival using multivariate Cox PH regression model that is adjusted for the effect of confounding variables [45].

In addition to aforementioned statistical methods, the use of machine learning (ML) techniques for survival analysis was proposed by researchers as early as mid 1990s [48, 49]. Yet, the results from these initial ML implementations were not revolutionary as they failed to outperform the standard Cox PH regression method on a breast cancer dataset [50]. Later, more advanced and capable ML algorithms have been proposed for survival analysis such as deep learning [51, 52], random survival forest [53], and support vector regressor for censored data [54]. Though proven to be efficient for the simple two-state survival analysis models, these ML-based approaches still fail to incorporate the effect of competing risk into the modeling framework and hence, is not currently applicable to the multi-state models yet. Three distinct method, Siamese survival analysis [55], Tree-Based Bayesian mixture model [56], and DeepHit [57], incorporates the effect of competing risks into the model. However, the application of these models in the real-world have been limited so far. Accordingly, their performance is needed to better investigated particularly before being applied to clinical datasets.

3.2 Clinical Review

Over the past two decades, there have been significant improvements in the overall survival (OS) and progression free survival of patients diagnosed with FL [45, 58, 59]. These improvements were associated with various factors including improvement in supportive care of patients, sequential application of treatment regimens, and introduction of more effective therapies such as Rituximab, Tositumomab, and Bendamustine [60–62]. Following the increase in survival rates and the availability of the sequential treatments, recent research in the last decade focused on studying the clinical course of FL under various treatments, analyzing short- and long-term disease patterns, and identifying risk factors associated with worse prognosis and survival. In this subsection, we provide an overview of some of the recent medical studies that are relevant to our research.

Earlier investigators studied the effect of different front-line strategies on the OS of FL patients [11, 12]. Comparing aggressive modality treatments (n = 60) with watchful-waiting approach (n = 44) for indolent lymphomas such as FL, *Young et al. (1988)* found no significant difference in overall survival amongst the two groups [12]. *Neri et al. (2001)*, studied the impact of interferon-alpha2b by comparing chemotherapy treatment with and without interferon-alpha2b and found no difference in the OS of patients in the two treatment arms [11]. Later, there have been several studies indicating improvement in event-, relapse- and progression-free survival of FL patients, no therapy has found to be statistically superior to another [11, 17, 18, 63].

More recently, *Friedberg et al. (2012)* studied the impact of various therapies on FL outcomes with a subset of stage I FL patients (n=471) in the NLCS dataset, and found no statistical difference in the OS amongst patients receiving different treatment regimens after adjusting for significant risk factors such as histology, LDH, and B-

symptoms [46]. On the other hand, studying 6,568 adult stage I-II FL patients from the Surveillance, Epidemiology, and End Results (SEER) database, *Pugh et al. (2010)* identified an improvement in the OS and lymphoma-specific survival amongst patients receiving radiation therapy (compared with patients receiving other treatments) [64]. Their multivariate analysis accounted for age, sex, race, stage, FL grade, and extra nodality.

Casulo et al. (2015a) investigated 588 patients with advanced FL stage, who were treated with first-line rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) therapy. In this cohort, the authors showed that progression of disease (PoD) < 2 years was associated with poor five year OS [8]. Examining 2,652 FL patients in the NLCS dataset, *Casulo et al. (2015b)* observed similar disease characteristics and outcomes for the young (i.e., age < 40) and middle-aged (i.e., age 40-60) patients and concluded that young adults might not necessarily benefit from more aggressive FL therapies [20]. Finally, leveraging a large SEER dataset (n = 5234), *Çağlayan et al. (2019)* studied the sequential first-, second- and third-line FL treatments for an older cohort (i.e., $>$ age 65) of high-risk FL patients [9]. The study identified that the use of R-CHOP at any line of treatment improved the OS after each treatment stage when compared with other treatment alternatives such as, rituximab, R-CVP (rituximab, cyclophosphamide, vincristine) and R-other (rituximab with other chemotherapy combinations).

3.3 Our Main Contributions and Key Differences

Currently, there is no comprehensive analysis of various sequential FL treatments over the general population with a large dataset. Existing studies are limited due to the use of simple survival analysis models, small sample size, restricted number

of treatment options, or the concentration on a specific sub-population (e.g., over age 65). In our study, we take the next step to assess the effectiveness of sequential FL therapies by leveraging NLCS dataset with a multi-state survival analysis framework. Our analysis covers the age range that corresponds to the whole adult population and up to six sequential therapies over the clinical course of FL while comparing the outcomes of eight different treatment regimens for each line of therapy.

In this study, we developed comprehensive continuous-time multi-state models to capture the clinical course of FL. We utilized Cox PH regression model to identify the significant risk factors affecting each clinical event, such as cause-specific death and the initiation of the next line of therapy, and Aalen-Johansen estimator to project the prognosis of FL over time as a function of critical factors. In particular, we developed two different multi state models, one with a single absorbing state named "all-cause death" and the other with two absorbing states corresponding to cause-specific deaths and compared the results from these two models. Further, utilizing a multivariable Cox regression model over our multi-state framework, we successfully quantified the impact of prognostic factors on FL-specific death after each line of therapy in the presence of other confounding variables while accounting for the risk of death due to non-lymphoma related causes. By utilizing the entire cohort of NLCS patients, we showed that none of existing treatment regimens have shown superior FL outcomes compared to one another. Based on our findings, we suggest that clinical research work should particularly focus on developing more advanced FL treatment regimens to further reduce the risk of FL-specific death.

Chapter 4

Method

Clinical "time-to-event" datasets are often censored, where the time of certain events are not recorded. For instance, a patient might be alive during the clinical trial and at the last follow-up, but the exact time of death for this patient might not be available in the dataset. The phenomena of censoring is categorized into three types; right-censored data, as explained in the above example, left-censored data, in which the event occurs before the participant enters the study period, and interval-censored data, in which an events occurs between two successive clinical visit, example recurrence of cancer. Fig 4.1 depicts the three types of censoring observed in a dataset in graphical form. In the NLCS dataset we only observe the most common type of censoring, i.e. right-censored data. This type of censoring mostly corresponds to the patients who were lost to follow-up or were alive at the last follow-up. We assume that the censoring of patients is independent and non-informative as there are no information indicating the otherwise.

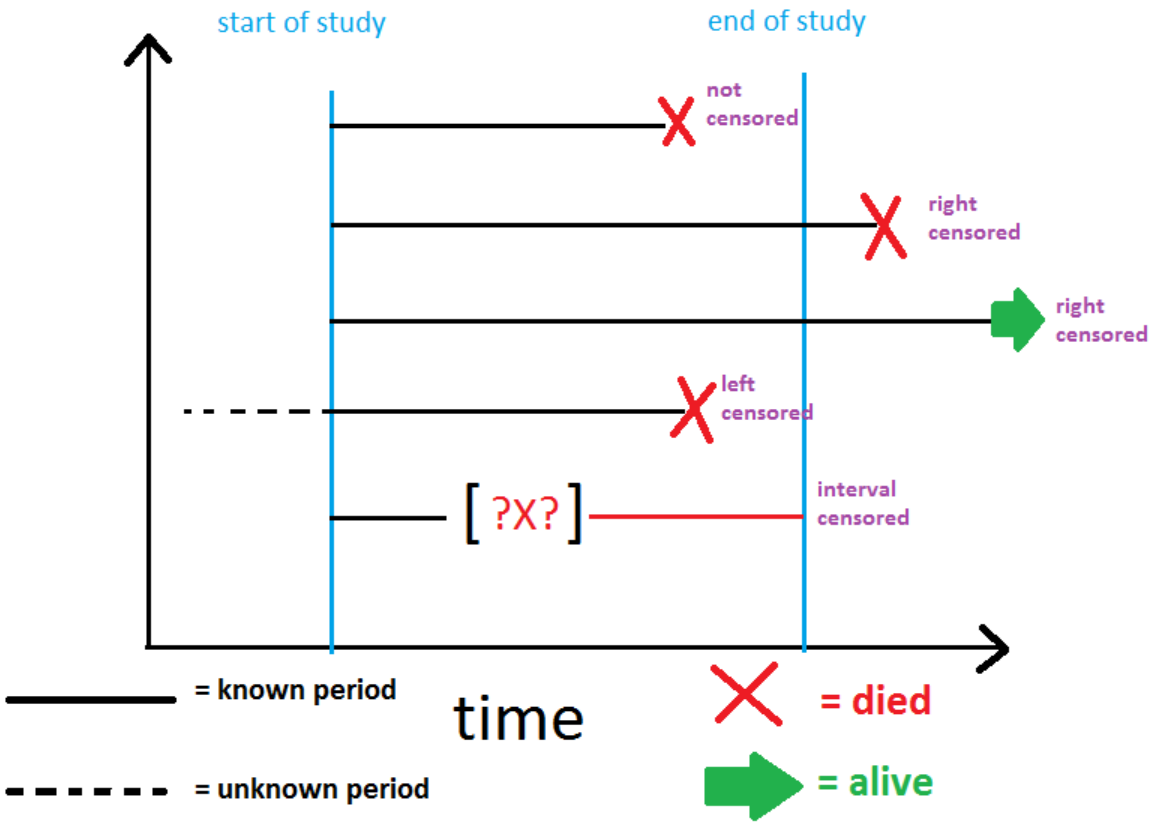

Figure 4.1: Types of Censoring [65]

4.1 Kaplan-Meier & Aalen Johansen

For survival analysis of two-state models, the underlying survival function $S(t)$, measuring the probability of being alive (i.e., not making a transition to the death state) beyond time t, is given by the following equation,

$$S(t) = P(T > t) \tag{4.1.1}$$

where T is a non-negative random variable representing the time until the occurrence of an event (e.g., death). We can also write the survival function in terms of the

cumulative density function (CDF) of survival time as follows:

$$\begin{aligned} S(t) &= P(T > t) \\ &= 1 - P(T \leq t) \\ &= 1 - F(t) \end{aligned} \tag{4.1.2}$$

In this equation, $F(t)$, the CDF of survival time, defines the risk of experiencing the event of interest at time point t.

Survival function of a two-state model is frequently estimated with Kaplan-Meier (KM) estimator. KM estimator, also known as the product limit estimator, is the maximizer of the nonparametric maximum likelihood of the survival function $S(t)$ [66], and hence, considered as the standard nonparametric estimator for two-state models [29]. We can formally formulate the KM estimator as follows: Let $t_1 < t_2 < t_3 ... < t_j$ represent the time t at which $d_1 < d_2 < d_3 ... < d_j$ number of people in the study experience the event of interest. Further, let $n_1 < n_2 < n_3 ... < n_j$ represent the number of people at the risk of experiencing an event at these specified time points. Then, the KM estimator of the survival function can be represented by $\hat{S}(t)$:

$$\hat{S}(t) = \prod_{j:t_j<t} (1 - \frac{d_j}{n_j}) \tag{4.1.3}$$

Yet, for multi-state model having multiple possible transitions from a model state, the use of KM estimator in isolation for each transition was shown to yield biased estimates [34, 67, 68]. Instead, Aalen and Johansen proposed a generalized matrix version of KM estimator for multi-state models [37]. In our analysis, we use the Aalen-Johansen (AJ) estimator over the multi-state model we develop to project

the course of FL over time. AJ estimator is represented as follows:

$$\widehat{P}(s,t) = \prod_{j:s<t_j\leq t} (I + \hat{a}_j) \tag{4.1.4}$$

In this formulation, the transition probability matrix $\widehat{P_{gh}}(s,t)$ denotes the probability that an individual in the model state g at time s will be in the model state h at a future time t. "I", and $\hat{a}_j$ are the identity, and transition intensity matrices, respectively, of order k $\times$ k, where k represents the total number of model states.

For transitions from a transient state g into an other transient or an absorbing state h, where $g \neq h$, the element (g,h) of the $\hat{a}_j$ matrix is calculated as follows:

$$\widehat{a_{ghj}} = \frac{d_{ghj}}{n_{gj}} \tag{4.1.5}$$

Here, d_{ghj} represents the number of people transitioned form a transient state g to another transient/absorbing state h, at time t_j, and n_{gj} represents the total number of people in the state g just prior to time t_j. On the other hand, the intensity values on the diagonal of the matrix $\hat{a}_j$, i.e when $g = h$, are calculated as follows:

$$\widehat{a_{ggj}} = -\frac{d_{gj}}{n_{gj}} \tag{4.1.6}$$

Here, d_{gj} represents the total number of people transition out of the state g, helping us to formulate the transition intensity matrix $\hat{a}_j$.

4.2 Cox- Proportional Hazard Regression

To identify significant risk factors and quantify their effect on the time and rate of transitions between model states corresponding clinical events,, we fit Cox PH

regression models [32, 69] to each transition in the multi-state model. This enables us to (i) express the rate of a clinical event (e.g., death or initiation of the next treatment) as a function of covariates at any given time and (ii) quantify the impact of each covariate (e.g., sex=female) on the corresponding transition compared to the baseline/reference value of that covariate (e.g., male). The Cox PH regression model is given by the following equation:

$$\lambda(t|z) = \lambda_0(t)e^{\beta z} \tag{4.2.1}$$

In this formulation, $\lambda(t|z)$ is the rate of transition at time t as a function of the vector of variables z. The rate $\lambda(t|z)$ is a function of covariate vector z, regression coefficient vector β and a non-parametric component $\lambda_0(t)$, known as the baseline hazard function. In this equation, $\lambda_0(t)$ depends on the value of time but not on the covariate vector z, whereas, the exponential function, $e^{\beta z}$, depends on the value of covariates z and is independent of time. Accordingly, the Cox regression model is known as a semi-parametric model due to the presence of non-parametric hazard function and the parametric covariate function.

In the given regression model, the hazard function $\lambda(t)$ can be mathematically expressed as follows:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t \mid T > t)}{\Delta t} \tag{4.2.2}$$

The numerator in the equation is a conditional probability of experiencing the event of interest in a small time interval $[t, t + \Delta t]$, given that the subject was event free until time t, and the denominator is the width of the time interval. On further simplification of equation 4.2.2, we can derive the relation between the hazard and

survival function as follows:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \times \frac{P(t < T \leq t + \Delta t \cap T > t)}{P(T > t)} \tag{4.2.3}$$

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \times \frac{P(t < T \leq t + \Delta t)}{P(T > t)} \tag{4.2.4}$$

Since, the numerator in the equation 4.2.4 is the difference between two CDFs at times $(t + \Delta t)$ and t, and the denominator is the survival function defined in equation 4.1.1, we can rewrite this equation as follows:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \times \frac{F(t + \Delta t) - F(t)}{S(t)} \tag{4.2.5}$$

where,

$$\lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dx} F(t) \tag{4.2.6}$$

By substituting the value of $F(t)$ from equation 4.1.2 in equation 4.2.6, we get,

$$\begin{aligned} \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} &= \frac{d}{dx}[1 - S(t)] \\ &= -S'(t) \end{aligned} \tag{4.2.7}$$

Further, by substituting equation 4.2.7 in equation 4.2.5, we get,

$$\lambda(t) = \frac{-S'(t)}{S(t)} \tag{4.2.8}$$

On integrating both sides from time 0 to t, the equation 4.2.8 can be expressed as

$$\ln S(t) = -\int_0^t \lambda(u)du$$
$$S(t) = \exp\{-H(t)\} \tag{4.2.9}$$

Here, $H(t)$ is cumulative hazard rate defined as $H(t) = \int_0^t \lambda(u)du$. Hence, the equation 4.2.9 shows the relation between the survival function and the hazard function, which allows us to capture the effect of covariates using the hazard function.

The interpretation of Cox PH model is made using hazard ratio (HR), defined as the ratios of predicted hazard function for two individuals with different set of covariates. For example, the HR for a patient with a vector of covariates z_1, compared to another patient with a vector of covariates z_2 is given by

$$\begin{aligned} HR(z_1, z_2) &= \frac{\lambda(t|z_1)}{\lambda(t|z_2)} \\ &= \frac{\lambda_0(t)e^{\beta z_1}}{\lambda_0(t)e^{\beta z_2}} \\ &= e^{\beta^T(z_1 - z_2)} \end{aligned} \tag{4.2.10}$$

Now, instead of comparing two patients with many different covariate values, let us consider two individuals with the same covariate vector, except for a single variable. For instance, let us consider two independent individuals, one male (z_{11} = 1) and other female (z_{21} = 0), with similar set of covariate vector, except for the variable "sex". Substituting the values of covariate vector in equation 4.2.10, we obtain

$$HR(z_1, z_2) = e^{\beta_1} \tag{4.2.11}$$

This term provides the relative effect of one-unit increase of the covariate "sex" (i.e.,

switching to being male=1 from the reference value female=0) on the risk of the event of interest. As demonstrated by this example, HR quantifies the relative increase or decrease in the risk of the event of interest due to the level of a risk factor compared to the reference value/level of that factor. Similarly, if we have multiple different covariates that affect a particular transition, then hazard ratio for two patients with different set of covariates is represented as

$$HR(z_1, z_2) = e^{(\beta_1(z_{11}-z_{12})+\beta_2(z_{21}-z_{22})+\ldots+\beta_n(z_{n1}-z_{n2}))} \tag{4.2.12}$$

The values of the regression coefficients (i.e., the components of the vector β) are estimated using Cox's partial likelihood estimate [32]. We considered the effect of a covariate as significant when the 95% confidence interval (CI) of the regression coefficient for a given covariate does not include the value zero (i.e., when $p < 0.05$). Note that the existence of a variable increases the relative risk when the associated coefficient of the β is positive. From the equation 4.2.10, it can be seen that HR is independent of the time function, as the the baseline hazard function cancels out. This implies that the proportional hazard (i.e., the relative risk due to a certain value of a factor compared to its reference value) is kept constant over time. In our study, we verified this assumption of the Cox PH models, known as proportional hazard assumption, for our final multivariable Cox models using Schoenfeld residuals [70].

4.3 Two-Step Analysis for Cox PH Regression

One is of the main objectives of our study is to identify which levels of each categorical variable, if any, are statistically significant predictors for the clinical events captured by our multi-state model. To do so, we utilized Cox PH regression models

and employed a two-step approach as described below.

In the first step, we separately conducted a univariate Cox PH regression analysis for each prognostic factor in the Tables 2.1 and 2.2. That is, we assumed that only a single prognostic factor played a role on each transition between the model states and attempted to identify the levels of the prognostic factor that are statically significant (compared to the reference level). This analysis, corresponding to the first step of our approach, provided us with the candidate variables (and their levels) that we later further used when we started building our multivariable models. An example of a univariate Cox PH model fitted for a dichotomous variable B-symptoms (present or absent) is provided in the Figure 4.2.

```
coxph(formula = Surv(Tstart, Tstop, status) ~ BSYMP.1 + BSYMP.2 +
    BSYMP.3 + BSYMP.4 + BSYMP.5 + BSYMP.6 + BSYMP.7 + BSYMP.8 +
    BSYMP.9 + BSYMP.10 + BSYMP.11 + BSYMP.12 + BSYMP.13 + BSYMP.14 +
    BSYMP.15 + strata(trans), data = ms_model_base,
    method = "breslow")
```

Figure 4.2: Univariate Cox PH Model

The output obtained form conducting a univariate analysis for the prognostic factor B-symptoms for each one of the 15 transitions between the model states, labeled as .1, .2, ..., and .15, is as shown in Figure 4.3 below.

	coef	exp(coef)	se(coef)	z	Pr(>|z|)	
BSYMP.1	-0.66532	0.51411	0.12713	-5.233	1.66e-07	***
BSYMP.2	0.39555	1.48521	0.04783	8.269	< 2e-16	***
BSYMP.3	0.40681	1.50201	0.15970	2.547	0.01085	*
BSYMP.4	0.85671	2.35541	0.41538	2.062	0.03916	*
BSYMP.5	0.08607	1.08988	0.06948	1.239	0.21540	
BSYMP.6	0.11581	1.12278	0.14697	0.788	0.43071	
BSYMP.7	0.30710	1.35948	0.09924	3.095	0.00197	**
BSYMP.8	0.46794	1.59671	0.17898	2.615	0.00894	**
BSYMP.9	0.37479	1.45468	0.13441	2.788	0.00530	**
BSYMP.10	0.69483	2.00337	0.22113	3.142	0.00168	**
BSYMP.11	0.03375	1.03433	0.17780	0.190	0.84945	
BSYMP.12	0.30454	1.35601	0.32470	0.938	0.34828	
BSYMP.13	-0.01017	0.98988	0.24260	-0.042	0.96655	
BSYMP.14	0.39157	1.47931	0.35301	1.109	0.26732	
BSYMP.15	0.50096	1.65030	0.32177	1.557	0.11950	

Figure 4.3: Output Obtained from Univariate Cox PH Model

These results suggest that the presence of B-symptoms among diagnosed FL patients might have a statistically significant impact on the 1^{st}, 2^{nd}, 3^{rd}, 4^{th}, 7^{th}, 8^{th}, 9^{th}, and 10^{th} transitions of the model presented in Figure 4.4. Accordingly, when we build our multivariable Cox PH models, corresponding to the second step of our approach, we include the presence of B-symptoms as a candidate predictor for these transitions.

In the second step, we build our multivariable Cox PH models and derive the final multivariable model for each transition by eliminating the insignificant variables. To do so, we begin with all the covariates identified by our univariate Cox models and develop an initial multivariable model for each transition. Later, by eliminating the insignificant variables over a stepwise approach, we derive the final models where all of the variables remained (i.e., not excluded) are statistically significant predictors of the corresponding transition. By using this two step analysis, we systematically construct

our multivariable Cox models for each transition while targeting to minimize the bias introduced in our model due to the presence of confounding variables.

4.4 Multi-State Model

We developed two separate multistate models to capture the clinical course of FL over time. The first model, as demonstrated in Figure 4.4, has nine model states, consisting of diagnosis, watchful waiting (WW), six treatment stages, and an absorbing health state "death". The total number of transition (n = 15) in the model is also listed in Figure 4.4. We refer this model as "*all-cause death model*". Model states are indicated by circles, and transitions from one model state to another is indicated by arrows.

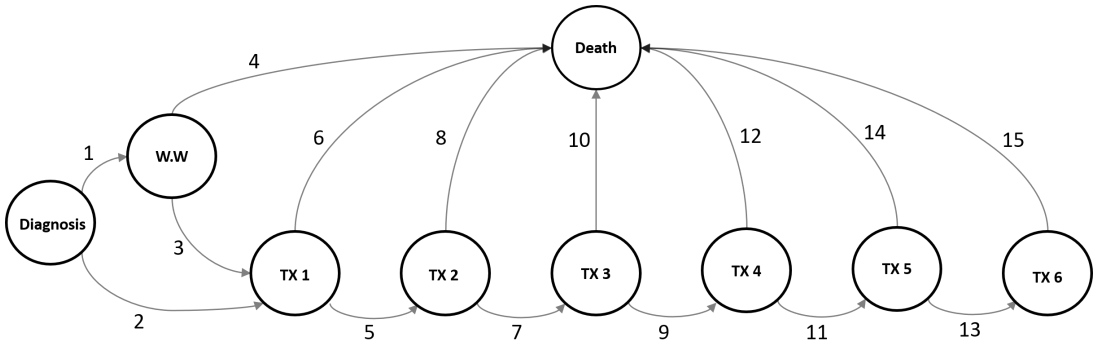

Figure 4.4: All-Cause Death Model

In the second model, we split the "death" state into two cause-specific mortality states (namely, death due to FL and death due to other causes) to better analyze the impact of patient characteristics and clinical factors on cause-related deaths. We refer this model as *cause-specific death model* and present its representation in Figure 4.5. In both models, a patient enters a treatment state with the initiation of the corresponding line of treatment and departs either when she dies, by transitioning to

(one) the "death" state(s) or when her next treatment is initiated, by transitioning to the next treatment state.

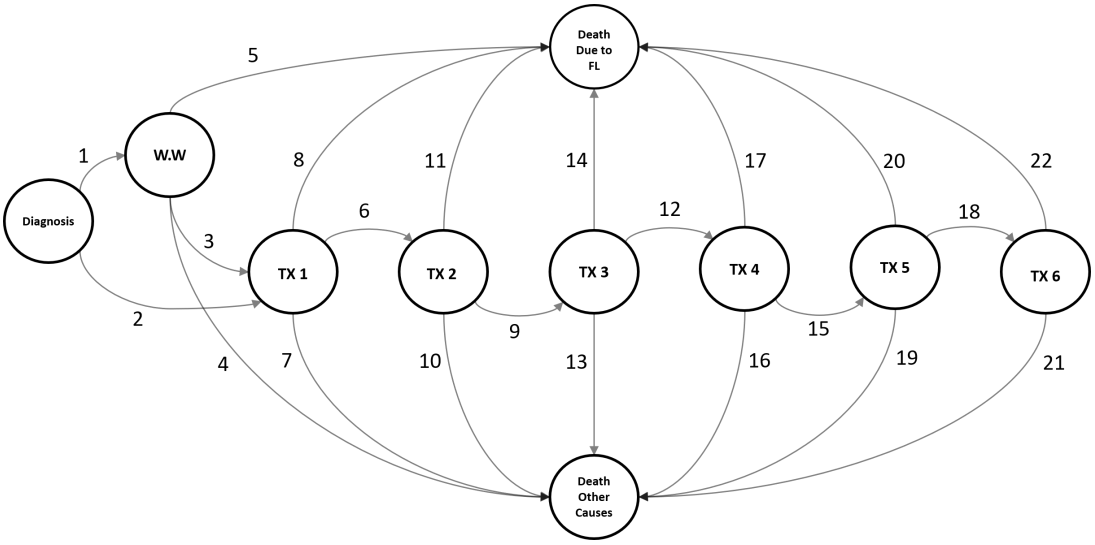

Figure 4.5: Cause-Specific Death Model

The observational period for a patient begins once he/she is diagnosed with FL and ends with a transition into the absorbing state. During this time period, patients make at least one transition into the subsequent transient state, Watchful Waiting (WW) or First-Line Treatment (TX 1), before being censored or entering the absorbing state. Patients transitioned into "WW" are kept under observation until symptoms develop that require medical intervention or therapy. But, for patients transitioning into "TX 1", therapies are initiated immediately by doctors/physicians. After this initial transition, a patient either enters the next transient state to receive a therapy or enters the absorbing state "death", where death is recorded, or the information is otherwise censored. Our model allows us to capture the effect of different covariates and therapies on each transition and quantify its effect. In the following section we present statistically significant risk factors that influence the transition to death, obtained by applying CPH regression model to each transition.

Chapter 5

Results

Our multi-state analysis is based on a cohort of 2,708 patients diagnosed with FL. These FL patients were enrolled in National LymphoCare Study (NLCS), a comprehensive clinical study with a 10-year follow-up period. A total of 2,475 patients received first line treatment after diagnosis. Amongst the remaining two hundred and thirty-three patients, death of forty-four patient's was recorded and one hundred and eighty-nine patients were censored following a period of "Watchful Waiting". Patient characteristics such as age, cancer type, B-symptoms, etc., were recorded at the time of diagnosis and are listed in Table 2.1

The number of patients receiving second-line, third-line, fourth-line, fifth line and sixth-line therapies were 1026, 472, 248, 144 and 77, respectively. The median duration was 0.92 months from diagnosis to first-line, 0.72 months from diagnosis to watchful waiting, and 5.85 months from watchful waiting to first-line. The median duration from first-line to second-line, second-line to third-line, third-line to fourth-line, fourth-line to fifth line, and fifth-line to sixth-line therapy was 6.64, 7.23, 6.21, 6.24, and 5.16 months, respectively. During the follow-up period of the NLCS, 621 total and 261 (42%) FL-related deaths were registered.

Utilizing multivariable Cox regression models, we identified the key risk factors for FL outcomes and quantified their impact on all-cause and FL-related deaths. Table 5.1 and Table 5.2 lists the significant risk factors identified from both the models. For patients undergoing "watchful waiting", the statistically significant factors associated with worse survival outcome, compared with the patients in the reference category (shown in Table 5.1), were the following [hazard ratio (HR) (95% confidence interval (CI)]: older age (age 71-80: 3.84 (1.7-8.9), age 81-89: 7.29 (3.0-17.5), age $\geq$ 90: 28.33 (3.4-237.1)), poor-risk FLIPI Score [2.63 (1.2-6.0)], presence of B-symptoms [2.26 (0.9-5.7)], low albumin level ($<$ 3.0 g/dL) [12.80 (3.2-50.5)], low platelet level ($<$ 100,000/mm3) [3.14 (1.0-10.1)] and performance status score (PSS) $<$ 60 [12.65 (3.3-47.8)]. Poor-risk FLIPI score [10.61 (1.7-65.7)] low platelet level ($<$ 100,000/mm3) [11.37 (1.1-115.18)] were significant predictors for *death due to FL*, the remaining factors, that is, age $>$ 70 [9.73 (2.1-45.8)], presence of B-symptoms [2.81 (1.0-7.6)], low albumin level 16.46 (3.8-70.9), and PSS $<$ 60 [13.81 (3.40-56.1)] were predictors for *death due to other causes*.

Following the first-line treatment, the risk factors associated with increased all-cause mortality risk were age (age 61-70: 4.33 (2.2-8.7), age 71-80: 10.03 (5.0-19.9), age 81-89: 18.79 (9.1-38.8), age $\geq$ 90: 20.97 (6.6-66.4)), Poor-risk FLIPI score (1.93 [1.0-3.9]), low albumin level ($<$ 3.0 g/dL) (1.79 [1.1-3.0]), elevated LDH level (1.50 [1.0-2.2]), PSS between 70-80 (1.66 [1.2-2.4]), and PSS less than 60 (2.22 [1.2-4.0]). In particular, age between 61-70 (2.45 [1.0-6.1]), age 71-80 (6.02 [2.5-14.2]), age 81-89 (11.72 [4.6-29.7]), poor-risk FLIPI score (6.0 [1.0-34.5]), low albumin level ($<$ 3.0 g/dL) (2.98 [1.4-6.4]), and elevated LDH level (1.80 [1.0-3.4]) were predictors for *death due to FL*, and age $\geq$ 90 (13.76 [4.2-44.7]), PSS between 70-80 (1.84 [1.2-2.8]), PSS less than 60 (2.27 [1.1-4.6]), and were predictor of *death due to other causes*. Bone marrow involvement (1.48 [1.0-2.2]) was also found to be a predictor for

other cause mortality. Additionally, investigational therapy (n=3) and radiotherapy (n=1) received during treatment 1 were significant predictors for *death due to other causes*, but this significance might be influenced by relatively low number of patients experiencing this transition after treatment 1.

Following the second-line treatment, the risk factors associated with increased all-cause mortality risk were age (age 51-60: 2.50 (1.1-5.6), age 61-70: 4.00 (1.8-8.9), age 71-80: 7.81 (3.6-17.1), age 81-89: 12.55 (5.3-29.5), age $\geq$ 90: 31.12 (8.8-109.8)), low albumin level ($<$ 3.0 g/dL) (3.82 [2.0-7.4]), presence of B-symptoms (1.52 [1.0-2.2]), elevated LDH (2.47 [1.5-4.1]) and PSS less than 60(2.42 [1.2-4.7]). In particular, age between 50-60 (4.78 [1.0-22.0]), age 61-70 (8.76 [1.9-40.3]), age 71-80 (15.81 [3.5-71.1]), age 81-89 (20.44 [3.9-106.1]), age greater than 90 (51.4 [6.6 - 401.7]), low albumin level ($<$ 3.0 g/dL) (4.51 [1.9-10.9]), elevated LDH level (7.61 [3.3-17.3]), and PSS less than 60 (4.16 [1.5-11.2]), were significant predictors of *death due to FL*. Additionally, two significant predictors, being female (0.46 [0.3-0.8]) and investigational therapy received during treatment 1 (4.66 [1.7-12.4]), were also found to be significant predictors of *death due to FL*.

Following the third-line of treatment, the risk factors associated with increased all-cause mortality risk were age (age 71-80: 2.58 (1.4-4.9), age 81-89: 6.27 (2.9-13.3), intermediate FLIPI score (2.53 [1.0-6.7]), Poor-risk FLIPI score (3.14 [1.2-8.2]), low platelet level ($<$ 100,000/mm3) (2.51 [1.4-4.5]), presence of B-symptoms (2.02 [1.2-3.3]), LDH level (>ULN) (3.62 [2.0-6.5]), PSS between 70-80 (1.82 [1.0-3.3]), and rituximab-monotherapy received during treatment 3 (0,44 [0.2-0.9]). In particular, age (age 71-80: 2.59 (1.3-5.3) , age 81-89: 2.83 (1.0-7.9)), Poor-risk FLIPI score (2.40 [1.0-5.9]), low platelet level ($<$ 100,000/mm3) (2.70 [1.2-5.8]), and elevated LDH level (3.47 [1.6-7.3]) are significant predictors of *death due to FL*. Further, patients receiving rituximab-monotherapy during treatment 1 are at a reduced risk (0.28 [0.1-0.7]) of

death due to other causes compared with patients receiving chemotherapy.

Following the forth-line of treatment, the risk factors associated with increased all-cause mortality risk were age (age 71-80: 2.92 (1.2-7.3), age 81-89: 8.71 (2.4-31.5)), and low platelet level (< 100,000/mm3) (4.85 [1.8-12.9]). Further, factors associated with reduced risk of all-cause mortality after treatment 4 were rituximab-monotherapy (n = 4) received during treatment 4 (0.20 [0.0-0.8]), radiotherapy (n=3) received during treatment 4 (0.13 [0.0-0.6]), BMT (n=1) received during treatment 4 (0.1 [0.0-0.8]), and rituximab-monotherapy (n=1) received during treatment 3 (0.10 [0.0-0.9]). In particular, low platelet level (< 100,000/mm3) (5.58 [1.8-17.6]), rituximab-monotherapy (n=2) received during treatment 4 (0.18 [0.0-0.9]), radiotherapy (n=1) received during treatment 4 (0.1 [0.0-0.7]), and bone marrow transplant (n=1) received during treatment 4 (0.1 [0.0-0.7]) were significant predictor of *death due to FL*. Age (age 71-80: 20.08 (3.8-106.5)) was found to be associated with increased risk of other-cause mortality.

Following the fifth line of treatment, the risk factors associated with increased all-cause mortality risk were age (age 71-80: 3.11 (1.1-8.9)), and radioimmunotherapy (n=4) received during treatment 2 (6.70 [1.8-24.9]). In particular, radioimmunotherapy (n=2) received during treatment 2 (5.51 [1.1-26.8]) was found to be a significant predictor of *death due to other causes*, whereas, low albumin level (< 3.0 g/dL) (3.63 [1.0-13.6]) was found to be the only significant predictor associated with increased risk of *death due to FL*.

Following the sixth line of treatment, the risk factors associated with increased all-cause mortality risk were radioimmunotherapy (n=5) during treatment 2 (6.70 [1.8-24.9]). Bone marrow involvement (2.82 [1.0 – 8.0]) is a significant predictor of *death due to FL*. Age was not found to be a significant predictor of cause-specific death after treatment 5. Hence, creating a cause-specific model enabled us to identify

and measure the factors affecting *death due to FL* more accurately.

Table 5.1: Significant Factors for Death due to Follicular Lymphoma ($\mathbf{D_{FL}}$): Hazard Ratio (95% CI)

Factors (Reference Group)	**Classification**	**WW → D_{FL}**	**TX 1 → D_{FL}**	**TX 2 → D_{FL}**	**TX 3 → D_{FL}**	**TX 4 → D_{FL}**	**TX 5 → D_{FL}**	**TX 6 → D_{FL}**
Age (≤ 50)	51 - 60			4.78 (1.0 - 22.0)*				
	61 -70		2.45 (1.0 - 6.1)*	8.79 (1.9 - 40.3)*				
	71 -80		6.02 (2.5 - 14.2)*	15.81 (3.5 - 71.1)*				
	81 - 89		11.72 (4.6 - 29.7)*	20.44 (3.9 - 106.1)*	2.83 (1.0 - 7.9)*			
	geq 90			51.4 (6.6 - 401.7)*	14.67 (1.2 - 186.7)			
Sex (Male)	Female			0.46 (0.3 - 0.8)				
FLIPI Score (Good)	Poor (3-5)	10.61 (1.7 - 65.7)*	6.0 (1.0 - 34.5)*		2.40 (1.0 - 5.9)*			
Hemoglobin (≥ 12 g/dL)	Not Done			3.92 (1.3 - 12.3)*				
Albumin (≥ 3 g/dL)	(< 3 g/dL)		2.98 (1.4 - 6.4)*	4.51 (1.9 - 10.9)*			3.63 (1.0 - 13.6)	
Platelet ($\geq 100,000/mm^3$)	($< 100,000/mm^3$)	11.37 (1.1 - 115.18)*		2.70 (1.2 - 5.8)*	5.58 (1.8 - 17.6)*			
BMT (No)	Yes							2.82 (1.0 - 8.0)
LDH (Normal)	> ULN		1.80 (1.0 - 3.4)*	7.61 (3.3 - 17.3)*	3.47 (1.6 - 7.3)*			
	Unknown			5.16 (2.5 - 10.8)*		3.43 (1.0 - 18.7)		

Factors (Reference Group)	**Classification**	**WW → D_{FL}**	**TX 1 → D_{FL}**	**TX 2 → D_{FL}**	**TX 3 → D_{FL}**	**TX 4 → D_{FL}**	**TX 5 → D_{FL}**	**TX 6 → D_{FL}**
ECOG (100-90)	< 60			4.16 (1.5 - 11.2)*				
	Unknown		1.88 (1.0 - 3.4)*					
TX 1 (Chemo)	Investigational Therapy			4.66 (1.7 - 12.4)				
TX 2 (Chemo)	R-Mono					0.10 (0.0 - 0.9)*		
TX 4 (Chemo)	R-Mono					0.18 (0.0 - 0.9)*		
	Radiotherapy					0.10 (0.0 - 0.7)*		
	BMT					0.10 (0.0 - 0.7)*		

* indicates that the factor is also a significant predictor of Death due to all-cause

Abbrevations: Chemo, Chemotherapy; R-Mono, Rituximab Monotherapy; ECOG, Eastern Cooperative Oncology Group; FLIPI, Follicular Lymphoma International Prognostic Index; BMT, Bone Marrow Transplant; LDH, Lectate Dehydrogenases; TX, Treatment

Table 5.2: Significant Factors for All-Cause Death ($\mathbf{D_a}$): Hazard Ratio (95% CI)

Factors (Reference Group)	**Classification**	**WW → D_{FL}**	**TX 1 → D_{FL}**	**TX 2 → D_{FL}**	**TX 3 → D_{FL}**	**TX 4 → D_{FL}**	**TX 5 → D_{FL}**	**TX 6 → D_{FL}**
Age (≤ 50)	51 - 60				2.50 (1.1 - 5.6)			
	61 -70		4.33 (2.2 - 8.7)	4.00 (1.8 - 8.9)				
	71 -80	3.84 (1.7 - 8.7)	10.03 (5.0 - 19.9)	7.81 (3.6 - 17.1)	2.58 (1.4 - 4.9)	2.92 (1.2 - 7.3)	3.11 (1.1 - 8.9)	
	81 - 89	7.29 (3.0 - 17.5)	18.79 (9.1 - 38.8)	12.55 (5.3 - 29.5)	6.27 (2.9 - 13.3)	8.71 (2.4 - 31.5)		
	≥ 90	28.33 (3.4 - 237.1)	20.97 (6.6 - 66.4)	31.12 (8.8 - 109.8)				
Stage (I & II)	NA			3.11 (1.0 - 9.2)				
FLIPI Score (Good)	Intermediate (2)				2.53 (1.0 - 6.7)			
	Poor (3-5)	2.63 (1.2 - 6.0)	1.93 (1.0 - 3.9)		3.14 (1.2 - 8.2)			
Hemoglobin (≥ 12 g/dL)	Not Done			2.83 (1.4 - 5.8)				
Albumin (≥ 3 g/dL)	(< 3 g/dL)		12.80 (3.2 - 50.5)	1.79 (1.1 - 3.0)	3.84 (2.0 - 7.4)			
Platelet ($\geq 100,000/mm^3$)	($< 100,000/mm^3$)	3.14 (1.0 - 10.1)			2.51 (1.4 - 4.5)	4.85 (1.8 - 12.9)		
BMT (No)	Unknown	2.29 (1.1 - 4.8)						
B-Symptoms (No)	Yes			1.52 (1.- 2.2)	2.02 (1.2 - 3.3)			

Factors (Reference Group)	**Classification**	**WW → D_{FL}**	**TX 1 → D_{FL}**	**TX 2 → D_{FL}**	**TX 3 → D_{FL}**	**TX 4 → D_{FL}**	**TX 5 → D_{FL}**	**TX 6 → D_{FL}**
LDH (Normal)	> ULN		1.50 (1.0 - 2.2)	2.47 (1.5 - 4.1)	3.62 (2.0 - 6.5)			
	Unknown			2.50 (1.6 - 4.8)				3.37 (1.0 - 11.2)
ECOG (100-90)	80 - 70	12.65 (3.3 - 47.8)	1.66 (1.2 - 2.4)	2.42 (1.2 - 4.7)	1.82 (1.0 - 3.3)			
	< 60		2.22 (1.2 - 4.0)					
	Unknown		1.91 (1.4 - 2.6)		1.77 (1.0 - 3.2)			
TX 2 (Chemo)	Radioimmunotherapy						6.92 (1.5 - 31.7)	6.70 (1.8 - 24.9)
TX 3 (Chemo)	R-Mono				0.44 (0.2 - 0.9)			
TX 4 (Chemo)	R-Mono					0.20 (0.0 - 0.8)		
	Radiotherapy					0.13 (0.0 - 0.6)		
	BMT					0.10 (0.0 - 0.8)		

Abbrevations: Chemo, Chemotherapy; R-Mono, Rituximab Monotherapy; ECOG, Eastern Cooperative Oncology Group; FLIPI, Follicular Lymphoma International Prognostic Index; BMT, Bone Marrow Transplant; LDH, Lectate Dehydrogenases; TX, Treatment

In a multi-state framework, transition probabilities estimated using A-J estimator are also known as cumulative incidence function (CIF). It is defined as the probability or the risk of moving from one model state to another during a specified time interval. In Fig 5.1 and Fig 5.2, we plot the CIF for all the transient and absorbing states in the two models, all-cause death and cause-specific death, respectively.

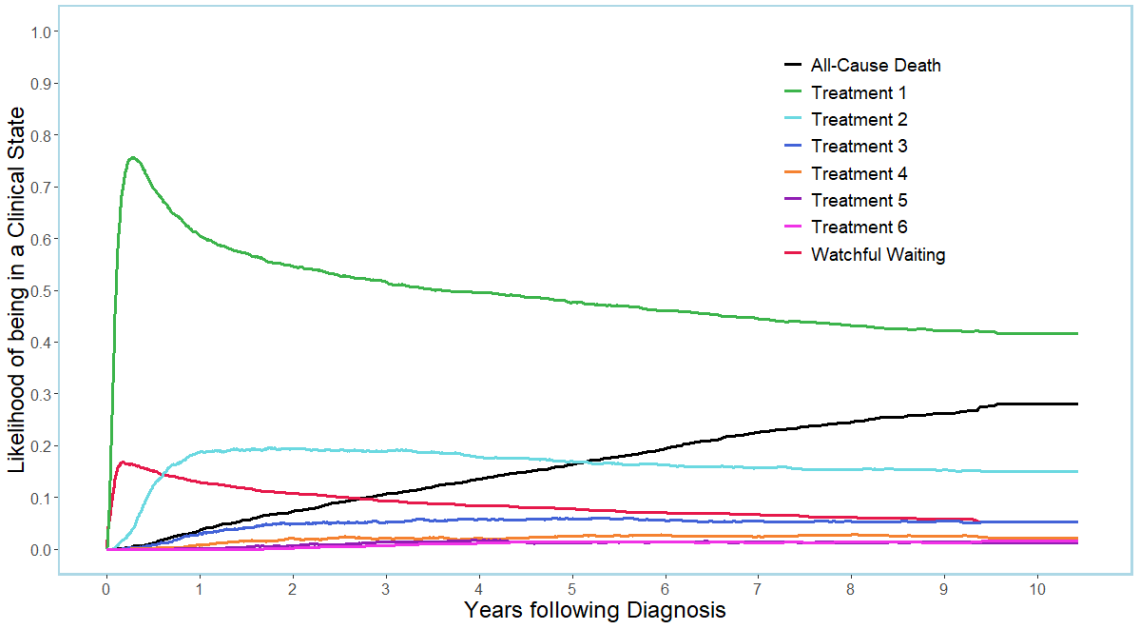

Figure 5.1: Likelihood of occupying a model state following diagnosis for the All-cause death model

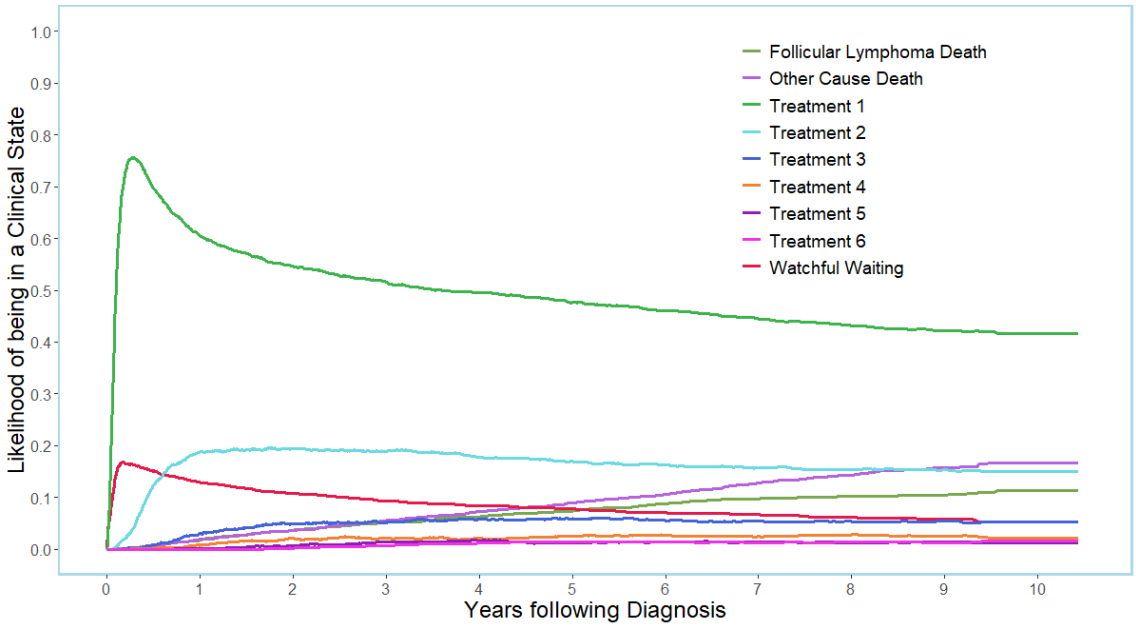

Figure 5.2: Likelihood of occupying a model state following diagnosis for the Cause-specific death model

Transient states in both our multi-state models are similar, therefore, the only difference observed in Fig 5.1 and Fig 5.2 is between the CIF curve for the absorbing death state. Wherein, the value of the CIF for all-cause death, at each time point, is the addition of CIF for FL-specific death and other-cause specific death. We isolate these three CIF curves in Fig 5.3 for better visualization.

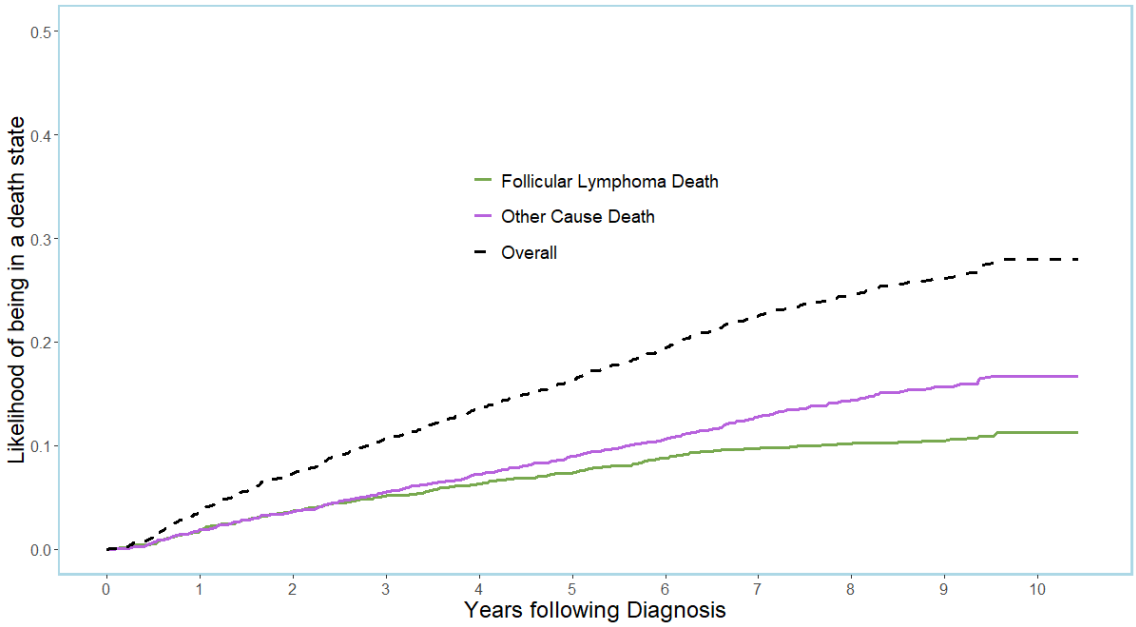

Figure 5.3: CIF for Cause-Specific death following diagnosis

The 2-, 5-, and 8-years probability of *death due to all cause*, and cause-specific death is presented in Table 5.3.

Table 5.3: Transition Probabilities for All-cause death following Diagnosis

	Cumulative Incidences, %		
Years	**FL-Death**	**Other Cause Death**	**All-Cause Death**
2	3.7	3.6	7.3
5	7.4	9.0	16.4
8	10.1	14.4	24.5

Abbreviation: FL, Follicular Lymphoma

Utilizing graphical and tabular result, we can make a generic statement that the risk of *death due to FL* is approximately equal to the risk of *death due to other-causes* up to the 3-year time point. After which, we observe that the risk of *death due to other-causes* exceeds the risk of *death due to FL*. On further investigation of

CIF for all-cause death, we found that the risk of all-cause death for patients with or without watchful waiting were almost identical (Fig 5.4 A). With WW patients being at reduced risk of all-cause death. Among WW patients experiencing death, it is observed that the risk of *death due to FL* is lower than the risk of *death due to other-causes* (Fig 5.4 B). On the other hand, the risk of *death due to FL* and *death due to other-causes* is identical up to the 5-years time point for patients receiving an induction treatment therapy (i.e without watchful waiting) (Fig 5.4 C). Later on, i.e. after the 5-year time point, the risk of *death due to other-causes* exceeds the risk of *death due to FL* for this group of patients.

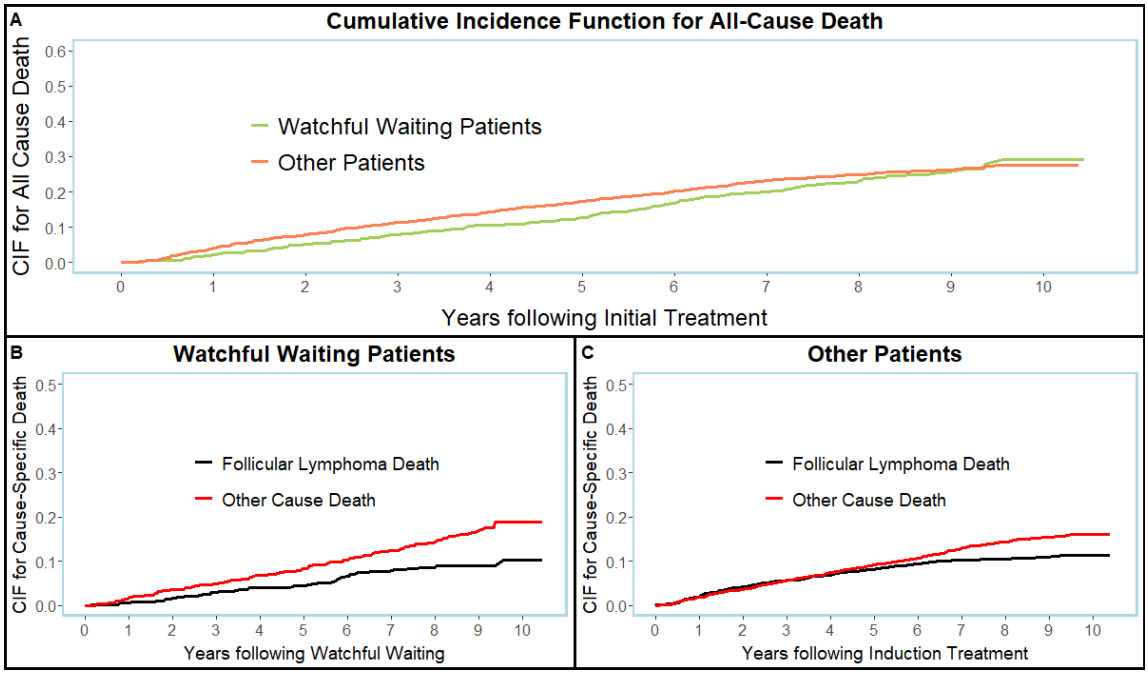

Figure 5.4: (A) All-Cause death CIF for Watchful Waiting and Other Induction treatment, (B) Death-specific CIF for Watchful Waiting patients and (C) Death-specific CIF for Other Patients

Table 5.4 provides death-specific transition probabilities for watchful waiting and other patients at 2-, 5-, and 8-year time points.

Table 5.4: Death-Specific Transition Probabilities following Initial Therapy

	Cumulative Incidences, %			
	WW Patients		**Other Patients**	
Years	**FL-Death**	**Other-Cause Death**	**FL-Death**	**Other-Cause Death**
2	1.6	3.6	4.3	3.7
5	4.5	8.3	8.1	9.2
8	8.9	14.4	10.6	14.3

Abbreviation: FL, Follicular Lymphoma; WW, Watchful Waiting

The risk of *death due to FL* associated with WW patients was 1.6%, 4.5% and 8.9% at 2-, 5-, and 8-years time point, respectively. Likewise, the risk of *death due to FL* among other patients, i.e. without watchful waiting, was 4.3%, 8.1% and 10.6% at the 2-, 5-, and 8-year time point. Results indicate that the risk of *death due to FL* is relatively less amongst WW patients compared to other patients. Whereas, the risk of *death due to other-cause* remains almost identical among the two subgroup of patients throughout the study time period. This further indicates that WW patients are at a higher risk of developing other chronic disease than dying due to FL, compared to patients without watchful waiting. Therefore, a different treatment strategy is required to control the growth of such disease amongst WW patients.

Chapter 6

Discussion and Conclusion

6.1 Discussion

Survival of patients with FL has been improving in the United States [45, 58–60] as more effective treatment regimens and strategies become available. However, currently, there is still no standard care for the first line and subsequent therapies of FL in the United States. Various management strategies, adopted by physicians to treat patients diagnosed with FL at different lines of treatment, including watchful waiting, rituximab(R)-chemotherapy, R-monotherapy, bone marrow transplant, investigational therapy, non-investigational therapy, and radiotherapy [16, 46]. The heterogeneity in FL prognosis and outcomes, and significant differences in OS among FL patients further complicate therapeutic decision-making [16, 71]. Accordingly, studying the efficacy of various sequential treatment strategies, identifying key factors contributing to worse FL outcomes, and projecting the prognosis of FL as a function of treatment and clinical factors remain important.

Including relatively large number of patients and various treatment regimens, the NLCS dataset provides a unique opportunity to formally assess the effect of

different treatment options and strategies over the clinical course of FL and identify the predictors for poor FL outcomes. Using National Lymphocare Study (NLCS) dataset, several studies have been conducted in recent years to evaluate the factors associated with high risk of death among patients diagnosed with FL [46, 72, 73]. The primary focus of these studies have been to evaluate the effect of single line treatment on clinical outcomes such as progression free survival (PFS) or OS [8, 9, 20]. In particular, *Martin et al.* identified age greater than 60 years, LDH level greater than upper limit of normal, hemoglobin level less than 12 g/dL, and ECOG performance status (PS) greater than or equal to 2 to be significant predictors of OS amongst patients receiving front-line combination therapies, i.e. rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP), rituximab plus cyclophosphamide, vincristine, and prednisone (R-CVP), and rituximab plus a fludarabine-containing regimen (R-Flu) [27]. Similarly, *Casulo et al.* in his study of high-risk NLCS patients, i.e. patients experiencing early progression of disease (POD), identified POD, age, and ECOG PS to be significant predictors of OS after front-line treatment with R-CHOP [8]. Another study by *Casulo et al.* utilizing NLCS dataset demonstrated that age greater than 60 is a significant predictor of OS for FL patients [20]. Yet, no study has analyzed the effect of sequential therapy among FL patients in NLCS dataset. Also, the focus of prior studies has been to identify the effect of prognostic factors on the OS of patients, i.e. death due to any cause, whereas, in our study we successfully evaluate the effect of prognostic factors on the cause-specific deaths, i.e. death due to FL and death due to other causes.

In our present study, we evaluate the effect of treatment regimens across the clinical course of FL with the NLCS data. We found that, amongst the significant factors affecting the survival of patients receiving watchful waiting and no further treatment, poor FLIPI score and low platelet level were the only two significant pre-

dictors of death due to FL. For the first-, third-, fifth- and sixth-line treatment, none of the treatment received during that treatment stage or prior treatments showed significant advantage over the baseline treatment group, i.e. chemotherapy, when evaluating death due to FL patients. On the other end, investigational therapy received during treatment stage 1 increased the risk of FL related mortality after treatment stage 2. In addition, R-monotherapy received during treatment 3 and 4, and radiotherapy therapy and BMT received at treatment stage 4, reduced the risk of death due to FL after treatment stage 4 among the NLCS patients. Another important finding from our analysis suggest that elevated LDH level is a significant predictor of FL mortality after first-, second-, third-, and fourth-line treatment. Likewise, low albumin and low platelet level have shown to increase the relative risk of death due to FL after treatment 1, 2 and 5, and treatment 3 and 4, respectively. These dynamic variables collected at the initiation of each treatment have not been studied in detailed before and therefore should be included in future studies for accurate analysis. A unique finding from our analysis suggests that females are at reduced risk of death due to FL after treatment stage 2 [HR: 0.46; 95% CI: 0.3-0.8] when compared with males. On the contrary, nodal and extra nodal sites, previously proven to be significant, are not identified as significant predictors of both, FL and all-cause mortality, after each of the treatment stage [22]. Also, the presence of B-symptoms at treatment stage 2 and 3 increases the risk of all-cause mortality but fails to show significant effect when analyzing FL mortality. Similarly, ECOG PS that shows significant effect on all-cause mortality after watchful waiting, treatment stage 1, 2 and 3, but does not increase/decrease the relative risk of FL mortality, except for treatment stage 2. Hence, we can say, that analyzing cause-specific model helps us better understand the effect of prognostic factors on the FL mortality and that, future studies should analyze FL patients using competing risk models, like the one studied in the present

thesis.

Previously, using another large dataset (N = 5234) obtained the Surveillance, Epidemiology, and End Results (SEER) registry, *Çağlayan et al.* demonstrated that the use of R-CHOP in first-, second- and third-line treatment increases the OS for high-risk older patients (i.e. age > 65) [9]. In contrast to the results from that study, treatment specific improvements, that reduces the risk of FL mortality, is not observed in the NLCS data set until treatment stage four. Even with specific treatments showing improved outcomes after treatment 4, the finding might be limited since very few (less than 10) patients experience FL death after treatment 4 in each of the treatment arm. Hence, more data would be required to confirm these findings. One explanation for this discrepancy in the results from two studies can the classification of rituximab-chemotherapy into treatment regimens, such as, R-CHOP, R-CVP, R-Other and Rituximab, along with the absence of patients aged less than 65 in the study published by *Çağlayan et al.*. Whereas, in our dataset of FL patients, where patient characteristics are similar to that observed in the SEER8 dataset, is a better generalization of the population in the United States. Patients in the current study are observed to receive various treatments including rituximab-chemotherapy. Another possible explanation could be the presence of high-risk patients, i.e. early progression of disease (POD), versus inclusion of patients with or without POD in our study.

6.2 Limitation

As any study, ours is not free of limitations. One of the study-related limitation is the ten-year time period, over which the data was collected on FL patients. To better investigate long-term prognosis of FL in the rituximab era it is necessary to increase the horizon of data collection process beyond 10-year period for future

studies. Another limitation in the NLCS dataset is the absence of information about unknown confounding variables, such as β2-microglobulin, which have proven to be significant prognostic factor in FL studies [74]. Further, the presence of the other chronic disease or conditions (e.g., breast cancer, lung cancer, secondary leukemia) might have influenced the selection of succeeding treatment or outcome following the current treatment. This information was not reported in the NLCS dataset.

In our study, we evaluated the effect of eight treatment regimes on the event of interest after each treatment stage. The NLCS data set also contains information about an additional maintenance therapy known as Rituximab(R)-Maintenance received in between two successive treatments. It is not an independent treatment by itself but an extension of the preceding treatment to control disease progression and extend the duration of disease remission. It has been utilized widely by doctors and physicians for the treatment of NHL patients. Although, studies have shown significant improvement in PFS of patients treated with rituximab-maintenance after induction therapy when compared with observation, no significant improvement is observed in the OS of patients [75, 76]. Since, the aim of our study was to evaluate the effect of prognostic factors on cause-specific death survival after each treatment line and not PFS, we adjusted the sequence of treatment by eliminating the intermediate R-maintenance therapy from the analysis data set. This limits our ability to quantify the effect of R-maintenance on the cause-specific deaths.

6.3 Summary and Conclusion

In this study, we conducted a multi-state survival analysis by developing a multi-state modeling framework and utilizing a large dataset. Our objectives were to (i) examine the performance of sequential FL treatment strategies over the adult U.S.

population and (ii) investigate the impact of sociodemographic and clinical factors recorded in the NCLS dataset on the clinical course of FL. The use of a comprehensive multi-state model allowed us to identify statistically significant prognostic factors affecting the mortality risk at each treatment stage. Further, using a cause-specific death model enabled us to accurately calculate the risk of FL-specific death over time together with the factors increasing this risk. Ours is the first study examining the outcomes related to FL-specific death involving a diverse group of patients from all age groups.

The main findings of our studies are as follows: We showed that (1) none of the treatment regimens yield superior FL outcomes in terms of FL-specific survival, (2) after 3 year time-point following the FL diagnosis, the risk of other-cause death surpassed the risk of FL-specific death, indicating that other mortality risk should be monitored as seriously as lymphoma for FL patients, (3) watchful waiting patients are a higher risk of other-cause specific death, and (4) dynamic variables such as low albumin level, low platelet and elevated LDH levels have significant impact on FL-specific death after certain treatment stage. In addition, we further verified that old age, poor FLIPI score, male sex, and poor ECOG have significant impact on lymphoma related mortality risk, as previously shown by other studies.

In the past few years, there have been significant improvements in the development of survival analysis specific machine learning algorithms that can cope with censored data and can capture the presence of competing events. Incorporating these methods with more traditional survival analysis methods or comparison of these two approaches appear to be a promising research area. In our future work, we plan to work on integrating machine learning algorithms with multi-state survival analysis models and apply this approach to lymphoma and other chronic diseases.

Bibliography

Follicular lymphoma.
https://rarediseases.org/rare-diseases/follicular-lymphoma/, 2019.

Staging of lymphoma, 2019. URL `\https://lymphoma-action.org.uk/about-lymphoma-tests-diagnosis-and-staging/stages-lymphoma`.

James O Armitage and Dennis D Weisenburger. New approach to classifying non-hodgkin's lymphomas: clinical features of the major histologic subtypes. non-hodgkin's lymphoma classification project. *Journal of Clinical Oncology*, 16(8): 2780–2795, 1998.

Non-Hodgkin's Lymphoma Classification Project. A clinical evaluation of the international lymphoma study group classification of non-hodgkin's lymphoma. *Blood, The Journal of the American Society of Hematology*, 89(11):3909–3918, 1997.

Cancer facts & figures 2019. `\https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2019/cancer-facts-and-figures-2019.pdf`.

Alessia Castellino, Elisa Santambrogio, Maura Nicolosi, Barbara Botto, Carola Boccomini, and Umberto Vitolo. Follicular lymphoma: the management of elderly patient. *Mediterranean journal of hematology and infectious diseases*, 9(1), 2017.

Carla Casulo, Michelle Byrtek, Keith L Dawson, Xiaolei Zhou, Charles Farber, Christopher R Flowers, John D Hainsworth, Brian K Link, Andrew D Zelenetz, and Jonathan W Friedberg. Early relapse of follicular lymphoma after r-chop uniquely defines patients at high risk for death: An analysis from the national lymphocare study, 2013.

Carla Casulo, Michelle Byrtek, Keith L Dawson, Xiaolei Zhou, Charles M Farber, Christopher R Flowers, John D Hainsworth, Matthew J Maurer, James R Cerhan, Brian K Link, et al. Early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone defines patients at high risk for death: an analysis from the national lymphocare study. *Journal of Clinical Oncology*, 33(23):2516, 2015a.

Çağlar Çağlayan, Hiromi Terawaki, Turgay Ayer, Jordan S Goldstein, Ashish Rai, Qiushi Chen, and Christopher Flowers. Assessing the effectiveness of treatment sequences for older patients with high-risk follicular lymphoma with a multistate model. *Clinical Lymphoma Myeloma and Leukemia*, 19(5):300–309, 2019a.

Sandra J Horning and Saul A Rosenberg. The natural history of initially untreated low-grade non-hodgkin's lymphomas. *New England Journal of Medicine*, 311(23): 1471–1475, 1984.

Natividad Neri, Agustin Avilés, Sergio Cleto, Noe Díaz, Alejandra Talavera, Edna L García, and José C Díaz-Maqueo. Chemotherapy plus interferon-α 2b versus chemotherapy in the treatment of follicular lymphoma. *Journal of hematotherapy & stem cell research*, 10(5):669–674, 2001.

RC Young, DL Longo, E Glatstein, DC Ihde, ES Jaffe, and VT DeVita Jr. The treatment of indolent lymphomas: watchful waiting v aggressive combined modality treatment. In *Seminars in hematology*, volume 25, pages 11–16, 1988.

Carla Casulo, Loretta Nastoupil, NH Fowler, JW Friedberg, and CR Flowers. Unmet needs in the first-line treatment of follicular lymphoma. *Annals of Oncology*, 28(9): 2094–2106, 2017.

Loretta J Nastoupil, Rajni Sinha, Michelle Byrtek, Ryan Ziemiecki, Xiaolei Zhou, Michael Taylor, Jonathan W Friedberg, Brian K Link, James R Cerhan, Keith Dawson, et al. Outcomes following watchful waiting for stage ii–iv follicular lymphoma patients in the modern era. *British journal of haematology*, 172(5):724–734, 2016.

Pauline Brice, Yves Bastion, Eric Lepage, Nicole Brousse, Corinne Haioun, Philippe Moreau, Nicole Straetmans, Herve Tilly, Isabelle Tabah, and P Solal-Celigny. Comparison in low-tumor-burden follicular lymphomas between an initial no-treatment policy, prednimustine, or interferon alfa: a randomized study from the groupe d'etude des lymphomes folliculaires. groupe d'etude des lymphomes de l'adulte. *Journal of Clinical Oncology*, 15(3):1110–1117, 1997.

Jonathan W Friedberg, Michael D Taylor, James R Cerhan, Christopher R Flowers, Hildy Dillon, Charles M Farber, Eric S Rogers, John D Hainsworth, Elaine K Wong, Julie M Vose, et al. Follicular lymphoma in the united states: first report of the national lymphocare study. *Journal of Clinical Oncology*, 27(8):1202, 2009.

Pier Luigi Zinzani, Alessandro Pulsoni, Alessio Perrotti, Simona Soverini, Francesco Zaja, Amalia De Renzo, Sergio Storti, Vito Michele Lauta, Luciano Guardigni, Patrizia Gentilini, et al. Fludarabine plus mitoxantrone with and without rituximab versus chop with and without rituximab as front-line treatment for patients with follicular lymphoma. *Journal of clinical oncology*, 22(13):2654–2661, 2004.

Luca Baldini, Maura Brugiatelli, Stefano Luminari, Marco Lombardo, Francesco Merli, Stefano Sacchi, Paolo Gobbi, Marina Liberati, Luigi Cavanna, Mariangela Colombi, et al. Treatment of indolent b-cell nonfollicular lymphomas: final results of the ll01 randomized trial of the gruppo italiano per lo studio dei linfomi. *Journal of clinical oncology*, 21(8):1459–1465, 2003.

AD Nichol, M Bailey, DJ Cooper, On behalf of the POLAR, et al. Challenging issues in randomised controlled trials. *Injury*, 41:S20–S23, 2010.

C Casulo, B Day, KL Dawson, X Zhou, CR Flowers, CM Farber, JD Hainsworth, James R Cerhan, BK Link, AD Zelenetz, et al. Disease characteristics, treatment patterns, and outcomes of follicular lymphoma in patients 40 years of age and younger: an analysis from the national lymphocare study. *Annals of Oncology*, 26 (11):2311–2317, 2015b.

G Buccheri, D Ferrigno, and M Tamburini. Karnofsky and ecog performance status scoring in lung cancer: a prospective, longitudinal study of 536 patients from a single institution. *European Journal of Cancer*, 32(7):1135–1141, 1996.

Philippe Solal-Céligny, Pascal Roy, Philippe Colombat, Josephine White, Jim O Armitage, Reyes Arranz-Saez, Wing Y Au, Monica Bellei, Pauline Brice, Dolores Caballero, et al. Follicular lymphoma international prognostic index. *Blood*, 104 (5):1258–1265, 2004.

Eduard Vrdoljak, Kornelija Miše, Damir Sapunar, Ante Rozga, and Matko Marušić. Survival analysis of untreated patients with non-small-cell lung cancer. *Chest*, 106 (6):1797–1800, 1994.

Abbas Rezaianzadeh, Janet Peacock, Daniel Reidpath, Abdolrasoul Talei, Seyed Vahid Hosseini, and Davood Mehrabani. Survival analysis of 1148 women diagnosed with breast cancer in southern iran. *BMC cancer*, 9(1):168, 2009.

Jalal Poorolajal, Elaheh Hooshmand, Hossein Mahjub, Nader Esmailnasab, and Ensiyeh Jenabi. Survival rate of aids disease and mortality in hiv-infected patients: a meta-analysis. *Public Health*, 139:3–12, 2016.

Olurotimi Bankole Ajagbe, Zubair Kabir, and Terry O'Connor. Survival analysis of adult tuberculosis disease. *PloS one*, 9(11):e112838–e112838, 2014.

Peter Martin, Michelle Byrtek, Keith Dawson, Ryan Ziemiecki, Jonathan W Friedberg, James R Cerhan, Christopher R Flowers, and Brian K Link. Patterns of delivery of chemoimmunotherapy to patients with follicular lymphoma in the united states: results of the national lymphocare study. *Cancer*, 119(23):4129–4136, 2013.

Tyler Smith and Besa Smith. Survival analysis and the application of cox's proportional hazards modeling using sas. In *Proceedings of the twenty-sixth annual SAS user's group international conference*, pages 244–246. SAS Institute Inc Cary (NC), 2001.

Edward L Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

Richard Peto and Julian Peto. Asymptotically efficient rank invariant test procedures. *Journal of the Royal Statistical Society: Series A (General)*, 135(2):185–198, 1972.

KOUHEI Akazawa, Tsuyoshi Nakamura, and Yuko Palesch. Power of logrank test and cox regression model in clinical trials with heterogeneous samples. *Statistics in medicine*, 16(5):583–597, 1997.

David R Cox. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202, 1972.

Hein Putter, Marta Fiocco, and Ronald B Geskus. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in medicine*, 26(11):2389–2430, 2007.

Per Kragh Andersen, Ronald B Geskus, Theo de Witte, and Hein Putter. Competing risks in epidemiology: possibilities and pitfalls. *International journal of epidemiology*, 41(3):861–870, 2012.

Gary T Deimling, Jacquelyn A Arendt, George Kypriotakis, and Karen F Bowman. Functioning of older, long-term cancer survivors: The role of cancer and comorbidities. *Journal of the American Geriatrics Society*, 57:S289–S292, 2009.

Evelyn Fix and Jerzy Neyman. A simple stochastic model of recovery, relapse, death and loss of patients. *Human Biology*, 23(3):205–241, 1951.

Odd O Aalen and Søren Johansen. An empirical transition matrix for non-homogeneous markov chains based on censored observations. *Scandinavian Journal of Statistics*, pages 141–150, 1978.

Çağlar Çağlayan, Jordan S Goldstein, Turgay Ayer, Ashish Rai, and Christopher R Flowers. A population-based multistate model for diffuse large b-cell lymphoma–specific mortality in older patients. *Cancer*, 125(11):1837–1847, 2019b.

Guillermo Marshall and Richard H Jones. Multi-state models and diabetic retinopathy. *Statistics in medicine*, 14(18):1975–1983, 1995.

Charles Craddock, Richard M Szydlo, John P Klein, Francesco Dazzi, Eduardo Olavarria, Frits van Rhee, Christopher Pocock, Kate Cwynarski, Jane F Apperley, and John M Goldman. Estimating leukemia-free survival after allografting for

chronic myeloid leukemia: a new method that takes into account patients who relapse and are restored to complete remission. *Blood, The Journal of the American Society of Hematology*, 96(1):86–90, 2000.

GH De Bock, H Putter, J Bonnema, JA Van Der Hage, H Bartelink, and CJ Van De Velde. The impact of loco-regional recurrences on metastatic progression in early-stage breast cancer: a multistate model. *Breast cancer research and treatment*, 117(2):401–408, 2009.

Rune Hoff, Hein Putter, Ingrid Sivesind Mehlum, and Jon Michael Gran. Landmark estimation of transition probabilities in non-markov multi-state models with covariates. *Lifetime Data Analysis*, 25(4):660–680, 2019.

Ana Moreira, Jacobo de Una-Alvarez, Luis Machado, et al. Presmoothing the aalen-johansen estimator in the illness-death model. *Electronic Journal of Statistics*, 7: 1491–1516, 2013.

Mariano Provencio, Pilar Sabín, Jose Gomez-Codina, Maria Torrente, Virginia Calvo, Marta Llanos, Josep Gumá, Cristina Quero, Ana Blasco, Miguel Angel Cruz, et al. Impact of treatment in long-term survival patients with follicular lymphoma: A spanish lymphoma oncology group registry. *PloS one*, 12(5), 2017.

Wade T Swenson, James E Wooldridge, Charles F Lynch, Valerie L Forman-Hoffman, Elizabeth Chrischilles, and Brian K Link. Improved survival of follicular lymphoma patients in the united states. *Journal of Clinical Oncology*, 23(22):5019–5026, 2005.

Jonathan W Friedberg, Michelle Byrtek, Brian K Link, Christopher Flowers, Michael Taylor, John Hainsworth, James R Cerhan, Andrew D Zelenetz, Jamie Hirata, and Thomas P Miller. Effectiveness of first-line management strategies for stage i follicular lymphoma: analysis of the national lymphocare study. *Journal of clinical oncology*, 30(27):3368, 2012.

Anita R Martin, DD Weisenburger, WC Chan, EI Ruby, JR Anderson, Julie Marie Vose, Philip Jay Bierman, MA Bast, DT Daley, and James Olen Armitage. Prognostic value of cellular proliferation and histologic grade in follicular lymphoma. 1995.

David Faraggi and Richard Simon. A neural network model for survival data. *Statistics in medicine*, 14(1):73–82, 1995.

Knut Liestbl, Per Kragh Andersen, and Ulrich Andersen. Survival analysis and neural nets. *Statistics in medicine*, 13(12):1189–1200, 1994.

L Mariani, D Coradini, E Biganzoli, P Boracchi, E Marubini, S Pilotti, B Salvadori, R Silvestrini, U Veronesi, R Zucali, et al. Prognostic factors for metachronous

contralateral breast cancer: a comparison of the linear cox regression model and its artificial neural network extension. *Breast cancer research and treatment*, 44(2): 167–178, 1997.

Jared L Katzman, Uri Shaham, Alexander Cloninger, Jonathan Bates, Tingting Jiang, and Yuval Kluger. Deep survival: A deep cox proportional hazards network. *stat*, 1050:2, 2016.

Margaux Luck, Tristan Sylvain, Héloïse Cardinal, Andrea Lodi, and Yoshua Bengio. Deep learning for patient-specific kidney graft survival analysis. *arXiv preprint arXiv:1705.10245*, 2017.

Hemant Ishwaran, Udaya B Kogalur, Eugene H Blackstone, Michael S Lauer, et al. Random survival forests. *The annals of applied statistics*, 2(3):841–860, 2008.

Pannagadatta K Shivaswamy, Wei Chu, and Martin Jansche. A support vector approach to censored targets. In *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, pages 655–660. IEEE, 2007.

Anton Nemchenko, Trent Kyono, and Mihaela Van Der Schaar. Siamese survival analysis with competing risks. In *International Conference on Artificial Neural Networks*, pages 260–269. Springer, 2018.

Alexis Bellot and Mihaela Schaar. Tree-based bayesian mixture model for competing risks. In *International Conference on Artificial Intelligence and Statistics*, pages 910–918, 2018.

Changhee Lee, William R Zame, Jinsung Yoon, and Mihaela van der Schaar. Deephit: A deep learning approach to survival analysis with competing risks. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

Arturo Molina. A decade of rituximab: improving survival outcomes in non-hodgkin's lymphoma. *Annu. Rev. Med.*, 59:237–250, 2008.

Qi Liu, Luis Fayad, Fernando Cabanillas, Fredrick B Hagemeister, Gregory D Ayers, Mark Hess, Jorge Romaguera, M Alma Rodriguez, Apostolia M Tsimberidou, Srdan Verstovsek, et al. Improvement of overall and failure-free survival in stage iv follicular lymphoma: 25 years of treatment experience at the university of texas md anderson cancer center. *J Clin Oncol*, 24(10):1582–9, 2006.

Stefano Sacchi, Samantha Pozzi, Luigi Marcheselli, Alessia Bari, Stefano Luminari, Francesco Angrilli, Francesco Merli, Daniele Vallisa, Luca Baldini, and Maura Brugiatelli. Introduction of rituximab in front-line and salvage therapies has improved outcome of advanced-stage follicular lymphoma patients. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 109(10):2077–2082, 2007.

Mathias J Rummel, Norbert Niederle, Georg Maschmeyer, G Andre Banat, Ulrich von Grünhagen, Christoph Losem, Dorothea Kofahl-Krause, Gerhard Heil, Manfred Welslau, Christina Balser, et al. Bendamustine plus rituximab versus chop plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. *The Lancet*, 381(9873):1203–1210, 2013.

Richard I Fisher, Mark S Kaminski, Richard L Wahl, Susan J Knox, Andrew D Zelenetz, Julie M Vose, John P Leonard, Stewart Kroll, Stanley J Goldsmith, and Morton Coleman. Tositumomab and iodine-131 tositumomab produces durable complete remissions in a subset of heavily pretreated patients with low-grade and transformed non-hodgkin's lymphomas. *Journal of clinical oncology*, 23(30):7565–7573, 2005.

Richard J Klasa, Ralph M Meyer, Chaim Shustik, Carol A Sawka, Anne Smith, Raymond Guévin, Andrew Maksymiuk, Morel Rubinger, Martin Samosh, Suzanne Laplante, et al. Randomized phase iii study of fludarabine phosphate versus cyclophosphamide, vincristine, and prednisone in patients with recurrent low-grade non-hodgkin's lymphoma previously treated with an alkylating agent or alkylator-containing regimen. *Journal of clinical oncology*, 20(24):4649–4654, 2002.

Thomas J Pugh, Ari Ballonoff, Francis Newman, and Rachel Rabinovitch. Improved survival in patients with early stage low-grade follicular lymphoma treated with radiation: a surveillance, epidemiology, and end results database analysis. *Cancer*, 116(16):3843–3851, 2010.

Survival data analysis. URL `\https://www.karlin.mff.cuni.cz/~pesta/NMFM404/survival.html`.

Ross L Prentice, John D Kalbfleisch, Arthur V Peterson Jr, Nancy Flournoy, Vernon T Farewell, and Norman E Breslow. The analysis of failure times in the presence of competing risks. *Biometrics*, pages 541–554, 1978.

Peter C Austin, Douglas S Lee, and Jason P Fine. Introduction to the analysis of survival data in the presence of competing risks. *Circulation*, 133(6):601–609, 2016.

F Campigotto, D Neuberg, and JI Zwicker. Biased estimation of thrombosis rates in cancer studies using the method of kaplan and meier. *Journal of Thrombosis and Haemostasis*, 10(7):1449–1451, 2012.

David R Cox. Partial likelihood. *Biometrika*, 62(2):269–276, 1975.

David Schoenfeld. Partial residuals for the proportional hazards regression model. *Biometrika*, 69(1):239–241, 1982.

Caglar Caglayan, Jesse Dixon, Anna Wall, Gilles A Salles, Eva Hoster, Wolfgang Hiddemann, Michael Herold, Franck Morschhauser, Catherine Sebban, Anton Hagenbeek, et al. A multistate survival analysis for patients with follicular lymphoma (fl) using 13 first-line randomized trials from fl analysis of surrogate hypothesis (flash) group, 2019.

Nina D Wagner-Johnston, Brian K Link, Michelle Byrtek, Keith L Dawson, John Hainsworth, Christopher R Flowers, Jonathan W Friedberg, and Nancy L Bartlett. Outcomes of transformed follicular lymphoma in the modern era: a report from the national lymphocare study (nlcs). *Blood, The Journal of the American Society of Hematology*, 126(7):851–857, 2015.

AK Nooka, C Nabhan, X Zhou, MD Taylor, M Byrtek, Thomas P Miller, JW Friedberg, AD Zelenetz, BK Link, James R Cerhan, et al. Examination of the follicular lymphoma international prognostic index (flipi) in the national lymphocare study (nlcs): a prospective us patient cohort treated predominantly in community practices. *Annals of oncology*, 24(2):441–448, 2013.

Massimo Federico, Cesare Guglielmi, Stefano Luminari, Caterina Mammi, Luigi Marcheselli, Umberto Gianelli, Antonino Maiorana, Francesco Merli, Monica Bellei, Samantha Pozzi, et al. Prognostic relevance of serum $\beta 2$ microglobulin in patients with follicular lymphoma treated with anthracycline-containing regimens. a gisl study. *haematologica*, 92(11):1482–1488, 2007.

Emmanuel Bachy, John F Seymour, Pierre Feugier, Fritz Offner, Armando López-Guillermo, David Belada, Luc Xerri, John V Catalano, Pauline Brice, François Lemonnier, et al. Sustained progression-free survival benefit of rituximab maintenance in patients with follicular lymphoma: long-term results of the prima study. *Journal of Clinical Oncology*, 37(31):2815–+, 2019.

Gilles Salles, John Francis Seymour, Fritz Offner, Armando López-Guillermo, David Belada, Luc Xerri, Pierre Feugier, Réda Bouabdallah, John Vincent Catalano, Pauline Brice, et al. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (prima): a phase 3, randomised controlled trial. *The Lancet*, 377(9759):42–51, 2011.