

Southern Methodist University

SMU Scholar

---

Statistical Science Theses and Dissertations

Statistical Science

---

Spring 2023

## Statistical Methods for Evaluating Treatment Effects from Surgical Registry Data

Hang Nguyen

*Southern Methodist University*, [hangnguyen@smu.edu](mailto:hangnguyen@smu.edu)

Follow this and additional works at: [https://scholar.smu.edu/hum\\_sci\\_statisticalscience\\_etds](https://scholar.smu.edu/hum_sci_statisticalscience_etds)

---

### Recommended Citation

Nguyen, Hang, "Statistical Methods for Evaluating Treatment Effects from Surgical Registry Data" (2023). *Statistical Science Theses and Dissertations*. 33.

[https://scholar.smu.edu/hum\\_sci\\_statisticalscience\\_etds/33](https://scholar.smu.edu/hum_sci_statisticalscience_etds/33)

This Dissertation is brought to you for free and open access by the Statistical Science at SMU Scholar. It has been accepted for inclusion in Statistical Science Theses and Dissertations by an authorized administrator of SMU Scholar. For more information, please visit <http://digitalrepository.smu.edu>.

STATISTICAL METHODS FOR EVALUATING TREATMENT EFFECTS  
FROM SURGICAL REGISTRY DATA

Approved by:

---

Daniel F. Heitjan, PhD  
Professor of Statistical Science, SMU  
Professor of Biostatistics, O'Donnell Jr.  
School of Public Health, UTSW

---

Haekyung Jeon-Slaughter, PhD  
Assistant Professor of Department of  
Internal Medicine, UTSW  
Biostatistician, VA North Texas Health  
Center

---

S. Lynne Stokes, PhD  
Professor of Statistical Science, SMU

---

Xinlei Wang, PhD  
Professor of Statistics and Data Science,  
UTA

STATISTICAL METHODS FOR EVALUATING TREATMENT EFFECTS  
FROM SURGICAL REGISTRY DATA

A Dissertation Presented to the Graduate Faculty of the  
Dedman College  
Southern Methodist University  
in  
Partial Fulfillment of the Requirements  
for the degree of  
Doctor of Philosophy  
with a  
Major in Biostatistics  
by  
Hang Nguyen

B.S., Mathematical Biology, University of Houston

May 13, 2023

Copyright (2023)

Hang Nguyen

All Rights Reserved

## ACKNOWLEDGMENTS

I would like to express my gratitude and appreciation to my humble advisor, Dr. Daniel Heitjan for accepting me into the program. I still remember vividly the joy when I received the call from him for a short interview. He has been a role model and shown me the passion and there is just so much to learn in this field. I also want to thank my co-advisor, Dr. Haekyung Jeon-Slaughter for supporting my PhD research journey, giving me the great opportunity to work at VA North Texas, and not just mentoring me in statistics and research but also other aspects of life. I also would like to thank Dr. Yulun Liu for supporting my research assistantship for the last year of the program, and Dr. Xinlei Wang and Dr. S. Lynne Stokes for being part of my dissertation committee and provided feedbacks for improvements in my research. Without their support and help throughout these five years, I wouldn't be able to succeed and complete my graduate degree.

Lastly, I would like to thank my family for always supporting and encouraging me from the beginning to end of this long process. Many things have happened along the way, and I truly doubt that it would have been possible to reach the end goal without their help. In particular, I thank my older brother, Nien, who guided me to this graduate degree and for always giving me helpful career advices. I also thank my dear husband, Nolan, for always supporting and believing in me, and my dear furry friend, Ki. He may not know it but he has always been there, being silly and helping me relieve so much stress for these past years.

Statistical Methods for Evaluating Treatment Effects  
from Surgical Registry Data

Advisor: Daniel F. Heitjan, PhD

Co-advisor: Haekyung Jeon-Slaughter, PhD

Doctor of Philosophy degree conferred May 13, 2023

Dissertation completed April 11, 2023

Thoracic endovascular aortic repair (TEVAR) is a commonly used, minimally invasive approach for treatment of thoracic aortic aneurysms (TAA). As there have been no randomized clinical trials, we often extract information for this treatment from observational data, such as the Vascular Quality Initiative registry of TEVAR patients. However, evaluating the effectiveness of the treatment using observational data can be challenging when there is no data on the untreated group (e.g TEVAR registry), and no reliable information on the cause of death. To address these issues, we propose using the relative survival approach, in which one estimates the excess mortality hazard attributable to a disease by comparing disease registry survival data to general-population control data.

Popular modeling methods of estimating the relative survival have two shortcomings that render them unsuitable in some cases. First, they assume the excess hazard is positive, which is undoubtedly true for cancer, but may not hold for diseases treated with potentially curative therapy. Second, they consider survival to be continuous, whereas population control data is often available only in discrete form, rounded to the nearest year. To address these concerns, in Chapter 1, we propose describing discrete mortality hazards with a flexible logistic regression model that permits the registry hazard to be either larger or smaller

than the population hazard. We apply our approach to analyze relative survival of patients who underwent thoracic endovascular aortic repair (TEVAR) for thoracic aortic aneurysm (TAA). Our results show that relative survival is favorable for the youngest and oldest TEVAR recipients and unfavorable for those in between.

We hypothesized that the superior survival at older ages occurred because surgeons were recommending only the hardest older TAA patients for TEVAR. Thus, some older patients may have been excluded from a treatment that could have increased their survival time. It is not possible to evaluate this bias directly because the registry includes only those TAA patients who underwent TEVAR. In Chapter 2, we address this bias by proposing the use of sensitivity analysis to investigate the extent at which results are sensitive to potential biases in sampling. Our model has two components: First, a one-parameter selection model posits a pool of “potential patients” who could have received the treatment but did not, with selection probability depending on age. Second, a mortality hazard model for the on-surgery potential outcomes of the excluded patients extends the hazard model that we used to describe mortality in the registry. We identify combinations of parameters of the models that eliminate the relative survival advantage in the older patients. The analysis confirms, and places a magnitude on, a “healthy screening effect”, which posits that older TEVAR patients are screened for hardness more critically than younger patients.

The size of the thoracic aneurysms is an important criterion for deciding whether and when to conduct the TEVAR. Although previous analyses have sought to identify the ideal TAA size at which to apply TEVAR, none have properly accounted for potential confounding of the size of the aneurysm with treatment outcome. In Chapter 3, we aim to study the marginal effect of aneurysm size on post-procedure survival in causal inference framework. We estimate a marginal structural model (MSM) using inverse propensity score weighting, and model nonlinearity in the effect of aneurysm size with a flexible fractional polynomial form for the MSM. We find that patients who are asymptomatic at presentation and undergo

elective surgery have better survival outcomes if the operation takes place when the TAA size is near 60mm.

# TABLE OF CONTENTS

LIST OF FIGURES .....	xi
LIST OF TABLES .....	xiv
CHAPTER	
1 Modeling relative survival: An application to thoracic endovascular aortic repair	1
1.1. Introduction .....	1
1.2. The TEVAR Data.....	3
1.3. Methods .....	3
1.3.1. Notation .....	3
1.3.2. Logistic discrete hazard model.....	5
1.3.3. Modeling the hazard relative to the general population.....	6
1.3.4. Estimating the hazard model .....	7
1.3.5. Relative survival .....	8
1.3.5.1. Definitions.....	8
1.3.5.2. Estimation.....	9
1.3.5.3. Standard errors .....	10
1.4. Simulation .....	11
1.4.1. Simulation design .....	11
1.4.2. Simulation results .....	12
1.5. Analysis of the TEVAR data .....	12
1.5.1. The data .....	12
1.5.2. Model selection .....	13
1.5.3. Relative survival .....	14

1.6.	Sensitivity analysis .....	17
1.6.1.	Effect of excluding older subjects .....	17
1.6.2.	Effect of spline knots .....	20
1.6.3.	Population hazard adjustment .....	20
1.7.	Discussion .....	22
1.7.1.	Statistical methods .....	22
1.7.2.	Substantive findings .....	23
2	Sensitivity analysis of potential sampling bias in registry data .....	25
2.1.	Introduction .....	25
2.2.	Methods .....	26
2.2.1.	Notation .....	26
2.2.2.	Models .....	26
2.2.2.1.	Model for selection probability ( $\psi$ ) .....	28
2.2.2.2.	Model for mortality hazard in the non-treatment group ( $\delta$ ) .....	29
2.2.2.3.	Model settings for sensitivity analyses .....	31
2.2.2.4.	Estimating relative survival for TAA patients post-TEVAR .....	32
2.3.	Results .....	33
2.3.1.	Varying $\delta$ .....	33
2.3.2.	Varying $\psi$ .....	34
2.4.	Discussion .....	34
3	Identifying the Optimal Size for Repair of a Thoracic Aortic Aneurysm .....	43
3.1.	Introduction .....	43
3.2.	Patients and methods .....	45
3.2.1.	The TEVAR registry .....	45
3.2.2.	Marginal structural model .....	45

3.2.3.	A fractional polynomial model for post-TEVAR survival .....	48
3.2.4.	A restricted cubic spline model for post-TEVAR survival .....	49
3.2.5.	Estimation of optimal TAA size for TEVAR .....	49
3.2.5.1.	Generalized propensity scores .....	49
3.2.5.2.	MSM for aneurysm size .....	51
3.3.	Results .....	51
3.3.1.	Patient characteristics .....	51
3.3.2.	Application to the TEVAR data .....	52
3.4.	Discussion .....	55
APPENDIX		
A	APPENDIX of CHAPTER 1 .....	59
A.1.	Figures .....	59
A.2.	Population hazard adjustment .....	59
A.2.1.	Diabetes .....	59
A.2.2.	COPD .....	61
B	APPENDIX of CHAPTER 3 .....	63
B.1.	Tables .....	63
C	Supplementary R codes .....	66
C.1.	Chapter 1 .....	66
C.1.1.	Maximum likelihood estimators .....	66
C.1.2.	Relative survival .....	67
C.2.	Chapter 2 .....	69
C.3.	Chapter 3 .....	71
BIBLIOGRAPHY .....		72

## LIST OF FIGURES

Figure	Page
1.1	TEVAR data extraction. . . . . 4
1.2	Relative survival $\hat{S}_R^{\text{age}}(t z, x)$ (Equation (1.10)), for (A) all subjects with 95% CI (black dashed curves) and (B) white male (WM), white female (WF), black male (BM), and black female (BF). The grey dashed line corresponds to a relative survival of 1.0. . . . . 16
1.3	Non-parametric relative survival $\hat{S}_R^{\text{NP,age}}(t z, k)$ (Equation (1.12)) for TEVAR patients at age $t$ , for each race-sex stratum $k$ . . . . . 16
1.4	Relative survival $\hat{S}_R^{\text{years}}(t z, x)$ (Equation (1.11)) with 95% CI, stratified by race and sex. . . . . 18
1.5	Relative survival $\hat{S}_R^{\text{age}}(t z, x)$ (Equation (1.10)) for patients with age at surgery ( $z$ ) limited to being less than 70, 75, 80, and 90. . . . . 19
1.6	Relative survival $\hat{S}_R^{\text{age}}(t z, x)$ from Equation (1.10) with various combinations of spline knots in $z$ and $t - z$ . . . . . 21
1.7	Relative survival $\hat{S}_R^{\text{age}}(t z, x)$ using adjusted population hazard ( $h_0$ ) based on prevalences of diabetes (DM) and COPD using model M1 and model M6. . . 21
2.1	The directed acyclic graph on the left shows random patient selection to have surgery; the graph on the right illustrates dependence between patient selection and age ( $Z$ ), thus confounding survival. . . . . 27
2.2	Left: Relative survival $\hat{S}_R^{\text{age}}(t z)$ for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.6), case (a). Right: One year post-TEVAR relative hazard for non-TEVAR white males defined in (2.6). Different curves correspond to different values of $\delta$ . . . . . 35

2.3	Left: Relative survival $\hat{S}_R^{\text{age}}(t z)$ for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.7), case (b). Right: One year post-TEVAR relative hazard for non-TEVAR white males defined in (2.7). Different curves correspond to different values of $\delta$ at the first knot of $z$ . Smaller values of $\delta$ indicate higher risk for TEVAR patients aged 60 and older.....	35
2.4	Left: Relative survival $\hat{S}_R^{\text{age}}(t z)$ for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.8), case (c). Right: One year post-TEVAR relative hazard for non-TEVAR white males defined in (2.8). Different curves correspond to different values of $\delta$ at the second knot of $z$ . Smaller values of $\delta$ indicate higher risk for TEVAR patients aged 74 or older. ....	36
2.5	Left: Relative survival $\hat{S}_R^{\text{age}}(t z)$ for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.9), case (d). Right: One year post-TEVAR relative hazard for non-TEVAR white males defined in (2.9). Different curves correspond to different values of $\delta$ for the $t - z$ term. ....	36
2.6	Left: Relative survival $\hat{S}_R^{\text{age}}(t z)$ for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.6), case (a). We specify different values of $\psi$ and $\delta$ to test sensitivity around the overall slope in the relative hazard for the non-TEVAR group. Larger $\psi$ indicates a lower selection probability for older subjects to undergo TEVAR. ....	37
2.7	Left: Relative survival $\hat{S}_R^{\text{age}}(t z)$ for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.7), case (b). We specify different values of $\psi$ and $\delta$ to test sensitivity around the first knot in $z$ in the relative hazard for the non-TEVAR group. ....	38
2.8	Left: Relative survival $\hat{S}_R^{\text{age}}(t z)$ for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.8), case (c). We specify different values of $\psi$ and $\delta$ to test sensitivity around the second knot in $z$ in the relative hazard for the non-TEVAR group. ....	39
2.9	Left: Relative survival $\hat{S}_R^{\text{age}}(t z)$ for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.9), case (d). We specify different values of $\psi$ and $\delta$ to test sensitivity around the $t - z$ term in the relative hazard for the non-TEVAR group. ....	40

3.1	Distribution of stabilized weights for aneurysm size at operation in each sample, estimated from Equation (3.9). Most weights are near 1 in all samples. . . . .	54
3.2	Comparison between hazard models estimated by fractional polynomials (FP) and restricted cubic splines (RCS) in the marginal structural Cox model, with 95% bootstrap confidence intervals for the size giving minimum hazard. . . . .	58
1.1	Overall relative survival for TEVAR patients at age $t$ , given TEVAR was performed at age $z$ prior to age $t$ . Model M1 was used in the main analysis, adjusting for sex and race and their interaction. Model M2 to M6 adjusted for additional baseline clinical covariates: aneurysm size, diabetes, smoking status, COPD, and CAD (see Table 1.4). . . . .	59

## LIST OF TABLES

Table		Page
1.1	Simulation results for parameter estimates in the discrete survival model (1.14).	12
1.2	Simulation results for estimates of $S_R^{\text{years}}(t z, x)$ , $t-z \in \{1, 3\}$ , $z \in \{40, 45, \dots, 85\}$ .	13
1.3	Baseline characteristics of the TEVAR patients, stratified by race and sex.....	14
1.4	Coefficient estimates (standard errors) for parameters in different models .....	15
2.1	Example data and potential outcomes of TEVAR patients. ....	27
3.1	Baseline variables in the TEVAR registry data .....	53
3.2	Estimated fractional polynomials in the marginal structural Cox model.....	55
3.3	Optimal TAA size in mm (95% CI) for TEVAR surgery .....	56
1.1	A subset of crude and adjusted prevalence of total diabetes in the US, NHANES 2011-2016 [69] .....	60
1.2	A subset of age-specific hazard ratios associated with diabetes, for age 30 and above [70] .....	62
2.1	Evaluating the fit of estimated fractional polynomial functions in the marginal structural Cox model .....	63
2.2	Likelihood ratio tests ( $\chi^2$ , $P$ value) for different FP models using significance level $\alpha = 0.05$ .....	64
2.3	AIC values of restricted cubic spline models for post-TEVAR survival .....	65

I dedicate this dissertation to my family and my furry best friend, Ki for always being there for me in this journey.

## CHAPTER 1

### Modeling relative survival: An application to thoracic endovascular aortic repair

#### 1.1. Introduction

In population-based studies of chronic diseases, it is challenging to estimate cause-specific survival because of the unavailability of reliable information on causes of death. The relative survival approach [1] addresses this problem by comparing death rates in the affected population to those in the general population, assuming that the disease is rare in the latter group. The relative survival is the ratio of observed survival of the patients to the expected survival of the general population, often matched on demographic factors.

One can estimate relative survival nonparametrically with life tables [1–3], but such estimates may be unreliable or inefficient in strata where data are sparse. For this reason, authors commonly use modeling approaches that define the all-cause mortality as a sum of two components: The expected hazard in the general population and the differential hazard from the disease. The first such models described the excess hazard in the disease population as a proportion of the population hazard [4–7]. Others later extended these models to incorporate non-linear, non-proportional hazards by including interaction between time and covariates. Smoothing tools such as restricted cubic splines [8–10], B-splines [11], and fractional polynomials [12] are also in use. A further extension includes random effects to address clustering [13–15]. Models permit the examination of effects of prognostic factors on excess mortality, but also provide estimates that are smooth and efficient when data are sparse.

Models used in relative survival analysis typically specify that the hazard in the disease population exceeds that in the general population. This assumption is undoubtedly valid for cancer, which has been the principal domain of application for these methods. In diseases where curative treatments are available, however, one might entertain the possibility that patients can be restored to “good as new” or “better than new” condition, in which case one would require models that admit the possibility of a reduction of the death hazard. A prominent example is hip replacement, which confers a relative survival above 1.0 in some diagnosis groups [16]. Among the principal models for relative survival, only that of Lambert et al [12] covers this situation by incorporating additive effects on the hazard scale.

A common method for modeling relative survival is to divide follow-up time into intervals with a parameter for each, as in a piecewise exponential survival model [4, 17]. In this approach, each subject contributes one row of data for each interval during the time at risk. Taking the number of events per interval as the outcome gives rise to a Poisson likelihood that one optimizes *via* generalized linear model calculations [9]. The choice of intervals is vulnerable to criticism for being arbitrary and subjective. In models using non-split time [8, 11], authors treat the time to event as continuous.

Because population life tables typically present hazard by year, it is reasonable to evaluate relative survival using year as the time element. In this paper, we take the outcome variable to be annual survival, which enables us to describe annual death hazards as probabilities and model them on the logistic scale. This practice eliminates the arbitrary choice of intervals and yields a straightforward interpretation of relative survival by age. The logistic hazard model allows us to flexibly quantify the departure of survival from the population standard.

We illustrate our method in an analysis of post-treatment survival data from a registry of patients who have undergone endovascular repair of a thoracic aortic aneurysm (TAA). We use splines to describe the dependence of the hazard on the age at surgery and the time since surgery. In subsequent sections, we describe the data, notation, and model; we

present results of simulation studies and an application to the registry data; and we discuss advantages and limitations of our approach.

## 1.2. The TEVAR Data

We used Vascular Quality Initiative (VQI) registry data on patients who underwent thoracic endovascular aortic aneurysm repair (TEVAR) at 182 participating US institutions between 2010 and 2020. Variables include age at TEVAR, sex, race, medical history, clinical characteristics, length of stay for the procedure, and vital status at a series of follow-ups. Dates of death from the Social Security Death Index are available for subjects who have valid 9-digit Social Security numbers.

Our analysis included only patients who underwent elective endovascular repair for an asymptomatic TAA of size 30–190 mm. We excluded patients with nominal age 90 years, as all patients with true age 90 or above were coded as 90. Finally, we excluded a modest number of subjects who listed race as neither white nor black, or who had missing data on sex or race. The final sample included 1357 subjects from 146 centers; Figure 1.1 displays the data extraction scheme.

## 1.3. Methods

### 1.3.1. Notation

Suppose subject  $i$ ,  $i = 1, 2, \dots, n$ , with baseline covariate vector  $X_i$  underwent TEVAR at discrete age  $Z_i$ , was followed to discrete censoring age  $C_i$ , and died at discrete age  $T_i$ . The observed survival is  $\tilde{T}_i = \min(T_i, C_i)$  with event indicator  $D_i = I(T_i \leq C_i)$ .

We denote the probability mass function of  $T$ , conditionally on  $Z = z$  and  $X = x$ , as  $f(t|z, x) = \Pr(T = t|z, x)$ . Similarly, the conditional survival function is  $S(t|t > z, x) =$

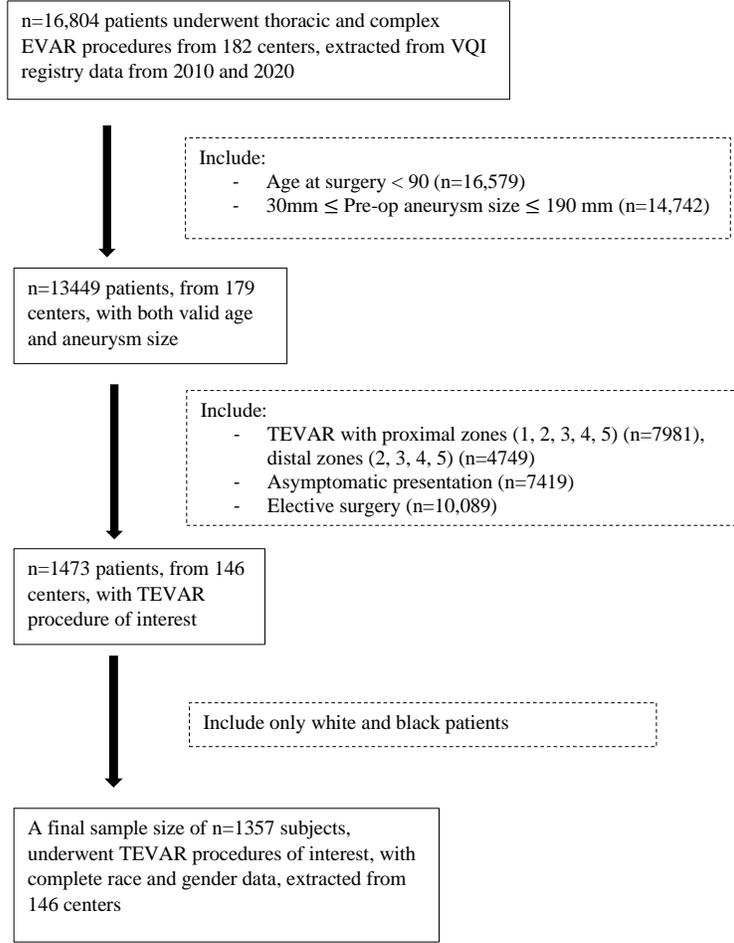


Figure 1.1: TEVAR data extraction.

$\Pr(T > t|z, x)$ . The conditional mortality hazard at age  $t$  is then

$$h(t|z, x) = \Pr(T = t|T \geq t, z, x) = \frac{f(t|z, x)}{S(t-1|z, x)},$$

$$t \in \{z, z+1, z+2, \dots\}, \quad (1.1)$$

the discrete conditional survival function is

$$S(t|t > z, X = x) = \prod_{u=z}^t [1 - h(u|x, z)], \quad (1.2)$$

and the conditional probability of death at age  $t$  after undergoing TEVAR at age  $z$  is

$$f(t|z, x) = h(t|z, x) \prod_{u=z}^{t-1} [1 - h(u|z, x)].$$

### 1.3.2. Logistic discrete hazard model

We model the post-TEVAR mortality hazard as a function of age at surgery ( $z$ ) and covariates ( $x$ ), including a potential stratification of the population based on a subset of demographic covariates. We denote  $k(x) \in \{1, \dots, K\}$  to be the stratum implied by the relevant elements of  $x$ . Because the discrete hazard in (1.1) is a probability, we describe it with the logistic model [18]

$$\ln \left[ \frac{h(t|z, x)}{1 - h(t|z, x)} \right] = \eta_{0k(x)}(t) + g(t, z, x; \theta),$$

$$t \in \{z, z + 1, z + 2, \dots\}. \quad (1.3)$$

We henceforth denote the baseline logit hazard at age  $t$  for individuals in stratum  $k$  to be  $\eta_{0k}(t)$ . Interpreting  $\eta_{0k}(t)$  as a known function derived from a reference population,  $g(t, z, x; \theta)$  represents the hazard odds ratio of a TEVAR patient who underwent the procedure at age  $z(\leq t)$  and has covariate vector  $x$ , relative to a patient in the same stratum of the general population. The parameter  $\theta$  governs the model.

We model survival by assuming the mortality hazard depends on the age at surgery  $z$  and time since surgery  $t - z$  as linear splines, potentially with interaction. The general model

is of the form

$$\begin{aligned}
g(t, z, x; \theta) &= \theta_0 \\
&+ \theta_{10}z + \sum_{j=1}^P \theta_{1j}(z - p_j)_+ \\
&+ \theta_{20}(t - z) + \sum_{l=1}^Q \theta_{2l}(t - z - q_l)_+ \\
&+ r(t, z; \theta_3) \\
&+ \theta_4 x,
\end{aligned} \tag{1.4}$$

where  $\theta_0$  denotes an intercept,  $\theta_1 = (\theta_{10}, \theta_{11}, \dots, \theta_{1P})$  denotes the coefficients for a spline in  $z$ ,  $\theta_2 = (\theta_{20}, \theta_{21}, \dots, \theta_{2Q})$  denotes the coefficients for a spline in  $t - z$ ,  $\theta_3$  is the coefficient vector for interactions between spline terms, and  $\theta_4$  is the coefficient vector for the covariates in  $x$ . The  $z$  and  $t - z$  splines have  $P$  and  $Q$  pre-selected knots  $p_1, \dots, p_P$  and  $q_1, \dots, q_Q$ , respectively, and  $(u)_+ = \max(0, u)$ . Thus,  $\theta = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ .

### 1.3.3. Modeling the hazard relative to the general population

Let  $h_{0k}(t)$  be the population hazard at age  $t$  of individuals in population stratum  $k$ , obtained here from 2017 US life tables [19]. Assuming the proportion of subjects who undergo TEVAR in the general population is negligible, we define the logit of the population hazard as an offset in the hazard model for the TEVAR patients. We express the population hazard at age  $t$  as

$$\ln \left( \frac{h_{0k}(t)}{1 - h_{0k}(t)} \right) = \eta_{0k}(t). \tag{1.5}$$

Substituting (1.4) and (1.5) into (1.3), and adding a subscript  $\theta$  to represent the parameter, the hazard becomes

$$h_{\theta}(t|z, x) = \text{expit} [\eta_{0k}(t) + g(t, z, x; \theta)],$$

$$t \in \{z, z + 1, z + 2, \dots\}, \quad (1.6)$$

where  $\text{expit}(u) = e^u / (1 + e^u)$ .

#### 1.3.4. Estimating the hazard model

Suppose  $n$  patients who underwent TEVAR have observed data  $(\tilde{t}, z, d, x)$ , where  $\tilde{t}$  is the vector of survival times,  $z$  is the vector of ages at surgery,  $d$  is the vector of death indicators, and  $x$  is the matrix of covariates. The log-likelihood is then

$$l(\theta|\tilde{t}, d, z, x) = \sum_{i=1}^n \left( d_i \ln[h_{\theta}(\tilde{t}_i|z_i, x_i)] \right. \\ \left. + \sum_{u=z_i}^{\tilde{t}_i-1} d_i \ln[1 - h_{\theta}(u|z_i, x_i)] \right. \\ \left. + \sum_{v=z_i}^{\tilde{t}_i} (1 - d_i) \ln[1 - h_{\theta}(v|z_i, x_i)] \right). \quad (1.7)$$

We use generic optimization methods to compute maximum likelihood estimates, and we select a best model by the Akaike Information Criterion (AIC). We then estimate all-cause mortality hazard by (1.6) and observed survival for TEVAR patients by (1.2). R code is available from the first author.

### 1.3.5. Relative survival

#### 1.3.5.1. Definitions

We define here two versions of *relative survival* [1] that convey complementary information about survival experience.

Let  $S(t|z, x)$  be the survival probability at age  $t$  for a patient with baseline covariate vector  $x$  who underwent TEVAR at age  $z$ . Similarly, let  $S_P(t|z, k(x))$  be the survival in the reference population, estimated by Ederer II [2] using the population survival of subjects beginning at time  $z$  and matching on stratum identity  $k(x)$ . We define the overall relative survival at age  $t$  of a set of subjects with age at surgery vector  $z$  and covariate matrix  $x$  to be the ratio of the sum of patient survival and the sum of survival of corresponding subjects in the population:

$$S_R^{\text{age}}(t|z, x) = \frac{\sum_{i=1}^n I(z_i \leq t \leq \tilde{t})S(t|z_i, x_i)}{\sum_{i=1}^n I(z_i \leq t \leq \tilde{t})S_P(t|z_i, k(x_i))}. \quad (1.8)$$

That is, the numerator is the sum of the estimated survival probabilities at age  $t$  for all subjects who underwent the procedure prior to age  $t$ , and were followed up until age  $\tilde{t}$ . The denominator is the analogous quantity for a corresponding set of subjects in the general population.

Alternatively, we can define relative survival by ages at procedure and death for a given covariate vector  $x$  and scalar surgery time  $z$ . The relative survival of a TEVAR patient at age  $t$  given the operation occurred at age  $z$  ( $< t$ ) is

$$S_R^{\text{years}}(t|z, x) = \frac{S(t|z, x)}{S_P(t|z, k(x))}. \quad (1.9)$$

For example,  $S_R^{\text{age}}(45|z, x)$  represents the relative survival at age 45 for *a set of TEVAR patients* with covariates  $x$  and surgery times  $z$ , whereas  $S_R^{\text{years}}(45|40, x)$  represents the relative survival at age 45 for *a single patient* with covariate vector  $x$  who underwent TEVAR at age 40.

### 1.3.5.2. Estimation

We can estimate  $S(t|z, x)$  using either our logistic discrete survival model (1.6) or a Kaplan-Meier curve stratified by  $z$  and  $k(x)$  [20]. The former approach avoids problems of sparsity but may introduce bias if model assumptions are incorrect. The latter is robust but may be unstable when data are sparse.

Let  $\hat{S}(t|z, x)$  be the model-estimated survival probability at age  $t$  for a patient who had TEVAR at age  $z$  with baseline covariate vector  $x$ . Assuming that  $S_P$  is known without error, and designating estimates with a circumflex, we define the model-based estimators of relative survival as

$$\hat{S}_R^{\text{age}}(t|z, x) = \frac{\sum_{i=1}^n I(z_i \leq t \leq \tilde{t}) \hat{S}(t|z_i, x_i)}{\sum_{i=1}^n I(z_i \leq t \leq \tilde{t}) S_P(t|z_i, k(x_i))} \quad (1.10)$$

and

$$\hat{S}_R^{\text{years}}(t|z, x) = \frac{\hat{S}(t|z, x)}{S_P(t|z, k(x))}. \quad (1.11)$$

To define the nonparametric estimate, let  $\hat{S}^{\text{KM}}(t|z, k)$  be the product-limit estimate of survival at age  $t$  for  $n_k$  individuals in stratum  $k$ , indexed  $i = 1, \dots, n_k$ , who underwent surgery at ages  $z_1, \dots, z_{n_k}$ . The nonparametric estimate of relative survival at age  $t$  for stratum  $k$  is

$$\hat{S}_R^{\text{NP,age}}(t|z, k) = \frac{\sum_{i=1}^{n_k} I(z_i \leq t \leq \tilde{t}) \hat{S}^{\text{KM}}(t|z_i, k)}{\sum_{i=1}^{n_k} I(z_i \leq t \leq \tilde{t}) S_P(t|z_i, k)}, \quad (1.12)$$

and the nonparametric estimate of relative survival at age  $t$  for a subject in stratum  $k$  who had the procedure at age  $z$  ( $< t$ ) is

$$\hat{S}_R^{\text{NP,years}}(t|z, k) = \frac{\hat{S}^{\text{KM}}(t|z, k)}{S_P(t|z, k)}. \quad (1.13)$$

One can also estimate the nonparametric relative survival for all patients, denoted as  $\hat{S}_R^{\text{NP,age}}(t|z)$  and  $\hat{S}_R^{\text{NP,years}}(t|z)$ , similarly to Equations (1.12) and (1.13) by collapsing all strata, thus obviating the  $k$  notation in survival estimators in the numerators. The expected survivals in the denominators are still estimated by best matching the general population to the patients under study through the covariates implied by  $k(x)$ .

#### 1.3.5.3. Standard errors

Because the model-based relative survival is a complicated function of the parameters and the values of  $x$  and  $z$ , we compute 95% confidence intervals for it by the parametric bootstrap. Denote  $\hat{\theta}$  to be the maximum likelihood estimate of the parameter vector  $\theta$  in the hazard function (1.6). Let  $W$  follow a multivariate normal distribution,  $W \sim N(\hat{\theta}, \Sigma)$ , with mean  $\hat{\theta}$  and covariance matrix  $\Sigma$  defined as the estimated covariance matrix of  $\hat{\theta}$ . First we draw 500 random samples of  $\theta$  from the distribution of  $W$ . For each sampled  $\theta$ , we calculate the relative survival (either (1.10) or (1.11)) for TEVAR patients with a designated  $x$ . We then compute the standard errors of the relative survival by taking the standard deviation of the 500 simulated relative survivals. We obtain a confidence interval by taking the estimated relative survival plus/minus the standard error times the relevant normal quantile.

We can compute standard errors for  $\hat{S}_R^{\text{NP,age}}(t|z, k)$  by the nonparametric bootstrap, and for  $\hat{S}_R^{\text{NP,years}}(t|z, k)$  by the delta method.

## 1.4. Simulation

### 1.4.1. Simulation design

We evaluated the performance of our method in a Monte Carlo experiment. We simulated data of patients from the TEVAR white male subgroup (the largest stratum) by assuming the true hazard function depends only on age at surgery with two spline knots at 60 and 74:

$$\begin{aligned} \ln \left[ \frac{h(t|z)}{1 - h(t|z)} \right] &= \ln \left[ \frac{h_0(t)}{1 - h_0(t)} \right] + 2.1 + 0.06z \\ &\quad - 0.1(z - 60)_+ - 0.05(z - 74)_+, \\ t &\in \{z, z + 1, z + 2, \dots\}, \end{aligned} \tag{1.14}$$

where  $h_0(t)$  is the population hazard at age  $t$  for white males. We simulated 1,000 replications at sample size  $n$  as follows:

1. Sample  $n$  data values for age at surgery ( $z$ ) with replacement from the TEVAR white males.
2. Simulate survival time and vital status under hazard (1.14), assuming censoring time  $c = z + 5$ .
3. Obtain the MLE and 95% confidence interval for each model parameter.
4. Estimate 1-year and 3-year relative survival given patients have TEVAR at age  $z \in \{40, 45, \dots, 85\}$ . Compute standard errors for relative survival as described above.

For parameters of interest, we estimated the bias, root mean squared error (rMSE), and coverage probability of nominal 95% confidence intervals.

### 1.4.2. Simulation results

Table 1.1 summarizes the simulated MLEs of the model coefficients: The intercept plus a spline in age at surgery ( $z$ ) with knots at 60 and 74, as in Equation (1.14). The MLEs have coverage probability close to the nominal level even with small  $n$ . Bias is appreciable for the intercept at  $n = 500$ , but not enough to impair coverage probability.

Table 1.1: Simulation results for parameter estimates in the discrete survival model (1.14).

Model Parameter	True value	$n = 500$			$n = 1000$		
		CP (%)	Bias	rMSE	CP (%)	Bias	rMSE
Intercept ( $\theta_0$ )	2.1	95.3	0.1620	2.15	94.7	0.0111	1.41
Surgery age $z$ ( $\theta_1$ )	0.06	95.0	-0.0023	0.04	94.7	-0.0001	0.03
Knot $z = 60$ ( $\theta_{11}$ )	-0.1	94.1	0.0013	0.06	94.5	0.0002	0.04
Knot $z = 74$ ( $\theta_{12}$ )	-0.05	94.9	0.0039	0.05	95.5	0.0017	0.04

CP: coverage probability; rMSE: root mean squared error

Table 1.2 shows results for simulated 1-year and 3-year relative survival. For the smaller sample size, the coverage probability is close to its nominal level for 1-year relative survival, but fluctuates slightly for 3-year relative survival. Increasing the sample size to  $n = 1,000$  stabilizes the coverage probability and further reduces bias and rMSE at most values of  $z$ .

## 1.5. Analysis of the TEVAR data

### 1.5.1. The data

The data set we extracted from the TEVAR registry on December 2, 2020 contains 1,357 patients (Figure 1.1). Table 1.3 shows patient baseline characteristics, stratified by race and sex. A plurality of the sample was white male. The mean age of surgery (SD) for all subjects was 72.1 (9.6), with the black groups slightly younger (67.3 for black male and 71.2 for black female). Black patients had TEVAR at a larger mean aortic diameter than whites. The highest proportions of diabetes and current smoking were observed in black men, while COPD was most common in white females, and CAD in males.

Table 1.2: Simulation results for estimates of  $S_R^{\text{years}}(t|z, x)$ ,  $t - z \in \{1, 3\}$ ,  $z \in \{40, 45, \dots, 85\}$ .

Years after TEVAR ( $t - z$ )	Surgery age ( $z$ )	True survival	$n = 500$			$n = 1000$		
			CP (%)	Bias	rMSE	CP (%)	Bias	rMSE
1 year	40	0.82	93.0	-0.0270	0.1078	93.0	-0.0091	0.0643
	45	0.72	92.2	-0.0193	0.0962	93.3	-0.0054	0.0609
	50	0.56	93.7	-0.0118	0.0732	94.4	-0.0023	0.0482
	55	0.37	93.2	-0.0064	0.0472	93.9	-0.0010	0.0323
	60	0.23	93.2	-0.0006	0.0450	95.0	0.0004	0.0320
	65	0.21	94.0	-0.0014	0.0266	94.7	-0.0006	0.0186
	70	0.18	94.3	-0.0005	0.0223	95.5	-0.0007	0.0149
	75	0.16	93.4	0.0005	0.0251	95.4	-0.0006	0.0167
	80	0.15	94.8	-0.0017	0.0210	94.2	-0.0018	0.0152
	85	0.14	95.1	-0.0010	0.0331	93.4	-0.0015	0.0249
3 years	40	0.53	88.5	-0.0252	0.1645	91.2	-0.0067	0.1153
	45	0.33	90.1	-0.0089	0.1159	92.5	0.0004	0.0819
	50	0.14	92.2	-0.0019	0.0504	93.7	0.0014	0.0355
	55	0.04	92.8	0.0001	0.0153	93.6	0.0007	0.0110
	60	0.01	93.6	0.0010	0.0066	93.9	0.0006	0.0047
	65	0.009	92.6	0.0002	0.0031	93.8	0.0001	0.0022
	70	0.006	93.3	0.0003	0.0023	96.1	0.00007	0.0015
	75	0.004	92.4	0.0002	0.0017	94.6	0.00004	0.0011
	80	0.003	93.6	0.0001	0.0016	92.6	-0.000001	0.0011
	85	0.003	96.0	0.0005	0.0025	93.3	0.0002	0.0018

CP: coverage probability; rMSE: root mean squared error

### 1.5.2. Model selection

We estimated models with different combinations of knots in the two spline functions in  $z$  (at  $z = 60, 74, 83$ , corresponding to the 10, 50, and 90 age centiles, respectively) and  $t - z$  (at  $t - z = 1, 2$ ), and their interactions. We also ran models that adjusted for pre-op TAA size, diabetes, smoking status, COPD, and CAD (Table 1.4). Likelihood ratio tests found no significant spline-spline interactions. Model comparisons based on AIC identified a best model that had knots in age of surgery at  $z = 60, 74$ , a slope but no spline terms in  $t - z$ , and no additional baseline covariates. The model included main effects of race and sex and their interaction, but we did not stratify the modeling by race and sex; that is, we assume

Table 1.3: Baseline characteristics of the TEVAR patients, stratified by race and sex.

	Total	White		Black	
		Male	Female	Male	Female
$n$ (% of total)	1,357 (100%)	636 (46.9%)	495 (36.5%)	112 (8.2%)	114 (8.4%)
Age at surgery	72.1 (9.6)	72.5 (9.1)	72.8 (9.7)	67.3 (10.2)	71.2 (10.6)
Pre-op TAA size (mm)	58.3 (11.0)	58.6 (11.2)	57.4 (10.2)	59.3 (12.5)	59.1 (11.6)
Diabetes (Yes)	228 (16.8)	113 (17.8)	64 (12.9)	28 (25.0)	23 (20.2)
Current smoking (Yes)	357 (26.3)	156 (24.5)	129 (26.1)	42 (37.5)	30 (26.3)
COPD (Yes)	523 (38.5)	246 (38.7)	201 (40.6)	39 (34.8)	37 (32.5)
CAD (Yes)	267 (19.7)	150 (23.6)	72 (14.5)	26 (23.2)	19 (16.7)

Mean (SD) for continuous variables;  $n$  (%) for categorical variables. TAA: thoracic aortic aneurysm; COPD: chronic obstructive pulmonary disease; CAD: coronary artery disease.

that the other parameters are the same across race-sex strata. We henceforth refer to this as model M1.

### 1.5.3. Relative survival

We used model M1 to estimate the relative survival for TEVAR patients by age, stratified by race and sex, defined above as  $S_R^{\text{age}}(t|z, x)$ . We observed that TEVAR patients under age 60 have a slightly higher survival than the general population (Figure 1.2A), whereas patients aged 60–80 experience slightly lower survival than the general population. Subjects aged 80 or above show markedly superior survival, with relative survival rising to almost 1.1 at age 90. The results change only slightly after adjustment for other baseline covariates (Figure 1.1). We observed similar trends within race and sex subgroups (Figure 1.2B), with relative survival climbing rapidly for patients over age 80. Relative survivals are similar across all groups for age under 70, and deviate at older ages. In particular, white males show higher relative survival than white females. Erratic changes in black males and females could be due to small samples in these subgroups. Nonparametric results echo model-based results (Figure 1.3), although the curves are less smooth.

We also estimated 1- to 5-year relative survival by age at surgery, defined as  $S_R^{\text{years}}(t|z, x)$ , stratified by race and sex (Figure 1.4). Similar to Figure 1.2, the graphs show that TEVAR

Table 1.4: Coefficient estimates (standard errors) for parameters in different models

Variable	Parameter	M1	M2	M3	M4	M5	M6
Intercept	$\theta_0$	-5.24 (6.76)	-5.23 (6.34)	-5.05 (6.36)	-4.28 (6.02)	-3.14 (5.63)	-3.89 (6.00)
Age at surgery ( $z$ )	$\theta_1$	0.11 (0.11)	0.10 (0.11)	0.09 (0.11)	0.08 (0.10)	0.06 (0.10)	0.07 (0.10)
Knot $z = 60$	$\theta_{11}$	-0.14 (0.12)	-0.12 (0.12)	-0.11 (0.12)	-0.10 (0.11)	-0.08 (0.11)	-0.09 (0.11)
Knot $z = 74$	$\theta_{12}$	-0.10 (0.04)	-0.10 (0.04)	-0.11 (0.04)	-0.11 (0.04)	-0.10 (0.04)	-0.09 (0.04)
Time since surgery ( $t - z$ )	$\theta_2$	-0.48 (0.07)	-0.48 (0.07)	-0.48 (0.07)	-0.48 (0.07)	-0.47 (0.07)	-0.47 (0.07)
Gender (Male)	$\theta_3$	-0.27 (0.39)	-0.29 (0.39)	-0.33 (0.39)	-0.37 (0.40)	-0.38 (0.40)	-0.40 (0.40)
Race (White)	$\theta_4$	0.13 (0.31)	0.17 (0.31)	0.19 (0.31)	0.19 (0.31)	0.16 (0.31)	0.17 (0.31)
Gender*Race	$\theta_5$	0.15 (0.43)	0.16 (0.43)	0.18 (0.43)	0.22 (0.43)	0.24 (0.43)	0.23 (0.43)
Pre-op TAA size (mm)	$\theta_6$	—	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)
Diabetes (Yes)	$\theta_7$	—	—	0.36 (0.19)	0.36 (0.19)	0.33 (0.19)	0.32 (0.19)
Current smoking (Yes)	$\theta_8$	—	—	—	0.26 (0.18)	0.20 (0.18)	0.20 (0.18)
COPD (Yes)	$\theta_9$	—	—	—	—	0.35(0.16)	0.33(0.16)
CAD (Yes)	$\theta_{10}$	—	—	—	—	—	0.30(0.18)
AIC	—	1389.35	1382.91	1381.47	1381.47	1378.86	1378.03

TAA: thoracic aortic aneurysm; COPD: chronic obstructive pulmonary disease; CAD: coronary artery disease; AIC: Akaike Information Criterion.

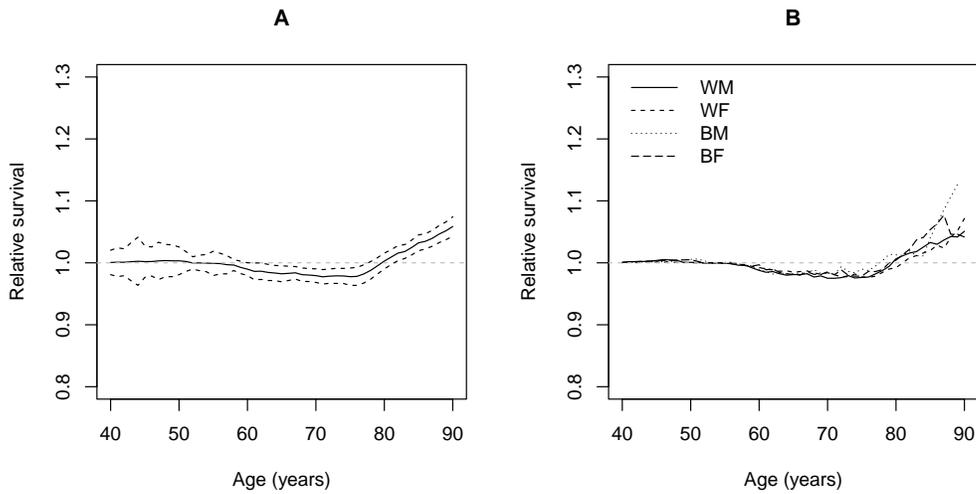


Figure 1.2: Relative survival  $\hat{S}_R^{\text{age}}(t|z, x)$  (Equation (1.10)), for (A) all subjects with 95% CI (black dashed curves) and (B) white male (WM), white female (WF), black male (BM), and black female (BF). The grey dashed line corresponds to a relative survival of 1.0.

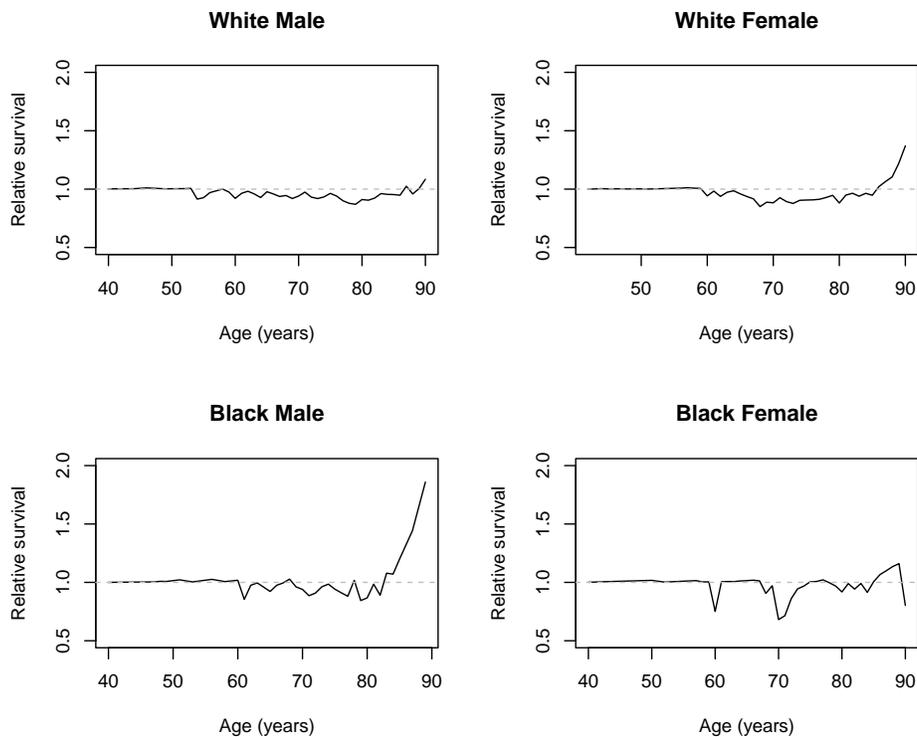


Figure 1.3: Non-parametric relative survival  $\hat{S}_R^{\text{NP,age}}(t|z, k)$  (Equation (1.12)) for TEVAR patients at age  $t$ , for each race-sex stratum  $k$ .

patients typically have survival comparable to the general population between surgery ages 40 and 50, before experiencing a slight drop and lower survival from 50 to 75 and a sharp increase thereafter. The difference in survival increases for larger  $t-z$ , with the 5-year relative survival curves revealing the most pronounced concavity. The effect is slightly stronger in males and blacks.

## 1.6. Sensitivity analysis

### 1.6.1. Effect of excluding older subjects

In Equation (1.10), we defined the set of patients for computing relative survival at age  $t$  to include all patients who had surgery prior to age  $t$  and were followed until age  $\tilde{t} \geq t$ . To study the influence of the older subjects on relative survival, we re-ran the analysis excluding patients with age at surgery greater than 80, 75, and 70, resulting in sample sizes  $n = 1061$ , 740, and 469, respectively. In each subgroup, we estimated the mortality hazard using a model similar to that used in the main analysis, with the logistic discrete hazard depending on age at surgery ( $z$ ) at two knots (10th and 50th centiles) and time since surgery ( $t-z$ ), and adjusting for race and sex. The range for age at surgery varies due to different sample sizes of subgroups, thus giving different knot locations in  $z$ . For each group, we estimated the relative survival  $S_R^{\text{age}}(t|z, x)$  (Equation (1.10)). Because the analyses exclude older subjects, the relative survival curves do not all extend to age 90.

Results appear in Figure (1.5). We observe that by removing patients who had TEVAR at older ages, the remaining part of the relative survival curve falls below 1.0. Evidently, unexpectedly good outcomes among the oldest TEVAR recipients drive the elevated relative survival at the oldest ages.

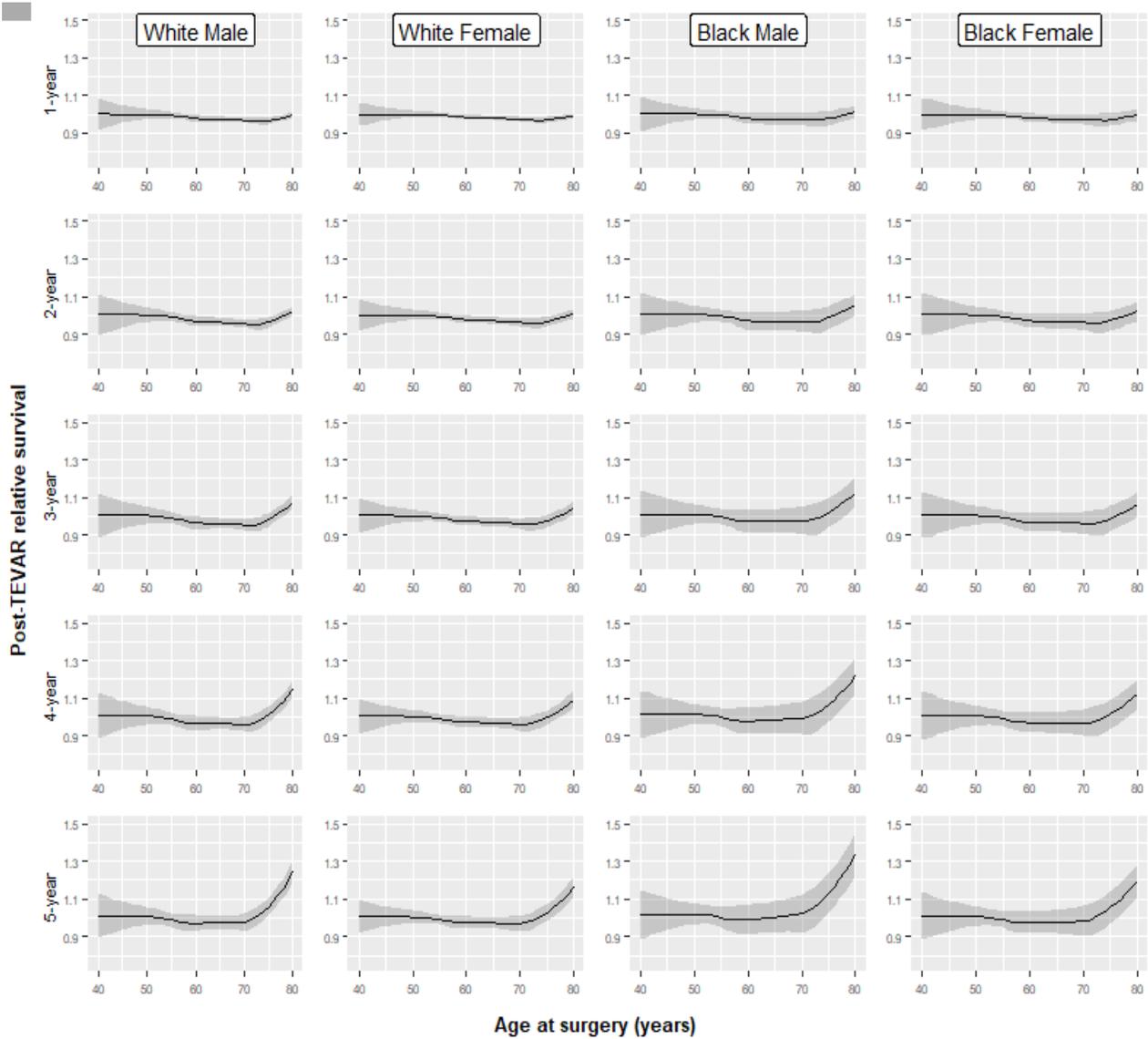


Figure 1.4: Relative survival  $\hat{S}_R^{\text{years}}(t|z, x)$  (Equation (1.11)) with 95% CI, stratified by race and sex.

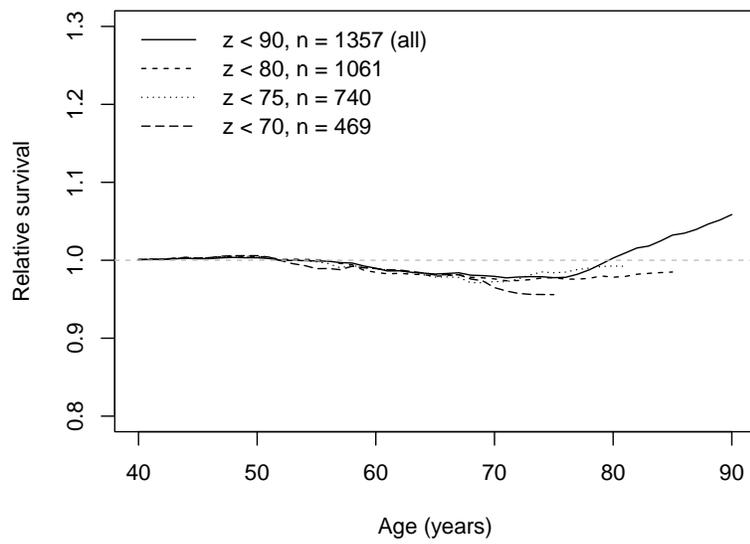


Figure 1.5: Relative survival  $\hat{S}_R^{\text{age}}(t|z, x)$  (Equation (1.10)) for patients with age at surgery ( $z$ ) limited to being less than 70, 75, 80, and 90.

### 1.6.2. Effect of spline knots

We examined the robustness of the estimated relative survival to the choice of knots in the hazard model. Recall that the final spline, selected by AIC, has knots at  $z = 60, 74$  and a slope but no spline term in  $t - z$ . In the first setting, we evaluated models with different numbers of knots selected from  $z \in \{55, 60, 65, 74, 80, 83\}$ , with no knot in  $t - z$ . In the second setting, we added knots in  $t - z$  at one and two years, testing various combinations of spline terms in both  $z$  and  $t - z$ . We then evaluated sensitivity of  $\hat{S}_R^{\text{age}}(t|z, x)$  (Equation (1.10)) for both settings.

Figure (1.6) shows that the estimated relative survival at age  $t$  is similar for all choices of knots. This suggests that our finding of increased relative survival at advanced ages is robust to the specifics of the spline model.

### 1.6.3. Population hazard adjustment

The age-specific population hazard from US national life tables represents a sum of hazards from a mixed population with different underlying conditions — for example, people who have diabetes or COPD. Knowing the prevalence and relative risk associated with a condition, one can adjust the population hazard to reflect the hazard for subjects who do or do not manifest the risk factor. Details appear in the Appendix.

We calculated the overall relative survival at age  $t$ ,  $\hat{S}_R^{\text{age}}(t|z, x)$ , using unadjusted and adjusted population hazards reflecting effects of diabetes and COPD. Adjusting for diabetes or COPD produces only minimal changes in relative survival (Figure 1.7). We also compared the results from model M1 (including only demographics) to model M6 (also including pre-op TAA size, smoking, and comorbidities; see Table 1.4) and found no difference.

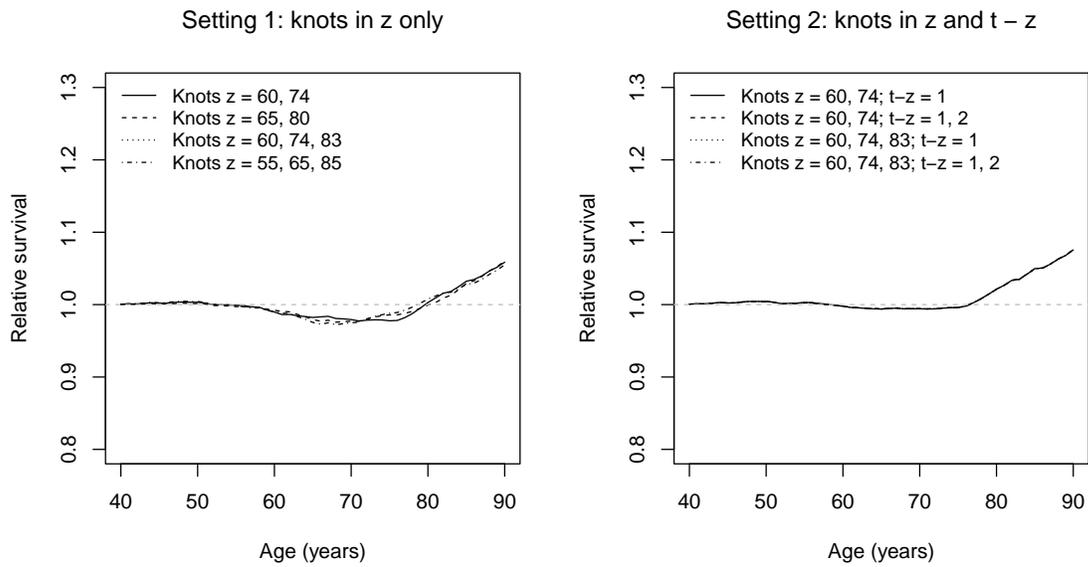


Figure 1.6: Relative survival  $\hat{S}_R^{\text{age}}(t|z, x)$  from Equation (1.10) with various combinations of spline knots in  $z$  and  $t - z$ .

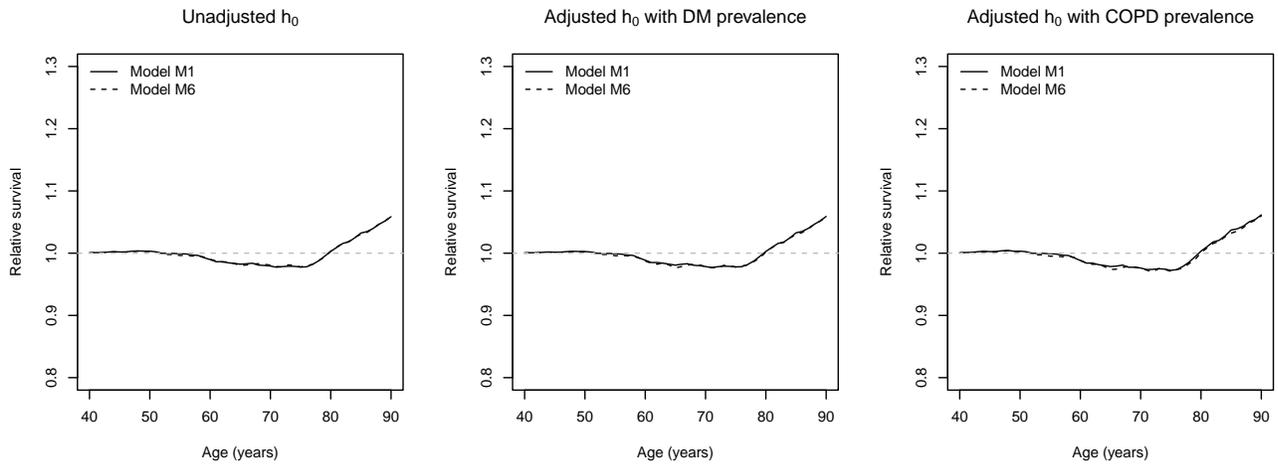


Figure 1.7: Relative survival  $\hat{S}_R^{\text{age}}(t|z, x)$  using adjusted population hazard ( $h_0$ ) based on prevalences of diabetes (DM) and COPD using model M1 and model M6.

## 1.7. Discussion

We have applied the relative survival framework to data from a registry of patients who underwent TEVAR for treatment of TAA. Our goal is to determine whether the endovascular repair of TAA renders patients effectively cured, in the sense that their survival is equivalent to that of demographically matched controls in the general population.

### 1.7.1. Statistical methods

Our hazard model relies on splines, [9, 11, 21–24] which offer considerable flexibility at some risk of over-fitting. Using model selection strategies, we obtained a satisfactory fit with simple models that yield smooth survival curves. Simulations confirmed that our model performs well even with a modest-sized sample, giving negligible bias with coverage probability close to nominal levels. Sensitivity analyses demonstrated robustness to the choice of knots.

An advantage of our approach is that it can model situations where the mortality hazard in the population of interest is reduced over the general population. Although such a feature is not needed for many chronic diseases, it can be valuable for modeling survival after curative therapy in conditions that are effectively reversible.

In order to incorporate life-table information effectively, it is reasonable to model the post-TEVAR mortality hazard as a discrete variate, which leads naturally to the use of a logistic model for annual death hazards. Whereas the model represents the effect of age at TEVAR and time since TEVAR on mortality hazard, our ultimate graphical analysis represents relative survival at each age, averaging across the distribution of ages at surgery in the registry prior to the age of interest. This approach leads to estimates that reflect the distribution of procedure times in the registry [25–27].

Relative survival analysis requires the proportion of diseased patients in the general population to be negligible. Studies on cancer have shown that the bias from violation of this assumption is modest [28], but no such studies exist in TAA. Relative survival also relies on external data from the general population. A weakness of such data is that they are seldom stratified by more than sex and race, whereas registries typically contain other relevant predictors of mortality [29, 30]. We attempted to adjust the population hazard with multiple covariates, but the robustness of our results after accounting for diabetes — a common comorbidity in TAA — suggest that this may be unnecessary.

### 1.7.2. Substantive findings

We observed that younger TEVAR patients experience superior survival compared to the general population, with those aged 60–80 often performing slightly worse and those over 80 substantially better. Our sensitivity analysis showed that this curious finding disappears when one excludes older subjects from the calculation. Past studies that estimated survival after TAA repair using Cox models reported higher risk in males and older patients [31, 32]. A retrospective case-control study with repaired TAA found that survival is lower than controls in the younger groups, but not different from the controls in octogenarians [31]. Studies in other diseases have also found superior relative survival in the elderly; examples include abdominal aortic aneurysm [33] and prostate cancer [34].

Previous authors have suggested that elderly patients should be considered candidates for TEVAR [35, 36]. In practice, physicians remain skeptical; a survey of 50 vascular surgeons in Europe and the US showed that they were uncertain about the benefits of TEVAR in patients aged 80–85 [37]. Thus, apparently some older patients who might benefit from the procedure are not receiving it.

A possible explanation for the upturn in relative survival among TEVAR recipients is that they receive additional attention from physicians, who are able to discover and treat other

conditions before they become fatal; this effect may be most pronounced among the oldest patients. Moreover, because we have included only elective TEVARs, it is possible that the oldest group is exceptionally fit — other than having a TAA — relative to comparably aged seniors in the general population [31,38]. This “healthy screening effect” was also discussed in the estimation of relative survival of prostate cancer patients [34].

A third possibility is that there is a bias in the collection of the death dates in our data, which are known only for subjects who have a complete, 9-digit Social Security number (SSN) and, of course, are dead. As our dataset does not contain an indicator of completeness of the SSN record, we do not know how large this bias is or in what direction it would tend. Further sensitivity analyses will help elucidate these effects.

## CHAPTER 2

### Sensitivity analysis of potential sampling bias in registry data

#### 2.1. Introduction

Registries provide valuable data for follow-up studies in clinical epidemiology. An advantage of registry studies is that they are not subject to selection bias in the description of outcomes of those who receive the treatment [39]. Limitations, however, can include incompleteness of covariate data, absence of extensive pre-treatment data, and the lack of data on untreated controls. Most importantly, because decisions about who will receive the treatment are not randomized, extrapolation of findings to the larger population of subjects who *could have* received the treatment, some of whom did not, is subject to bias.

In a previous article [40], we used procedure registry data to show that older thoracic aortic aneurysm (TAA) patients who underwent thoracic endovascular aortic repair (TEVAR) had superior survival relative to age-, sex-, and race-matched members of the general US population. This motivates us to investigate whether this is a real effect or a result of a bias in the selection of patients to treat. With the limitations in registry data, we generally cannot determine whether “hidden bias” exists because we cannot or did not measure the confounding variables. Certain additional data are required to fully analyze the biases, but this is not possible with limitations in observational studies. Methods to adjust for controllable confounding — such as stratification, adjustment, or inverse probability weighting — are not applicable in this context.

To address these issues, we need to resort to sensitivity analysis, as others have proposed in various epidemiologic contexts [41–45]. Specifically, we describe our data using a causal

model within the potential outcomes framework. We then conduct analyses to determine the potential impact of sampling bias on post-treatment relative survival from a registry of TEVAR recipients. We describe sensitivity parameters that govern biases in sampling based on age and fitness. A specific goal is to identify the degree of bias necessary to abrogate the putative advantage in relative survival that accrues to patients who are older at diagnosis. We conduct multiple such analyses and discuss advantages and limitations of our approach.

## 2.2. Methods

### 2.2.1. Notation

Suppose TAA subject  $i$ ,  $i = 1, 2, \dots, n$  who has baseline covariate vector  $X_i$  is diagnosed at age  $Z_i$  and undergoes TEVAR ( $A_i = 1$ ). In addition, we assume there is an unknown number  $m$  of “potential patients” who were diagnosed with TAA but did not undergo TEVAR ( $A_i = 0$ ). Denote  $T_i(0)$  to be the age at death of subject  $i$  should he not receive TEVAR, and  $T_i(1)$  to be his age at death should he receive TEVAR. Thus  $T_i = (T_i(0), T_i(1))$  is the vector of potential outcomes for subject  $i$ , only one of which we can observe.

Because our data are from a registry that includes only TAA subjects who undergo TEVAR, we do not know the number of subjects who were not offered TEVAR or declined it, let alone their covariates or survival times. Table 2.1 is a schematic of the data structure.

### 2.2.2. Models

We henceforth assume that the discrete on-treatment survival time  $T(1)$  is affected by the age at diagnosis ( $Z$ ) and the baseline covariate ( $X$ ). We further assume that physicians recommend a patient for treatment based on the patient’s age at diagnosis, and that the operation follows diagnosis quickly. We assume independence between  $Z$  and  $X$ . When the selection mechanism is nonignorable (i.e., the measured covariates do not include all

Table 2.1: Example data and potential outcomes of TEVAR patients.

Unit ( $i$ )	$A_i$	$T_i(0)$	$T_i(1)$	$Z_i$	$X_i$
1	1	?	2226	82	$x_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n$	1	?	302	59	$x_n$
$n + 1$	0	?	?	?	?
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n + m$	0	?	?	?	?

confounders), we must address selection bias. We illustrate these relationships in the directed acyclic graphs in Figure 2.1.

Our sensitivity analysis seeks to examine the effect of applying various models for the selection mechanism and effect of age on TEVAR survival on the distribution of survival post-surgery given age, could all subjects be induced to undergo TEVAR. We begin by making the following definitions (omitting the subscript for subjects):

- $p_\psi(z) = \Pr[A = 1|Z = z]$ , the fraction of subjects diagnosed at age  $z$  who receive TEVAR.
- $f_1(t|z, a = 1) = \Pr[T(1) = t|Z = z, A = 1]$  is the probability of death at age  $t$  for a TEVAR patient.

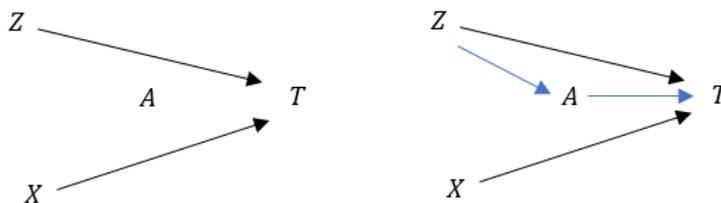


Figure 2.1: The directed acyclic graph on the left shows random patient selection to have surgery; the graph on the right illustrates dependence between patient selection and age ( $Z$ ), thus confounding survival.

- $f_1(t|z, a = 0) = \Pr[T(1) = t|Z = z, A = 0]$  is the probability of death on TEVAR at age  $t$  for a non-TEVAR patient.

From these, we derive the distribution of the age at death post-TEVAR given age at diagnosis, including both those who received TEVAR and those who did not:

$$f(t|z) = f_1(t|z, a = 1)p_\psi(z) + f_1(t|z, a = 0)(1 - p_\psi(z)), \quad t \in \{z, z + 1, z + 2, \dots\}. \quad (2.1)$$

That is, this is the distribution of age at death given age at diagnosis if all subjects undergo TEVAR, not just those who ultimately received it under current practice. We are able to directly estimate only  $f_1(t|z, a = 1)$  from the available TEVAR data. The sensitivity analysis consists of positing forms for  $p_\psi(z)$  and  $f_1(t|z, a = 0)$  and assessing their effects on the distribution  $f(t|z)$  of age at death given  $z$  could all subjects be induced to undergo TEVAR.

### 2.2.2.1. Model for selection probability ( $\psi$ )

We hypothesize that the observed superior relative survival of older TEVAR patients reflects bias in the selection of patients to receive the surgery — essentially, the bar for suitability for TEVAR is higher in older TAA patients. To examine how large such an effect would have to be to abrogate the observed result, we add hypothetical patients according to a selection probability model governed by a sensitivity parameter  $\psi$ . We represent the probability that a subject diagnosed at age  $z$  receives TEVAR as  $p_\psi(z)$ , which we assume has the following form:

$$\Pr[A = 1|Z = z] = p_\psi(z) = \begin{cases} 1, & z \leq 40 \\ \psi^{-(z-40)}, & z > 40, \end{cases} \quad (2.2)$$

with  $\psi \geq 1$ . This model implies that patients diagnosed at ages up to 40 all receive TEVAR, whereas older patients are increasingly unlikely to receive TEVAR. In the boundary case  $\psi = 1$ , all subjects receive TEVAR, and the registry equals the entire TAA population. When  $\psi > 1$ , the registry systematically excludes older subjects.

*2.2.2.2. Model for mortality hazard in the non-treatment group ( $\delta$ )*

Denote the discrete hazard function at age  $t$  for a TEVAR patient to be

$$h_1(t|z, a = 1) = \Pr[T(1) = t | T(1) \geq t, Z = z, A = 1], \quad t \in \{z, z + 1, z + 2, \dots\}. \quad (2.3)$$

We can rewrite the probability of death at age  $t$  for a TEVAR patient, as a function of discrete death hazards:

$$f_1(t|z, a = 1) = h_1(t|z, a = 1) \prod_{u=z}^{t-1} [1 - h_1(u|z, a = 1)], \quad t \in \{z, z + 1, z + 2, \dots\}$$

where we define the hazard function of TEVAR patients depending on age as [40]

$$h_1(t|z, a = 1) = \text{expit}[\eta_{0k}(t) + g(t, z; \theta)], \quad t \in \{z, z + 1, z + 2, \dots\} \quad (2.4)$$

and

$$\begin{aligned} g(t, z; \theta) &= \theta_0 \\ &+ \theta_{10}z + \theta_{11}(z - p_1)_+ + \theta_{12}(z - p_2)_+ \\ &+ \theta_{20}(t - z). \end{aligned}$$

Similarly, we denote the discrete hazard function at age  $t$  for the on-TEVAR potential outcome in a non-TEVAR patient to be

$$h_1(t|z, a = 0) = \Pr[T(1) = t | T(1) \geq t, Z = z, A = 0]$$

and rewrite his probability of death at age  $t$  as

$$f_1(t|z, a = 0) = h_1(t|z, a = 0) \prod_{u=z}^{t-1} [1 - h_1(u|z, a = 0)], \quad t \in \{z, z + 1, z + 2, \dots\}.$$

Extending from the hazard model (2.4) for TEVAR patients, we assume the non-TEVAR patients follow a mortality hazard model governed by an additional unknown parameter,  $\delta$ :

$$h_1(t|z, a = 0) = \text{expit} [\eta_{0k}(t) + l(t, z; \theta, \delta)], \quad t \in \{z, z + 1, z + 2, \dots\}, \quad (2.5)$$

with the relative hazard  $l(t, z; \theta, \delta)$  following taking four different forms as follows:

- Case (a) — overall effect on the linear predictor:

$$l(t, z; \theta, \delta) = \theta_0 + \delta[\theta_{10}z + \theta_{11}(z - p_1)_+ + \theta_{12}(z - p_2)_+ + \theta_{20}(t - z)]. \quad (2.6)$$

- Case (b) — effect on the first knot in  $z$ :

$$l(t, z; \theta, \delta) = \theta_0 + \theta_{10}z + \delta\theta_{11}(z - p_1)_+ + \theta_{12}(z - p_2)_+ + \theta_{20}(t - z). \quad (2.7)$$

- Case (c) — effect on the second knot in  $z$ :

$$l(t, z; \theta, \delta) = \theta_0 + \theta_{10}z + \theta_{11}(z - p_1)_+ + \delta\theta_{12}(z - p_2)_+ + \theta_{20}(t - z). \quad (2.8)$$

- Case (d) — effect on  $t - z$ :

$$l(t, z; \theta, \delta) = \theta_0 + \theta_{10}z + \theta_{11}(z - p_1)_+ + \theta_{12}(z - p_2)_+ + \delta\theta_{20}(t - z). \quad (2.9)$$

In every case, when  $\delta = 1$  we have  $g(\cdot) = l(\cdot)$ , implying that the patients who did not receive TEVAR have on-TEVAR mortality risk equal to patients who did receive TEVAR. In case (a), we assess sensitivity of the results to the overall slope governed by all terms: age at surgery ( $z$ ) and its knots, and time since surgery ( $t - z$ ) (Equation 2.6). In cases (b–d), we target individual terms in the relative hazard linear predictor. First, in case (b), we place  $\delta$  next to the coefficient of the first knot of  $z$ , which is at the 10th centile (Equation 2.7). This will tell us how the results change specifically for the people who are older than the 10th centile age. Similarly, in case (c), we evaluate the sensitivity around the second knot, which we have placed at the median age at surgery (Equation 2.8). In case (d), we evaluate the effect of time since surgery to the overall result (Equation 2.9). We hypothesize that the shorter the survival time since surgery, the lower the relative survival compared to the general population. Different values of  $\delta$  in cases (a) and (d) might change the curve overall, and part of the curves after the first and second knot of  $z$  in case (b) and (c). In the next section, we propose a series of settings with different combinations of sensitivity parameters  $(\delta, \psi)$  to visualize their effects on relative survival results.

### 2.2.2.3. Model settings for sensitivity analyses

When  $\delta = 1$  and  $\psi = 1$ , we obtain the null model from our previous paper. [40] The effect of  $\delta$  in each case on the mortality hazard of the on-TEVAR potential outcome differs depending on the sign of the MLEs obtained in the previous analysis. Here we focus on the subgroup of white males, in which the MLE for  $(\theta_0, \theta_{10}, \theta_{11}, \theta_{12}, \theta_{20})$  is  $(-5.24, 0.11, -0.14, -0.10, -0.48)$ . We seek to identify sensitivity parameters that are suf-

ficient to abrogate the superior relative survival seen at older ages, thereby explaining the observed advantage for older patients as a selection bias.

We consider five settings:

- Setting 1: Fix  $\psi = 1.1$  and set  $\delta = 1.02, 1.05, 1.1$  in case (a).
- Setting 2: Fix  $\psi = 1.1$  and set  $\delta = 0.9, 0.8, 0.7, 0.6$  in case (b).
- Setting 3: Fix  $\psi = 1.1$  and set  $\delta = 0.1, 0.01, -0.1, -0.3$  in case (c).
- Setting 4: Fix  $\psi = 1.1$  and set  $\delta = 0.8, 0.5, 0.1$  in case (d).
- Setting 5: Combinations of  $\delta = 0.8, 0.5, 0.1, 0.01$  and  $\psi = 1.01, 1.05, 1.1$  in four cases (a–d).

Because both  $\hat{\theta}_{11}$  and  $\hat{\theta}_{12}$  are negative, reducing  $\delta$  in cases (b) and (c) will increase the death hazard for the untreated patients. In Settings 2 and 3, we propose decreasing values of  $\delta$  and a fixed value for  $\psi$ . This is slightly trickier in cases (a) and (d), because we set  $\delta$  such that the changes in both  $t$  and  $z$  can affect the overall results. Thus, in Settings 1 and 4 we vary values and direction of  $\delta$ . In setting 5, we examine how the changes in both selection probability and mortality hazard for the potential patients affect the overall results.

#### 2.2.2.4. Estimating relative survival for TAA patients post-TEVAR

From equation (2.1), we can similarly derive the hazard function at age  $t$  post-TEVAR given age at diagnosis ( $z$ ), including patients who received TEVAR and who did not:

$$\begin{aligned} h(t|z) &= \Pr[T(1) = t | T(1) \geq t, Z = z] \\ &= h_1(t|z, a = 1)p_\psi(z) + h_1(t|z, a = 0)(1 - p_\psi(z)), \quad t \in \{z, z + 1, z + 2, \dots\} \end{aligned}$$

where  $h_1(t|z, a = 1)$  and  $h_1(t|z, a = 0)$  are defined in (2.4) and (2.5).

With the hazard function defined above, let the new discrete conditional survival function for post-TEVAR patients be

$$\bar{S}(t|t > z) = \Pr(T(1) > t|Z = z) = \prod_{u=z}^t [1 - h(u|z)], \quad t \in \{z, z + 1, z + 2, \dots\}. \quad (2.10)$$

Suppose  $\tilde{t}$  is the vector of observed survival times for  $n$  patients who underwent TEVAR and  $z$  is their vector of observed ages at diagnosis. Using the method from our previous paper [40], we estimate the relative survival at age  $t$  for post-TEVAR patients, given age at diagnosis for a particular subgroup (e.g. white male), as

$$\hat{S}_R^{\text{age}}(t|z) = \frac{\sum_{i=1}^n I[z_i \leq t \leq \tilde{t}] \hat{S}(t|z_i)}{\sum_{i=1}^n I[z_i \leq t \leq \tilde{t}] S_P(t|z_i)}, \quad (2.11)$$

where  $\hat{S}(t|z)$  is the model-estimated survival probability at age  $t$  defined above in (2.10) and  $S_P(t|z)$  is the survival of the corresponding patients (white males with matching ages) from the general population, estimated by Ederer II [2].

## 2.3. Results

### 2.3.1. Varying $\delta$

We varied  $\delta$  to identify potential effects of the addition of patients who do not receive the treatment on the relative survival. Increasing or decreasing  $\delta$  has a different effect in each case, depending on the signs of the MLE regression coefficients. If the MLE is negative, increasing  $\delta$  lowers the relative hazard for the control group. In case (a), by increasing  $\delta$  we assume that the potential patients have higher mortality risk than those who received TEVAR. Figure (2.2) shows that with higher values of  $\delta$ , we obtained lower relative survival for patients at age 55 or above. In case (b), lowering  $\delta$  affects the MLE at the first knot in  $z$ , resulting in lower post-TEVAR relative survival for patients aged 65 and older (Figure 2.3). We observe a similar pattern in case (d), where we examined the sensitivity around term

$t - z$ , such that the overall relative survival curve is also lower than the one in null model as  $\delta$  decreases (Figure 2.5). In case (c), where we checked sensitivity of results to the second knot in  $z$ , lowering  $\delta$  implies that the potential TEVAR patients aged 74 and older have higher hazard, because the MLE for  $\theta_{12}$  is negative. Thus, if these individuals were given the treatment, the overall relative survival would be lower. This effect is shown in Figure (2.4) by the relative survival decline at the tail as  $\delta$  decreases.

### 2.3.2. Varying $\psi$

Figure (2.6)–(2.9) present sensitivity analysis results in case (a)–(d) respectively, with varying combinations of  $\psi$  and  $\delta$ . Increasing  $\psi$  results in higher relative survival in case (a) only, and lower relative survival in cases (b–d). The results are similar across different values of  $\delta$  when increasing  $\psi$  in case (a), but show larger changes in other cases.

Higher  $\psi$  values indicate lower selection probability for older patients in the observed treatment group. The change is largest in case (b) with small values of  $\delta = 0.1, 0.01$ . The potential patient group consists of a larger portion of older patients as  $\psi$  increases, and these patients also have higher hazard as  $\delta$  at the first knot in  $z$  is small. By removing the older patients in the observed treatment group, the reduction in relative survival at age 60 implies that the old individuals in our data are healthy and fit. We observe similar patterns in cases (c) and (d), but with moderate changes between different values of  $\psi$ . The relative survival at older ages declines as we remove more older patients in the observed treatment group. The results in case (a) are reversed in that higher  $\psi$  values lead to higher relative survival.

## 2.4. Discussion

We have investigated the effect of the suspected exclusion from TEVAR surgery of elderly TAA patients on the survival of TEVAR patients relative to matched population controls. First, we included a proportion of patients who could have been treated, but are excluded,

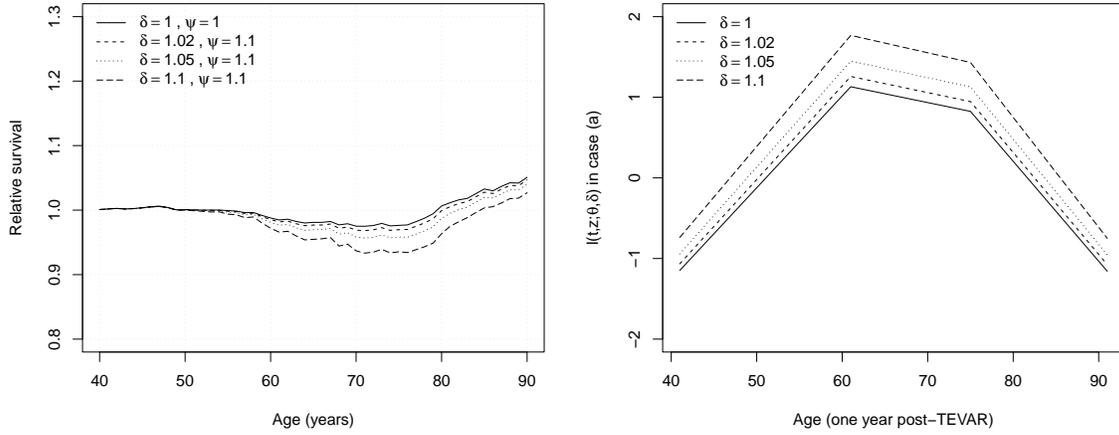


Figure 2.2: Left: Relative survival  $\hat{S}_R^{\text{age}}(t|z)$  for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.6), case (a). Right: One year post-TEVAR relative hazard for non-TEVAR white males defined in (2.6). Different curves correspond to different values of  $\delta$ .

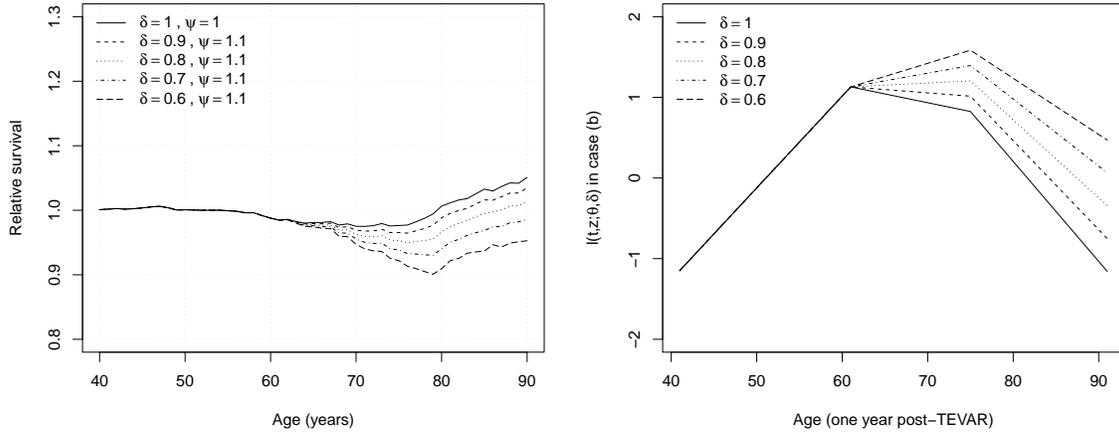


Figure 2.3: Left: Relative survival  $\hat{S}_R^{\text{age}}(t|z)$  for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.7), case (b). Right: One year post-TEVAR relative hazard for non-TEVAR white males defined in (2.7). Different curves correspond to different values of  $\delta$  at the first knot of  $z$ . Smaller values of  $\delta$  indicate higher risk for TEVAR patients aged 60 and older.

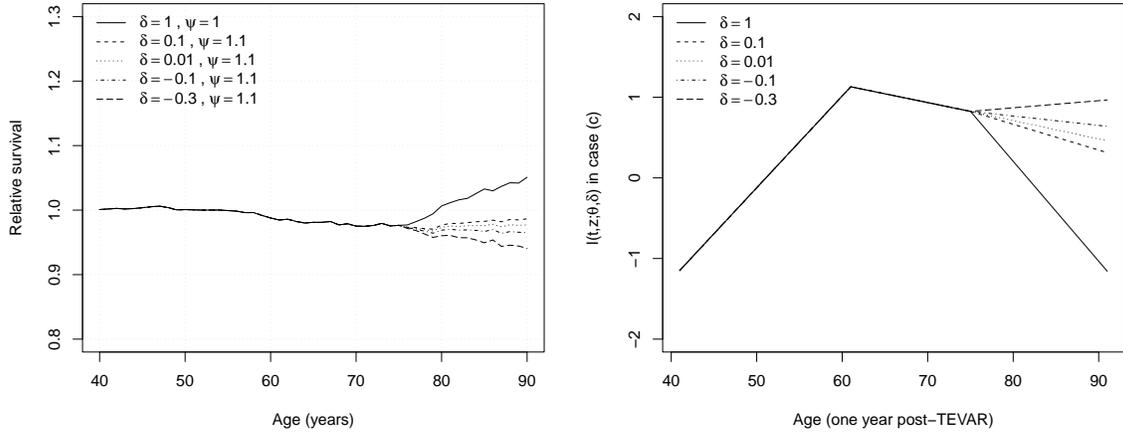


Figure 2.4: Left: Relative survival  $\hat{S}_R^{\text{age}}(t|z)$  for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.8), case (c). Right: One year post-TEVAR relative hazard for non-TEVAR white males defined in (2.8). Different curves correspond to different values of  $\delta$  at the second knot of  $z$ . Smaller values of  $\delta$  indicate higher risk for TEVAR patients aged 74 or older.

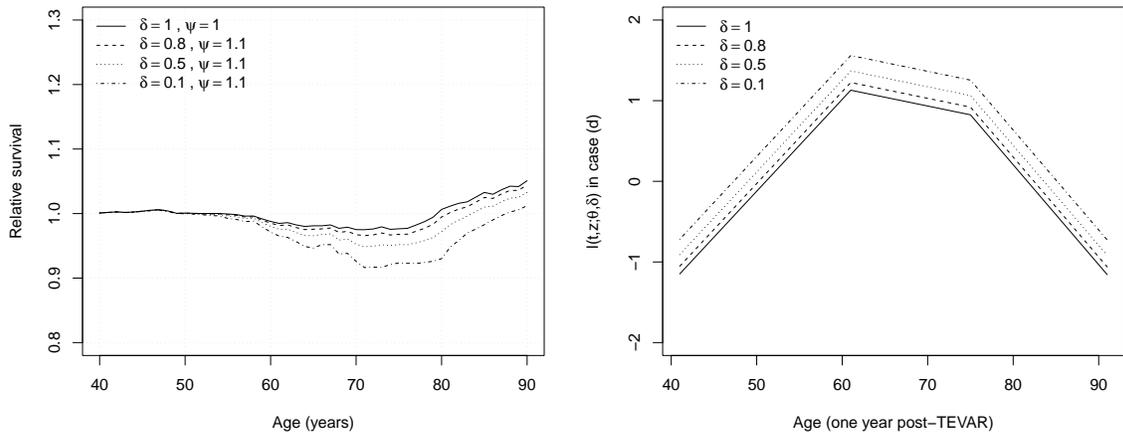


Figure 2.5: Left: Relative survival  $\hat{S}_R^{\text{age}}(t|z)$  for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.9), case (d). Right: One year post-TEVAR relative hazard for non-TEVAR white males defined in (2.9). Different curves correspond to different values of  $\delta$  for the  $t - z$  term.

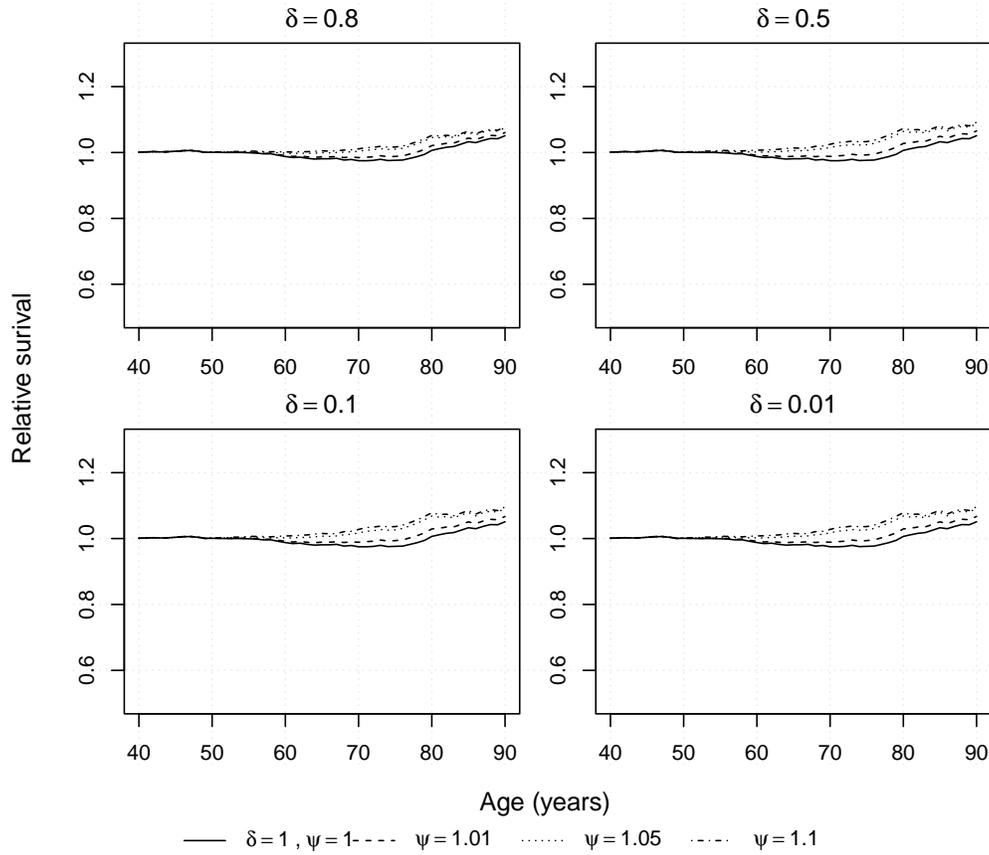


Figure 2.6: Left: Relative survival  $\hat{S}_R^{\text{age}}(t|z)$  for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.6), case (a). We specify different values of  $\psi$  and  $\delta$  to test sensitivity around the overall slope in the relative hazard for the non-TEVAR group. Larger  $\psi$  indicates a lower selection probability for older subjects to undergo TEVAR.

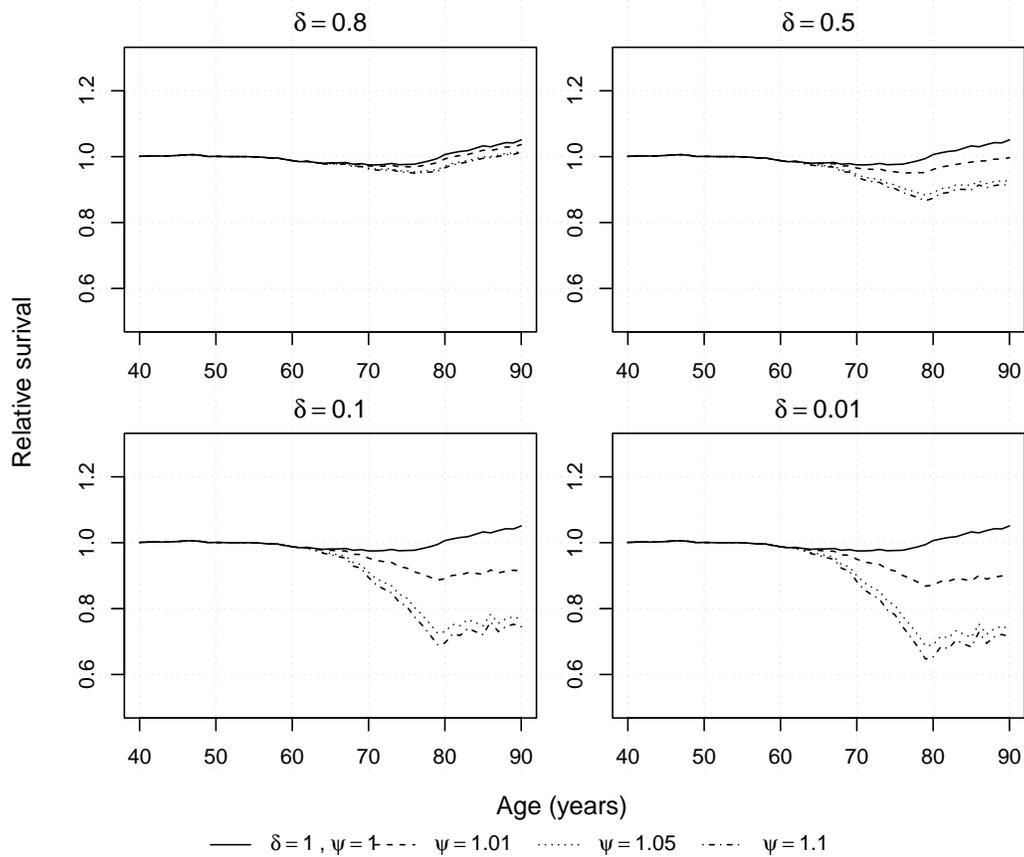


Figure 2.7: Left: Relative survival  $\hat{S}_R^{\text{age}}(t|z)$  for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.7), case (b). We specify different values of  $\psi$  and  $\delta$  to test sensitivity around the first knot in  $z$  in the relative hazard for the non-TEVAR group.

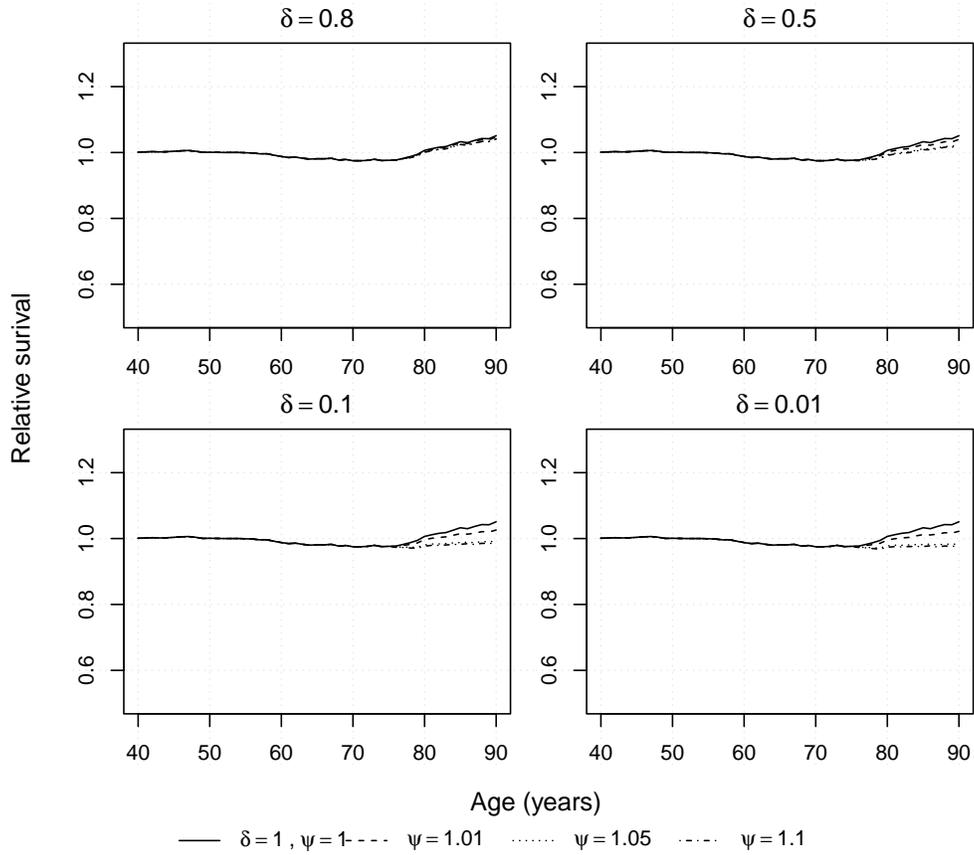


Figure 2.8: Left: Relative survival  $\hat{S}_R^{\text{age}}(t|z)$  for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.8), case (c). We specify different values of  $\psi$  and  $\delta$  to test sensitivity around the second knot in  $z$  in the relative hazard for the non-TEVAR group.

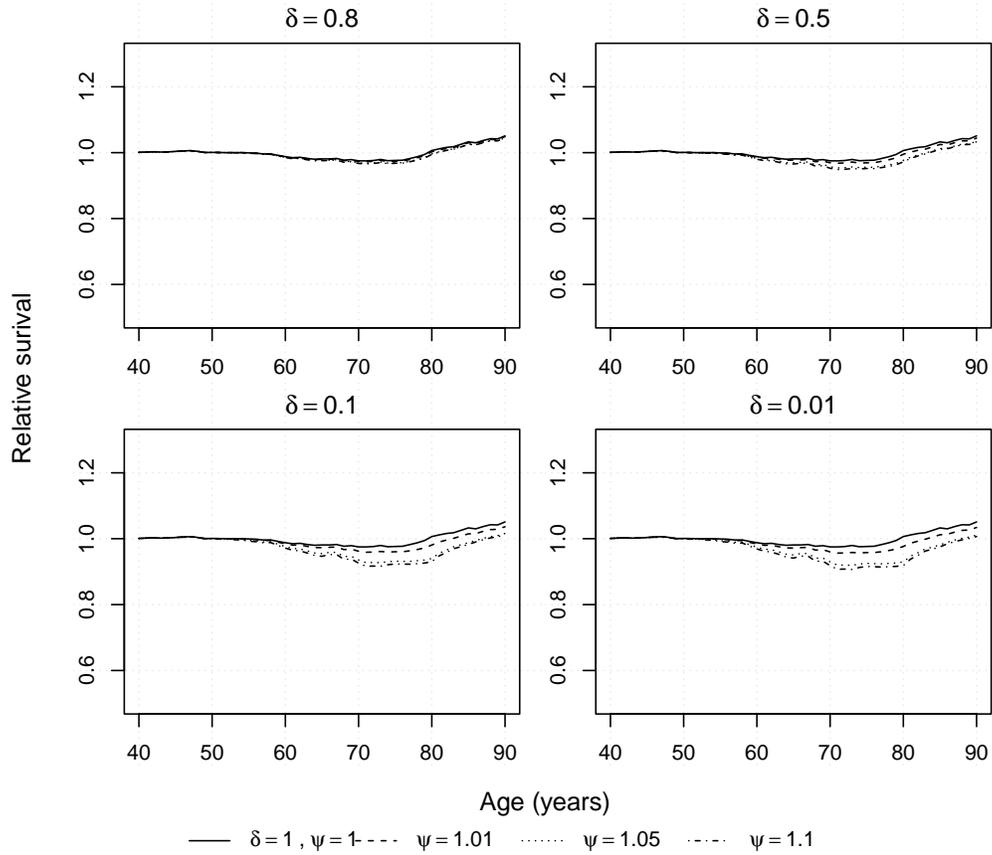


Figure 2.9: Left: Relative survival  $\hat{S}_R^{\text{age}}(t|z)$  for white males, defined in equation (2.11), including both TEVAR and non-TEVAR patients, with the mortality hazard for non-TEVAR patients is defined in (2.9), case (d). We specify different values of  $\psi$  and  $\delta$  to test sensitivity around the  $t - z$  term in the relative hazard for the non-TEVAR group.

to the treated group with a selection probability depending on age. We achieved this by introducing a one-parameter selection model that assumes that older TAA patients are less likely to receive TEVAR. We then modeled the mortality hazard for the potential patients by multiplying terms in the linear predictor for mortality hazard by a scale factor. We examined the changes in survival by varying values of these two parameters. Although the underlying hazard model is complex, the sensitivity analysis is revealing and simple to conduct.

The results show that adding the potential patients to the treated group can substantially affect the overall results. In particular, adding frail, elderly patients results in lower relative survival at older ages. This confirms our hypothesis that the oldest patients selected in our data are likely to be in more robust health, apart from having a TAA, than those in equivalent strata in the general population.

The choice of selection model can influence the sensitivity assessments. We used a monotone exponential decay function for the selection probability, restricting the selection probability to 1 for subjects diagnosed at age 40 or before and to decline with age thereafter. The results are more robust to the selection parameter  $\psi$  when the hazard of the potential patients is similar to the treated ones in the data.

We have examined only age as a potential confounder of the association between selection for surgery and mortality. This is because we observed the unexpected superior survival at older ages, which could be due to the “healthy screening effect” in older patients that we mentioned in our earlier analysis of the registry patients only. One can investigate other potential confounding variables, or incorporate prior quantitative knowledge regarding the unknown parameters, but sensitivity analysis results become complicated and difficult to interpret if we attempt to evaluate many parameters at once. Moreover, in this analysis, we create pseudo-data for the potential patients who could have been treated and included in the study. One might debate the suitability of this approach, but the goal of sensitivity

analysis is not to correctly identify the exact bias, but to add a further element to objectivity in understanding the results.

## CHAPTER 3

### Identifying the Optimal Size for Repair of a Thoracic Aortic Aneurysm

#### 3.1. Introduction

Thoracic aortic aneurysm (TAA) is a potentially lethal, often asymptomatic condition that is commonly detected during medical visits for other reasons. Its incidence is roughly 5.3/100,000/year [46], and its mortality rate is over 90% [47,48]. Most of these deaths could be prevented by early repair of the TAA. Thoracic endovascular aortic repair (TEVAR), a minimally invasive approach, has gradually replaced open repair as treatment for thoracic aortic disease. Although generally safer than open repair, the procedure can be complicated depending on factors such as the patient's overall health and the location and size of the aneurysm.

Clinical research has identified aneurysm size as having a significant impact on rates of aneurysm rupture and dissection and mortality. Larger aneurysms have an increased risk of rupture: The mean rupture rate is 2% for small aneurysms, rising to 6% for aneurysms of size 6cm or greater. The rate of serious complications also rises significantly as aneurysm size increases [49]. Several surgical guidelines recommend 6 cm to be the critical diameter [49] and suggest 5.5 to 6 cm as the ideal diameter for prophylactic surgical aortic repair [50,51].

These pioneering studies assessed size as a risk factor for aneurysm rupture regardless of whether patients underwent surgical treatment. In this paper, we investigate the question further by examining the effect of aneurysm size on post-surgery survival for TAA patients who underwent elective TEVAR. In particular, we want to determine the dependence of survival on size at repair and identify the optimal size at which to conduct TEVAR surgery.

As there have been no randomized clinical trials of TAA diameter at operation, we extract information for these treatments from observational data — specifically the Society for Vascular Surgery Vascular Quality Initiative registry of TEVAR patients. A significant deficiency of observational data is the potential for confounding bias due to covariate imbalance. That is, if there are covariates that are predictive of both size at surgery and post-surgical outcome, straightforward comparisons of mortality by size at surgery may be biased. One can reduce this bias by applying statistical methods designed to extract causal inferences from observational data. In this paper, we propose using propensity score analysis and the marginal structural model to estimate the marginal effect of aneurysm size on post-surgery survival.

Because aneurysm size at surgery is a continuous variable, a causal analysis of its effect on mortality must specify a functional form for this relationship. As the true form is unknown and possibly nonlinear, we seek to model the effect of size on mortality using a flexible function that will represent the data well without substantial bias. Common approaches to flexible modeling of this kind include piecewise polynomial models such as restricted cubic splines [52], B-splines [53], and penalized splines [54]. An alternative approach that does not require segmentation of the range of the predictor is the fractional polynomial (FP) [55]. With this method, one describes potential nonlinear effects using a multiple regression model that includes a limited number of terms involving logs and low powers of the predictor variable. The choice of function balances quality of fit with model complexity. We propose using the FP approach to ensure that models are parsimonious while providing consistent and meaningful results. To our knowledge, this paper represents the first use of fractional polynomials in a marginal structural model.

In subsequent sections, we describe the TEVAR data and methods, present the results of the analysis and discuss advantages and limitations of our study.

## 3.2. Patients and methods

### 3.2.1. The TEVAR registry

The Society for Vascular Surgery Vascular Quality Initiative (SVS VQI) prospectively enrolled patients who underwent TEVAR from 2010 to 2020 at 182 centers into a database of thoracic and complex EVAR procedures. The registry records patient demographics, baseline clinical characteristics, and follow-up outcomes. Baseline variables include pre-op maximum aortic diameter of the aneurysm, age at surgery, sex, race, smoking status, and indicators for chronic obstructive pulmonary disease (COPD), coronary artery disease (CAD), and diabetes. The outcome is post-procedure survival. Patient mortality status is identified by electronic health records at the centers, or by the Social Security Death Index for subjects who have valid 9-digit Social Security numbers.

We limited our analysis to patients aged less than 90 years, as those with true age 90 or above were coded as 90. We also restricted attention to patients who underwent TEVAR procedures with initial aneurysm sizes 30–190 mm. In addition to the full sample (including all eligible subjects), we examined two subgroups: (1) Those who had elective surgery and (2) those who had elective surgery with asymptomatic presentation.

### 3.2.2. Marginal structural model

The marginal structural model (MSM) [56,57] is a causal model that allows adjustment for confounding in potentially complex observational studies. The MSM is “marginal” in that it seeks to represent the relationship of potential outcomes to treatments averaged over covariates in the population of interest. It is “structural” in that it aims to explore causal effects, a common notion in the econometric literature. [58]

Suppose subject  $i \in \{1, \dots, n\}$  with baseline covariate vector  $X_i$  underwent TEVAR at aneurysm size  $S_i$ . We denote  $T_i$  to be the post-TEVAR survival outcome of patient  $i$ ,

measuring survival time since the operation. To reflect that physicians had a choice in TAA size at surgery, we adopt a potential outcomes notation. Given that we treat aneurysm size  $S \in \mathbb{S}$  as the continuous exposure/treatment of interest, we denote the potential outcomes to be  $\{T_i(s) : s \in \mathbb{S}\}$ , which represent the set of post-TEVAR survival outcomes that are possible for subject  $i$ . Only the factual survival outcome at size  $s$ ,  $T_i(S_i = s)$  is observed for patient  $i$ ; outcomes at other sizes,  $T_i(S_i \neq s)$ , are unobserved *counterfactual* potential outcomes.

Consider the following model for potential outcome  $T(s)$

$$T(s) \sim \mathbb{F}_{T(s)} = g(s; \beta),$$

which describes the distribution  $\mathbb{F}_{T(s)}$  of  $T(s)$  as a function  $g(\cdot; \cdot)$  of  $s$  and causal parameter  $\beta$ . It is an MSM because it is a structural model for the marginal distribution of the counterfactual variable  $T(s)$ . In addition to correct model specification, valid estimation of the MSM requires four main assumptions: Stable unit treatment values, exchangeability, consistency, and positivity. The stable unit treatment value assumption (SUTVA) states that there is no interference between units, in that the potential outcome of subject  $i$  is not affected by the potential outcomes of other subjects. Exchangeability, often referred to as no unmeasured confounding, states that the set of potential outcomes  $\{T(s) : s \in \mathbb{S}\}$  are independent of the exposure (aneurysm size), conditioning on the measured confounders. Consistency, which relates the counterfactual outcome to the observed outcome, states that the potential outcome for a subject who experiences treatment  $s$  under his observed covariates is precisely his observed outcome,  $S_i = s \Rightarrow T_i = T_i(s)$ . Positivity states that the probability of being treated at each aneurysm size is positive for each combination of covariates; that is, there are no impossible treatment conditions.

In the TAA analysis, we estimate the mean causal effect of baseline aneurysm size on the post-TEVAR survival outcome using a marginal structural Cox proportional hazards model. Let  $\lambda_{T(s)}(t|s, \beta)$  be the hazard of death at time  $t$ , in a hypothetical study in which all subjects underwent TEVAR at aneurysm size  $s$ :

$$\lambda_{T(s)}(t|s, \beta) = \lambda_0(t) \exp[\Phi(s; \beta)], \quad (3.1)$$

where  $\lambda_0(t)$  is a baseline hazard at time  $t$  post-TEVAR, and  $\Phi(s; \beta)$  is a smooth function of aneurysm size with corresponding coefficient vector  $\beta$ , which quantifies the causal effect of aneurysm size on post-TEVAR survival time.

Because of the potential for confounding, estimating the model by standard Cox regression generally gives a biased estimate of  $\beta$  in (3.1). One can eliminate the bias by estimating the Cox model with a subject weight proportional to the inverse predicted density function of the observed aneurysm size that the subject was treated at — so-called *inverse probability of treatment weighting (IPTW)* [58]. This method gives regression parameter estimates that are consistent for  $\beta$  in the MSM.

Two types of weights are in regular use: *Unstabilized* weights  $w_i$  are calculated as

$$w_i = \frac{1}{f(S_i|X_i)} = \frac{1}{b(s_i, x_i)}, \quad (3.2)$$

where the denominator is also referred as the *generalized propensity score* [59, 60], defined as the conditional density of treatment at aneurysm size  $S_i$  given covariate  $X_i$ . Generalized propensity score is an extension of the traditional propensity score [61], to apply to continuous exposure/treatment.

These unstabilized weights can vary substantially and lead to an undesirable dependence of the estimate on a few individuals with small propensity scores. To avoid this problem,

one can alternatively use *stabilized* weights, defined as

$$sw_i = \frac{f(S_i)}{f(S_i|X_i)} = \frac{a(s_i)}{b(s_i, x_i)}, \quad (3.3)$$

where the numerator is the marginal density function for aneurysm size [58]. In this application we treat aneurysm size  $S$  as a continuous exposure, which implies that we can obtain a finite variance only by using the stabilized weights. The estimated weights create a pseudo-population that effectively contains  $sw_i$  copies of patient  $i$ . Thus the final estimated MSM represents the relationship between exposure and outcome in a notional population from which the subjects represent a probability-weighted sample.

### 3.2.3. A fractional polynomial model for post-TEVAR survival

We describe the effect of aneurysm size  $S$  on post-TEVAR survival  $T$  by the marginal structural Cox proportional hazard model defined in (3.1). Particularly, we specify the smoothing function  $\Phi(s; \beta)$  in (3.1) with fractional polynomials [55]. In a model that includes  $m$  powers of  $s$ ,  $p_1 \leq \dots \leq p_m$ , the fractional polynomial function [55] is defined as

$$\Phi_{\text{FP}}(s; \beta) = \sum_{j=0}^m \beta_j H_j(s) \quad (3.4)$$

where  $H_0(s) = 1, p_0 = 0$  and

$$H_j(s) = \begin{cases} s^{(p_j)} & \text{if } p_j \neq p_{j-1} \\ H_{j-1}(s) \ln s & \text{if } p_j = p_{j-1}. \end{cases} \quad (3.5)$$

One typically restricts  $p_j$  to lie in the set  $\{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$ . The simplest fractional polynomial is the first-degree (FP1), which includes only one term. The power  $p = 1$  refers to  $H(s) = s$ ; all others are nonlinear functions, with  $p = 0$  referring to the natural logarithm

in  $s$ . A second-degree fractional polynomial (FP2) contains two terms, with equal or unequal power transformation.

### 3.2.4. A restricted cubic spline model for post-TEVAR survival

Another common smoothing tool to model nonlinearity is the restricted cubic spline [52], which incorporates cubic spline functions with tails constrained to be linear. Instead of fractional polynomials, we define  $\Phi(s; \beta)$  in (3.1) with restricted cubic splines. Suppose there are  $Q$  fixed knots,  $k_1 < k_2 < k_3 < \dots < k_{Q-1} < k_Q$ , and  $k_1$  and  $k_Q$  are two boundary knots. The restricted cubic spline function is defined as

$$\Phi_{\text{RCS}}(s; \beta) = \beta_{00} + \beta_0 s + \sum_{j=1}^{Q-2} \beta_j v_j(s), \quad (3.6)$$

where  $v_j(\cdot)$  is the  $j$ th basis function defined as

$$v_j(s) = (s - k_j)_+^3 - \frac{(s - k_{Q-1})_+^3 (k_Q - k_j)}{k_Q - k_{Q-1}} + \frac{(s - k_Q)_+^3 (k_{Q-1} - k_j)}{k_Q - k_{Q-1}}, \quad j = 1, \dots, Q-2, \quad (3.7)$$

with  $(u)_+ = \max(0, u)$  and  $\beta$  as the coefficient vector for the spline terms. In this analysis, we fit different models by specifying 3, 4, or 5 fixed knots at different centiles of aneurysm size (Table 2.3).

### 3.2.5. Estimation of optimal TAA size for TEVAR

#### 3.2.5.1. Generalized propensity scores

To estimate the marginal effect of aneurysm size on post-TEVAR mortality using a marginal structural model [58] defined in (3.1), we first estimated a propensity score model that predicts aneurysm size at operation from the available covariates. We fit a linear

regression model as

$$S = \gamma X + \epsilon$$

where the aneurysm size  $S$  is a continuous exposure,  $X$  is the vector of potential confounders, and  $\epsilon \sim N(0, \sigma_1^2)$ . The propensity score for subject  $i$  is estimated as

$$\hat{b}(s_i, x_i) = \frac{1}{\sqrt{2\pi\hat{\sigma}_1^2}} \exp\left(-\frac{1}{2\hat{\sigma}_1^2}(s_i - \hat{s}_i)^2\right),$$

where  $\hat{\sigma}_1^2$  is the estimated variance obtained from the regression model, and  $\hat{s}_i = \hat{\gamma}x_i$ . Thus, the resulting estimated unstabilized weight for subject  $i$  is

$$\hat{w}_i = \frac{1}{\hat{b}(s_i, x_i)}. \quad (3.8)$$

Next we compute stabilized weights (3.3) from the estimated propensity scores [58]. Because the exposure — aneurysm size at operation — is effectively continuous, we calculate the stabilized weight as a ratio of density functions. We estimate the numerator in (3.3) by fitting a model with only intercept:

$$S = \gamma_0 + \epsilon,$$

where  $\epsilon \sim N(0, \sigma_2^2)$ . The numerator,  $a(s_i)$  in (3.3) is estimated as

$$\hat{a}(s_i) = \frac{1}{\sqrt{2\pi\hat{\sigma}_2^2}} \exp\left(-\frac{1}{2\hat{\sigma}_2^2}(s_i - \hat{\gamma}_0)^2\right),$$

with  $\hat{\sigma}_2^2$  and  $\hat{\gamma}_0$  are the variance and mean aneurysm size in the sample. Then the estimated stabilized weight for subject  $i$  as

$$\hat{sw}_i = \frac{\hat{a}(s_i)}{\hat{b}(s_i, x_i)}. \quad (3.9)$$

### 3.2.5.2. MSM for aneurysm size

We modeled the nonlinear effect of aneurysm size using fractional polynomials and restricted cubic splines. We estimated the fractional polynomial models with R package *mfp* [62], which uses a backfitting algorithm to search for the best transformation of continuous variables from the set of powers. At a pre-specified significance level (we used  $\alpha = 0.05$ ), the algorithm uses the closed testing selection procedure to maintain the Type I error rate [63]. In our case, the algorithm tests for non-linearity by comparing the best-fitting FP2 against a straight-line model. It then tests the best-fitting FP2 against the best-fitting FP1 to identify a final best-fitting model. Because *mfp* does not accommodate weights, we estimated the FP function using a pseudo-dataset that included copies of each observation in proportion to its stabilized weight.

In estimating spline models, we used function *coxph* in R package *coxphw* and *rcs* in *rms*, selecting the model that minimizes the Akaike Information Criterion (AIC). We ran this analysis separately for the full sample and the subgroups.

From each final model, either using FP or restricted cubic splines, we identified the optimal aneurysm size by plotting the estimated log hazard against size and finding its minimum. We then computed bootstrap 95% quantile confidence intervals for the optimal aneurysm size with 2,000 random re-samples.

We executed all computations in R (R Foundation for Statistical Computing; Vienna, Austria).

## 3.3. Results

### 3.3.1. Patient characteristics

Table 3.1 shows that roughly half of the sample underwent elective surgery, and about 2/3 had no symptoms at presentation. The sample is 70% white and includes slightly more

males than females. Across all groups, the proportion of smokers is 27%; 36% have COPD; 21% have CAD; and 18% are diabetic. The covariates are reasonably well balanced between the overall sample and the subgroups.

### 3.3.2. Application to the TEVAR data

For each sample, we first estimated the stabilized weight in (3.3) for TEVAR aneurysm size. Specifically, we estimated the numerator in (3.3) by computing the marginal density of selection of the observed aneurysm size, and the denominator by the conditional density of the observed aneurysm size given age at surgery, sex, race, smoking status, COPD, diabetes, and CAD. Figure 3.1 plots the distribution of stabilized weights by aneurysm size for patients in each sample.

Table 2.1 presents fit statistics from various estimated FP models. The *mfp* algorithm conducts test for non-linearity (FP against a straight line) and simplification (FP2 vs FP1) to select the best fitting model. Table 3.2 displays the best-fitting univariate FP models for aneurysm size in each sample. In the unweighted version, we estimated the Cox regression with equal weights. In the weighted version, we estimated the same model using the stabilized weights in Equation (3.3). In the full-sample and elective-only groups, there was little difference between the unweighted and weighted estimates, with power terms of  $p = 1$  and  $p = 2$ , respectively. In the elective/asymptomatic group, the best-fitting model is FP1 with power  $p = 3$  for the unweighted sample, and FP2 with two powers (3, 3) for the weighted model.

Table 2.3 shows AIC values for restricted cubic spline models for aneurysm size in each sample. We tested different combinations of knots at various locations based on the centiles of aneurysm size in the samples. For the full sample, we selected the final model with knots at the 10, 30, 40 centiles, which has the lowest AIC. Using same criteria, we selected the model with knots at the 5, 50, 95 centiles of aneurysm size for the elective-TEVAR subgroup.

Table 3.1: Baseline variables in the TEVAR registry data

Variable	Value	All ( <i>n</i> = 2240)	Elective ( <i>n</i> = 1802)	Elective/Asymptomatic ( <i>n</i> = 1472)
Discrete: <i>n</i> (%)				
Sex	Male	1213 (54.2)	992 (55.0)	830 (56.4)
	Female	1027 (45.8)	810 (45.0)	642 (43.6)
Race	Non-Hispanic White	1585 (70.8)	1296 (71.9)	1093 (74.3)
	Non-Hispanic Black	396 (17.7)	305 (16.9)	223 (15.1)
	Hispanic	106 (4.7)	78 (4.3)	58 (3.9)
	Other	153 (6.8)	123 (6.8)	98 (6.7)
Current smoking	Yes	607 (27.1)	476 (26.4)	369 (25.1)
COPD	Yes	815 (36.4)	681 (37.8)	559 (38.0)
CAD	Yes	470 (21.0)	377 (20.9)	291 (17.8)
Diabetes	Yes	394 (17.6)	308 (17.1)	254 (17.3)
Elective surgery	Yes	1802 (80.4)	—	—
Presentation	Asymptomatic	1513 (67.5)	—	—
	Symptomatic	567 (25.3)	—	—
	Ruptured	160 (7.1)	—	—
Continuous: mean (SD)				
Age at surgery		72.0 (10.7)	71.8 (10.5)	72.2 (9.8)
Max pre-op TAA diameter (mm)		59.6 (13.3)	58.7 (12.0)	58.5 (11.1)

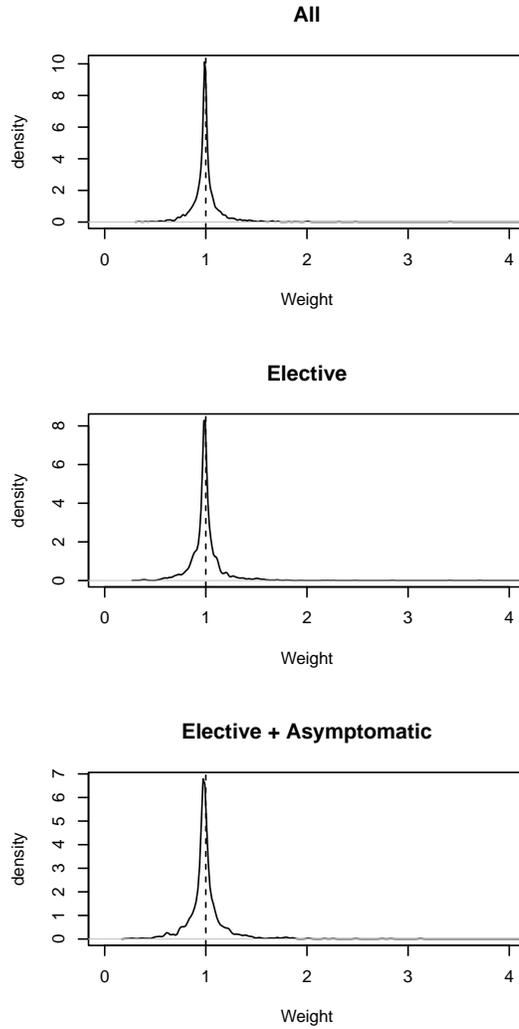


Figure 3.1: Distribution of stabilized weights for aneurysm size at operation in each sample, estimated from Equation (3.9). Most weights are near 1 in all samples.

For the elective/asymptomatic subgroup, we selected the model with knots at the 5, 50, 95 centiles. Though its AIC value is not the lowest, it has fewer parameters than the model that gives the lowest AIC.

Figure 3.2 shows the estimated log mortality hazard in each group, estimated from the marginal structural models with fractional polynomials (FP) and restricted cubic splines (RCS) in aneurysm size. For comparison purposes, we scaled the log relative hazards to have value 0 at the minimum. All curves show a similar trend of increasing hazard as

Table 3.2: Estimated fractional polynomials in the marginal structural Cox model

Sample	Model	Power ( $p$ )	Coefficient	SE
All (n=2240)	Unweighted	$p_1 = 1$	0.019	0.0034
	Weighted	$p_1 = 1$	0.019	0.0030
Elective (n=1802)	Unweighted	$p_1 = 2$	$1.5 \times 10^{-4}$	$2.7 \times 10^{-5}$
	Weighted	$p_1 = 2$	$1.4 \times 10^{-4}$	$2.1 \times 10^{-5}$
Elective/Asymptomatic (n=1472)	Unweighted	$p_1 = 3$	$1.7 \times 10^{-6}$	$3.8 \times 10^{-7}$
	Weighted	$p_1 = 3$	$-2.5 \times 10^{-5}$	$8.1 \times 10^{-6}$
		$p_2 = 3$	$5.7 \times 10^{-6}$	$1.7 \times 10^{-6}$

aneurysm size increases. RCS curves, however, show more upward concavity at the specified knots than FP curves. For the full sample, the chosen hazard model is the first-degree FP with power  $p = 1$ , giving a linear increase in the hazard with aneurysm size. Meanwhile, the RCS curve shows a minimum hazard at 52 mm, with 95% CI 30–53. In the elective sample, the hazard is concave upward in the FP model, while the RCS curve shows a dip in the 50–65 mm range. We identified the optimal aneurysm size to be approximately 55 mm, with bootstrap 95% CI (30–59). Similarly, in the E+A subgroup, the FP and RCS curves both show dips near 60mm. The estimated optimal TAA size in the FP model is 60 (95% CI: 30–69), and in the RCS model it is 57 (95% CI: 50–61).

### 3.4. Discussion

We have estimated marginal structural models to identify the causal effect of aneurysm size on the post-surgery survival outcome for patients with TAA who underwent TEVAR. First, we estimated a propensity score model predicting aneurysm size at operation from a panel of available covariates. We then estimated a Cox model predicting survival from aneurysm size at operation, using weights calculated from the propensity score analysis. The survival model estimated in this way has the interpretation of describing the average survival hazard as a function of aneurysm size in the population of interest, and therefore is a basis for causal inferences. Best-fitting models reflected non-linearity in the effect of

Table 3.3: Optimal TAA size in mm (95% CI) for TEVAR surgery

Model	All ( $n = 2240$ )	Elective ( $n = 1802$ )	Elective/Asymptomatic ( $n = 1472$ )
Fractional polynomial	30 (30–49)	30 (30–58)	60 (30–69)
Restricted cubic spline	52 (30–53)	55 (30–59)	57 (50–61)

aneurysm size on mortality hazard. Our findings largely support medical guidelines that recommend conducting TEVAR when the aneurysm is slightly smaller than 60mm [50, 51].

An alternative way to model aneurysm size would be to categorize it into broad ranges. Although this would simplify modeling to some extent, it introduces other problems: The need to choose the number and bounds of the categories, and some potential loss of efficiency compared to retaining size as continuous [64]. When analyzing the continuous aneurysm size variable, it is advisable to use a flexible functional form. Fractional polynomials are flexible and straightforward to apply, and avoid the need for arbitrary selection of knots [55, 65]. In our analysis, we limited the number of FP regression terms to two, as recommended by Royston & Altman (1994), to simplify model selection. We note that in Figure 3.2, the FP models are generally smoother than restricted cubic spline curves estimated from the same data. With RCS, it appears that there is a substantial chance of overfitting, in the sense that the estimated model prominently reflects artifacts in the data.

The validity of our results depends on the five assumptions indicated above. That the models “fit” the data is plausible thanks to our use of flexible curve-fitting approaches. SUTVA is also plausible, because each patient is selected for surgery based on the size of his own aneurysm only and not the sizes of others’ aneurysms. Positivity also seems plausible, as — in principle, at least — operation is possible at any aneurysm size. The exchangeability assumption, which posits that we have all measured confounding covariates and modeled them properly, is more speculative, as it would be in any analysis of an observational study.

Deviation from correctly specifying model for continuous exposure can result in extreme weights and biased causal inference [66]. Such assumptions are not testable from the data, but can be examined using sensitivity analysis [43, 67]. For instance, we specified a standard approach based on normal distribution to construct the inverse propensity scores for aneurysm size, assuming constant variance for the exposure. One can explore other distributional forms for the exposure to examine for heteroscedasticity or outliers that could potential lead to extreme weights [68].

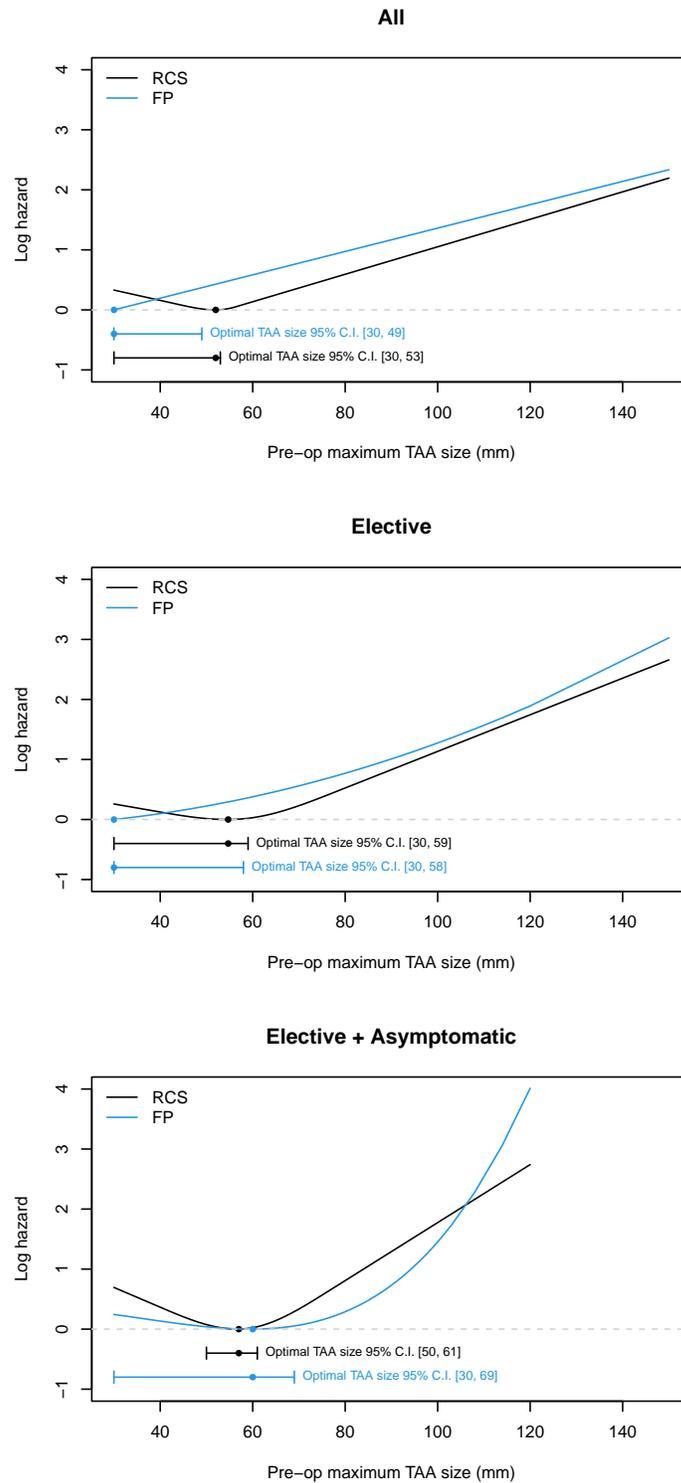


Figure 3.2: Comparison between hazard models estimated by fractional polynomials (FP) and restricted cubic splines (RCS) in the marginal structural Cox model, with 95% bootstrap confidence intervals for the size giving minimum hazard.

APPENDIX A  
APPENDIX of CHAPTER 1

**A.1. Figures**

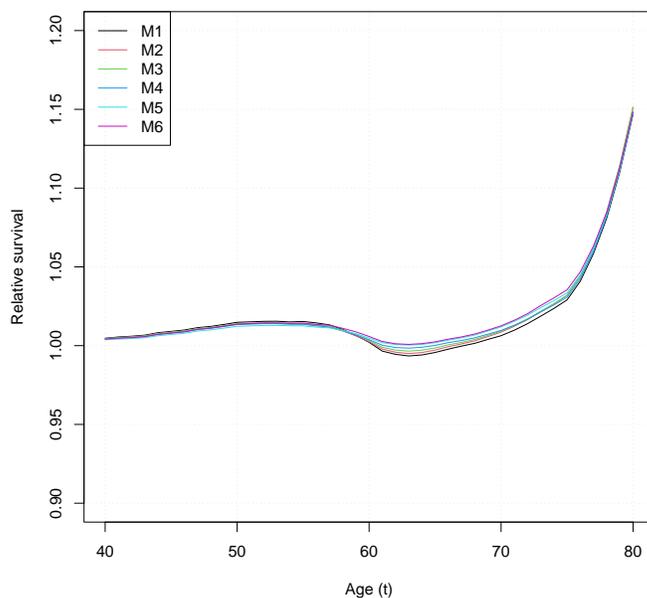


Figure 1.1: Overall relative survival for TEVAR patients at age  $t$ , given TEVAR was performed at age  $z$  prior to age  $t$ . Model M1 was used in the main analysis, adjusting for sex and race and their interaction. Model M2 to M6 adjusted for additional baseline clinical covariates: aneurysm size, diabetes, smoking status, COPD, and CAD (see Table 1.4).

**A.2. Population hazard adjustment**

A.2.1. Diabetes

Denote  $h_{0k}(t)$  to be the population mortality hazard at age  $t$  from national life tables, stratified on a subset of demographic covariates,  $k(x)$ , such as race and sex. Suppose this

population hazard can be further decomposed based on other clinical characteristics. Considering diabetes (DM), for example, we can express the total population hazard as the sum of hazards at age  $t$  of diabetics, denoted as  $h_{0k}(t|\text{DM})$ , and non-diabetics, denoted as  $h_{0k}(t|\text{DM}^c)$ ,

$$\begin{aligned} h_{0k}(t) &= h_{0k}(t|\text{DM})\text{Pr}(\text{DM}) + h_{0k}(t|\text{DM}^c)\text{Pr}(\text{DM}^c) \\ &= \text{HR}_{\text{DM},k}h_{0k}(t|\text{DM}^c)\text{Pr}(\text{DM}) \\ &\quad + h_{0k}(t|\text{DM}^c)(1 - \text{Pr}(\text{DM})). \end{aligned} \tag{1.1}$$

where

- $\text{Pr}(\text{DM})$  and  $\text{Pr}(\text{DM}^c)$  are the age- and sex-adjusted prevalences of diabetes and non-diabetes, respectively, stratified by race (Table 1.1), calculated using multiple logistic regression [69]; and
- $\text{HR}_{\text{DM},k}$  is the mortality hazard ratio for diabetes in stratum  $k$  (here, age, race, and sex; see Table 1.2). We use estimated hazard ratios obtained from Cox proportional hazard models [70].

Table 1.1: A subset of crude and adjusted prevalence of total diabetes in the US, NHANES 2011-2016 [69]

Race/Ethnicity	Crude prevalence (% (95% CI))	Age-sex-adjusted prevalence (% (95% CI))
Non-Hispanic white	13.3 (12.1, 14.6)	12.1 (11.0, 13.4)
Non-Hispanic black	18.3 (16.6, 19.9)	20.4 (18.8, 22.1)
.	.	.
.	.	.

Thus, the adjusted population hazard at age  $t$  for non-diabetics follows as

$$h_{0k}(t|DM^c) = \frac{h_{0k}(t)}{HR_{DM,k}Pr(DM) + (1 - Pr(DM))}. \quad (1.2)$$

For example, we estimate the adjusted population hazard at age 40 for a non-Hispanic white male non-diabetic as

$$\frac{0.0024}{(2.886)(0.204) + (1 - 0.204)} \approx 0.0017,$$

which is 30% lower than the one from national life tables with the same age, sex, and race ( $h_{0k}(40) = 0.0024$ ). The adjusted population hazard for a diabetic with the same age, sex, and race is calculated by multiplying that of a non-diabetic by a factor  $HR_{DM,k}$ , resulting in  $(0.0017)(2.886) \approx 0.0049$ , 100% higher than the unadjusted population hazard from the life tables.

#### A.2.2. COPD

We have also recalculated the population hazard using COPD statistics. The age-adjusted prevalences of COPD in men and women are 5.3 and 6.5 respectively, collected from self-reported and physician diagnosed COPD cases; the paper also reported separately the age-adjusted prevalence of COPD by race [71]. The hazard ratio for moderate COPD was 1.6, calculated from a multivariate Cox regression [72]. For example, the adjusted population hazard at age 40 for a non-COPD male is

$$\frac{0.0024}{(1.6)(0.053) + (1 - 0.053)} \approx 0.0023,$$

and the adjusted population hazard for corresponding individual with COPD is  $(0.0023)(1.6) \approx 0.0037$ .

Table 1.2: A subset of age-specific hazard ratios associated with diabetes, for age 30 and above [70]

Age interval	Non-Hispanic White		Non-Hispanic Black	
	Men	Women	Men	Women
30–34	3.549	3.9993	2.815	3.167
35–39	3.201	3.601	2.538	2.856
40–44	2.886	3.248	2.289	2.576
.	.	.	.	.
.	.	.	.	.
80–84	1.263	1.421	1.002	1.127
85 +	1.139	1.281	1.000	1.016

APPENDIX B

APPENDIX of CHAPTER 3

**B.1. Tables**

Table 2.1: Evaluating the fit of estimated fractional polynomial functions in the marginal structural Cox model

Sample	Model	Unweighted		Weighted <sup>a</sup>	
		Power ( $p_1, p_2$ )	Deviance	Power ( $p_1, p_2$ )	Deviance
All ( $n = 2240$ )	Linear	1,—	5329.17	1,—	5458.95
	FP1	2,—	5329.10	2,—	5457.18
	FP2	-1,-1	5322.76	-1,-0.5	5451.16
Elective ( $n = 1802$ )	Linear	1,—	3398.01	1,—	3486.43
	FP1	2,—	3394.27	2,—	3482.17
	FP2	-1, 0	3389.44	-0.5, 0	3477.65
Elective/Asymptomatic ( $n = 1472$ )	Linear	1,—	2586.09	1,—	2710.91
	FP1	3,—	2578.89	3,—	2693.70
	FP2	-2, 3	2574.26	3, 3	2683.32

Deviance =  $-2$  times the maximized log-likelihood.

<sup>a</sup>Adjusted for sex, race, smoking status, COPD, CAD, and diabetes.

Table 2.2: Likelihood ratio tests ( $\chi^2$ ,  $P$  value) for different FP models using significance level  $\alpha = 0.05$

Sample	Unweighted		Weighted <sup>a</sup>	
	FP2 (-1, -1) vs. Linear	FP2 (-1, -1) vs. FP1 (2)	FP2 (-1, -0.5) vs. Linear	FP2 (-1, -0.5) vs. FP1 (2)
All ( $n = 2240$ )	$\chi^2 = 6.41$ $P = 0.0933$	—	$\chi^2 = 7.79$ $P = 0.0506$	—
Elective ( $n = 1802$ )	FP2 (-1, 0) vs. Linear $\chi^2 = 8.57$ $P = 0.0356$	FP2 (-1, 0) vs. FP1 (2)	FP2 (-0.5, 0) vs. Linear $\chi^2 = 4.83$ $P = 0.0894$	FP2 (-0.5, 0) vs. FP1 (2)
Elective/Asymptomatic ( $n = 1472$ )	FP2 (-2, 3) vs. Linear $\chi^2 = 11.83$ $P = 0.0080$	FP2 (-2, 3) vs. FP1 (3) $\chi^2 = 4.63$ $P = 0.0988$	FP2 (3, 3) vs. Linear $\chi^2 = 27.59$ $P < 0.0001$	FP2 (3, 3) vs. FP1 (3) $\chi^2 = 10.38$ $P = 0.0056$

<sup>a</sup>Adjusted for sex, race, smoking status, COPD, CAD, and diabetes.

$\chi^2$  statistics is computed as difference in deviances between two models.

Likelihood Ratio test between FP2 and straight line (linear) uses  $\chi^2$  distribution with d.f. = 3; FP2 and FP1 uses  $\chi^2$  with d.f. = 2.

Steps of selecting final models:

1. If the test between FP2 vs. straight line (linear) is not significant ( $P \geq 0.05$ ), then we reject FP2 model and select linear model as final model.
2. If  $P < 0.05$  in step (1), then test again between FP2 vs FP1.
3. If  $P < 0.05$  in step (2), then select FP2 as final model, otherwise select the best-fitting FP1.

Table 2.3: AIC values of restricted cubic spline models for post-TEVAR survival

Model	Location of knots* (size centiles)	All ( $n = 2240$ )	Elective ( $n = 1802$ )	Elective/Asymptomatic ( $n = 1472$ )
1	5, 50, 95	5384.29	<b>3413.04</b>	<b>2649.37</b>
2	10, 50, 90	5383.99	3413.39	2650.89
3	20, 50, 80	5383.80	3413.96	2652.20
4	25, 50, 75	5383.71	3414.10	2652.34
5	30, 50, 70	5384.00	3414.20	2652.64
6	10, 30, 70	5382.35	3413.39	2653.76
7	10, 30, 50	5381.97	3413.34	2654.37
8	10, 30, 40	<b>5381.80</b>	3413.31	2654.74
9	30, 50, 70, 90	5385.13	3415.90	2649.66
10	10, 30, 70, 90	5383.66	3415.21	2651.28
11	5, 27.5, 50, 72.5, 95	5385.66	3415.02	2647.41

\*Percentiles of aneurysm sizes.

Bolded AIC values belong to the final selected model in each sample.

## APPENDIX C

### Supplementary R codes

#### C.1. Chapter 1

##### C.1.1. Maximum likelihood estimators

We first created dataset *dd*, in which each subject has multiple rows, with each row representing the data for every age from start of surgery until age at last follow-up. This was done to incorporate the mortality hazard by age of general population from the national life tables. The function **fun.mod4b** maximizes the log-likelihood for the survival of TEVAR patients and computes the maximum likelihood estimators for the vector of coefficients,  $\theta$ , corresponding to the regression terms in the hazard model (Equation 1.3 & 1.4). To obtain the optimal MLEs, we iterate this function until the MLEs converge.

```
fun.mod4b <- function(b){
  beta0 <- b[1]
  beta1 <- b[2]
  beta2 <- b[3]
  beta3 <- b[4]
  beta4 <- b[5]
  beta5 <- b[6]
  beta6 <- b[7]
  beta7 <- b[8]
  d1 <- subset(dd, dd$event==1 & dd$alive.age == floor(dd$surv.age))
  d2 <- subset(dd, dd$event==1 & dd$alive.age < floor(dd$surv.age))
  d3 <- subset(dd, dd$event==0 & dd$alive.age <= floor(dd$surv.age))
  p1 <- exp(beta0 + beta1*d1$AGE + beta2*d1$V2 + beta3*d1$V3 +
    ↪ beta4*d1$tminusz + beta5*d1$gender1 + beta6*d1$race1 +
    ↪ beta7*d1$gender1*d1$race1)*d1$census.h0
  p2 <- exp(beta0 + beta1*d2$AGE + beta2*d2$V2 + beta3*d2$V3 +
    ↪ beta4*d2$tminusz + beta5*d2$gender1 + beta6*d2$race1 +
    ↪ beta7*d2$gender1*d2$race1)*d2$census.h0
```

```

p3 <- exp(beta0 + beta1*d3$AGE + beta2*d3$V2 + beta3*d3$V3 +
  → beta4*d3$tminusz + beta5*d3$gender1 + beta6*d3$race1 +
  → beta7*d3$gender1*d3$race1)*d3$census.h0
logl <- sum(log(p1/(1+p1))) + sum(log(1 - p2/(1+p2))) + sum(log(1 -
  → p3/(1+p3)))
  return(-logl)
}

### Iterations to find optimal MLEs
model <- fun.mod4b # ML function to estimate the maximum likelihood
kk <- 8
dd <- temp2
est.method="Nelder-Mead"
opt <- optim(rep(0.1,kk), model, hessian=T, method=est.method, control=ctr)
int.values <- opt$par

iterCount <- 0
repeat {
  int.values <- opt$par
  opt <- optim(int.values, model, hessian=T, method=est.method, control=ctr)
  iterCount = iterCount + 1
  print(iterCount)
  print(round(opt$par,4))
  print(opt$value)
  if(sum(int.values - opt$par)==0) {
    break
  }
}

```

### C.1.2. Relative survival

Function **fun.RS.tt** estimates the overall relative survival at age  $t$ ,  $\hat{S}_R^{\text{age}}(t|z, x)$  in Equation (1.10). Function **fun.t.z** computes  $\hat{S}_R^{\text{years}}(t|z, x)$ , relative survival of a TEVAR patient at age  $t$  given the operation occurred at age  $z$  ( $< t$ ) (Equation 1.11).

```

### Overall relative survival at age t:

fun.RS.tt <- function(pop.data, tt) {
  for (j in 1:length(tt)){
    obs.data <- subset(TEVARdata1, AGE <= tt[j] & surv.age >=tt[j])
    temp.h0.1 <- subset(pop.data, AGE <= (tt[j]-1))
    for (i in 1:length(obs.data$AGE)) {

```

```

# observed surv
rrace[i] <- obs.data$RACE[i]
ggender[i] <- obs.data$GENDER[i]
temp.h0 <- subset(temp.h0.1, RACE==rrace[i] & GENDER==ggender[i] & AGE
  ↪ >= (obs.data$AGE[i]))
zz <- temp.h0$AGE
h0 <- temp.h0$h0
race1 <- ifelse(temp.h0$RACE==5,1,0)
gender1 <- ifelse(temp.h0$GENDER==1,1,0)
exp.beta.m8 <- exp(bb[1] + bb[2]*(zz) + bb[3]*ifelse(zz-knot[1] > 0,
  ↪ zz-knot[1], 0) + bb[4]*ifelse(zz-knot[2] > 0, zz-knot[2], 0) +
  ↪ bb[5]*(tt[j]-zz) + bb[6]*gender1 + bb[7]*race1 +
  ↪ bb[8]*gender1*race1)
haz.est.m8 <-
  ↪ ((h0/(1-h0))*(exp.beta.m8)/(1+(h0/(1-h0))*(exp.beta.m8)))
obs.surv <- 1-haz.est.m8
obs.surv.t[i] <- prod(obs.surv)

# expected surv
exp.surv.t[i] <- prod(1-h0)
}
sum.obs.surv.t[j] <- sum(obs.surv.t)
sum.exp.surv.t[j] <- sum(exp.surv.t)
rs[j] <- sum.obs.surv.t[j]/sum.exp.surv.t[j]
}
tbl <- cbind(tt, sum.obs.surv.t, sum.exp.surv.t, rs)
return(tbl)
}

### Relative survival at age t given TEVAR at age z:

fun.t.z <- function(pop.data, aage, rrace, ggender) {
tbl.obs <- matrix(rep(NA, length(aage)*5), length(aage), 5)
tbl.exp <- matrix(rep(NA, length(aage)*5), length(aage), 5)
for (j in 1:length(aage)){
tt <- aage[j] + 1:5
for (i in 1:length(tt)) {

## observed surv
temp.h0 <- pop.data %>% filter(RACE==rrace & GENDER==ggender & AGE <=
  ↪ tt[i]-1 & AGE >= aage[j])
zz <- temp.h0$AGE
zz.star <- temp.h0$AGE - 70
h0 <- temp.h0$h0

```

```

race1 <- ifelse(temp.h0$RACE==5,1,0)
gender1 <- ifelse(temp.h0$GENDER==1,1,0)
exp.beta.m8 <- exp(bb[1] + bb[2]*(zz) + bb[3]*ifelse(zz-knot[1] > 0,
  ↪ zz-knot[1], 0) + bb[4]*ifelse(zz-knot[2] > 0, zz-knot[2], 0) +
  ↪ bb[5]*(tt[i]-zz) + bb[6]*gender1 + bb[7]*race1 +
  ↪ bb[8]*gender1*race1)
haz.est.m8 <-
  ↪ ((h0/(1-h0))*(exp.beta.m8)/(1+(h0/(1-h0))*(exp.beta.m8)))
obs.surv <- 1-haz.est.m8
obs.surv.t[i] <- prod(obs.surv)

## expected surv
exp.surv.t[i] <- prod(1-h0)
}
tbl.obs[j,] <- obs.surv.t
tbl.exp[j,] <- exp.surv.t
}
colnames(tbl.obs) <- c("obs1", "obs2", "obs3", "obs4", "obs5")
colnames(tbl.exp) <- c("exp1", "exp2", "exp3", "exp4", "exp5")
rs1 <- tbl.obs[,1]/tbl.exp[,1]
rs2 <- tbl.obs[,2]/tbl.exp[,2]
rs3 <- tbl.obs[,3]/tbl.exp[,3]
rs4 <- tbl.obs[,4]/tbl.exp[,4]
rs5 <- tbl.obs[,5]/tbl.exp[,5]
final.tbl <- cbind(aage, rs1, rs2, rs3, rs4, rs5)
return(final.tbl)
}

```

## C.2. Chapter 2

Similar to function **fun.RS.tt**, function **fun.RS.tt.sen** also estimates the overall relative survival at age  $t$  for TEVAR patients. However, it is slightly different because now we consider the hazard of the “potential patients” governed by  $\delta$ , and incorporate parameter  $\delta$  governs the patient selection probability ( $\psi$ ). The function below shows the mortality hazard of untreated TEVAR patients in case (a) (Equation 2.6). One can move  $\delta$  to appropriate locations to represent the hazard models in other cases (b–d). When  $\delta = 1$  and  $\psi = 1$ , we obtain the results from the main model in Chapter 1.

```

fun.RS.tt.sen <- function(pop.data, tt, delta, psi) {
  for (j in 1:length(tt)) {
    obs.data <- subset(TEVARdata1, AGE <= tt[j] & surv.age >= tt[j])
    temp.h0.1 <- subset(pop.data, AGE <= (tt[j]-1))
    for (i in 1:length(obs.data$AGE)) {

```

```

# observed surv
rrace[i] <- obs.data$RACE[i]
ggender[i] <- obs.data$GENDER[i]
temp.h0 <- subset(temp.h0.1, RACE==rrace[i] & GENDER==ggender[i] & AGE
  ↪ >= (obs.data$AGE[i]))
zz <- temp.h0$AGE
h0 <- temp.h0$h0
race1 <- ifelse(temp.h0$RACE==5,1,0)
gender1 <- ifelse(temp.h0$GENDER==1,1,0)

exp.beta1 <- exp(bb[1] + bb[2]*(zz) + bb[3]*ifelse(zz-knot[1] > 0,
  ↪ zz-knot[1], 0) + bb[4]*ifelse(zz-knot[2] > 0, zz-knot[2], 0) +
  ↪ bb[5]*(tt[j]-zz) + bb[6]*gender1 + bb[7]*race1 +
  ↪ bb[8]*gender1*race1)
exp.beta0 <- exp(bb[1] + delta*(bb[2]*(zz) + bb[3]*ifelse(zz-knot[1]
  ↪ > 0, zz-knot[1], 0) + bb[4]*ifelse(zz-knot[2] > 0, zz-knot[2], 0)
  ↪ + bb[5]*(tt[j]-zz)+ bb[6]*gender1 + bb[7]*race1 +
  ↪ bb[8]*gender1*race1))

haz1 <- ((h0/(1-h0))*(exp.beta1)/(1+(h0/(1-h0))*(exp.beta1)))
haz0 <- ((h0/(1-h0))*(exp.beta0)/(1+(h0/(1-h0))*(exp.beta0)))

p1 <- psi^(zz-40)
p0 <- 1 - p1

haz.total <- haz1*p1 + haz0*p0

obs.surv <- 1-haz.total
obs.surv.t[i] <- prod(obs.surv)

# expected surv
exp.surv.t[i] <- prod(1-h0)
}
sum.obs.surv.t[j] <- sum(obs.surv.t)
sum.exp.surv.t[j] <- sum(exp.surv.t)
rs[j] <- sum.obs.surv.t[j]/sum.exp.surv.t[j]
}
tbl <- cbind(tt, sum.obs.surv.t, sum.exp.surv.t, rs)
return(tbl)
}

```

### C.3. Chapter 3

```
### Estimate stabilized weights for continuous exposure (aneurysm size)

wt.1 <- ipwpoint(exposure = PREOP_MAXAAADIA,
  family = "gaussian",
  numerator = ~ 1,
  denominator = ~ AGE + I(gender_grp=="M") +
  ↪ as.factor(race_grp) + as.factor(current_smoking) +
  ↪ as.factor(COPD_1) + as.factor(DM) + as.factor(CAD),
  ↪ data = temp)

### Create pseudo-population using estimated weights above

temp1 <- temp %>% dplyr::select(PRIMPROCID, DEAD, PROC_SURVIVALDAYS,
  ↪ PREOP_MAXAAADIA, AGE, gender_grp, race_grp, current_smoking, COPD_1, DM,
  ↪ CAD) %>% dplyr::mutate(wt = wt.1$ipw.weights, wt.r = round(wt))
temp2 <- temp1 %>% tidyr::uncount(wt.r)

### Run marginal structural Cox model with fractional polynomials

f <- mfp(Surv(PROC_SURVIVALDAYS, DEAD) ~ fp(PREOP_MAXAAADIA), family = cox,
  ↪ method="breslow", alpha=0.05, select=1, data = temp2, verbose=T)

### Run marginal structural Cox model with restricted cubic splines

cph.rcs <- coxph(Surv(PROC_SURVIVALDAYS, DEAD) ~ rcs(PREOP_MAXAAADIA,
  ↪ quantile(PREOP_MAXAAADIA, c(0.1, 0.3, 0.4))), weights =
  ↪ wt.1$ipw.weights, data=temp)
```

## BIBLIOGRAPHY

- [1] Ederer F, Axtell L, Cutler SJ. The relative survival rate: A statistical methodology. *Natl Cancer Inst Monogr* 1961; 6: 101–121.
- [2] Ederer F, Heise H. Instructions to IBM 650 programmers in processing survival computations. *National Cancer Institute* 1959.
- [3] Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 1982; 38: 933–942.
- [4] Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. *J Appl Stat* 1987; 36: 309–317.
- [5] Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: Elements for further discussion. *Stat Med* 1990; 9: 529–538.
- [6] Cortese G, Scheike TH. Dynamic regression hazards models for relative survival. *Stat Med* 2008; 27: 3563–3584.
- [7] Perme MP, Henderson R, Stare J. An approach to estimation in relative survival regression. *Biostatistics* 2009; 10: 136–146.
- [8] Bolard P, Quantin C, Esteve J, Faivre J, Abrahamowicz M. Modelling time-dependent hazard ratios in relative survival: Application to colon cancer. *J Clin Epidemiol* 2001; 54: 986–996.
- [9] Nelson CP, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. *Stat Med* 2007; 26: 5486–5498.
- [10] Remontet L, Bossard N, Belot A, Esteve J. An overall strategy based on regression models to estimate relative survival and model the effects of prognostic factors in cancer survival studies. *Stat Med* 2007; 26: 2214–2228.
- [11] Giorgi R, Abrahamowicz M, Quantin C, et al. A relative survival regression model using B-spline functions to model non-proportional hazards. *Stat Med* 2003; 22: 2767–2784.
- [12] Lambert PC, Smith LK, Jones DR, Botha JL. Additive and multiplicative covariate regression models for relative survival incorporating fractional polynomials for time-dependent effects. *Stat Med* 2005; 24: 3871–3885.
- [13] Dupont C, Bossard N, Remontet L, Belot A. Description of an approach based on maximum likelihood to adjust an excess hazard model with a random effect. *Cancer Epidemiol* 2013; 37: 449–456.

- [14] Charvat H, Remontet L, Bossard Nea. A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. *Stat Med* 2016; 35: 3066–3084.
- [15] Kuss O, Blankenburg T, Haerting J. A Relative Survival Model for Clustered Responses. *Biom J* 2008; 50: 408–418.
- [16] Cnudde P, Rolfson O, Timperley AJ, et al. Do patients live longer after THA and is the relative survival diagnosis-specific?. *Clin Orthop Relat Res* 2018; 476: 1176–1175.
- [17] Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med* 2004; 23: 51–64.
- [18] Tutz G, Schmid M. *Modeling Discrete Time-to-Event Data*. New York: Springer . 2016.
- [19] Arias E, Xu J. United States life tables, 2017. *Natl Vital Stat Rep* 2019; 68.
- [20] Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958; 53: 457–481.
- [21] Durrleman S, Simon R. Flexible regression models with cubic splines. *Stat Med* 1989; 8: 551–561.
- [22] Gray RJ. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *J Am Stat Assoc* 1992; 87: 942–951.
- [23] Royston P. Flexible parametric alternatives to the Cox model, and more. *Stata J* 2001; 1: 1–28.
- [24] Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* 2002; 21: 2175–2197.
- [25] Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: Choice of the time-scale. *Am J Epidemiol* 1997; 145: 72–80.
- [26] Lamarca R, Alonso J, Gómez G, Muñoz A. Left-truncated data with age as time scale: An alternative for survival analysis in the elderly population. *J Gerontol A Biol Sci Med Sci* 1998; 53(5): M337–M343.
- [27] Kim M, Paik MC, Jang J, et al. Cox proportional hazards models with left truncation and time-varying coefficient: Application of age at event as outcome in cohort studies. *Biom J* 2017; 59: 405–419.
- [28] Talbäck M, Dickman PW. Estimating expected survival probabilities for relative survival analysis – Exploring the impact of including cancer patient mortality in the calculations. *Eur J Cancer* 2011; 47: 2626–2632.
- [29] Grafféo N, Jooste V, Giorgi R. The impact of additional life-table variables on excess mortality estimates. *Stat Med* 2012; 31: 4219–4230.
- [30] Blakely T, Soeberg M, Carter K, Costilla R, Atkinson J, Sarfati D. Bias in relative survival methods when using incorrect life-tables: Lung and bladder cancer by smoking status and ethnicity in New Zealand. *Int J Cancer* 2012; 131: E974–E982.

- [31] Karthikesalingam A, Bahia SS, Patterson BO, et al. The shortfall in long-term survival of patients with repaired thoracic or abdominal aortic aneurysms: Retrospective case-control analysis of hospital episode statistics. *Eur J Vasc Endovasc Surg* 2013; 46: 533–541.
- [32] Wang T, Desai MY. Thoracic aortic aneurysm: Optimal surveillance and treatment. *Cleve Clin J Med* 2020; 87: 557–568.
- [33] Norman PE, Semmens J, Lawrence-Brown MM, Holman CD. Long term relative survival after surgery for abdominal aortic aneurysm in Western Australia: Population-based study. *BMJ* 1998; 317: 852–826.
- [34] Orrason AW, Garmo H, Styrke J, Dickman PW, Stattin P. Comparison of Relative Survival and Cause-Specific Survival in Men With Prostate Cancer According to Age and Risk Category: A Nationwide, Population-Based Study. *Am J Epidemiol* 2021; 190: 2053–2063.
- [35] Kpodonu J, Preventza O, Ramaiah VG, et al. Endovascular repair of the thoracic aorta in octogenarians. *Eur J Cardiothorac Surg* 2008; 34: 630–634.
- [36] Dakour-Aridi H, Yin K, Hussain F, Locham S, Azizzadeh A, Malas MB. Outcomes of intact thoracic endovascular aortic repair in octogenarians. *J Vasc Surg* 2021; 74: 882–892.
- [37] Rudarakanchana N, Reeves BC, Bicknell CD, Heatley FM, Cheshire NJ, Powell JT. Treatment decisions for descending thoracic aneurysm: Preferences for thoracic endovascular aneurysm repair or surveillance in a discrete choice experiment. *Eur J Vasc Endovasc Surg* 2014; 48: 13–22.
- [38] Schaffer JM, Lingala B, Miller DC, Woo YJ, Mitchell RS, Dake MD. Midterm survival after thoracic endovascular aortic repair in more than 10,000 Medicare patients. *J Thorac Cardiovasc Surg* 2015; 149: 808–823.
- [39] Thygesen LC, Ersbøll AK. When the entire population is the sample: strengths and limitations in register-based epidemiology. *Eur J Epidemiol* 2014; 29: 551–558.
- [40] Nguyen H, Heitjan DF, Jeon-Slaughter H. Modeling relative survival: An application to thoracic endovascular aortic repair. *Unpublished* 2023; 0: 00–00.
- [41] Greenland S. Basic methods for sensitivity analysis of bias. *Int J Epidemiol* 1996; 25: 1107–1116.
- [42] Copas JB, Li HG. Inference for nonrandom samples (with discussion). *J R Stat Soc Series B* 1997; 59: 55–95.
- [43] Robins JM. Association, causation, and marginal structural models. *Synthese* 1999; 121: 151–179.
- [44] Little RJA, Rubin DB. *Statistical Analysis with Missing Data, 2nd edn*. New York: Wiley . 2002.
- [45] Troxel AB, Ma G, Heitjan DF. An index of local sensitivity to nonignorability. *Stat Sin* 2004; 14: 1221–1237.
- [46] Melo RG, Duarte GS, Lopes A, et al. Incidence and prevalence of thoracic aortic aneurysms: A systematic review and meta-analysis of population-based studies. *Semin Thorac Cardiovasc Surg* 2021; 34: 1–16.

- [47] Bickerstaff LK, Pairolero PC, Hollier LH, et al. Thoracic aortic aneurysms: A population-based study. *Surgery* 1982; 92: 1103–1108.
- [48] Johansson G, Markstrom U, Swedenborg J. Ruptured thoracic aortic aneurysms: A study of incidence and mortality rates. *J Vasc Surg* 1995; 21: 985–988.
- [49] Davies RR, Goldstein LJ, Coady MA, et al. Yearly rupture or dissection rates for thoracic aortic aneurysms: Simple prediction based on size. *Ann Thorac Surg* 2002; 73: 17–28.
- [50] Perko MJ, Norgaard M, Herzog TM, Skov Olsen P, Schroeder TV, Pettersson G. Unoperated aortic aneurysm: A survey of 170 patients. *Ann Thorac Surg* 1995; 59: 1204–1209.
- [51] Riambau V, Böckler D, Brunkwall J, et al. Editor’s Choice — Management of descending thoracic aorta diseases: Clinical practice guidelines of the European Society for Vascular Surgery. *Eur J Vasc Endovasc Surg* 2017; 53: 4–52.
- [52] Herndon JE, Harrell FE. The restricted cubic spline as baseline hazard in the proportional hazards model with step function time-dependent covariables. *Stat Med* 1995; 14: 2119–2129.
- [53] Abrahamowicz M, MacKenzie T, Esdaile JM. Time-dependent hazard ratio: Modeling and hypothesis testing with application in lupus nephritis. *J Am Stat Assoc* 1996; 91: 1432–1439.
- [54] Eilers PHC, Marx BD. Flexible smoothing with B-Splines and penalties. *Statist Sci* 1996; 11(2): 89–121.
- [55] Royston P, Altman D. Regression using fractional polynomials of continuous covariates: Parsimonious parametric modelling. *J R Stat Soc Ser C Appl Stat* 1994; 43: 429–453.
- [56] Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000; 11: 561–570.
- [57] Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC . 2020.
- [58] Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; 11: 550–560.
- [59] Imai K, Van Dyk D. Causal inference with general treatment regimes: Generalizing the propensity score. *J Am Stat Assoc* 2004; 99 (467): 854–866.
- [60] Hirano K, Imbens GW. The propensity score with continuous treatments. In: Gelman A, Meng XL., eds. *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*. New York: Wiley. 2004 (pp. 73–84).
- [61] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983; 70: 41–55.
- [62] Benner A. Multivariable fractional polynomials. *R News* 2005; 5: 20–23.
- [63] Ambler G, Royston P. Fractional polynomial model selection procedures: Investigation of Type I error rate. *J Stat Comput Simul* 2001; 69: 89–108.
- [64] Greenland S. Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis. *Epidemiology* 1995; 6: 450–454.

- [65] Sauerbrei W, Perperoglou A, Schmid M, et al. State of the art in selection of variables and functional forms in multivariable analysis — outstanding issues. *Diagn Progn Res* 2020; 4.
- [66] Fong C, Hazlett C, Imai K. Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements. *Ann Appl Stat* 2018; 12: 156–177.
- [67] Robins JM, Rotnitzky A, Scharfstein DO. *Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models*. New York: Springer . 1999.
- [68] Naimi AI, Moodie EEM, Auger N, Kaufman JS. Constructing inverse probability weights for continuous exposures: A comparison of methods. *Epidemiology* 2004; 25(2): 292–299.
- [69] Cheng YJ, Kanaya AM, Araneta ea. Prevalence of diabetes by race and ethnicity in the United States, 2011–2016. *JAMA* 2019; 322: 2389–2398.
- [70] Preston SH, Choi D, Elo IT, Stokes A. Effect of diabetes on life expectancy in the United States by race and ethnicity. *Biodemography Soc Biol* 2018; 64: 139–151.
- [71] Sullivan J, Pravosud V, Mannino DM, Siegel K, Choate R, Sullivan T. National and state estimates of COPD morbidity and mortality — United States, 2014–2015. *Chronic Obstr Pulm Dis* 2018; 5: 324–333.
- [72] Mannino DM, Buist AS, Petty TL, Enright PL, Redd SC. Lung function and mortality in the United States: Data from the first National Health and Nutrition Examination Survey follow-up study. *Thorax* 2003; 58: 388–393.