20

Survival Analysis of Case-Control Data: A Sample Survey Approach

Norman Breslow and Jie Hu

University of Washington, Seattle

CONTENTS

20.1 Introduction

Large cohort (follow-up) studies, including those of patients in randomized clinical trials, provide the most definitive evidence of the effect of treatments on clinical outcomes and of exposure to risk factors on disease outcomes. The Atherosclerosis Risk in Communities (ARIC) study, for example, ascertained nearly 16,000 subjects to investigate the impact of environmental and genetic risk factors on cardiovascular disease.(Williams, 1989) The Women's Health Initiative (WHI) randomized over 26,000 women to hormone therapy or placebo in clinical trials for women with and without an intact uterus.(Anderson et al., 2003) In view of logistic and financial constraints associated with measurment of biomarkers on tens of thousands of subjects, both ARIC and WHI selected random samples, termed a cohort random sample and a subcohort, respectively, for whom routine bioassays of selected biomarkers were performed. Sampling was stratified on demographic factors in order to achieve desired minority representation. In a series of substudies, biomarkers were assayed using stored serum samples for additional subjects who developed one of several disease outcomes of interest. Data from the cohort sample and from disease cases outside the sample were combined using specialized techniques based on the Cox model to estimate disease risks.(Prentice, 1986) However, this approach ignored large quantities of information on baseline factors available for main cohort members who were not also a disease case or in the cohort sample.

This chapter presents a general sample survey approach to survival analysis of case-control data. By considering properties of the sampling design separately from those of the assumed model, it accomodates a variety of parametric and semiparametric models, in particular the Cox (1972) proportional and Lin-Ying (1994) additive hazards models. Stratification of the case-control sample may be based on any information available for the entire cohort, with the goal to select the most informative subjects. For example, stratification on a correlate of exposure can increase the variation of the true exposure in the sample and thus enhance the precision of the corresponding regression coefficient.(Borgan et al., 2000) Sampling of cases as well as controls is permitted, which is important when studying more common disease outcomes or when damaged serum samples or failed bioassays result in missing biomarker data for both cases and controls. Recovery of information lost by ignoring baseline covariates for non-sampled subjects is facilitated.

20.2 Example: Radiation and breast cancer (BC)

Women treated for tuberculosis between 1930 and 1956 at one of two Massachusetts sanitoria were followed for the occurrence of breast cancer (BC) through 1980.(Hrubec et al., 1989) A version of the data from this study, originally distributed with the Epicure computer program (Preston et al., 1993), is used for illustration throughout this chapter. Most women received radiation exposure to the breast from multiple fluoroscopies used for diagnostic X-ray in conjuction with lung collapse therapy for their tuberculosis; those not examined by fluoroscopy were considered unexposed. Cumulative doses of radiation to the breast (Gy) were estimated from the number of fluoroscopies, a reconstruction of exposure conditions, absorbed dose calculations and other available data. Although the radiation doses were in fact known for all women in the cohort, two random samples were drawn to simulate case-control studies in which the actual dose, the "expensive" covariate, was measured only on a sample. The fact that dose actually was known for the entire cohort allowed results obtained from the case-control analyses to be compared, for illustrative purposes, with results of fitting the same models to the entire cohort.

Note that we have referred to the two simulated studies as "case-control" studies rather than as "case-cohort" studies. From the viewpoint of survey sampling we regard the two designs as equivalent, and both terms are used in the sequel. The cases in a case-cohort study consist of all cases ascertained by the time of data analysis, whether or not they happened to have been sampled for the cohort random sample (subcohort). The controls consist of all non-cases sampled for the subcohort. Since the subcohort was randomly sampled (possibly with stratification) from the main cohort, the controls consist of a random sample of non-cases in the main cohort. All cases are sampled. The main difference between the two designs is that, in some applications involving a partially missing time-dependent covariate, the value of this covariate may be known only at the time of occurrence of cases that occur outside the subcohort, which would rule out the sample survey approach we describe. Other methods of analysis that include such cases only in the "risk set" corresponding to the time of its occurrence are required.(Prentice, 1986; Borgan et al., 2000)

20.2.1 Description of the study cohort

Table 20.1 shows the frequency distributions of the 75 cases and 1645 controls according to three major risk factors. There was keen interest in whether women first irradiated early in life had higher rates of BC, at the same dose level and attained age, compared with those irradiated later.

Participants were followed from the date of discharge following their first hospitalization for tuberculosis, which was usually shortly after the onset of treatment, until the earliest of death, BC diagnosis or study closure. These

TABLE 20.1
Risk factors in the breast cancer study

	Cases (BC)	Controls	Totals
Totals	75	1645	1720
Attained age (years)			
0-39	10	244	254
40-59	51	435	486
60-93	14	966	980
Radiation dose (Gy)			
0	21	677	698
< 2	46	858	904
≥ 2	8	110	118
Age at first treatment (years)			
0-19	33	621	654
20-29	28	609	637
30+	14	415	429

events were recorded as of the ages at which they occurred. From the viewpoint of survival analysis, therefore, age at BC was left truncated at entry to the cohort and right censored at exit. Age was the fundamental time variable in the models; hence baseline hazards correspond to baseline age-specific rates.

A simple random sample (subcohort) of 150 of the 1720 women, including by chance 7 of the 75 BC cases, was first drawn to simulate the standard case-cohort design. In an attempt to increase design efficiency for estimation of the effects of age at first treatment, a second stratified random subcohort was drawn of 50 women from each of the three categories of age at first treatment shown in Table 20.1. Since, again by chance, 6 of the 75 BC cases were included in the stratified sample, the simple and stratified studies differed only slightly in terms of total sample size: 218 and 219, respectively. Radiation doses were considered unknown for subjects who were neither cases nor in the respective sample. Comparisons of the results obtained using the full cohort data with those for the two case-cohort samples, simple and stratified, are presented in sections 20.5 and 20.7 for the proportional and additive hazards models. We consider first, however, how these data are best viewed from a sample survey perspective and outline the methods of analysis used to make the comparisons.

20.3 Two-phase sampling

Two-phase (or double) sampling was originally proposed for estimation of the finite population total of some target variable that is difficult to measure. A Phase I sample is first drawn from the finite population. Readily available information on some correlate of the target variable is ascertained for all Phase I subjects and used to stratify the sampling of a Phase II subsample. Multiplying together the Phase I and Phase II sampling probabilities for each Phase II subject, the population total is estimated by inverse probability weighting of the Phase II observations.(Lumley, 2012, §8.1)

20.3.1 The case-control study as a two-phase design

The case-control study is profitably viewed as a two-phase design. Here the population is a probability model (superpopulation) and the goal is estimation of model parameters. The Phase I sample (cohort) is regarded as a series of independent and identically distributed (i.i.d) observations drawn from the model. If these were completely observed, inference would proceed simply by fitting the model to the cohort data following the usual paradigm of model based inference. However, a portion of the observations, *e.g.*, the biomarkers, is in general only available for subjects in a Phase II subsample.

Let $\mathbf{Z}$ denote a generic observation from the probability model. For survival analyses with left truncation or a series of intermittent observation periods, $\mathbf{Z} = (Y, T, \Delta, \mathbf{X})$ where $Y = Y(t)$ is the "at risk" indicator of whether or not the subject is under observation at t , Δ is a censoring indicator for the survival time T and $\mathbf{X}$ is a vector of covariates, of which only a portion $\widetilde{\mathbf{X}}$ is fully observed at Phase I. When time is simply time on study, $Y(t) = \mathbf{1}(T \geq t)$.

The parameter is $(\boldsymbol{\beta}, \Lambda_0)$ where $\boldsymbol{\beta}$ are regression coefficients, interpretable as log hazard ratios, excess hazards or otherwise depending on the survival model, and Λ_0 is the baseline cumulative hazard function. We denote by $\mathbf{V} = (Y, T, \Delta, \widetilde{\mathbf{X}}, \mathbf{W}) \in \mathcal{V}$ the variables that are fully observable for the entire cohort, including possibly a vector $\mathbf{W}$ of auxiliary variables, *i.e.*, variables that are not wanted in the model but may be useful for stratification of the Phase II subsample or otherwise to improve efficiency. Let $\mathcal{V} = \mathcal{V}_1 \cup \ldots \cup \mathcal{V}_J$ denote a partition of the range of the fully observed variables into J strata. For each of N Phase I subjects, define $R_i = 1/0,\ i = 1, \ldots, N$ to be the indicator of whether or not the i^{th} subject is selected for the Phase II subsample. There are two possibilities.

TABLE 20.2
Two-phase stratified sampling

	Stratum				
	1	2	$\cdots$	J	Total
Phase I (cohort)	N_1	N_2	$\cdots$	N_J	N
Phase II (subsample)	n_1	n_2	$\cdots$	n_J	n
Sampling fractions	$\frac{n_1}{N_1}$	$\frac{n_2}{N_2}$	$\cdots$	$\frac{n_J}{N_J}$	$\frac{n}{N}$

20.3.2 Finite population stratified sampling (FPSS)

If the $\mathbf{V}$'s were available simultaneously for everyone in the cohort, we could count the numbers N_j in each of the J strata and use simple random sampling without replacement to sample n_j of them for Phase II (Table 20.2). Since the sampling indicators satisfy $\sum_{i=1}^{N} R_i \mathbf{1}(\mathbf{V}_i \in \mathcal{V}_j) = n_j \leq N_j$, where the n_j are fixed, they are dependent random variables, although exchangeable within strata and independent from stratum to stratum. The dependence complicates the asymptotic theory, for which reason most researchers prefer to develop the theory to preserve the i.i.d. structure of the data.

20.3.3 Bernoulli sampling

Alternatively, suppose the $\mathbf{V}_i$ became available sequentially and that a sampling indicator was drawn independently for each one according to $\Pr(R_i = 1) = \pi_0(\mathbf{V}_i)$. Here $\pi_0(\mathbf{v})$ is a known function that is bounded away from zero; for stratified sampling, $\pi_0(\mathbf{v}) = p_j$ for $\mathbf{v} \in \mathcal{V}_j$ for p_j known *a priori*. In this case the Phase II sample sizes are random variables and n_j/N_j would converge in probability to p_j as $N \uparrow \infty$. In the sequel we explain how apparent differences between the two sampling schemes can be resolved by calibration of the samplng weights to the stratum totals N_j.

20.3.4 Example: radiation and breast cancer

The two-phase design corresponding to the standard case-cohort study had just two strata, one comprising the cases sampled at 100% ($n_1 = N_1 = 75$) and the second the controls with $n_2 = 143, N_2 = 1645$. The four strata for the stratified study were the cases ($n_1 = N_1 = 75$) and the three control samples stratified by age at first treatment with, respectively, $n_2 = 44, N_2 = 621$, $n_3 = 50, N_3 = 609$ and $n_4 = 50, N_4 = 415$. Compare Tables 20.1 and 20.2.

20.4 Estimation of parameters in the Cox model

This section summarizes statistical methods that have been developed to estimate both relative (log hazard ratio) and absolute (cumulative hazard) risks when fitting the Cox model to full cohort data and to two-phase samples. The theoretical basis for the methodology relies heavily on empirical processes and semiparametric inference, for which a detailed discussion is beyond the scope of this Handbook.(van der Vaart and Wellner, 1996; van der Vaart, 1998) Hence some readers may find it advisable to proceed directly to the next section, on applications of the methods to the radiation and breast cancer study, referring back to the formulas presented in this section as needed to understand how the various estimates, standard errors and related quantities shown there were actually calculated from the data.

Under the Cox model, the cumulative hazard at t for the i$^{\text{th}}$ subject with covariates $\mathbf{X}_i$ is $\Lambda_i(t;\boldsymbol{\beta},\Lambda_0) = e^{\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}}\Lambda_0(t)$, where the $\boldsymbol{\beta}$ are log hazard ratios corresonding to unit changes in the covariates and Λ_0 is the baseline cumulative hazard. This is assumed to be continuously differentiable on a finite time interval $[0,\tau]$ such that $\Pr[Y(\tau)=1] > 0$. Were the covariates fully observable for all Phase I subjects, $\boldsymbol{\beta}$ would be estimated by solving the (Cox, 1972) partial likelihood score equations $\widetilde{U}(\widetilde{\boldsymbol{\beta}}_N) = \sum_{i=1}^N \widetilde{U}(\mathbf{Z}_i;\boldsymbol{\beta}) = 0$ for $\widetilde{\boldsymbol{\beta}}_N$, where

$$\begin{aligned}\widetilde{U}(\mathbf{Z};\boldsymbol{\beta}) &= \Delta\left[\mathbf{X} - \widetilde{\mathbf{m}}(T;\beta)\right], \qquad \text{and} && (20.1)\\ \widetilde{\mathbf{m}}(t;\boldsymbol{\beta}) &= \sum_{i=1}^N \mathbf{X}_i e^{\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}} Y_i(t) \Bigg/ \sum_{i=1}^N e^{\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\beta}} Y_i(t) && (20.2)\end{aligned}$$

denotes the weighted mean of the covariates of subjects “at risk” at t, weighting each by its hazard ratio. Asymptotic properties of estimators like $\widetilde{\boldsymbol{\beta}}_N$ are usually derived from asymptotic expansions for its normalized difference from the true value of the parameter, here $\sqrt{N}\left(\widetilde{\boldsymbol{\beta}}_N - \boldsymbol{\beta}\right)$.

20.4.1 Expansion via the efficient influence function

Denote by $\mathbb{N}_i(t) = \mathbf{1}[T_i \le t, \Delta_i = 1]$ the standard counting process, by

$$\widetilde{\Lambda}_N(t;\boldsymbol{\beta}) = \frac{1}{N}\sum_{i=1}^N \frac{\Delta_i \mathbf{1}[T_i \le t]}{\sum_{j=1}^N e^{\mathbf{X}_j^{\mathrm{T}}\boldsymbol{\beta}} Y_j(T_i)}$$

the Breslow estimate of $\Lambda_0(t;\boldsymbol{\beta})$, by

$$\widetilde{U}^*(\mathbf{Z};\boldsymbol{\beta}) = \int_0^\tau \left[\mathbf{X} - \widetilde{\mathbf{m}}(t;\boldsymbol{\beta})\right] Y(t)\left(d\mathbb{N}(t) - e^{\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}} d\Lambda_0(t;\boldsymbol{\beta})\right), \qquad (20.3)$$

the semiparametric efficient score that corrects for estimation of Λ_0, and by

$$\begin{aligned}\widetilde{\mathbf{I}}(\boldsymbol{\beta}) &= -\left.\frac{\partial \mathrm{E}\left[\widetilde{U}^*(\mathbf{Z};\widetilde{\boldsymbol{\beta}})\right]}{\partial \widetilde{\boldsymbol{\beta}}^{\mathrm{T}}}\right|_{\widetilde{\boldsymbol{\beta}}=\boldsymbol{\beta}} && (20.4)\\ &= \int_0^\tau \left[\frac{\mathrm{E}\mathbf{X}^{\otimes 2} e^{\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}} Y}{\mathrm{E}e^{\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}} Y} - \left(\frac{\mathrm{E}\mathbf{X} e^{\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}} Y}{\mathrm{E}e^{\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}} Y}\right)^{\otimes 2}\right] \mathrm{E}e^{\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}} Y d\Lambda_0,\end{aligned}$$

the efficient information. Then

$$\sqrt{N}\left(\widetilde{\boldsymbol{\beta}}_N - \boldsymbol{\beta}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^N \widetilde{\ell}(\mathbf{Z}_i;\boldsymbol{\beta}) \ + o_p(1) \qquad (20.5)$$

with $\widetilde{\ell}(\mathbf{Z}_i;\boldsymbol{\beta}) = \widetilde{\mathbf{I}}^{-1}(\boldsymbol{\beta})\widetilde{U}_i^*(\mathbf{Z}_i;\boldsymbol{\beta})$ denoting the contribution to the efficient influence function for the i$^{\text{th}}$ subject.(Cox, 1972; van der Vaart, 1998, §25.12.1)

20.4.2 Inverse probability weighted (IPW) estimate

The survey estimator $\widehat{\boldsymbol{\beta}}_N$ (Binder, 1992; Lin, 2000) solves an IPW version of the Phase II score equations (20.1, 20.2), namely

$$\begin{aligned} \widehat{U}(\widehat{\boldsymbol{\beta}}_N) &= \sum_{i=1}^{N} \frac{R_i}{\pi_i} \int_0^{\tau} \left[\mathbf{X}_i - \widehat{\mathbf{m}}(t; \widehat{\boldsymbol{\beta}}_N) \right] Y_i(t) d\mathbb{N}_i(t) \ = 0, \quad \text{with} \\ \widehat{\mathbf{m}}(t, \boldsymbol{\beta}) &= \sum_{i=1}^{N} \frac{R_i}{\pi_i} \mathbf{X}_i e^{\mathbf{X}_i^{\mathrm{T}} \boldsymbol{\beta}} Y_i(t) \Big/ \sum_{i=1}^{N} \frac{R_i}{\pi_i} e^{\mathbf{X}_i^{\mathrm{T}} \boldsymbol{\beta}} Y_i(t). \end{aligned}$$

Here the π_i are the sampling probabilities used to select the Phase II sample ($R_i = 1$) based on the variables $\mathbf{V}_i$ known at Phase I. An asymptotic expansion for $\widehat{\boldsymbol{\beta}}_N$ that follows from the work of Lin (2000) and others is

$$\begin{aligned} \sqrt{N}\left(\widehat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}\right) &= \sqrt{N}\left(\widetilde{\boldsymbol{\beta}}_N - \boldsymbol{\beta}\right) + \sqrt{N}\left(\widehat{\boldsymbol{\beta}}_N - \widetilde{\boldsymbol{\beta}}_N\right) \\ &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[\frac{R_i}{\pi_i} \widetilde{\ell}(\mathbf{Z}_i; \boldsymbol{\beta}) \right] + o_p(1) \qquad (20.6) \\ &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[\widetilde{\ell}(\mathbf{Z}_i; \boldsymbol{\beta}) + \left(\frac{R_i - \pi_i}{\pi_i} \right) \widetilde{\ell}(\mathbf{Z}_i; \boldsymbol{\beta}) \right] + o_p(1). \end{aligned}$$

These equations lead to several important insights:

1. The normalized difference between the IPW (Phase II) estimator and the true value equals the normalized difference between the (unobservable) Phase I estimator $\widetilde{\boldsymbol{\beta}}_N$ and the true value plus the normalized difference between the Phase I and Phase II estimators. The first term in the expansion shown on the bottom line is the term corresponding to the Phase I estimator (20.5).
2. Since $\mathrm{E}(R_i|\mathbf{Z}_i) = \pi_i$, the two components of this expansion are uncorrelated, hence asymptotically independent. Thus

   $$\text{Total Variance} = \text{Phase I Variance} + \text{Phase II Variance}.$$

3. The Phase I variance Var $\widetilde{\ell}(\mathbf{Z}; \boldsymbol{\beta})$, due to i.i.d. sampling from the superpopulation, is not affected by the choice of Phase II design. The Phase II variance, which represents the loss of information from not having complete data at Phase I, *is* affected, *e.g.*, by stratification. It is the normalized error arising from IPW estimation of an unknown finite population total, namely, the total of the influence function contributions $\widetilde{\ell}(\mathbf{Z}_i; \boldsymbol{\beta})$ for all Phase I subjects. It is best viewed as *design based*, with the randomness stemming exclusively from the sampling indicators R_i, with the Phase I observations $\mathbf{Z}_i$ being regarded as fixed.

For FPSS of the Phase II sample from the cohort (Table 20.2), $\pi_i = n_j/N_j$ if the i^{th} subject is in stratum j, *i.e.*, if $\mathbf{V}_i \in \mathcal{V}_j$. Given the Phase I data

$$\text{Phase II Variance } = \sum_{j=1}^{J} \frac{N_j}{N} \frac{N_j - n_j}{n_j} \widehat{\text{Var}}_j \left[\widetilde{\ell}(\mathbf{Z};\boldsymbol{\beta}) \right], \qquad (20.7)$$

where $\widehat{\text{Var}}_j$ denotes the finite population variance of the influence function contributions $\widetilde{\ell}(\mathbf{Z}_i;\boldsymbol{\beta})$ observed for stratum j in the Phase II sample. Strata sampled at 100% ($n_j = N_j$) do not contribute to the Phase II variance. See Chen and Lo (1999) for the simple case-cohort design ($J = 1$), Borgan et al. (2000) for the stratified case-cohort design and Breslow and Wellner (2007) for the general two-phase stratified design. The generality of the latter is important since it covers situations were the cases may also be selected by stratified sampling with sampling fractions less than 100%.

20.4.3 Calibration of the sampling weights

Survey samplers (Deville and Särndal, 1992) improve estimates of finite population totals by calibrating the design weights to known population totals of auxiliary variables that are highly correlated with the variable whose total is to be estimated. This technique may also improve the efficiency of IPW estimators. Suppose $\mathbf{C} = \mathbf{C}(\mathbf{V}) = (C_1(\mathbf{V}), \ldots, C_K(\mathbf{V}))^{\text{T}}$, where the calibration variables $C_k(\mathbf{V})$ are known for the entire Phase I sample and $\text{E}\left(\mathbf{C}\mathbf{C}^{\text{T}}\right)$ is nonsingular. We consider calibration for Bernoulli sampling (20.3.3). This implies the Phase II sample is an i.i.d. random sample from the superpopulation based on known selection probabilities $\pi_0(\mathbf{V}_i)$. The calibrated weights w_i are chosen to be as close as possible to the design weights $d_i = 1/\pi_0(\mathbf{V}_i)$ in terms of a distance measure G and also to satisfy the calibration equations $\sum_{i=1}^{N} C_k(\mathbf{V}_i) = \sum_{i=1}^{N} R_i w_i C_k(\mathbf{V}_i)$ whereby the IPW estimator exactly estimates the Phase I totals for $k = 1, \ldots, K$. The standard constrained optimization procedure involves calculation of a K-vector $\widehat{\lambda}_N$ of Lagrange multipliers for the calibration equations. Choosing the Poisson deviance $G(w, d) = w \log(w) - w + d$ as the distance measure leads to a procedure known as "raking" that guarantees non-negative weights: $w_i = \exp\left[-\widehat{\lambda}_N^{\text{T}} \mathbf{C}(\mathbf{V}_i)\right] \Big/ \pi_0(\mathbf{V}_i)$.(Lumley, 2012, §7.4).

Calibration generally results in a reduction in the limiting Phase II variance that may or may not be realized in finite samples. With $\widehat{\boldsymbol{\beta}}_N(\widehat{\lambda}_N)$ denoting the estimator with calibrated weights, $\text{Var}_{\text{A}}\sqrt{N}(\widehat{\boldsymbol{\beta}}_N(\widehat{\lambda}_N) - \boldsymbol{\beta}) =$

$$\text{Var}\left[\widetilde{\ell}(Z;\boldsymbol{\beta})\right] + \text{E}\left\{ \left(\frac{1-\pi_0(\mathbf{V})}{\pi_0(\mathbf{V})} \right) \left[\widetilde{\ell}(\mathbf{Z};\boldsymbol{\beta}) - \Pi_{\mathbf{C}}\widetilde{\ell}(\mathbf{Z};\boldsymbol{\beta}) \right]^{\otimes 2} \right\} \qquad (20.8)$$

where $\Pi_{\mathbf{C}}\widetilde{\ell} = \text{E}(\widetilde{\ell}\mathbf{C}^{\text{T}})(\text{E}\mathbf{C}\mathbf{C}^{\text{T}})^{-1}\mathbf{C}$ denotes population least squares projection of the influence function on $[\mathbf{C}_1, \ldots, \mathbf{C}_K]$.(Breslow and Lumley, 2009) Since we are effectively attempting to estimate the unknown Phase I totals of this influence function, a good choice for the calibration variables would be $\mathbf{C} = \mathrm{E}(\widetilde{\ell}|\mathbf{V})$. This yields the optimal member of the class of augmented inverse probability weighted (AIPW) estimators of (Robins et al., 1994); see Lumley et al. (2011). One method of approximating this optimal $\mathbf{C}$ is considered in the example in the next section.

It is instructive to consider calibration to the Phase I stratum totals by selecting $\mathbf{C}_j(\mathbf{V}) = \mathbf{1}[\mathbf{V} \in \mathcal{V}_j]$ for $j = 1, \ldots, J$. The calibrated weights are inverses of the actual sampling fractions n_j/N_j, the design weights inverses of the *a priori* p_j. Since the projection onto $[\mathbf{C}_1, \ldots, \mathbf{C}_J]$ yields stratum specific means, the limiting Phase II variance of the calibrated estimator, *i.e.*, the second term in (20.8), equals

$$\sum_{j=1}^{J} \Pr(\mathcal{V}_j) \frac{1-p_j}{p_j} \mathrm{Var}_j \left[\widetilde{\ell}(\mathbf{Z};\boldsymbol{\beta})\right] \tag{20.9}$$

where Var_j denotes the variance based on the restriction of the probability distribution of $\mathbf{Z}$ to the j^{th} stratum. Without calibration the terms $\mathrm{Var}_j[\widetilde{\ell}(\mathbf{Z};\boldsymbol{\beta})]$ are replaced by $\mathrm{E}_j\left[\widetilde{\ell}(\mathbf{Z};\boldsymbol{\beta})^{\otimes 2}\right]$ where E_j denotes the corresponding expectation. Hence the gain from calibration would be greatest if the inluence function contributions varied strongly by strata. Note that (20.9) is the limit of (20.7). Hence calibration allows one to reconcile the apparent differences between the FPSS and Bernoulli sampling schemes. After fitting the model assuming Bernoulli sampling with *a priori* weights, one could calibrate to the stratum frequencies to obtain correct standard errors under the more realistic FPSS sampling scheme.

In actual practice, as illustrated in the next section, one would start with design weights for FPSS and calibrate using variables $\mathbf{C}(\mathbf{V})$ thought to be highly correlated with $\widetilde{\ell}(\mathbf{Z};\boldsymbol{\beta})$ to increase the precision of the standard survey estimator described in 20.4.2. Consideration of the related Bernoulli sampling scheme and (20.8) serves primarily to understand why $\mathbf{C}(\mathbf{V}) = \mathrm{E}(\widetilde{\ell}|\mathbf{V})$ might be a good choice. See, however, Saegusa and Wellner (2013).

20.4.4 Estimation of the cumulative hazard

The survey estimator of the baseline cumulative hazard function is the IPW version of the Breslow estimator

$$\widehat{\Lambda}_N(t;\boldsymbol{\beta}) \;=\; \frac{1}{N}\sum_{i=1}^{N} \frac{\frac{R_i}{\pi_0(\mathbf{V}_i)}\Delta_i \mathbf{1}[T_i \le t]}{\sum_{j=1}^{N} \frac{R_j}{\pi_0(\mathbf{V}_j)} e^{\mathbf{X}_j^{\mathrm{T}}\boldsymbol{\beta}} Y_j(T_i)}, \tag{20.10}$$

which depends on the log hazard ratios $\boldsymbol{\beta}$.(Lin, 2000; Breslow et al., 2015) This is combined with $\widehat{\boldsymbol{\beta}}_N$ to estimate the cumulative hazard over the interval between times t_0 and t_1 for a subject with covariates $\mathbf{x}_0$, assuming the subject is "at risk" for the entire interval. The estimate is

$$e^{\mathbf{x}_0^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_N}\left(\widehat{\Lambda}_N(t_1;\widehat{\boldsymbol{\beta}}_N) - \widehat{\Lambda}_N(t_0;\widehat{\boldsymbol{\beta}}_N)\right). \tag{20.11}$$

Therneau and Grambsch (2000, §10.2) discussed how the cumulative hazard may be interpreted as an estimate of "expected" survival over the interval.

20.4.5 Model misspecification

Most of the reults stated in this section remain valid even if the Cox model does not strictly hold. As usual under general misspecification, the parameters $(\boldsymbol{\beta}, \Lambda_0)$ being estimated are those for the Cox model that is closest, in the sense of Kullback-Leibler distance, to the probability distribution generating the data. The only change is that the limiting Phase I variance, *i.e.*, the asymptotic variance of $\sqrt{N}(\widehat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$, rather than being the inverse of the efficient information $\widetilde{\mathbf{I}}$ (20.4), is instead the "robust sandwich" version $\widetilde{\mathbf{I}}^{-1}\mathrm{E}\left[\widetilde{U}^*\left(\widetilde{U}^*\right)^{\mathrm{T}}\right]\widetilde{\mathbf{I}}^{-1}$, where $\widetilde{U}^*$ denotes the efficient score (20.3) for a generic subject. Expressions derived under the model for the asymptotic expansion of the IPW Breslow estimator (20.10), which allow calculation of its asymptotic variance, continue to hold off the model.(Breslow et al., 2015, §6.1)

20.4.6 Time-dependent covariates

The expressions shown in this section for scores, efficient scores, efficient information *etc.* are easily extended to accomodate time-dependent covariates, replacing $\mathbf{X}$ by $\mathbf{X}(t)$ or $\mathbf{X}(T)$ as the case may be. Interpretation of the results of analyses of such time-dependent data, however, is another matter. It is relatively straightforward when the covariates are "external" in the sense of Kalbfleisch and Prentice (2002, §6.3); results may be highly subject to misinterpretation otherwise.

In practice, time-dependent covariates are most easily handled by splitting the follow-up record for each subject at the times at which the covariate values change and using the "counting process" form of the Cox model.(Therneau and Grambsch, 2000, §§3.7, 5.6) Contributions to the scores and information from each of the split records for a given subject are aggregated into a single contribution per subject before being inserted into the expressions shown here. An example of such a time-dependent analysis is presented in the next section.

20.5 Cox model analysis of radiation and breast cancer

Here we apply the methods outlined in the last section to the analysis of the data on radiation and cancer described in Section 20.2. The data and computer code needed to perform these analyses are available from the Handbook website. The code is written for the freely available statistical programming language R. It includes an updated version of the cch command found in the R **survival** package. This suffices for estimation of hazard ratios using data from simple and stratified case-cohort studies, where cases are sampled at 100%. For estimation of cumulative hazards we used commands from Lumley's (2012) **survey** package, which applies more generally to two-phase study designs where not all cases need be sampled.

20.5.1 Preliminary graphical analyses

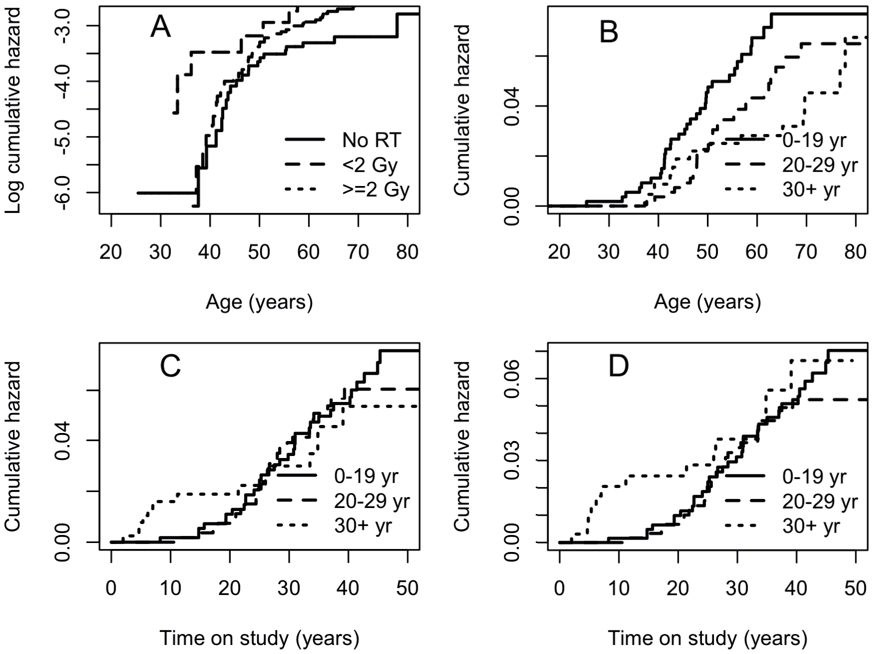

FIGURE 20.1
Cumulative hazards by radiation dose and age at at first treatment. See text for interpretation of each panel.

To provide some insight into the basic structure of the data, we first estimated in Figure 20.1, panels A-C, cumulative hazards from the full cohort data using the survfit command in the **survival** package. Risk categories for radiation dose and age at first treatment correspond to those in Table 20.1. The estimates in panel C, with time-on-study rather than age as the basic time variable, were also estimated in panel D from the simple case-cohort data using the svykm command in the **survey** package. The two show good agreement. As of this writing, svykm was not implemented for "counting process" survival data as needed to handle the left truncation of age at entry into the cohort (panels A-B). In principle, however, these cumulative hazards are also estimable from two-phase study data.

The cumulative hazards in panel A were plotted on the log scale. If the hazards for the three dose groups were exactly proportional, one would therefore expect to see a constant vertical distance between the three curves. While the middle curve for the 0-2 Gy group did not fall precisely between the other two, overall the assumption of proportionality seems reasonable. This is not so clear for age at first treatment (panel B), where the plot of the cumulative hazard itself offered a slightly different perspective. While the curves for the first two age at treatment categories (0-19 and 20-29 years) appeared proportional, that for the last category (30+ years) was not proportional to either of the others. This was due to a handfull of BC cases that occurred within 20 years among women who started tuberculosis treatments after 30 years of age. For women irradiated earlier in life, time since first radiation largely determined the evolution of BC risk during the next decade or two. For women irradiated later, however, additional age-dependent risk factors may have started to exert their influence. Radiation treatment may have accelerated the appearance of cancers that would have occurred later anyway. This idea is reinforced by the plots in panel C, where the cumulative hazards according to time since start of radiation overlapped for those who were treated early, but started to rise sooner for those first treated at a later age. These remarks are of course highly speculative, given the relatively few BC events observed, especially for women first treated after 30 years of age. With a larger sample, it is possible that the three curves would overlap when plotted against time on study.

In spite of the questionable proportionality for age at first treatment, the analyses that follow assume that the proportional hazards model holds for underlying *continuous* versions of these two risk factors.

20.5.2 Model for radiation and breast cancer

In view of radiobiological evidence of low dose linearity of radiation effects, radiation epidemiologists often choose the excess relative risk version of the proportional hazards model, where $\mathrm{ERR}(\mathbf{X};\boldsymbol{\beta}) = (1 + \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta})$ rather than the standard exponential term $\mathrm{RR}(\mathbf{X};\boldsymbol{\beta}) = \exp(\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta})$ multiplies the baseline cumulative hazard. Due to limitations in software available for two-phase studies, however, and to be able to illustrate the standard Cox model, we instead followed Borgan and Samuelson (2013) and used a log transform of radiation dose with the standard model. The covariates were $\mathbf{x} = (x_1, x_2)$ with $x_1 = \log_2(\text{dose} + 1)$ and $x_2 = \text{ageRx}/10$, where ageRx denotes age at first treatment. This implies the age-specific BC risk was multiplied by a *power* of (dose+1). Division by 10 ensured that the regression coefficient for ageRx represented the change in log hazard associated with a decade long rather than yearly change. It also had the effect of equalizing the magnitudes of the regression coefficients for the two covariates. The model fitted to the data for the cumulative hazard of a subject with t = attained age and covariates $\mathbf{x}$ was thus

$$\Lambda(t, \mathbf{x}; \boldsymbol{\beta}, \Lambda_0) \ = \ \exp(x_1\beta_1 + x_2\beta_2)\Lambda_0(t), \qquad (20.12)$$

with $\Lambda_0(t)$ the baseline ($\mathbf{x} = 0$) cumulative hazard at age t. We fit model (20.12) to three different data sets as described in Section 20.2.1: the entire set of cohort data, the simple case-cohort sample and the stratified (on ageRx) case-cohort sample. The case-cohort analyses, however, failed to take advantage of all the information available for subjects in the main cohort. Hence we considered the total number of fluoroscopies (nof) per subject, which was used to reconstruct the radiation dose, to be an auxiliary variable known for all Phase I subjects that served as a surrogate for dose. There was a very high correlation of 0.98 between log transforms of the two measurements (Figure 20.2, panel A) and the availability of such a strong surrogate exposure would be rare in actual practice. Nonetheless, its use here illustrated the principle that efficiency gains are possible by using a good surrogate exposure to calibrate the weights.

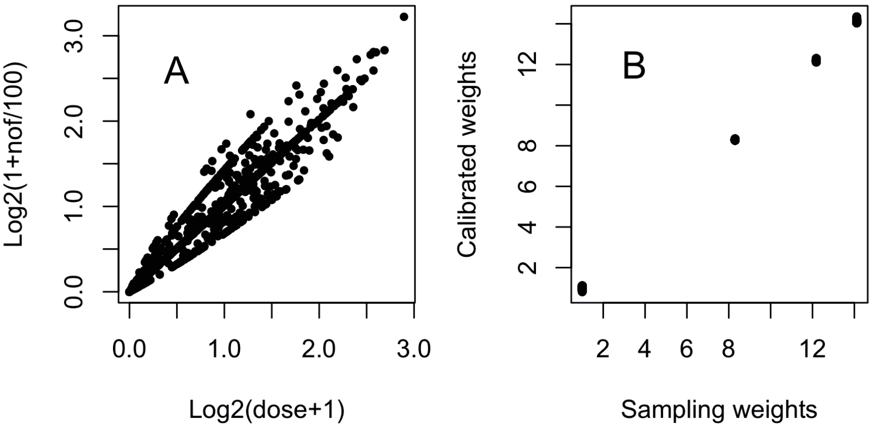

FIGURE 20.2
Results of calibration: A. Calibration variable by dose variable. B. Calibrated weights by sampling weights

20.5.3 Calibration and post-stratification

As described in Section 20.4.3, the idea behind the calibration method we employed was to modify the IPW design weights so that the Phase I totals of the calibration variable(s) $\mathbf{C} = \mathbf{C}(\mathbf{V})$ were exactly estimated by the IPW estimator, keeping the two sets of weights as close together as possible according to the Poisson deviance distance measure. The "insights" developed in Section 20.4.2 suggested that the optimal calibration variable was the conditional expectation of the influence function given all the main cohort variables $\mathbf{V}$. Since this was unknown, we needed to approximate it. Given the availability of a strong surrogate, one way to do this was simply to fit model (20.12) to the Phase I data, substituting $x_1 = \log_2(1+\mathrm{nof}/100)$ for $x_1 = \log_2(\mathrm{dose}+1)$. Estimates of the influence function contributions are extracted with the dfbeta option for the residual function applied to the model fitt resulting from the call to coxph in the **survival** package. There are two components to vector of influence function contributions, one for each covariate, and these are used as the calibration variables. Calibration resulted in only minor changes to the weights (Figure 20.2, panel B), but substantial gains in efficiency (see next section).

Gains from calibration depend heavily on the correlations between the calibration variables and those included in the model. They will be nil for biomarkers like genotype, known only at Phase II, that are not at all related to the Phase I variables. When the model contains adjustment variables, such as baseline risk factors known for all, calibration to the dfbeta's for these variables can increase substantially the precision of their regression coefficients. It may also have a modest effect on the precision of estimation of interaction effects between exposures known only at Phase II and adjustment variables known for everyone at Phase I.

If a strong exposure surrogate is not available, one may first fit an imputation model using IPW to the Phase II data.(Kulich and Lin, 2004) Regression coefficients from the imputation model are then used to predict exposures for Phase I subjects who lack them, *i.e.*, for those not included in the Phase II sample. The exposures for all Phase I subjects, whether observed or imputed, are next used together with the adjustment variables to fit a calibration model from which the dfbeta's are extracted. Finally, the **dfbetas** are used as the calibration variables to adjust the design weights in an IPW analysis of the two-phase data. In other circumstances, it may be more efficacious simply to try to bring additional Phase I information to bear on the analysis by calibrating to variables, such as frequencies in a cross-classification, that summarize the marginal association between auxiliary Phase I variables and outcomes. See Breslow et al. (2009); Breslow and Lumley (2009); Breslow et al. (2013) for examples of each of these approaches.

Post stratification is an alternative to calibration for bringing into the analysis additional Phase I information. Rather than restricting the partition of the Phase I sample to the strata actually used for sampling, a finer partition based on additional (necessarily discrete) Phase I variables is used for the analysis. For example, the partition we chose based on categories of age at first treatment, which was used to stratify the sample, could have been based instead on the cross-classification of age at first treatment by number of fluoroscopies, *e.g.*, in the three categories (i) none, (ii) 1-149, and (iii) 150+ considered by Borgan and Samuelson (2013). For a non-stratified Phase II design, post-stratification to a set of stratum frequencies is equivalent to calibration using those same frequencies and avoids having to solve the calibration equations, an advantage since the need for specialized software may be avoided.

20.5.4 Results of the analyses: regression coefficients

TABLE 20.3
Comparison of results for four designs. Robust standard errors

Model term	Coef.	SE1	SE2	SE	Z	p
			A. Entire cohort			
$\log_2$(dose + 1)	0.469	0.153	NA	0.153	3.057	0.0022
ageRx/10	-0.242	0.144	NA	0.144	-1.677	0.0936
			B. Simple case-cohort sample			
$\log_2$(dose + 1)	0.519	0.146	0.142	0.204	2.541	0.0111
ageRx/10	-0.111	0.159	0.106	0.191	-0.579	0.5629
			C. Stratified case-cohort sample			
$\log_2$(dose+1)	0.559	0.185	0.162	0.246	2.273	0.0230
ageRx/10	-0.244	0.143	0.069	0.159	-1.537	0.1244
			D. Stratified case-cohort sample with calibrated weights			
$\log_2$(dose+1)	0.440	0.182	0.0311	0.184	2.387	0.0170
ageRx/10	-0.250	0.140	0.0082	0.140	-1.781	0.0749

Regression coefficients and standard errors estimated by each model fit are shown in Table 20.3. Here Z denotes the equivalent normal deviate used by the Wald test, on which p is based. The column labeled SE1 contains the robust estimate of the Phase I standard error, that due to sampling of the cohort data from the population, as described above in Sections 20.4.2 and 20.4.5. Entries in column SE2 are the estimated Phase II standard errors, found by taking the square root in equation (20.7), and those in column SE the total standard errors. Note that $\text{SE1}^2 + \text{SE2}^2 = \text{SE}^2$.

Results for the entire cohort (Table 20.3, Part A) were obtained using survfit in the **survival** package, something made possible because radiation dose was in fact known for all Phase I subjects. They suggest that the age-specific risk of BC increased by a factor of exp(0.469)=1.60 with each doubling of (dose+1), whereas the risk decreased by a factor of exp(-0.242)=0.79 with each ten year increase in age at first treatment. The latter result, however, was subject to substantial sampling error. Considerably greater uncertainty, even in the coefficient for dose, was apparent when fitting the model to the simple case-cohort sample using the ChenLo option with the cch function (Table 20.3, Part B). The uncertainty in the coefficient for dose due to sampling from the cohort (SE2=0.142) was nearly as large as that due to sampling the cohort from the population (SE1=0.146). Even for age at first treatment the two standard errors were of comparable magnitude. A substantial reduction in the Phase II standard error for ageRx was achieved by using the II.Borgan option for cch with data from the stratified case-cohort design (Table 20.3, Part C). The main difference between the two designs was that the relatively small number of controls (29) in the highest ageRx category (30+ years) that resulted from simple random sampling was increased to 50 by stratification on ageRx, with concomitant reductions in controls for the first two categories. Only 14 BC cases were observed for ageRx of 30 years or more. The additional controls in this category increased the information regarding risk at high ageRx and this lowered the standard error of the regression coefficient.

The most dramatic changes in Phase II standard errors followed calibration of the stratified sampling weights (see Section 20.5.3) using the estimated influence function contributions, which had components both for the surrogate dose measurement $\log_2(1+\text{nof}/100)$ and for ageRx (Table 20.3, Part D). Phase II contributions to the total SE were negligible in comparison to those for Phase I and results obtained from the calibrated design closely paralleled those obtained using the entire cohort.

One surprising feature of Table 20.3 was the variability in SE1 for $\log_2(\text{dose}+1)$ depending on the method of analysis, with values ranging from 0.146 to 0.185. All were estimating the same Phase I standard error. Since robust standard errors are notoriously variable, we reanalyzed the data using the options for model based standard errors and found SE1 for $\log_2(\text{dose}+1)$ of, respectively, 0.161, 0.157 and 0.178 for the models fitted to the entire cohort, to the simple case-cohort sample and to the stratified case-cohort sample. (The **survey** package provides only robust standard errors.) While the model based SE1 were indeed closer together, the high value of 0.178 for the stratified case-cohort sample is a reminder of the variability to be expected with relatively small Phase II samples.

20.5.5 Results of the analyses: expected BC risk

Figure 20.3 shows the expected numbers of BC between 40 and 60 years of age estimated for four configurations of covariate values. Two subjects received no radiation and two a dose of 3 Gy ($\log_2(\text{dose}+1)=2$) while at each radiation dose one started tuberculosis treatments (ageRx) at 15 years of age and one at age 35. The expected numbers were estimated from Equation (20.11) with $(t_0, t_1) = (40, 60)$ and $\mathbf{x}_0$, respectively, (0, 1.5), (0, 3.5), (2, 1.5) and (2, 3.5). The calculations were implemented using the predict command with type option expected, which is available in both the **survival** (for the full cohort analysis) and **survey** (for the case-cohort analyses) packages. Hence they represent cumulative hazards for these four covariate configurations, assuming that the women were at risk for the entire interval between 40 and 60 years of age. The four designs, designated Cohort, Case-cohort, Stratified and Calibrated, correspond to those described in detail in Section 20.5.4. The cor-

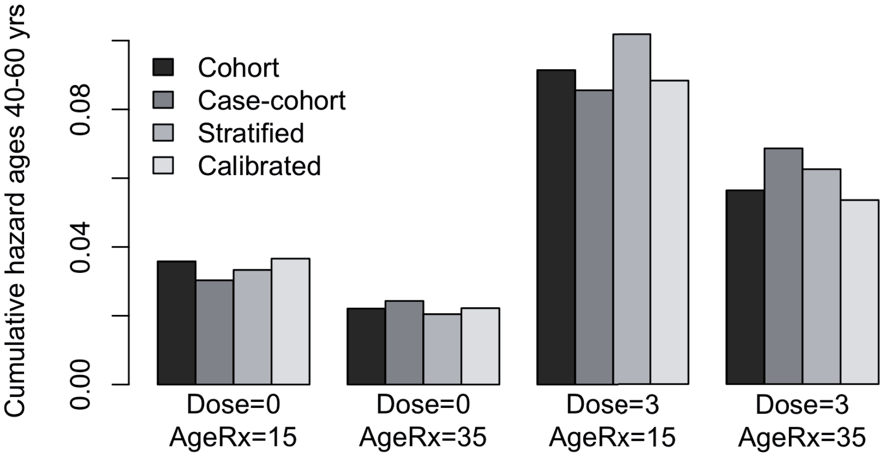

FIGURE 20.3
Expected numbers of breast cancers (cumulative hazards) between 40 and 60 years of age depending on covariate values.

responding regression coefficients and standard errors are as shown in Table 20.3, Parts A-D.

Using data from the full cohort, the expected number of BC cases between ages 40 and 60 for a woman who started treatment for tuberculosis at age 15 but received no radiation was 3.6%. This declined to 2.2% for a non-irradiated woman who started tuberculosis treatment at age 35. The estimated risks were considerably higher (9.1% and 5.6%, respectively) at a radiation dose of 3 Gy. Taking the full cohort estimates as the "gold standard", overall those for the case-cohort analyses were as anticipated: worst for the simple case-cohort design, better for the stratified case-cohort design and best for the stratified design with calibration of the weights (Figure 20.3). In fact, the estimated risks using calibration were quite close to those estimated from the full cohort data, at 3.7%, 2.2%, 8.8% and 5.4% for the four covariate configurations.

20.5.6 A time-dependent covariate

20.6 The additive hazards model

20.7 Concluding remarks

Bibliography

Anderson, G. L., J. Manson, R. Wallace, B. Lund, D. Hall, S. Davis, S. Shumaker, C.-Y. Wang, E. Stein, and R. L. Prentice (2003). Implementation of the Women's Health Initiative study design. *Annals of Epidemiology 13*(9), S5–S17.

Binder, D. A. (1992). Fitting Cox's proportional hazards model from survey data. *Biometrika 79*(1), 139–147.

Borgan, Ø., B. Langholz, S. O. Samuelsen, L. Goldstein, and J. Pogoda (2000). Exposure stratified case-cohort designs. *Lifetime Data Analysis 6*(1), 39–58.

Borgan, Ø. and S. O. Samuelson (2013). Nested case-control and case-cohort studies. In J. P. Klein, van Houwelingen H. C., J. G. Ibrahim, and T. H. Scheike (Eds.), *Handbook of Survival Analysis*, pp. 343–368. Boca Raton: Chapman and Hall/CRC.

Breslow, N. E., G. Amorim, M. B. Pettinger, and J. Rossouw (2013). Using the whole cohort in the analysis of case-control data. Application to the Women's Health Initiative. *Statistics in Biosciences* (2), 232–249.

Breslow, N. E., J. Hu, and J. A. Wellner (2015). Z-estimation and stratified samples. Application to survival models. *Lifetime Data Analysis 21*, 493–516.

Breslow, N. E. and T. Lumley (2009). Improved Horvitz-Thompson estimation of model parameters from two-phase stratified samples: Applications in epidemiology. *Statistics in Biosciences 1*, 32–49.

Breslow, N. E., T. Lumley, C. M. Ballantyne, L. E. Chambless, and M. Kulich (2009). Using the whole cohort in the analysis of case-cohort data. *American Journal of Epidemiology 169*(11), 1398–1405.

Breslow, N. E. and J. A. Wellner (2007). Weighted likelihood for semiparametric models and two-phase stratified samples, with application to Cox regression. *Scandinavian Journal of Statistics 34*(1), 86–102.

Chen, K. N. and S. H. Lo (1999). Case-cohort and case-control analysis with Coxs model. *Biometrika 86*, 755–764.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society (Series B) 34*(2), 187–220.

Deville, J. C. and C. E. Särndal (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association 87*(418), 376–382.

Hrubec, Z., J. D. Boice, R. R. Monson, and M. Rosenstein (1989). Breast-cancer after multiple chest fluoroscopies - 2nd follow-up of Massachusetts women with tuberculosis. *Cancer Research 49*(1), 229–234.

Kalbfleisch, J. D. and R. L. Prentice (2002). *The Statistical Analysis of Failure Time Data* (Second ed.). Hoboken, NJ: Wiley.

Kulich, M. and D. Y. Lin (2004). Improving the efficiency of relative-risk estimation in case-cohort studies. *Journal of the American Statistical Association 99*(467), 832–844.

Lin, D. Y. (2000). On fitting Cox's proportional hazards model to survey data. *Biometrika 87*(1), 37–47.

Lin, D. Y. and Z. Ying (1994). Semiparametric analysis of the additive risk model. *Biometrika 81*(1), 61–71.

Lumley, T. (2012). *Complex Surveys: A Guide to Analysis Using R.* Wiley Series in Survey Methodology. Hoboken, N.J.: John Wiley & Sons, Inc.

Lumley, T., P. A. Shaw, and J. Y. Dai (2011). Connections between survey calibration estimators and semiparametric models for incomplete data. *International Statistical Review 79*(2), 200–220.

Prentice, R. L. (1986). A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika 73*(1), 1–11.

Preston, D. L., J. H. Lubin, D. A. Pierce, and M. D. McConney (1993). *Epicure Users Guide.* Seattle, WA: Hirosoft International Corporation.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1994). Estimation of regression-coefficients when some regressors are not always observed. *Journal of the American Statistical Association 89*(427), 846–866.

Saegusa, T. and J. A. Wellner (2013). Weighted likelihood estimation under two-phase sampling. *Annals of Statistics 41*(1), 269–295.

Therneau, T. M. and P. M. Grambsch (2000). *Modeling Survival Data: Extending the Cox model.* New York: Springer.

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge, UK: Cambridge University Press.

van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes with Applications in Statistics.* New York: Springer.

Williams, O. D. (1989). The Atherosclerosis Risk in Communities (ARIC) study - design and objectives. *American Journal of Epidemiology 129*, 687–702.