Analysis of Interval-Censored Failure Time Data with Application to Studies of HIV Infection

Thesis submitted to University College London
for the degree of

DOCTOR OF PHILOSOPHY

by

Ruth Louise Goodall

MRC Clinical Trials Unit,
Department of Primary Care & Population Sciences,
Royal Free & University College Hospital

May 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Ruth Louise Goodall, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

I would like to thank to following:

Abdel Babiker and David Dunn for their time and support.

The Delta and CHIPS Steering Committees for allowing me to use their data.

My family and friends for their love, encouragement and understanding - particularly those who listened, tactfully stopped asking me when I was going to finish or shared a house with me during this time.

Abstract

In clinical trials and cohort studies the event of interest is often not observable, and is known only to have occurred between the visit when the event was first observed and the previous visit; such data are called interval-censored. This thesis develops three pieces of research that build upon published methods for interval-censored data. Novel methods are developed which can be applied via self-written macros in the standard packages, with the aim of increasing the use of appropriate methods in applied medical research.

The non-parametric maximum likelihood estimator [1,2] (NPMLE) is the most common statistical method for estimating of the survivor function for interval-censored data. However, the choice of method for obtaining confidence intervals for the survivor function is unclear. Three methods are assessed and compared using simulated data and data from the MRC Delta trial [3].

Non- or semi- parametric methods that correctly account for interval-censoring are not readily available in statistical packages. Typically the event time is taken to be the right endpoint of the censoring interval and standard methods (e.g. Kaplan-Meier) for the analysis of right-censored failure time data are used, giving biased estimates of the survival curve. A simulation study compared simple imputation using the right endpoint and interval midpoint to the NPMLE and a proposed smoothed version of the NPMLE that extends the work of Pan and Chappell [4]. These methods were also applied to data from the CHIPS study [5].

Different approaches to the estimation of a binary covariate are compared: (i) a proportional hazards model [6], (ii) a piecewise exponential model [7], (iii) a simpler proportional hazards model based on imputed event times, and (iv) a proposed approximation to the piecewise exponential model that is a more rigorous alternative to simple imputation methods whilst simple to fit using standard software.

References

1. Peto R. Experimental survival curves for interval-censored data. *Applied Statistics* 1973, **22**:86-91.

2. Turnbull BW. The empirical distribution function with arbitrary grouped, censored and truncated data. *Journal of the Royal Statistical Society B* 1976, **38**:290-295.

3. Delta Coordinating Committee. Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. *Lancet* 1996, **348**:283-291.

4. Pan W, Chappell R. Estimating survival curves with left truncated and interval censored data via the EMS algorithm. *Communications in Statistics-Theory and Methods* 1998, **27**(4):777-793.

5. Gibb DM, Duong T, Tookey PA, Sharland M, Tudor-Williams G, Novelli V, on behalf of the National Study of HIV in Pregnancy and Childhood (NSHPC) and of the Collaborative HIV Paediatric Study (CHIPS). Decline in mortality, AIDS and hospital admissions in perinatally HIV-1 infected children in the UK and Ireland. *BMJ* 2003, **327**:1019-1025.

6. Finkelstein DM. A proportional hazards model for interval-censored failure time data. *Biometrics* 1986, **42**(4):845-854.

7. Lindsey JC, Ryan LM. Tutorial in biostatistics - Methods for interval-censored data. *Statistics in Medicine* 1998, **17**(2):219-238.

List of Contents

List of Figures

Chapter 4

List of Tables

CHAPTER 1: INTRODUCTION

The motivation for the work in this thesis comes from the analysis of follow-up studies of HIV infection. After defining interval-censored data in section 1.1, a brief description of the background to the problem is given in section 1.2. An overview of methods for the analysis of interval-censored data from the published literature is given in section 1.3. Finally, the structure of the thesis in subsequent chapters is outlined in section 1.4

1.1 INTERVAL-CENSORED DATA

In clinical trials and observational follow-up studies individuals are followed over time for the occurrence of a specific event. When either the precise time of the event is known or individuals do not experience the event during follow-up (right censoring), standard survival analysis methods are applied [1]. The survival data are often summarised using the survivor or hazard functions. The survivor function at time t is defined as the probability that the time to the event (survival time) is greater than t. The hazard function is the instantaneous risk of death at time t, conditional on having survived to time t.

Often, however, the precise time of onset of the event is unobservable, e.g. the recurrence of cancer for a patient in remission or an increase in a biological marker to above a certain level. In this instance the event is known only to have occurred within the interval defined by the time the event was first observed and the previous event free visit, the *censoring interval.* Data of this type are called *interval-censored.*

Individuals may have been either monitored at irregular times over the study period or observed at pre-specified times as defined by the study protocol. In the latter case many ties will exist within the data, which are considered to be *grouped.*

1.2 HIV INFECTION

Human immunodeficiency virus (HIV) is a retrovirus that infects cells in the human immune system, primarily CD4+ T lymphocytes. Infected cells are depleted through three main mechanisms: direct viral killing, increased rates of apoptosis, and destruction

by CD8 cytotoxic lymphocytes. When CD4+ T cell numbers decline below a critical level, cell-mediated immunity is lost, and the individual becomes progressively more likely to develop opportunistic infections and other pathologies, collectively referred to as Acquired Immune Deficiency Syndrome (AIDS) [2]. Without potent treatment, the median survival after HIV-1 seroconversion has been estimated to be 12.5 years for people aged 15-24 years and 7.9 years for those aged 45-54 years [2].

Infection with HIV occurs by the transfer of blood, semen, vaginal fluid, or breast milk. Within these body fluids HIV is present as both free virus particles and virus within infected immune cells. The three major routes of transmission are unprotected sexual intercourse, contaminated needles and transmission from an infected mother to her baby at birth or through breast milk. Screening of blood products for HIV in the developed world has largely eliminated transmission through blood transfusions or infected blood products in these countries.

HIV infection in humans is now pandemic. An estimated 38.6 million (33.4-46.0 million) people worldwide were living with HIV at the end of 2005, an estimated 4.1 million (3.4-6.2 million) became newly infected with HIV and an estimated 2.8 million (2.4-3.3 million) died from AIDS during the year [3]. Antiretroviral treatment reduces both the mortality and the morbidity of HIV infection, but routine access to antiretroviral medication is not available in all countries.

Current treatment for HIV infection consists of highly active antiretroviral therapy (HAART) which first became available in 1996. At the same time, assays for quantification of the virus (measured by the number of HIV RNA copies in the plasma or serum) were developed and used to monitor the response to treatment. Current HAART options are combinations consisting of at least three drugs belonging to at least two classes of anti-retroviral agents. The three main classes are nucleoside analogue reverse transcriptase inhibitors (NRTIs), protease inhibitor (PIs), and non-nucleoside reverse transcriptase inhibitor (NNRTIs). With currently available antiretroviral agents, eradication of HIV infection is not likely to be possible. The main aim of treatment is thus to prolong life and improve quality of life by maintaining suppression of virus replication for as long as possible. In adults, treatment is usually started when the disease becomes symptomatic or the risk of clinical progression is high [4,5]. Because AIDS progression in infants is more rapid and less predictable than in adults guidelines

generally recommend treatment in children less than a year old. Decisions about treatment in older children are usually based on immune function [6].

The first clinical trials of antiretroviral drugs for the treatment of HIV assessed efficacy on the clinical endpoints of AIDS or death [7,8]. The advent of HAART, together with improved prophylaxis and diagnosis of opportunistic infections, has led to increased AIDS-free and overall survival (Figure 1.1) [9]. Consequently clinical endpoint studies are now very difficult to conduct with larger and longer trials required to detect an improvement in clinical outcomes. In a dynamic area such as HIV where new therapies and new data from small studies have a major influence on patient management this can be problematic, and is compounded by issues of cost, retaining trial participants throughout follow-up, and adherence to allocated treatment strategies. Thus, endpoints based on biological markers are frequently used as surrogates for clinical outcome [10,11].

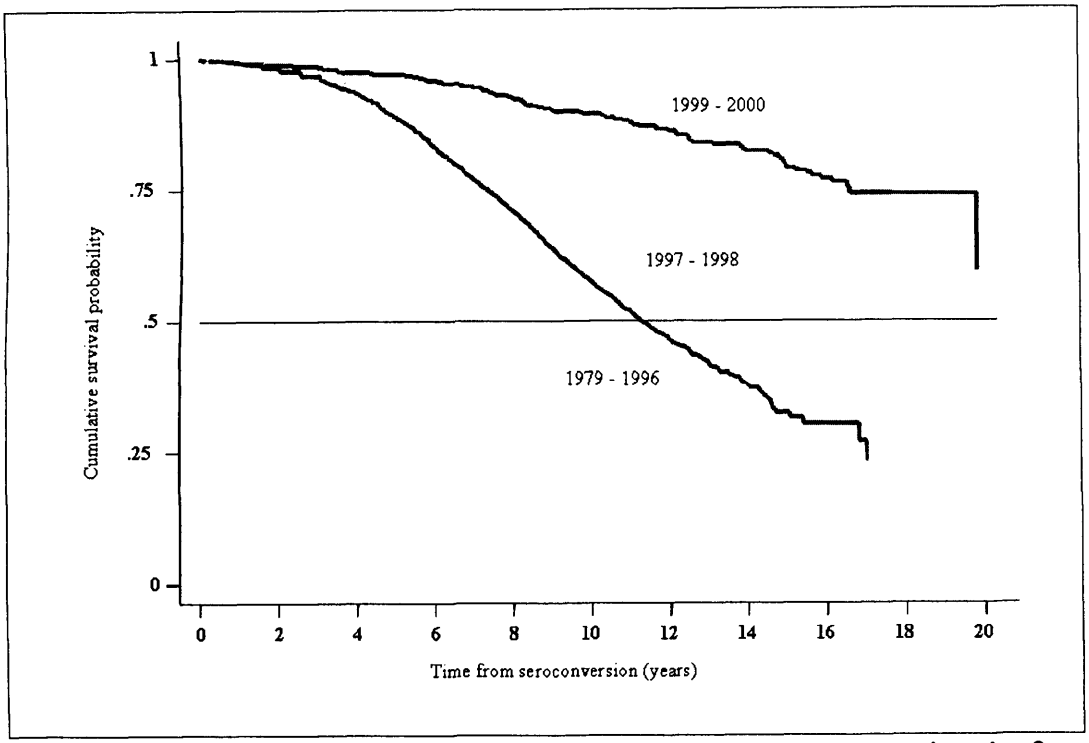

Figure 1.1 Estimated survival from time of HIV-1 seroconversion in 3 calendar periods [9]

Both CD4 cell count and HIV RNA have been shown to be strong prognostic markers for clinical progression but neither is an ideal surrogate marker [12-14]. Nonetheless, regulatory agencies license new drugs on the basis of these markers [15,16]. HIV RNA

provides the more direct measure of the potency of antiretroviral drugs. Virological endpoints include quantitative change in HIV RNA levels, time to undetectable viraemia, the proportion of subjects with undetectable viraemia at a fixed time point, and the durability of virological control. For time to undetectable viraemia, each individual's event time is censored by the interval defined by the time of the first sample where virus concentration was below the detection limit of the assay and the previous sample (where it was not).

1.3 OVERVIEW OF ANALYTICAL APPROACHES TO INTERVAL-CENSORED DATA

This section gives a brief overview of different analytical approaches to interval-censored data that have been proposed. A fuller description and a mathematical exposition of selected methods are given at relevant points in subsequent chapters.

1.3.1 Imputation

Interval-censored data in HIV infection are typically analysed by simply taking the time the event is observed (the right endpoint of the censoring interval) as the time of the actual event. An alternative and potentially less biased approach is to use the midpoint of the observed censoring interval [17]. It is then possible to apply standard methods available for the analysis of exact and right-censored observations in statistical software packages. For example, estimates of the survivor function using the Kaplan-Meier estimator [18], or the effect of covariates upon the hazard using a Cox proportional hazards model [19] are then easily obtainable. Imputation may be a reasonable approach when the width of the censoring interval is small relative to the total follow-up. However, simple imputation leads to underestimation of standard errors, which can be overcome by using multiple imputation [20].

1.3.2 Interval-censoring methods

There is an expanding statistical literature on methods for the analysis of interval-censored data [21-23]. These can be broadly classified into parametric methods requiring specification of a family of survivor functions, semi-parametric methods requiring specification of a model for the effect of covariates only, and assumption-free non-parametric methods.

a) Parametric methods

Most standard statistical software packages can accommodate interval-censored data in routines for parametric models. Also, the estimation of a single survivor function can be easily extended to model the effect of covariates. However, these advantages must be balanced against a potentially severe bias if an inappropriate family of distributions is chosen [24].

b) Non-parametric methods

Unlike parametric methods, only limited routines are provided within standard packages for non-parametric or semi-parametric analysis of interval-censored data. The most intensively studied non-parametric method was first proposed by Peto [25] and Turnbull [26]. This non-parametric maximum likelihood estimator (NPMLE), obtained by an iterative procedure such as the EM algorithm [27], allocates probability mass within intervals defined by the observed data. The resulting step function often has large regions where the survival curve is flat or undefined as shown by examples in later chapters, and several authors have proposed techniques that achieve a smoother estimate to overcome this problem.

Pan & Chappell propose a Expectation Maximisation Smoothing (EMS) algorithm, which introduces a smoothing step to the Expectation Maximisation (EM) algorithm described [28]. Other approaches include kernel smoothing the probability density function [29], modelling the log density [30] or the hazard function [31] using regression splines. Alternatively, Bebchuk et al. describe a multiple imputation approach [32], and Betensky et al. discuss a procedure based on local likelihood [33]. Several of these methods can be extended to allow estimation of covariate effects within a proportional hazards framework [34-37].

c) Semi-parametric methods

The piecewise exponential model assumes a constant hazard within predefined time intervals [38], and can be considered a special case of the method proposed by Rosenberg [31,34]. Carsenten [39], Farringdon [40] and Smith et al. [41] present a generalised linear model approach to fitting a piecewise exponential model, whereas Lindsey & Ryan [21] use an EM algorithm.

The Cox proportional hazards model [19] is commonly used in survival analysis as it allows analysis of covariate effects without specification of the baseline hazard function. However, construction of the partial likelihood is precluded when the data are interval-censored, since the ordered rankings of failure and censoring times cannot be determined. In an early approach by Finkelstein [42], a full likelihood proportional hazards model was fitted using a parameterisation of the log cumulative hazard where covariates effects and nuisance parameters associated with the baseline hazard were estimated simultaneously. This approach is similar to the piecewise exponential method, except that the intervals are determined by the data rather than pre-specified. Satten has proposed two rank-based methods, one which estimates the covariate effects without specification of the baseline hazard by maximising the marginal likelihood using Gibb's sampling [43], the other imputes failure times by assuming a parametric model for the baseline hazard which are used to determine ranks [44].

Other regressions models, including the accelerated failure time model [23,45,46] have been proposed for interval-censored data. In addition, some methods have been extended to truncated and interval-censored data or doubly censored data, where both the time origin and the event time are interval-censored. Methods have also been proposed to allow for informative examination times or informative dropout. These types of data are not considered further.

1.4 SCOPE OF THESIS

This thesis develops three pieces of work relating to the analysis of interval-censored data. Within each chapter, a simulation study is used to examine the effect of different factors, e.g. sample size and degree of censoring on the relative performance of the different methods under study. Data from the MRC Delta trial [8] and the Collaborative HIV Paediatric Study (CHIPS) [47] are used to exemplify these methods.

Chapter 2 examines three different methods for obtaining (pointwise) confidence intervals for the NPMLE proposed by Peto [25] and Turnbull [26]. The method proposed by Peto and Turnbull in their original papers is based on the inversion of a submatrix of the information matrix, without clear justification. This method is assessed and compared with inverting the full information matrix, and a profile likelihood approach that has not previously been considered for interval-censoring.

As noted above, except in the case of parametric models, methods that correctly account for interval-censoring are not readily available in statistical packages, and the use of right-endpoint, and to a lesser extent, mid-point imputation are commonly employed. Chapter 3 examines the accuracy of these simple methods in the context of estimating the survivor function of a single sample. Results are compared with the NPMLE and a smoothed version of the NPMLE based on the work of Pan & Chappell [28]. A modification to the smoothed estimator is proposed and assessed.

This work was extended in chapter 4 to consider the comparison of two samples, specifically the estimation of the hazard ratio within a proportional hazards model. Simple imputation methods are compared to the proportional hazard model for interval-censored data proposed by Finkelstein [42], a piecewise exponential model [38] and a proposed approximation to the piecewise exponential model that can be fitted using standard software.

Finally, the conclusions to be drawn from this work are discussed in chapter 5, along with suggestions of how the work could be extended.

References

1. Collett D. *Modelling Survival Data in Medical Research.* Chapman and Hall: London, 2003.

2. Collaborative Group on AIDS Incubation and HIV Survival. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative reanalysis. *Lancet* 2000, **355**:1131-1137.

3. UNAIDS. 2006 report on the global AIDS epidemic. 2006. Geneva, UNAIDS.

4. Gazzard B, BHIVA Writing Committee. British HIV Association (BHIVA) guidelines for the treatment of HIV-infected adults with antiretroviral therapy (2005). *HIV Medicine* 2005, **6**(Suppl 2):1-61.

5. Hammer SM, Saag MS, Schechter M, et al. Treatment for adult HIV infection: 2006 recommendations of the International AIDS Society-USA panel. *JAMA* 2006, **296**:827-843.

6. Sharland M, Blanche S, Castelli G, Ramos J, Gibb D, on behalf of the PENTA Steering Committee. PENTA guidelines for the use of antiretroviral therapy, 2004. *HIV Medicine* 2004, **5**:61-86.

7. Concorde Coordinating Committee. Concorde: MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. *Lancet* 1994, **343**:871-882.

8. Delta Coordinating Committee. Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. *Lancet* 1996, **348**:283-291.

9. Porter K, on behalf of CASCADE collaboration. Changes in HIV survival in a European cohort of persons with well estimated dates of seroconversion. Oxford. 2001. Joint conference of the International Epidemiological Association and the Society for Social Medicine, Oxford 2001.

10. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat.Med.* 1989, **8**:431-440.

11. Babiker AG, Darbyshire JH, Peto TEA, Walker AS. Issues in the design and analysis of therapeutic trials in human immunodeficiency virus infection. *Journal of the Royal Statistical Society A* 1998, **161**:239-249.

12. Babiker AG, Hughes M, Darbyshire JH. The role of HIV-1 RNA as a surrogate marker for clinical outcome in clinical trials evaluating new anti-HIV therapies. *International Antiretroviral News* 1998, **6**:10-11.

13. Delta Coordinating Committee and Virology Group. An evaluation of HIV RNA and CD4 cell count as surrogates for clinical outcome. *AIDS* 1993, **13**:565-573.

14. HIV Surrogate Marker Collaboration Group. Human immunodeficiency virus type I RNA level and CD3 count as prognostic markers and surrogate end points: a meta-analysis. *AIDS research and human retroviruses* 2000, **16**:1123-1133.

15. Food and Drug Administration. Guidance for industry: Antiretroviral drugs using plasma RNA measurements - clinical considerations for accelerated and traditional approval. 2002.

16. Committee for Medicinal Products for Human Use (CHMP), EMEA. Guideline on the clinical development of medicinal products for the treatment of HIV infection. 2005.

17. Dorey FJ, Little RJA, Schenker N. Multiple imputation for threshold-crossing data with interval censoring. *Statistics in Medicine* 1993, **12**:1589-1603.

18. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 1958, **53**:457-481.

19. Cox DR. Regression models with lifetables (with discussion). *Journal of the Royal Statistical Society B* 1972, **34**:187-220.

20. Rubin DB. *Multiple Imputations for Non response in Surveys.* Wiley: New York, 1987.

21. Lindsey JC, Ryan LM. Tutorial in biostatistics - Methods for interval-censored data. *Statistics in Medicine* 1998, **17**(2):219-238.

22. Sun J. Interval censoring. *Encyclopedia of Biostatistics*. John Wiley & Sons Ltd: 2005: pp 2603-2609.

23. Lesaffre E, Komarek A, Declerck D. An overview of methods for the analysis of interval censored data with an emphasis on applications in dentistry. *Statistical Methods in Medical Research* 2005, **14**:539-552.

24. Miller RG. What price Kaplan-Meier? *Biometrics* 1983, **39**:1077-1081.

25. Peto R. Experimental survival curves for interval-censored data. *Applied Statistics* 1973, **22**:86-91.

26. Turnbull BW. The empirical distribution function with arbitrary grouped, censored and truncated data. *Journal of the Royal Statistical Society B* 1976, **38**:290-295.

27. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B* 1977, **39**:1-38.

28. Pan W, Chappell R. Estimating survival curves with left truncated and interval censored data via the EMS algorithm. *Communications in Statistics-Theory and Methods* 1998, **27**(4):777-793.

29. Pan W. Smooth estimation of the survival function for interval censored data. *Statistics in Medicine* 2000, **19**(19):2611-2624.

30. Kooperberg C, Stone CJ. Logspline density estimation for censored data. *Journal of computational and graphical statistics* 1992, **1**:301-328.

31. Rosenberg PS. Hazard function estimation using B-splines. *Biometrics* 1995, **51**:874-887.

32. Bebchuk JD, Betensky RA. Multiple imputation for simple estimation of the hazard function based on interval-censored data. *Stat.Med.* 2000, **19**:405-419.

33. Betensky RA, Lindsey JC, Ryan LM, Wand MP. Local EM estimation of the hazard function for interval-censored data. *Biometrics* 1999, **55**:238-245.

34. Rosenberg PS, Goedert JJ, Giggar RJ. Effect of age at seroconversion on the natural AIDS incubation distribution. *AIDS* 1994, **8**:803-810.

35. Kooperberg C, Clarkson DB. Kooperberg C, Hazard regression with interval-censored data. *Biometrics* 1997, **53**:1485-1494.

36. Betensky RA, Lindsey JC, Ryan LM, Wand MP. A local likelihood proportional hazards model for interval-censored data. *Stat.Med.* 2002, **21**:263-275.

37. Pan W. A multiple imputation approach to Cox regression with interval-censored data. *Biometrics* 2000, **56**:199-203.

38. Friedman M. Piecewise exponential models for survival data with covariates. *Annals of Statistics* 1982, **10**:101-109.

39. Carsenten B. Regression models for interval-censored survival data: application to HIV infection in Danish homosexual men. *Stat.Med.* 1996, **15**:2177-2189.

40. Farrington CP. Interval censored survival data: a generalized linear modelling approach. *Statistics in Medicine* 1996, **15**(3):283-292.

41. Smith PJ, Thompson TJ, Jereb JA. A model for interval-censored tuberculosis outbreak data. *Stat.Med.* 2006, **16**:485-496.

42. Finkelstein DM. A proportional hazards model for interval-censored failure time data. *Biometrics* 1986, **42**(4):845-854.

43. Satten GA. Rank-based inference in the proportional hazards model for interval-censored data. *Biometrika* 1996, **83**:355-370.

44. Satten GA, Datta S, Williamson JM. Inference based on imputed failure times for the proportional hazards model with interval-censored data. *Journal of the American Statistical Association* 1998, **93**:318-327.

45. Odell PM, Anderson KM, D'Agostino RB. Maximum likelihood estimation for interval censored data using a Weibull-based accelerated failure time model. *Biometrics* 1992, **48**:951-959.

46. Betensky RA, Rabinowitz D, Tsiatis AA. Computationally simple accelerated failure time regression for interval-censored data. *Biometrika* 2006, **88**:703-711.

47. Gibb DM, Duong T, Tookey PA, Sharland M, Tudor-Williams G, Novelli V, on behalf of the National Study of HIV in Pregnancy and Childhood (NSHPC) and of the Collaborative HIV Paediatric Study (CHIPS). Decline in mortality, AIDS and hospital admissions in perinatally HIV-1 infected children in the UK and Ireland. *BMJ* 2003, **327**:1019-1025.

CHAPTER 2: CONFIDENCE INTERVALS FOR THE NON-PARAMETRIC MAXIMUM LIKELIHOOD ESTIMATOR

2.1 INTRODUCTION

Of the interval-censoring methods discussed in section 1.3.2, the non-parametric maximum likelihood estimator (NPMLE) [1,2] is the most regularly used method to obtain an estimate of the survivor distribution function. A variety of fitting methods [1-4] and the properties of the estimator have been discussed in the literature [4-7] and are described in more detail in section 2.2.2. However, there has been limited research on methods for obtaining confidence intervals for the survivor function. Peto and Turnbull described the derivation of (pointwise) confidence intervals based on the observed information matrix. However, several authors have questioned the applicability of standard asymptotic theory since the number of parameters that define the survivor function increases with the number of observations [8-10]. In typical applications, many of these parameters are estimated as zero. Peto and Turnbull suggest that the rows and columns of the information matrix corresponding to these zero elements should be deleted prior to matrix inversion, although the theoretical justification for this is unclear [1,2]. An alternative approach, also based on asymptotic theory, is to obtain confidence intervals by inverting a likelihood ratio test at each time point of interest, as has been examined in the case of right-censored data [11] but not previously considered when the data are interval-censored.

The aim of this chapter is to examine the accuracy of confidence intervals for the non-parametric survivor function based on the three methods described above which have not been previously assessed or compared. In section 2.2, the underlying theory for each approach is described. In section 2.3, an illustrative data set is used to exemplify these methods. Section 2.4 describes a simulation study examining the coverage of the confidence intervals and how this is affected by selected factors, including sample size and the width of the individual censoring intervals. Section 2.5 summarises the main

findings from this analysis, considers the practical implications of these findings for applied research, and suggests possible areas for further research.

2.2 NOTATION AND THEORY

2.2.1 Observed data

Let the failure times T arise from the survivor distribution $S(t) = Pr(T \geq t)$. Each individual has a sequence of examination times which, to ensure that censoring is non-informative, are assumed to be independent of the failure time. The observed data consist of censoring intervals $I_i=(L_i,R_i)$ for individual i=1,...,n. In the special cases of left-censored and right-censored observations, $L_i=0$ and $R_i=\infty$, respectively.

2.2.2 Non-parametric maximum likelihood estimator (NPMLE)

If no parametric assumptions about the distribution of S(t) can be made the methods of Peto [1] or Turnbull [2] can be applied to obtain the analogue of the Kaplan-Meier product limit estimator for right-censored data. The Kaplan-Meier method can be easily adapted to deal with left censored data, but no simple algebraic method can be implemented in the more general cases where data are a mixture of left and right censored observations, or when any observation is censored into an interval. The log-likelihood is expressed as a function of the interval endpoints

$$\ell = \sum_{i=1}^{n} \log\left(S(L_i) - S(R_i -)\right) \tag{2.1}$$

The search for a function S that maximises (2.1) is facilitated by the following reduction of the problem. Let $\{u_j; j=1,...,m+1\}$ denote the unique ordered values of $\{0,\{L_i\},\{R_i\}\}$ and define an indicator variable $\alpha_{ij}=1$ if $(u_j, u_{j+1}) \subseteq I_i$ and 0 otherwise. Let $f_j=S(u_j)-S(u_{j+1})$ for j=1,...,m. The log-likelihood can then be rewritten as

$$\ell(\mathbf{f}) = \sum_{i=1}^{n} \log\left(\sum_{j=1}^{m} \alpha_{ij} f_j \right)$$

$$\text{subject to } f_j \geq 0, \ \sum_{j=1}^{m} f_j = 1 \tag{2.2}$$

One parameter is redundant due to the linear constraint (2.2). Conventionally, f_m is re-expressed as $1 - \sum_{j=1}^{m-1} f_j$, although inference is unaffected by choice of the parameter that is eliminated [1,2]. It can be shown that $\hat{f}_j \neq 0$ only if $u_j \in \{L_i\}$ and

$u_{j+1} \in \{R_i\}$ (Appendix 1) [12]; if this condition is not satisfied then $\hat{f}_j = 0$ *a priori* and these elements can be eliminated from the model. All subsequent analyses in this thesis are predicated on this fact.

$\hat{S}(t)$ is constant outside all distinct intervals with endpoints $u_j \in \{L_i\}$ and $u_{j+1} \in \{R_i\}$, and undefined within any interval for which $\hat{f}_j$ is strictly greater than zero. Probability mass tends to be concentrated in a small number of intervals, giving rise to large regions where the survivor function is flat and a number of regions where it is undefined. These characteristics become less pronounced with increasing sample size, although only gradually – due to cube root asymptotics of the estimator [4,6,13].

In general, there is no closed form solution for the maximum likelihood estimate (MLE) $\hat{\mathbf{f}} = (\hat{f}_1 ... \hat{f}_m)$ and an iterative method is required. Peto proposes a constrained Newton-Raphson approach, while Turnbull proposes an expectation maximisation (EM) algorithm [14] that is easy to implement but slow to converge. Define an indicator function $I_{ij} = 1$ if $T_j \in I_i$, 0 otherwise, which constitutes the "complete data". The E-step calculates the conditional expectation of person i, i=1,…,n, experiencing an event in interval j, j=1,…,m, given the observed data and the current estimates of $\hat{\mathbf{f}} = (\hat{f}_1 ... \hat{f}_m)$, namely

$$E[I_{ij}] = \alpha_{ij} \hat{f}_j \Big/ \left(\sum_{j=1}^{m} \alpha_{ij} \hat{f}_j \right) \tag{2.3}$$

Considering $E[I_{ij}]$ as an observed rather than expected frequency, the M-step is simply to estimate the proportion of events in interval j,

$$\hat{f}_j = \frac{1}{n} \sum_{i=1}^{n} E[I_{ij}] \tag{2.4}$$

The algorithm iterates between (2.3) are (2.4) until convergence. This is equivalent to a self-consistency approach with iteration of the equation

$$f_j^{new} = \frac{1}{n} \sum_{i=1}^{n} \left[\alpha_{ij} f_j^{old} \Big/ \left(\sum_{j=1}^{m} \alpha_{ij} f_j^{old} \right) \right]$$

until convergence. Gentleman & Geyer show how the Kuhn-Tucker conditions can be used to verify that the self-consistent estimator is the MLE [5]. A more complex, convex minorant algorithm that is faster to converge than the EM algorithm has been proposed [4]. In addition, Böhning has described how algorithms suggested for mixture

model problems can be applied [3]. In practice, increased computer power has facilitated the use of standard optimization techniques and eliminated real concern over speed of convergence for the EM algorithm.

Wald type confidence intervals are easily derived via the observed information

$$\text{var}(\hat{\mathbf{f}}) \approx \left[-\frac{\partial^2 \ell}{\partial f_s \partial f_l} \right]^{-1} = \left(A^T D A \right)^{-1} \tag{2.5}$$

where A is the n x m matrix with elements α_{ij} and D the n x n diagonal matrix with elements

$$d_{ii} = -1 \Big/ \left(\sum_{j=1}^{m} \alpha_{ij} \hat{f}_j \right)^2 .$$

The estimated survivor function at u_{j+1}, $\hat{S}(u_{j+1}) = 1 - \sum_{k=1}^{j} \hat{f}_k$, can be written as $1 - \left[\mathbf{1}_j^{\,T} \;\; \mathbf{0}_{m-j}^{\,T} \right] \hat{\mathbf{f}}$ where $\mathbf{1}_j$ is the unit vector of length j and $\mathbf{0}_{m-j}$ the null vector of length (m-j). From (2.5) the approximate variance of $\hat{S}(u_{j+1})$ is

$$\left[\mathbf{1}_j^{\,T} \;\; \mathbf{0}_{m-j}^{\,T} \right] \left(A^T D A \right)^{-1} \begin{bmatrix} \mathbf{1}_j \\ \mathbf{0}_{m-j} \end{bmatrix}.$$

Asymptotic pointwise (1-α) confidence intervals are obtained by $\hat{S}(u_{j+1}) \pm z_{\alpha/2} \text{SE}\left[\hat{S}(u_{j+1})\right]$ where SE denotes standard error and z_α is the upper α point of the standard normal distribution. This will be referred to as the Wald-1 method.

It has been suggested, although without formal justification, that the rows and columns of the information matrix corresponding to zero elements of $\hat{\mathbf{f}}$ should be deleted prior to matrix inversion [1,2]. This will be referred to as the Wald-2 method. A key objective of this chapter is to examine the validity of this procedure.

Alternatively, confidence intervals can be constructed via the likelihood ratio test (also called profile likelihood confidence intervals). The likelihood ratio statistic for testing $H_O : S(u_{j+1}) = S_O$ is $W(S_O) = 2\left[\ell(\hat{\mathbf{f}}) - \ell(\tilde{\mathbf{f}})\right]$ where $\tilde{\mathbf{f}}$ is the MLE subject to $\left[\mathbf{1}_j^{\,T} \;\; \mathbf{0}_{m-j}^{\,T} \right] \tilde{\mathbf{f}} = S_O$. A (1-α) confidence interval consists of the set of values $\left\{ S_O : W(S_O) < \chi^2_\alpha \right\}$ where χ^2_α is the upper α point of the χ^2 distribution with 1 degree of freedom.

Although the likelihood has a simple form estimation of the survivor function or pointwise confidence intervals is not a straightforward problem. Except in the special case when the data are grouped, each new observation can potentially add up to two additional parameters in (2.2), depending on the precision of the time scale. The MLE will typically have several parameters identically zero i.e. on the boundary of the parameter space so that standard asymptotic theory may not apply [15]. Although the MLE has been shown to be consistent [4-7] the validity of standard methods for determining confidence intervals for the survivor function has not previously been examined. In particular, the NPMLE does not satisfy the Central Limit Theorem with the usual $n^{1/2}$-rate convergence and use of the observed information to construct confidence intervals for the NPMLE does not have large sample justification [6]. Also, the likelihood ratio statistic does not necessarily follow a χ^2 distribution with the nominal number of degrees of freedom if parameter estimates, including nuisance parameters, lie on the boundary of the parameter space [16].

2.2.3 Implementation

Throughout this chapter the NPMLE has been fitted using a quasi-Newton optimisation procedure implemented in SAS Version 8.2 [17]. A program available from the SAS/IML sample library fits the NPMLE to interval-censored data but gives confidence intervals using the Wald-1 method only [18]. This SAS code was adapted to allow estimation of Wald-2 and likelihood ratio confidence limits (see Appendices 2A-C).

2.3 EXAMPLE: The Delta Virology Substudy

Delta was a double blind randomised trial comparing zidovudine (AZT) monotherapy with AZT combined with either didanosine (ddI) or zalcitabine (ddC) in 3207 HIV-1 infected individuals [19]. Here, the focus is on the 420 subjects allocated to AZT+ddC who participated in the virology substudy [20]. Estimation of the survivor function is considered for the endpoint of time to achieving undetectable levels of plasma HIV RNA, measured by the NASBA assay with limit of detection 800 copies/ml [21].

The individual censoring intervals are shown in Figure 2.1. For the purpose of presentation, R_i was set to 60 weeks for subjects who exhibited right censoring i.e. never achieved undetectable HIV RNA. This was commonly observed, occurring in 247 (59%) subjects, of whom 177 (42%) were last assessed at 48 weeks. Note also the large number of left-censored subjects who had undetectable HIV RNA at the first

measurement at 4 weeks or who became undetectable between their first and second measurements at 4 and 8 weeks.

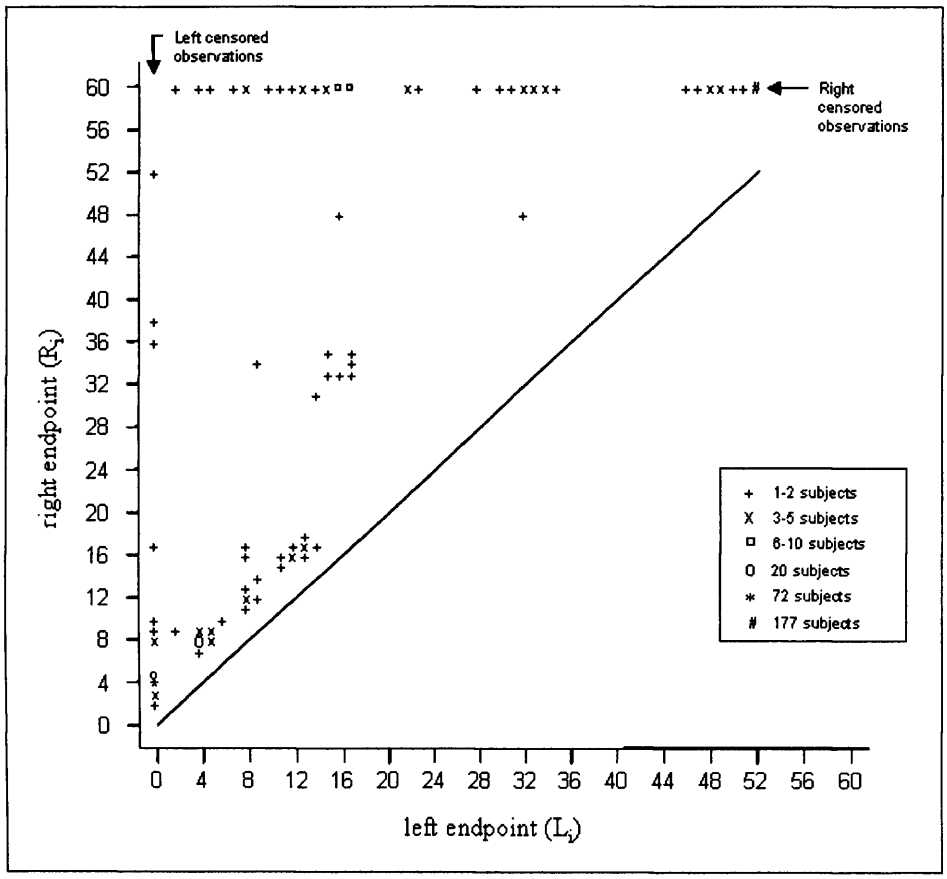

Figure 2.1: Individual censoring intervals in the Delta Virology Substudy

The data induced 22 intervals, in 13 of which the probability density was estimated as zero (Table 2.1, Figure 2.2). The proportion of subjects who achieved HIV RNA <800 copies/ml increased steadily to approximately 40% by week 16, with little change thereafter. This example illustrates the highly discontinuous behaviour, with probability density concentrated in a few intervals, which is characteristic of the non-parametric estimator.

Several points emerge from a comparison of the confidence intervals in Table 2.1. First, for the Wald-2 method, a confidence interval is by definition identical to the preceding confidence interval if $\hat{f}_j = 0$. Although this is not true for the likelihood ratio method, the changes between adjacent intervals are generally very small, particularly for the upper confidence limit; for example, this is identical to the fourth decimal place (0.6988) in intervals 9-10. Second, the Wald-2 and the likelihood ratio confidence intervals are generally very similar, except for interval 1, where the latter is much

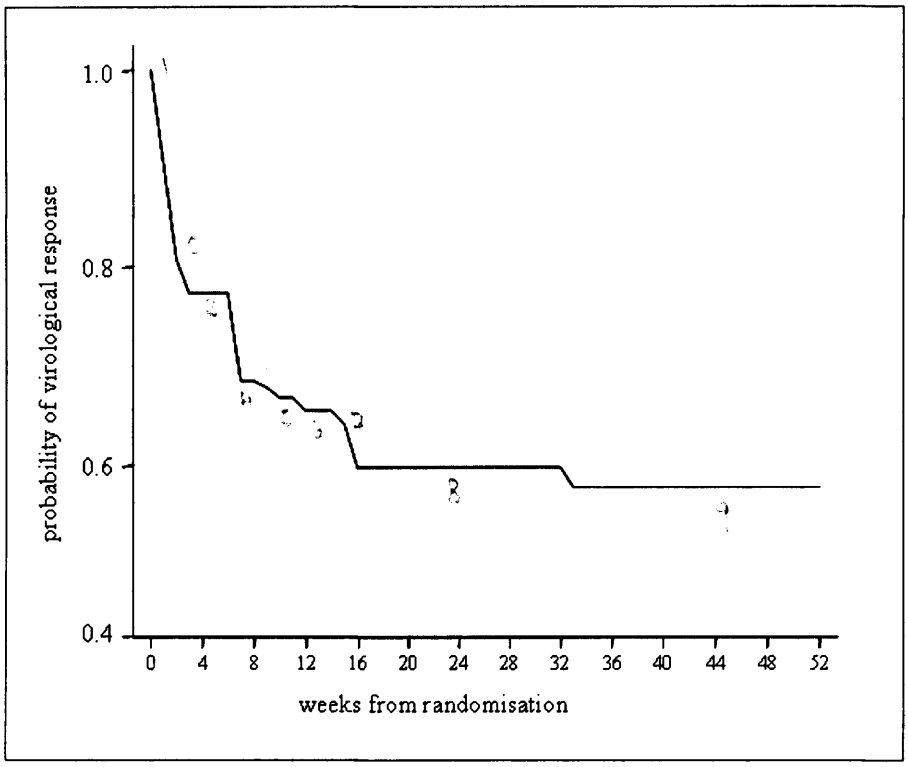

Figure 2.2: Non-parametric survivor function in Delta Virology Substudy

Although the NPMLE is strictly only defined within time intervals corresponding to where the curve is flat, it is shown with interpolated values (diagonal lines) for the intervening intervals.

narrower. Third, the Wald-1 confidence intervals are consistently wider than those obtained by the other two methods, often considerably so. Fourth, the Wald-1 confidence intervals are pathologically wide at some time points, in comparison with adjacent time points (e.g. intervals 7 and 20), due to ill-conditioning of the information matrix.

The ill-conditioning was observed to only affect confidence intervals for intervals (u_j, u_{j+1}) that satisfied all of the following conditions:

(i) $u_j \leq R_i < u_{j+1} \Rightarrow S(L_i)=1$, for all i

(ii) $u_j < L_i \leq u_{j+1} \Rightarrow S(R_i)=0$, for all i

(iii) $\hat{f}_j = 0$.

Consider only those intervals (u_j, u_{j+1}) that are not known *a priori* to have zero mass (section 3.2.1). Points (i) and (ii) can be determined by looking at all individuals with censoring intervals that do not cover both (u_j, u_{j+1}) and the next interval. Point (i) states that the censoring interval I_i for all individuals with $(u_j, u_{j+1}) \subseteq I_i$ does not contain any later interval and contains ALL previous intervals; if the censoring interval for all other

Table 2.1: Proportion of subjects who have not achieved viral load <800 copies/ml in Delta Virology Substudy

J	(u_j, u_{j+1}) [a]	$\hat{f}_j$	$\hat{S}(u_{j+1})$	95% confidence interval for $S(u_{j+1})$		
				Wald-1	Wald-2	Likelihood ratio
1	(0,2)	0.1951	0.8049	0.6202-0.9897	0.6210-0.9889	0.7325-0.9565
2	(2,3)	0.0315	0.7734	0.7279-0.8189	0.7324-0.8145	0.7294-0.8128
3	(4,5)	0.0000	0.7734	0.6963-0.8505		0.7290-0.8128
4	(6,7)	0.0894	0.6840	0.5527-0.8154	0.6383-0.7298	0.6371-0.7515
5	(7,8)	0.0000	0.6840	0.6309-0.7372		0.6372-0.7287
6	(8,9)	0.0078	0.6762	0.6152-0.7372	0.6243-0.7281	0.6251-0.7243
7	(9,10)	0.0092	0.6670	0.4589-0.8750	0.6138-0.7201	0.6123-0.7206
8	(10,11)	0.0000	0.6670	0.6062-0.7277		0.6120-0.7162
9	(11,12)	0.0140	0.6530	0.5959-0.7101	0.6061-0.6999	0.6051-0.6988
10	(12,13)	0.0000	0.6530	0.5902-0.7158		0.6053-0.6988
11	(13,14)	0.0000	0.6530	0.5848-0.7212		0.5981-0.6983
12	(14,15)	0.0125	0.6405	0.5674-0.7136	0.5745-0.7064	0.5725-0.6862
13	(15,16)	0.0430	0.5975	0.5351-0.6599	0.5497-0.6453	0.5490-0.6448
14	(16,17)	0.0000	0.5975	0.5347-0.6603		0.5491-0.6367
15	(17,18)	0.0000	0.5975	0.5003-0.6948		0.5455-0.6438
16	(30,31)	0.0000	0.5975	0.5201-0.6749	0.5497-0.6453	0.5451-0.6441
17	(32,33)	0.0205	0.5770	0.5120-0.6420	0.5287-0.6253	0.5364-0.6242
18	(33,34)	0.0000	0.5770	0.5038-0.6502		0.5281-0.6246
19	(34,35)	0.0000	0.5770	0.5046-0.6494		0.5278-0.6257
20	(35,36)	0.0000	0.5770	0.1042-1.0498		0.5282-0.6246
21	(47,48)	0.0000	0.5770	0.5061-0.6480		0.5304-0.6241
22	(51,52)	0.0000	0.5770	0.4924-0.6616		0.5284-0.6272

(a) Time since randomisation (weeks)

individuals does not contain (u_j, u_{j+1}) but contains ALL the subsequent later intervals then point (ii) holds. It was not possible to determine if this rule holds in the general case. However, conditions (i) and (ii) will hold only under extensive left and right censoring. It should be noted that condition (iii) is never satisfied under the Wald-2 method, and may explain why this approach is numerically stable.

2.4 SIMULATION STUDY

In light of the concerns discussed in section 2.2.2 regarding the asymptotic validity of standard methods for confidence interval estimation, a simulation study was carried out to assess the coverage of confidence intervals obtained by the Wald-1, Wald-2, and likelihood ratio methods.

2.4.1 Methods

Data were generated from a hypothetical prospective study of 48 weeks duration in which visits were scheduled every 4 weeks (Figure 2.3). Individuals were assumed to miss each visit at random with a fixed probability, *p*. In addition, to allow for the fact that a patient may not attend on the exact scheduled date, the actual visit time was assumed to be normally distributed, centred at the scheduled visit time with standard deviation 1.33 weeks. Survival times were generated from a Weibull distribution and the individual censoring intervals determined.

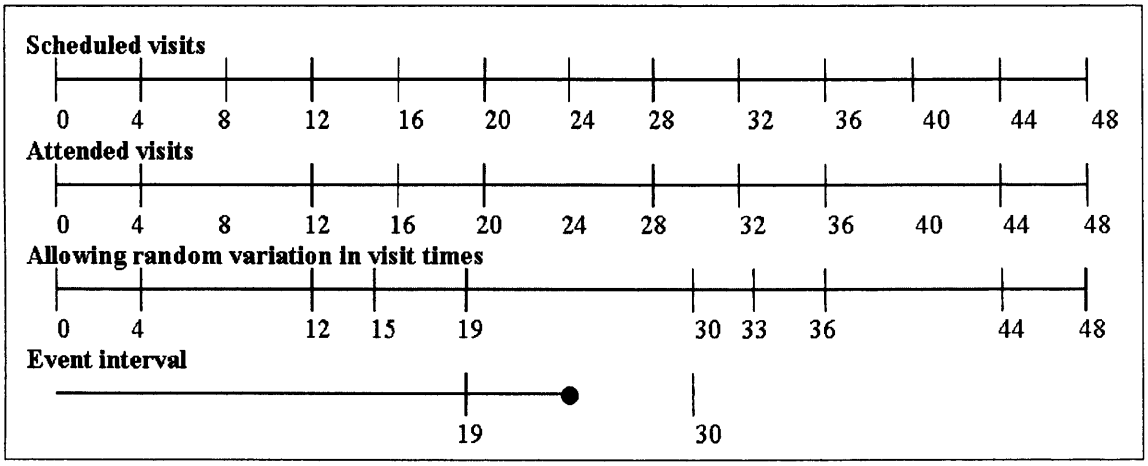

Figure 2.3: Scheme illustrating generation of individual data in simulation study (hypothetical case)

1,000 hypothetical studies were simulated for each combination of the following factors (*scenario*) and pointwise confidence intervals for the survivor function computed by

each of the three methods at t=12, 24 or 48 weeks:

(a) Sample size, n=100, 200, 400, 800 subjects.

(b) Probability of missing a scheduled visit, p=0.2, 0.4.

(c) Shape parameter of Weibull distribution, γ=0.6, 1.0, 1.5, representing a decreasing, constant, and increasing hazard function, respectively. For each value of the shape parameter, the scale parameter, λ, was selected so that S(12)=0.75, S(24)=0.5, and S(48)=0.25. These values are consistent with estimates of the frequency of HIV-1 RNA undetectability following initiation of antiretroviral therapy [22,23].

(d) Proportion lost to follow up, d=0%, 10%. This was achieved by assuming individuals were lost from the study at random with a fixed probability at each visit, such that the expected loss by 48 weeks was d%.

As was noted in section 2.2.2, $\hat{S}$ is undefined at all times within any interval (u_j,u_{j+1}) for which $\hat{f}_j \neq 0$. When this occurred at t=12, 24, or 48 weeks, $\hat{S}(t)$ was estimated by linear interpolation, and a confidence interval derived from the lower limit that pertained at u_{j+1} and the upper limit that pertained at u_j. This approach yields conservative confidence limits but was required in relatively few simulations (Tables 2.2a & b), never exceeding 5.6% or 5.0% in simulations with 0% and 10% loss to follow-up rates respectively, and was always less than 1.1% when n>100.

For each simulation, the Wald-1 and Wald-2 methods give an estimate of the standard error (of the estimated survivor function) at each time point. These estimates were averaged over the 1000 simulations, and compared with the Monte Carlo estimates of standard error i.e. the empirical standard deviation of the point estimates from each of the 1000 simulations. The coverage of the confidence intervals for each of the different scenarios was then calculated and compared.

2.4.2 Results

a) Number of parameters

Because the data induce the parameterisation (section 2.2.2), the number of parameters (probability densities) to be estimated varies from simulation to simulation. With no loss to follow up and a sample size of 100, the mean number of parameters was 28 (range 22 - 33) and the mean number of non-zero parameters was 13 (range 9 – 16).

There was an observed trend of increasing number of parameters (total and number estimated as non-zero) with increasing sample size. For a sample size of 800, the corresponding values were 51 (range 50 – 51) and 29 (21 – 35). Results were very similar when considering a 10% loss to follow-up rate (data not shown).

b) Consistency

First the consistency of the point estimates (against the known values of 0.75, 0.50, and 0.25) was examined, since any such bias would affect the coverage achieved by any method of confidence interval estimation (Tables 2.2a & b). Bias was found to be minimal relative to the Monte Carlo standard error, particularly for the larger sample size (n=800) in line with theoretical results on the consistency of the non-parametric estimator [5].

c) Coverage

The Wald-1 method clearly over-estimated the standard error (Figure 2.4) for all combinations of the simulation factors (Table 2.3), particularly for larger sample sizes. The standard error was over-estimated by a factor of 1.34-2.40 for n=100, decreasing to a factor of 1.08-1.56 for n=800, compared with the Monte Carlo (empirical) estimate.

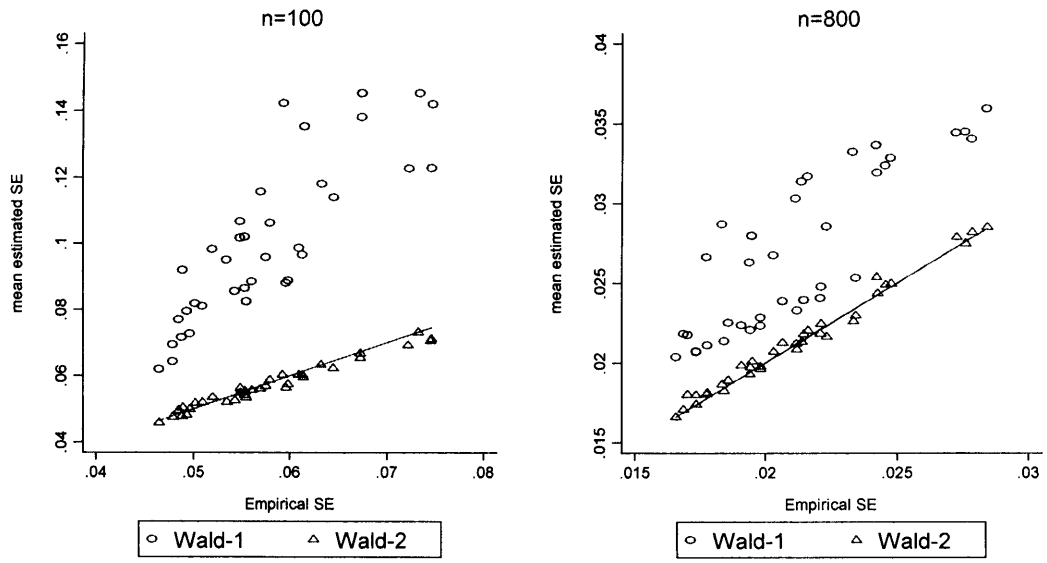

Figure 2.4: Comparison of mean estimated standard error and empirical standard error.

Each point represents the mean of 1000 replicates from a single scenario.

Table 2.2a: Results of simulation study – no loss to follow-up

						Standard error			Coverage (%)[a]		
t	γ, λ	n	p	I[b]	Mean bias ($x10^{-3}$)	Monte Carlo	Wald-1[c]	Wald-2[c]	Wald-1	Wald-2	Likelihood ratio
12	0.6, 0.02	100	0.2	32	2.3266	0.0488	0.0717	0.0478	98.4*	93.2	95.4
			0.4	46	1.3315	0.0549	0.1066	0.0550	99.5*	93.2	95.6
		200	0.2	5	-0.4830	0.0354	0.0530	0.0348	98.7*	93.9	94.9
			0.4	6	-0.1957	0.0408	0.0882	0.0409	99.7*	93.0*	95.8
		400	0.2	0	0.2161	0.0245	0.0333	0.0251	99.4*	94.0	95.1
			0.4	1	-1.9146	0.0289	0.0529	0.0298	100.0*	94.7	96.4
		800	0.2	0	0.3732	0.0174	0.0207	0.0180	98.1*	95.2	95.9
			0.4	0	0.5112	0.0212	0.0304	0.0212	99.6*	94.9	96.4
	1.0, 0.003	100	0.2	9	-0.6665	0.0503	0.0818	0.0517	98.9*	93.7	95.6
			0.4	31	1.7978	0.0645	0.1139	0.0623	99.4*	90.1*	94.4
		200	0.2	1	1.8979	0.0378	0.0549	0.0377	99.5*	93.2	95.9
			0.4	2	0.5416	0.0469	0.0863	0.0454	99.5*	92.6*	95.6
		400	0.2	0	-1.2848	0.0268	0.0344	0.0274	98.7*	95.2	96.2
			0.4	0	-1.3801	0.0332	0.0523	0.0337	99.6*	93.6	96.8
		800	0.2	0	0.5585	0.0195	0.0221	0.0198	97.5*	94.6	96.1
			0.4	0	-0.9088	0.0243	0.0320	0.0244	99.0*	94.0	95.5
	1.5, 0.0004	100	0.2	3	-1.3456	0.0596	0.0880	0.0565	98.3*	91.0*	94.0
			0.4	9	-0.3309	0.0722	0.1225	0.0692	98.2*	90.0*	94.3
		200	0.2	0	-0.3264	0.0420	0.0573	0.0410	98.8*	91.5*	94.7
			0.4	0	2.6451	0.0524	0.0848	0.0520	99.3*	91.2*	95.6
		400	0.2	0	-0.3514	0.0298	0.0358	0.0301	98.1*	93.4	94.7
			0.4	0	-0.2276	0.0383	0.0532	0.0381	99.4*	93.8	96.1
		800	0.2	0	0.9397	0.0215	0.0240	0.0219	97.7*	94.3	95.2
			0.4	0	0.3688	0.0279	0.0341	0.0282	98.4*	94.2	95.9

*Significantly different from nominal 95% coverage at P<0.01 (99% confidence intervals for 950/1000 binomial proportion = (93.2, 96.8))

(a) SE of coverage estimate is approximately 0.6892 % (100 x sqrt (.95 x .05/1000)), (b) number of times interpolation was required because $\hat{S}(t)$ was undefined , (c) mean over 1000 simulations

						Standard error			Coverage (%)[a]		
t	γ, λ	n	p	I[b]	Mean bias ($x10^{-3}$)	Monte Carlo	Wald-1[c]	Wald-2[c]	Wald-1	Wald-2	Likelihood ratio
24	0.6, 0.03	100	0.2	56	1.8485	0.0555	0.0824	0.0535	98.2*	93.5	94.2
			0.4	40	-2.3557	0.0614	0.1353	0.0597	100.0*	92.8*	95.8
		200	0.2	11	0.2460	0.0392	0.0586	0.0384	98.6*	93.4	94.4
			0.4	5	0.3991	0.0434	0.1044	0.0427	99.9*	93.6	95.9
		400	0.2	0	-0.0608	0.0281	0.0359	0.0275	98.3*	93.8	94.4
			0.4	0	-0.5972	0.0306	0.0596	0.0309	99.9*	93.4	95.3
		800	0.2	0	-1.0658	0.0198	0.0224	0.0196	96.6	94.5	94.7
			0.4	0	0.0103	0.0217	0.0317	0.0221	99.3*	94.6	96.0
	1.0, 0.004	100	0.2	21	1.7916	0.0561	0.0884	0.0557	98.7*	93.4	95.0
			0.4	25	0.4057	0.0673	0.1380	0.0654	99.7*	91.3*	93.9
		200	0.2	0	-0.5041	0.0400	0.0594	0.0404	99.3*	94.0	95.1
			0.4	1	-0.7707	0.0484	0.0992	0.0479	100.0*	92.9*	95.8
		400	0.2	0	0.5852	0.0284	0.0364	0.0291	98.6*	94.6	95.6
			0.4	0	1.9465	0.0342	0.0559	0.0344	99.6*	93.8	95.6
		800	0.2	0	-0.6001	0.0212	0.0233	0.0209	97.0*	93.9	94.2
			0.4	0	0.6758	0.0248	0.0329	0.0250	99.1*	94.0	95.6
	1.5, 0.0003	100	0.2	8	-0.5794	0.0613	0.0966	0.0603	99.2*	91.4*	94.1
			0.4	9	3.2438	0.0746	0.1420	0.0711	99.7*	90.6*	95.3
		200	0.2	0	-0.9476	0.0438	0.0602	0.0431	99.1*	94.3	96.0
			0.4	0	2.8872	0.0540	0.0955	0.0524	99.6*	92.0*	95.3
		400	0.2	0	-1.1902	0.0302	0.0373	0.0312	98.0*	95.4	96.2
			0.4	0	2.7303	0.0381	0.0565	0.0383	99.8*	93.4	95.4
		800	0.2	0	-0.3388	0.0222	0.0248	0.0225	97.3*	94.9	95.4
			0.4	0	-0.2320	0.0276	0.0346	0.0275	98.1*	93.0*	94.9

*Significantly different from nominal 95% coverage at P<0.01 (99% confidence intervals for 950/1000 binomial proportion = (93.2, 96.8))

(a) SE of coverage estimate is approximately 0.6892 % (100 x sqrt (.95 x .05/1000)), (b) number of times interpolation was required because $\hat{S}(t)$ was undefined , (c) mean over 1000 simulations

						Standard error			Coverage (%)[a]		
t	γ, λ	n	p	I[b]	Mean bias ($x10^{-3}$)	Monte Carlo	Wald-1[c]	Wald-2[c]	Wald-1	Wald-2	Likelihood ratio
48	0.6, 0.04	100	0.2	1	3.5975	0.0465	0.0621	0.0459	98.0*	94.0	95.1
			0.4	1	2.9881	0.0490	0.0919	0.0504	99.7*	94.9	96.0
		200	0.2	0	1.4420	0.0327	0.0472	0.0326	98.8*	94.7	94.3
			0.4	0	2.9987	0.0341	0.0673	0.0357	99.7*	95.8	96.8
		400	0.2	0	2.4064	0.0231	0.0331	0.0233	98.8*	95.7	95.9
			0.4	0	2.9577	0.0243	0.0446	0.0254	99.9*	95.8	96.5
		800	0.2	0	1.5887	0.0166	0.0204	0.0166	98.3*	95.3	95.6
			0.4	0	1.9508	0.0178	0.0266	0.0180	99.5*	94.7	94.9
	1.0, 0.004	100	0.2	0	4.7199	0.0479	0.0694	0.0476	98.8*	93.5	94.4
			0.4	2	6.0347	0.0520	0.0983	0.0536	98.9*	94.7	95.4
		200	0.2	0	3.3311	0.0340	0.0513	0.0341	99.4*	94.5	95.0
			0.4	0	4.1345	0.0390	0.0692	0.0378	99.8*	93.8	95.0
		400	0.2	0	2.2663	0.0234	0.0326	0.0244	99.1*	95.3	95.6
			0.4	0	3.4265	0.0266	0.0430	0.0271	99.6*	95.2	95.5
		800	0.2	0	2.2453	0.0174	0.0208	0.0174	98.4*	95.5	95.4
			0.4	0	1.5476	0.0194	0.0263	0.0193	99.2*	94.4	95.1
	1.5, 0.0002	100	0.2	2	5.5986	0.0486	0.0770	0.0496	98.8*	94.9	94.8
			0.4	1	3.7177	0.0549	0.1017	0.0563	99.3*	94.8	96.2
		200	0.2	0	2.5844	0.0355	0.0518	0.0355	99.6*	94.4	95.6
			0.4	0	6.1361	0.0385	0.0685	0.0403	99.8*	95.0	96.2
		400	0.2	0	1.0230	0.0248	0.0330	0.0255	98.4*	95.3	95.7
			0.4	0	2.3032	0.0279	0.0429	0.0289	99.6*	95.2	95.5
		800	0.2	0	0.8830	0.0185	0.0214	0.0183	98.0*	94.7	95.0
			0.4	0	2.0059	0.0203	0.0268	0.0207	98.7*	94.6	95.4

*Significantly different from nominal 95% coverage at P<0.01 (99% confidence intervals for 950/1000 binomial proportion = (93.2, 96.8))

(a) SE of coverage estimate is approximately 0.6892 % (100 x sqrt (.95 x .05/1000)), (b) number of times interpolation was required because $\hat{S}(t)$ was undefined , (c) mean over 1000 simulations

Table 2.2b: Results of simulation study – 10% loss to follow-up

						Standard error			Coverage (%)[a]		
t	γ, λ	n	p	I[b]	Mean bias ($x10^{-3}$)	Monte Carlo	Wald-1[c]	Wald-2[c]	Wald-1	Wald-2	Likelihood ratio
12	0.6, 0.02	100	0.2	29	0.5089	0.0494	0.0795	0.0482	99.2*	93.3	95.5
			0.4	45	-3.0113	0.0570	0.1156	0.0560	99.0*	92.4*	94.7
		200	0.2	2	-0.1783	0.0364	0.0585	0.0346	98.9*	93.0*	93.9
			0.4	4	-0.6327	0.0422	0.0908	0.0409	99.7*	92.2*	94.7
		400	0.2	0	1.0763	0.0253	0.0338	0.0253	98.5*	94.7	95.0
			0.4	0	-1.0131	0.0294	0.0554	0.0298	99.8*	93.6	95.9
		800	0.2	0	0.3204	0.0178	0.0211	0.0182	98.2*	95.5	96.1
			0.4	0	-1.4577	0.0214	0.0314	0.0214	99.6*	94.0	95.2
	1.0, 0.003	100	0.2	14	-1.2786	0.0543	0.0856	0.0525	98.9*	92.1*	94.4
			0.4	29	-1.9429	0.0632	0.1180	0.0633	98.7*	92.5*	95.3
		200	0.2	0	-1.0773	0.0383	0.0574	0.0381	99.3*	93.2	95.0
			0.4	0	0.9020	0.0463	0.0878	0.0467	99.7*	91.8*	95.7
		400	0.2	0	0.2702	0.0268	0.0345	0.0278	98.5*	94.7	96.0
			0.4	0	0.9412	0.0340	0.0542	0.0338	99.9*	92.9*	95.8
		800	0.2	0	-0.8238	0.0191	0.0224	0.0199	97.9*	95.0	95.4
			0.4	0	-0.0032	0.0246	0.0324	0.0249	99.1*	93.7	95.9
	1.5, 0.0004	100	0.2	2	-0.0653	0.0598	0.0888	0.0575	98.4*	91.5*	94.9
			0.4	9	-1.4748	0.0745	0.1226	0.0705	98.5*	89.0*	94.3
		200	0.2	0	0.6152	0.0429	0.0583	0.0417	99.2*	91.7*	96.1
			0.4	0	0.8071	0.0522	0.0891	0.0526	99.4*	91.9*	95.8
		400	0.2	0	1.1038	0.0310	0.0361	0.0303	97.0*	92.3*	94.0
			0.4	0	-1.4781	0.0384	0.0538	0.0384	99.4*	92.4*	95.3
		800	0.2	0	0.1008	0.0221	0.0241	0.0219	96.6	93.1*	94.6
			0.4	0	1.4471	0.0273	0.0345	0.0279	98.6*	93.6	95.9

*Significantly different from nominal 95% coverage at P<0.01 (99% confidence intervals for 950/1000 binomial proportion = (93.2, 96.8))

(a) SE of coverage estimate is approximately 0.6892 % (100 x sqrt (.95 x .05/1000)), (b) number of times interpolation was required because $\hat{S}(t)$ was undefined , (c) mean over 1000 simulations

						Standard error			Coverage (%)[a]		
t	γ, λ	n	p	I[b]	Mean bias ($x10^{-3}$)	Monte Carlo	Wald-1[c]	Wald-2[c]	Wald-1	Wald-2	Likelihood ratio
24	0.6, 0.03	100	0.2	50	-2.1486	0.0553	0.0866	0.0542	98.2*	93.0*	93.8
			0.4	45	-1.1216	0.0593	0.1422	0.0603	99.9*	93.8	96.2
		200	0.2	3	1.6409	0.0379	0.0621	0.0387	99.3*	95.7	96.2
			0.4	9	-0.3633	0.0440	0.1079	0.0436	99.9*	93.4	94.9
		400	0.2	1	-0.9358	0.0264	0.0374	0.0277	99.4*	96.1	96.8
			0.4	0	1.4075	0.0304	0.0621	0.0314	99.7*	94.4	95.4
		800	0.2	0	0.9563	0.0199	0.0229	0.0198	97.5*	94.9	95.4
			0.4	0	0.6274	0.0234	0.0333	0.0227	98.7*	93.2	94.0
	1.0, 0.004	100	0.2	16	-1.2262	0.0575	0.0958	0.0568	99.3*	93.9	94.8
			0.4	25	1.4478	0.0673	0.1451	0.0668	99.9*	92.5*	95.6
		200	0.2	0	2.3278	0.0427	0.0622	0.0412	99.1*	92.2*	94.2
			0.4	3	5.3057	0.0484	0.1033	0.0484	99.7*	92.1*	94.7
		400	0.2	0	-0.2802	0.0308	0.0373	0.0296	98.1*	93.0*	93.9
			0.4	0	0.9933	0.0363	0.0582	0.0349	99.4*	93.0*	94.8
		800	0.2	0	1.5331	0.0207	0.0239	0.0213	97.5*	95.3	95.7
			0.4	0	0.2247	0.0243	0.0337	0.0254	99.3*	95.0	96.2
	1.5, 0.0003	100	0.2	9	2.6189	0.0609	0.0986	0.0603	99.0*	92.3*	94.7
			0.4	13	-0.9153	0.0732	0.1452	0.0731	99.8*	92.4*	96.0
		200	0.2	0	-0.2413	0.0432	0.0627	0.0440	99.2*	93.9	95.8
			0.4	0	2.6864	0.0537	0.1022	0.0542	100.0*	93.1*	95.6
		400	0.2	0	0.1738	0.0307	0.0385	0.0319	98.6*	94.3	95.2
			0.4	0	-1.0763	0.0411	0.0590	0.0389	99.3*	91.1*	94.5
		800	0.2	0	-0.4325	0.0235	0.0254	0.0230	96.6	93.8	95.3
			0.4	0	0.7865	0.0285	0.0360	0.0285	98.8*	93.5	96.0

*Significantly different from nominal 95% coverage at P<0.01 (99% confidence intervals for 950/1000 binomial proportion = (93.2, 96.8))

(a) SE of coverage estimate is approximately 0.6892 % (100 x sqrt (.95 x .05/1000)), (b) number of times interpolation was required because $\hat{S}(t)$ was undefined , (c) mean over 1000 simulations

						Standard error			Coverage (%)[a]		
t	γ, λ	n	p	I[b]	Mean bias ($x10^{-3}$)	Monte Carlo	Wald-1[c]	Wald-2[c]	Wald-1	Wald-2	Likelihood ratio
48	0.6, 0.04	100	0.2	0	4.1859	0.0479	0.0644	0.0475	99.0*	94.0	94.5
			0.4	6	3.5864	0.0535	0.0950	0.0521	99.3*	94.1	95.6
		200	0.2	0	2.4558	0.0327	0.0497	0.0337	99.2*	94.7	95.4
			0.4	0	2.7828	0.0373	0.0720	0.0370	99.7*	94.4	95.0
		400	0.2	0	3.4518	0.0234	0.0349	0.0240	99.1*	95.3	95.6
			0.4	0	3.7130	0.0255	0.0473	0.0263	99.7*	95.0	95.6
		800	0.2	0	1.3025	0.0169	0.0218	0.0171	98.5*	95.1	95.1
			0.4	0	1.7187	0.0184	0.0287	0.0187	99.5*	94.5	94.7
	1.0, 0.004	100	0.2	0	5.2698	0.0497	0.0727	0.0499	98.6*	95.0	95.5
			0.4	3	7.7437	0.0554	0.1020	0.0554	99.4*	94.1	95.3
		200	0.2	0	4.6489	0.0350	0.0537	0.0353	99.2*	95.2	95.3
			0.4	0	6.2061	0.0373	0.0728	0.0395	99.9*	95.1	96.5
		400	0.2	0	2.5611	0.0246	0.0352	0.0251	98.9*	95.3	95.8
			0.4	0	3.7911	0.0270	0.0461	0.0281	99.4*	94.9	96.3
		800	0.2	0	1.4833	0.0171	0.0218	0.0180	98.4*	95.7	95.8
			0.4	0	2.3375	0.0195	0.0280	0.0201	99.6*	95.5	95.9
	1.5, 0.0002	100	0.2	0	6.1132	0.0510	0.0810	0.0520	99.2*	94.6	95.2
			0.4	1	5.5128	0.0580	0.1061	0.0588	99.4*	94.5	96.1
		200	0.2	0	3.2411	0.0369	0.0558	0.0371	99.6*	94.9	95.9
			0.4	0	5.1546	0.0414	0.0732	0.0422	99.8*	95.0	95.9
		400	0.2	0	2.3482	0.0260	0.0353	0.0265	98.9*	94.5	95.1
			0.4	0	1.2184	0.0281	0.0458	0.0301	99.7*	95.7	96.8
		800	0.2	0	0.7850	0.0186	0.0226	0.0189	98.0*	94.1	94.7
			0.4	0	3.0781	0.0223	0.0286	0.0217	98.8*	94.5	95.2

*Significantly different from nominal 95% coverage at P<0.01 (99% confidence intervals for 950/1000 binomial proportion = (93.2, 96.8))

(a) SE of coverage estimate is approximately 0.6892 % (100 x sqrt (.95 x .05/1000)), (b) number of times interpolation was required because $\hat{S}(t)$ was undefined , (c) mean over 1000 simulations

Consequently, the coverage achieved by the Wald-1 method substantially exceeded the nominal coverage of 95% and, overall, the true point estimate lay within the confidence interval in 99.0% of simulations. As the percentage of parameters in the model estimated as zero rises, the Wald-1 method becomes increasingly conservative (Figure 2.5).

Table 2.3: Average coverage according to factors considered in simulation study

Factor	Value	Average coverage (%)		
		Wald-1	Wald-2	Likelihood ratio
Probability of missing a visit (p):		*		
	0.2	98.5	94.1	95.2
	0.4	99.4	93.5	95.5
Time (weeks):		*	*	
	12	98.9	93.1	95.4
	24	99.0	93.5	95.2
	48	99.1	94.8	95.5
Shape parameter (γ):			*	
	0.6	99.1	94.2	95.3
	1.0	99.1	93.9	95.4
	1.5	98.8	93.5	95.3
No. of subjects (n):		*	*	
	100	99.0	92.9	95.1
	200	99.4	93.5	95.4
	400	99.1	94.3	95.6
	800	98.4	94.5	95.4
Percent lost to follow up (d):				
	0	99.0	93.9	95.4
	40	99.0	93.8	95.3

* significant variation at $p<0.01$ (global χ^2 test)

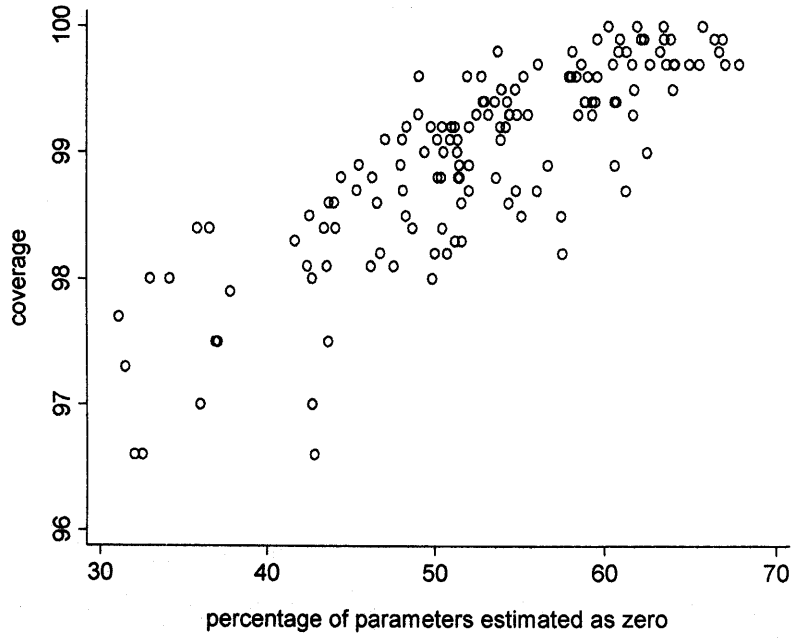

Figure 2.5: Relationship between percentages of parameters estimated as zero and coverage when using the Wald-1 method

In contrast, the mean estimated standard errors from the Wald-2 method were in close agreement with the empirical standard error (Figure 2.4) for all combinations of the simulation factors (Table 2.3). However, coverage was significantly less than 95% (at $p<0.01$) for 14 of the combinations with $d=0$ (Table 2.2a), and the overall coverage for this method was 93.9% (Table 2.3). With a 10% loss to follow-up rate this rose to 24 combinations but the overall coverage was similar (93.8%) (Tables 2.2b & 2.3). Because this method estimated the standard error consistently, this finding at first seemed paradoxical. To clarify this, a combination of factors that gave particularly inaccurate coverage (90.0%; $n=100$, $\gamma=1.5$, $t=12$, $p=0.4$, $d=0$) was examined in more detail. The distribution of estimated survival probabilities were found to be approximately normally distributed but the distribution of estimated standard errors was highly positively skewed (Figure 2.6).

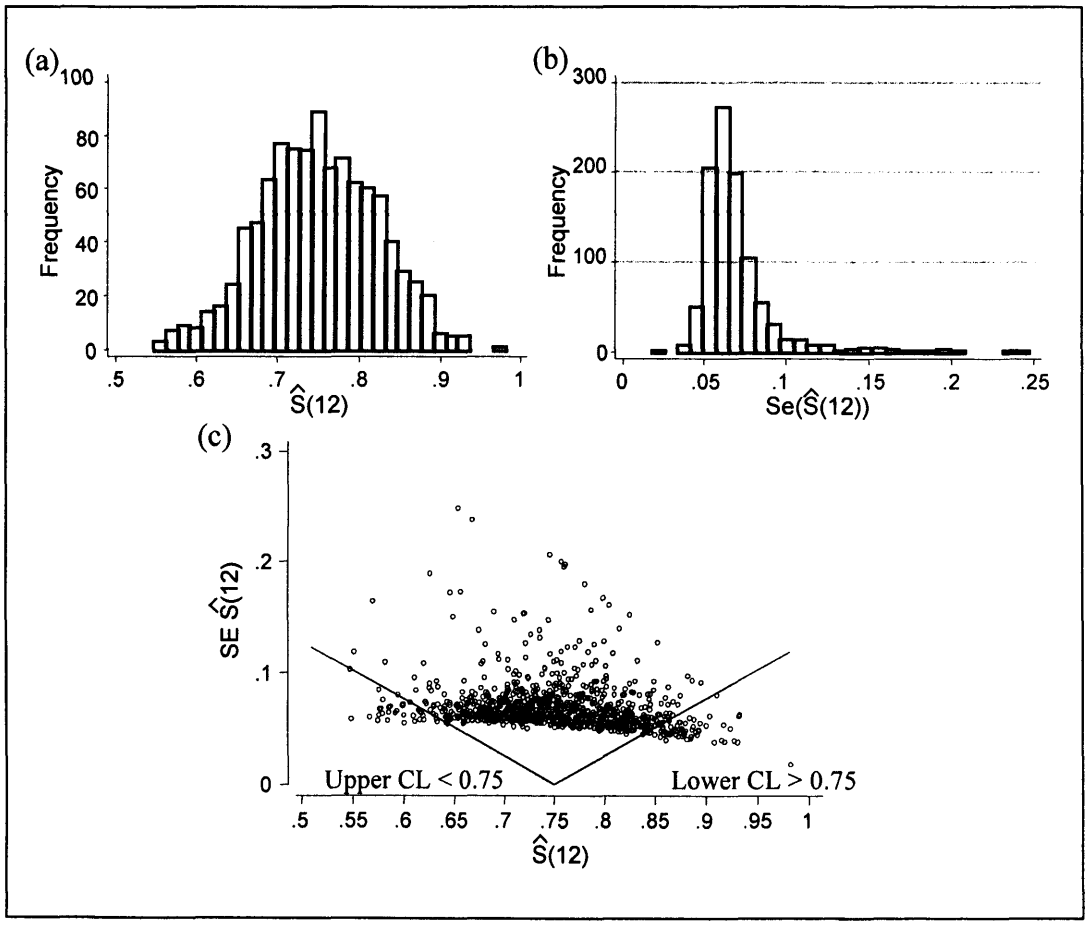

Figure 2.6: Relationship between Wald-2 standard error and point estimate over 1000 simulations ($\gamma=1.5$, $n=100$, $t=12$, $p=0.4$, $d=0$). Note that S(12)=0.75

Figure 2.6c shows the estimated standard error plotted against the point estimate from

each simulation. The left line represents values of the upper 95% confidence limit for a point estimate of 0.75 for different values of the standard error of the estimate. Similarly, the right line plots values of the lower 95% confidence limit at different values of the standard error. This figure shows that the inaccurate coverage was mainly due to the lower confidence limit exceeding the true survival probability (6.4% of simulations), rather than the upper confidence limit being too small (3.6% of simulations). This was a consequence of estimating smaller standard errors at high values of the point estimate.

The coverage achieved by the likelihood ratio method was consistently close to the nominal value of 95%, and even the most extreme values observed (93.9% and 96.8%) are compatible with expected binomial variation given parameters 1000 and 0.95. Overall coverage using the profile likelihood approach was 95.4% (Table 2.3). In addition, the distribution of the likelihood ratio statistic was examined as part of the simulation study. The likelihood ratio statistics from four different scenarios (with a range of estimated coverage probabilities) were found to follow closely the nominal χ_1^2-distribution up to values of approximately 10 (Figure 2.7). This implies that correct coverage would be achieved at any reasonable choice of significance level.

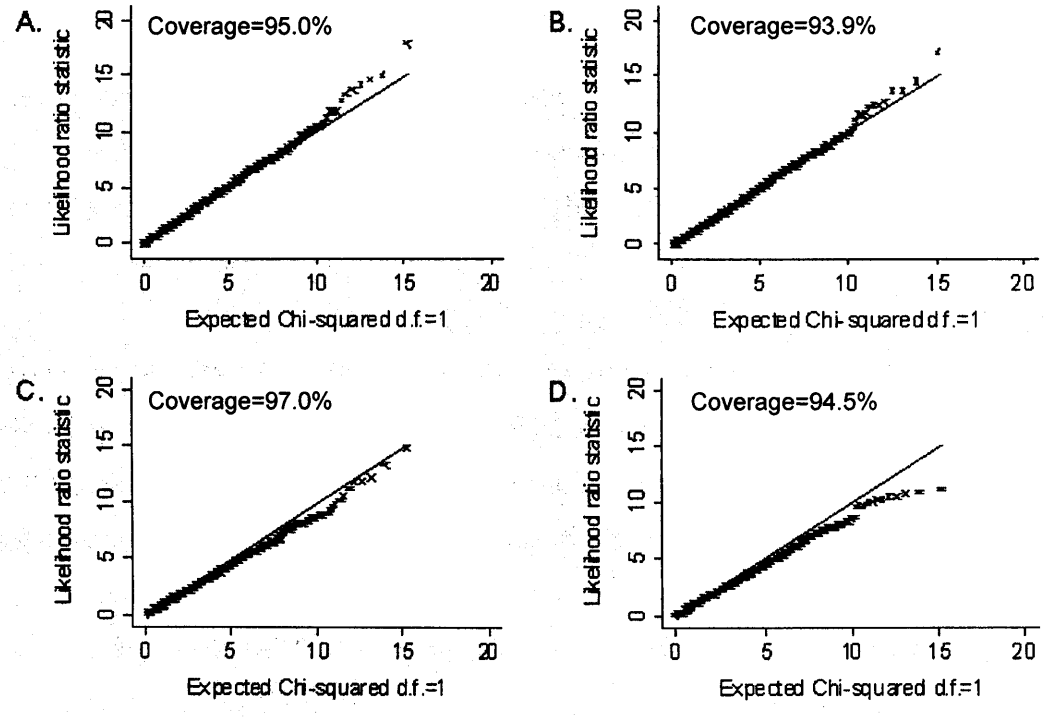

Figure 2.7: Chi-squared Q-Q plot of likelihood ratio statistics for 4 different simulated scenarios.

There was no evidence of systematic variation in coverage with respect to sample size, frequency of missing visits, Weibull shape parameter, time point and loss to follow-up rate for the likelihood ratio method (Table 2.3). However, for the Wald-2 method, coverage was more accurate as sample size increased e.g. 94.5% (n=800) compared to 92.9% (n=100). In addition, there was a significant effect of both time point and shape of the hazard distribution, with closer to nominal coverage at 48 weeks or with a decreasing hazard. The impact of a 10% loss to follow-up rate was minimal.

2.5 DISCUSSION

In this chapter, three methods for calculating pointwise confidence intervals for the non-parametric survivor function estimated from interval-censored data have been described and assessed: the first based on the full information matrix (Wald-1), the second a modification of this approach involving deletion of rows and columns of the information matrix corresponding to zero estimates prior to inversion (Wald-2), and the third based on likelihood ratio inference. The Wald-1 and Wald-2 methods were based on the observed rather than expected (Fisher) information, the latter being highly complex because the parameterisation is data dependent.

The simulation study showed clearly that the Wald-1 method substantially and consistently over-estimated the standard error (of the estimated survivor function at fixed time points), and therefore resulted in confidence intervals that were too wide, achieving an average coverage of 98.6% against a nominal value of 95%. Furthermore, this approach may give pathological results under extensive left and right censoring, as highlighted by the example in section 2.3. The Wald-1 method cannot therefore be generally recommended.

In contrast, the Wald-2 method produced accurate standard errors, using the Monte Carlo estimates as a benchmark. Despite this, confidence interval coverage was slightly less than the nominal value of 95%, a consequence of correlation between the estimated standard error and the point estimate – standard errors were smaller at higher values of the estimated survival probability. An underlying weakness of the Wald-2 approach, which also applies to the Wald-1 approach, is the use of symmetric confidence intervals for a probability [24]. However, this may be overcome by using a log-log transformation. For most of the combinations of factors considered, the coverage was, for practical purposes, acceptably close to 95%. However, for some combinations, the

coverage was unacceptably low (minimum value 90.0%), understating the true uncertainty in the point estimates. The key practical issue is identifying the circumstances when the Wald-2 method can and cannot be validly used. Since this method is based on asymptotic theory, coverage was more accurate for a sample size of 800 than for a sample size of 100. However, the information in a set of interval-censored observations reflects the width of the censoring intervals as well as sample size. Unexpectedly the "probability of missing a visit" parameter had little effect, although this was in the right direction with standard errors slightly larger when 40% of visits were missed compared to 20%. Adding to the complexity, coverage was also dependent on the shape of the survivor function and the centile of the survivor function. This suggests it may not be possible to identify a simple rule to guide when the Wald-2 method can be used reliably, and that the use of this approach should always be attached with caveats.

The likelihood ratio method gave the most accurate confidence intervals with coverage consistently close to the nominal level of 95%. This reassuring result was not predictable from a theoretical standpoint since a high proportion of the estimates lay on the boundary of the parameter space [16]. In addition, the distribution of the likelihood ratio statistic was found to follow closely the nominal χ_1^2-distribution up to values of approximately 10. This implies that correct coverage would be achieved at any reasonable choice of significance level.

The likelihood ratio approach also offers several theoretical advantages over Wald confidence intervals, for example, ensuring that confidence limits do not extend beyond the range of the parameter space [25]. The drawback with the likelihood ratio method is computational. It is not currently implemented in any of the standard statistical packages and therefore requires the use of computationally intensive ad hoc programs (see Appendix II) [26]. Also, in the simulations performed here it was found that "tolerance" parameters needed to be set to very small values to ensure convergence. Finally, in contrast to the methods based on the information matrix, the program has to be re-applied at each time point of interest.

Two alternative approaches based on re-sampling methods have been described by Sun [8]. The first was a simple non-parametric bootstrap, re-sampling with replacement from the original set of observations. The second approach employed multiple

imputation of the exact survival time for each individual, apart from those who were right-censored, given the individual censoring intervals and the maximum likelihood estimate of the survivor function. This produced a set of right-censored data to which the Greenwood formula [27] was applied. Both methods performed reasonably well in a small simulation study. However, the simplicity of re-sampling methods is compromised by the fact that the estimated survivor function is not uniquely determined at the time point of interest in some of the replications. The work in this chapter could be extended to compare the methods of confidence interval estimation discussed here with those proposed by Sun.

This chapter focuses on uncertainty in the estimated survivor function at fixed time points. However, there is often interest in the reverse problem – estimation of the time points that correspond to selected quantiles of the survivor function. For right-censored data, one approach calculates approximate confidence intervals using the first derivative of the survival distribution but this would not generalise easily to interval-censored data where the non-parametric distribution function is undefined at certain time points [27].

An undesirable feature of the non-parametric estimator for interval-censored data is sharp discontinuities in the survivor function. Smoothing techniques, allied to penalised likelihood inference, which have been developed to reduce this phenomenon, are likely to be increasingly used. It might be expected that this would result in significantly tighter confidence intervals, although at the cost of the assumptions underlying the smoothing algorithm. Little research has been conducted on confidence intervals for this approach, although re-sampling methods may offer the only tractable solution. Alternatively, the possibility of using full parametric survival models should not be overlooked. Inference for such models, including the derivation of confidence intervals, is relatively straightforward. Both these methods are compared to the NPMLE in chapter 3. Nevertheless, non-parametric methods will continue to play an important role in informing the selection of an appropriate family of distributions and in assessing goodness-of-fit [28].

References

1. Peto R. Experimental survival curves for interval-censored data. *Applied Statistics* 1973, **22**:86-91.

2. Turnbull BW. The empirical distribution function with arbitrary grouped, censored and truncated data. *Journal of the Royal Statistical Society B* 1976, **38**:290-295.

3. Bohning S, Schlattmann P, Dietz E. Interval censored data: A note on the nonparametric maximum likelihood estimator of the distribution function. *Biometrika* 1996, **83**(2):462-466.

4. Groeneboom P, Wellner JA. *Information bounds and nonparametric maximum likelihood estimation.* Birkhauser Verlag: Basel, 1992

5. Gentleman R, Geyer CJ. Maximum likelihood for interval censored data - consistency and computation. *Biometrika* 1994, **81**(3):618-623.

6. Huang J, Wellner JA. Interval censored survival data: a review of recent progress. *Lecture Notes in Statistics: Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*. Springer-Verlag: New York, 1997: pp 123-169.

7. Yu Q, Li L, Wong GWC. On consistency of the self-consistent estimator of survival functions with interval-censored data. *Scandanavian Journal of Statistics* 2000, **27** (1):35-44.

8. Sun J. Variance estimation of a survival function for interval-censored survival data. *Statistics in Medicine* 2001, **20**(8):1249-1257.

9. Lindsey JC, Ryan LM. Tutorial in biostatistics - Methods for interval-censored data. *Statistics in Medicine* 1998, **17**(2):219-238.

10. Finkelstein DM. A proportional hazards model for interval-censored failure time data. *Biometrics* 1986, **42**(4):845-854.

11. Thomas DR, Grunkemeier GL. Confidence Interval Estimation of Survival Probabilities for Censored Data. *Journal of the American Statistical Association* 1975, 663-669.

12. Alioum A, Commenges D. A proportional hazards model for arbitrarily censored and truncated data. *Biometrics* 1996, **52**(2):512-524.

13. Sun J. Interval censoring. *Encyclopedia of Biostatistics*. John Wiley & Sons Ltd: 2005: pp 2603-2609.

14. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B* 1977, **39**:1-38.

15. Neyman J, Scott EJ. Consistent estimates based on partially consistent observations. *Econometrica* 1948, **16**:1-32.

16. Self SG, Liang KY. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non-standard conditions. *Journal of the American Statistical Association* 1987, **82**(398):605-610.

17. SAS Institute Inc., Cary, North Carolina, USA. 1999, (8.2).

18. ICE.sas. *http://support.sas.com/ctx/samples/index.jsp?sid=477&tab=about* 2006. 6-17-2006.

19. Delta Coordinating Committee. Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. *Lancet* 1996, **348**:283-291.

20. Delta Coordinating Committee and Delta Virology Committee. HIV-1 RNA response to antiretroviral treatment in 1280 participants in the Delta Trial: an extended virology study. *AIDS* 1999, **13**(1):57-65.

21. The AVANTI Steering Committee. Analysis of HIV-1 clinical trials: statistical magic? *Lancet* 1999, **353**(9169):2061-2064.

22. Ledergerber B, Egger M, Opravil M, Telenti A, Hirschel B, Battegay M, Vernazza P, Sudre P, Flepp M, Furrer H, Francioli P, Weber R. Clinical progression and virological failure on highly active antiretroviral therapy in HIV-1 patients: a prospective cohort study. Swiss HIV Cohort Study. *Lancet* 1999, **353**(9156):863-868.

23. Paredes R, Mocroft A, Kirk O, Lazzarin A, Barton SE, van Lunzen J, Katzenstein TL, Antunes F, Lundgren JD, Clotet B. Predictors of virological success and ensuing failure in HIV-positive patients starting highly active antiretroviral therapy in Europe: results from the EuroSIDA study. *Archives of Internal Medicine* 2000, **160**:1123-1132.

24. Vollset SE. Confidence intervals for a binomial proportion. *Statistics in Medicine* 1993, **12**:809-824.

25. Cox DR, Snell EJ. Appendix 1. *Analysis of Binary Data.* Chapman and Hall: London, 1989

26. Alho JM. On the computation of likelihood ratio and score test based confidence intervals in generalized linear models. *Statistics in Medicine* 1992, **11**(7):923-930.

27. Collett D. *Modelling Survival Data in Medical Research.* Chapman and Hall: London, 2003

28. De Gruttola V, Lagakos SW. Analysis of doubly-censored survival data, with application to AIDS. *Biometrics* 1989, **45**(1):1-11.

CHAPTER 3: ESTIMATING THE SURVIVOR FUNCTION FOR A SINGLE SAMPLE

3.1 INTRODUCTION

The non-parametric maximum likelihood estimator (NPMLE, section 2.2.2) for interval-censored data is a step function that tends to have a small number of jumps and hence large jump sizes. Smooth estimators may therefore be more desirable when the underlying survivor function can be assumed to be smooth. Non-parametric smoothing techniques used in the estimation of the hazard or survivor function include kernel smoothers, splines, and local or penalized likelihood estimates.

Kernel smoothing has been well studied when the data are right censored and Pan [1] applies this method to interval-censored data. The kernel estimator of the density function of the NPMLE is integrated to give the kernel estimator of the survivor function. The main difficulty with this approach is the choice of an appropriate bandwidth, especially since some of the proposed bandwidth selection methods for right-censored data can no longer be implemented as there is generally no explicit form of the NPMLE when data are interval-censored. Pan uses a 10-fold cross-validation technique and the Parzen kernel in a small simulation study, where he compares the kernel estimator to the NPMLE and the log-spline density estimator proposed by Kooperberg and Stone [2]. Here the density function is directly modeled as a smooth function using cubic splines, which may be more efficient than smoothing the discrete NPMLE. Kooperberg and Stone provide S-plus [3] functions to fit the log-spline density which include algorithms for selection of the number and placement of knots that use the Akaike or Bayesian Information Criterion for final model selection. Results from the simulations study (where sample size=100, number of follow up visits=1 or 3, and initially 4 knots are used for the log-spline density) indicate that all 3 methods have small bias, but that both smoothed estimators have reduced variability compared to the NPMLE. The calculation of confidence intervals for either method is not discussed, but Rosenberg [4] uses bootstrap percentile confidence intervals for spline estimation of the

hazard function for interval-censored data. However, Pan notes that software for the log-spline density model does not support small to medium or heavy censored samples.

Betensky et al. [5] describe local likelihood estimation of the hazard function when data are interval-censored. Here, the local likelihood estimate of the hazard function at time t is based on a polynomial approximation to the log hazard in a smoothing window with defined bandwidth around t. As estimation is relatively straightforward when the data are right-censored a local EM algorithm is used to facilitate estimation when data are interval-censored. However, the method is restricted to using locally constant approximations to the log hazard function with a Gaussian kernel weighting function due to computation complexities. Since this may lead to biases at the boundaries, Bebchuck et al. [6] propose an iterative multiple imputation algorithm that treats the interval-censored observations as missing data, imputes values for them and then obtains estimates for the more tractable right-censored data problem. This process is iterated until convergence. Both methods are highly computationally intensive.

Lesaffre et al [7] describe a maximum likelihood based approach for the accelerated failure time model for interval-censored data that exploits penalised smoothing of the baseline density. A much simpler approach that is closely related to maximum penalised likelihood estimation was proposed by Pan and Chappell [8]. By introducing a smoothing or S-step into the EM algorithm when estimating the NPMLE, they obtain an estimator that is not smooth in the usual sense but is "smoother" in that the jump sizes do not vary so rapidly. The resulting EMS algorithm [9], being the simplest of the above smoothing methods to apply, is discussed further in section 3.2.5.

Despite the wide range of approaches to the analysis of interval-censored failure-time data, in practice the use of the simple imputation methods predominate, partly due to the fact they are conceptually easier for the non-expert and as few of the alternative approaches are implemented in the major statistical packages. This chapter considers the use of the standard Kaplan-Meier method based on right-endpoint (RKM) and the less commonly employed midpoint (MKM) imputation for the estimation of the survivor function for a single sample. These methods are compared with the NPMLE, the smoothed estimator of Pan and Chappell and a fully parametric model. In addition, a new smoothed estimator that is obtained through a modification of method proposed by Pan and Chappell is assessed.

One objective of the work in this chapter is to identify circumstances under which the relatively straightforward mid-point approach achieves sufficiently accurate estimates of the survivor function compared with more complex approaches. Although mid-point imputation for interval-censored failure-time data has been investigated in a number of studies, this question has not been directly addressed.

Odell et al [10] compared an accelerated failure time Weibull model based on the full interval-censored data and a Weibull model based on midpoint imputation. Simulation studies focus mainly on estimation of a single continuous covariate, which will be discussed further in chapter 4. However, the authors state that their results indicate that if interest lies in estimation of the location or scale parameter, the interval-based model is almost always superior, especially for decreasing hazards. The paper did not consider any non-parametric approaches.

Dorey et al [11] considered a continuous outcome where the endpoint is the time of crossing a certain threshold value. Their main approach was based on multiple imputation of the threshold-crossing time with use of models that took into account the continuous nature of the measurements. These were found to be superior to using midpoint imputation in two example datasets.

Law and Brookmeyer [12] considered the case of doubly censored data, where both the time origin (e.g. date of HIV infection) and the event time (e.g. dates of AIDS diagnosis) are interval-censored. They investigate the impact of using midpoint imputation for the time of infection on: the asymptotic bias and coverage of Kaplan-Meier estimates of the latency distribution, the bias in estimation of hazard ratios, and the size/power of the log-rank test. From results of a simulation study (where infection distribution = exponential or log-logistic; latency distribution = Weibull, median 10 years; sample size=100; time between visits=1, 2, 4 or 8 years) the authors conclude that midpoint imputation is reasonably accurate when the gap between visits is no more than 2 years.

In addition to examining the performance of the simple imputation approaches, a simulation study, in which a number of factors – the true survivor function, sample size, and the frequency and pattern on inspection times – are allowed to vary (section 3.3), was used to investigate the following questions: (i) the right endpoint method is clearly

biased, but by how much? (ii) how much accuracy is gained, if any, by using a smoothed rather than a non-smoothed non-parametric approach? (iii) if the correct parametric model is adopted, what is the gain in accuracy compared with the optimal non-parametric approach? An illustrative data set is used to exemplify the different methods (section 3.4), before a summary of the main findings and their practical implications (section 3.5).

3.2 NOTATION AND THEORY

3.2.1 Observed data

The same notation as in section 2.2.1 is used. In addition, let $M_i = (L_i+R_i)/2$ denote the midpoint of the observed censoring interval

3.2.2 Imputing event times

By replacing the observed censoring interval by a single time point it is possible to implement methods available for the analysis of right-censored data. Non-parametric estimators of the survivor function are obtained by applying the standard Kaplan-Meier method [13] to data where the censoring intervals are replaced by i) the censoring interval right endpoint, R_i (RKM), or ii) the midpoint, M_i (MKM).

3.2.3 Parametric Weibull model (W)

Parametric models are easily fitted to interval-censored data. The family of Weibull distributions has a monotonic hazard function that is defined by a shape parameter γ, and a scale parameter λ, and is widely used in the parametric analysis of survival data since the hazard function can take on a wide variety of forms depending upon the value of the shape parameter (Figure 3.1). The hazard function is decreasing, constant or increasing when $\gamma<1$, $\gamma=1$ or $\gamma>1$ respectively.

The survivor function is given by $S(t) = \exp(-\lambda t^{\gamma})$.The log-likelihood function is then given by

$$\ell(\lambda,\gamma) = \sum_{i=1}^{n} \log\left(S(L_i) - S(R_i -)\right)$$

$$= \sum_{i=1}^{n} \log\left(\exp(-\lambda L_i^{\gamma}) - \exp(-\lambda R_i^{\gamma})\right).$$

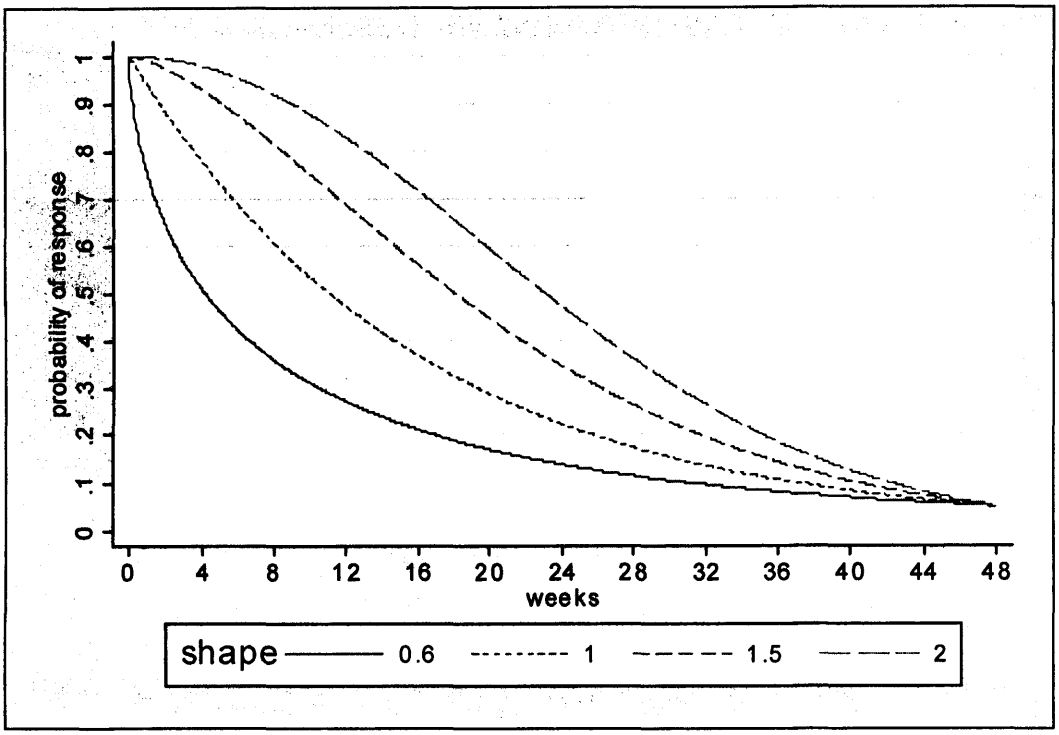

Figure 3.1: Underlying Weibull survivor functions used in simulation study

3.2.4 Non-parametric maximum likelihood estimator (NPMLE)

As described in section 2.2.2, and returning to that notation, let $\{u_j; j=1,...,m+1\}$ denote the unique ordered values of $\{0,\{L_i\},\{R_i\}\}$ and $f_j=S(u_j)-S(u_{j+1})$ for $j=1,...,m$. Define an indicator variable $\alpha_{ij}=1$ if $(u_j, u_{j+1})\subseteq I_i$ and 0 otherwise. The log-likelihood function is

$$\ell(\mathbf{f}) = \sum_{i=1}^{n} \log\left(\sum_{j=1}^{m} \alpha_{ij} f_j\right) \quad \text{subject to } f_j \geq 0,\ \sum_{j=1}^{m} f_j = 1 \qquad (3.1)$$

As discussed in section 2.2.2, there is no closed form solution for the maximum likelihood estimate (MLE) $\hat{\mathbf{f}} = (\hat{f}_1 ... \hat{f}_m)$ and an iterative method is required, such as the EM algorithm [14,15]. Define an indicator function $I_{ij} = 1$ if $T_j \in I_i$, 0 otherwise, which constitutes the "complete data". The E-step calculates the conditional expectation of person i, $i=1,\ldots,n$, experiencing an event in interval j, $j=1,\ldots,m$, given the observed data and the current estimates of $\hat{\mathbf{f}} = (\hat{f}_1 ... \hat{f}_m)$, namely

$$E[I_{ij}] = \alpha_{ij}\hat{f}_j \Big/ \left(\sum_{j=1}^{m} \alpha_{ij}\hat{f}_j\right)$$

Considering $E[I_{ij}]$ as an observed rather than expected frequency, the M-step is simply to estimate the proportion of events in interval j,

$$\hat{f}_j = \frac{1}{n}\sum_{i=1}^{n} E[I_{ij}]$$

The algorithm iterates between (2.3) are (2.4) until convergence. It can be shown that $\hat{f}_j \neq 0$ only if $u_j \in \{L_i\}$ and $u_{j+1} \in \{R_i\}$ if this condition is not satisfied then $\hat{f}_j = 0$ *a priori* and these elements can be eliminated from the model [16].

3.2.5 Smoothed NPMLE (SNP & SNP2)

The EMS algorithm [9,17] is a simple modification of the EM algorithm which adds a smoothing (S) step after the expectation (E) and maximisation (M) steps. The S-step is often realised as a weighted average of neighbouring parameter estimates. Both Pan and Chappell [8] and Leung and Elashoff [18], who use an EMS algorithm for the analysis of doubly censored data, propose 3-point smoothing as this provides good estimates without sacrificing efficiency. Pan & Chappell obtained similar results using 5-point and 3-point smoothing in simulations and 3-point smoothing is used throughout this thesis.

Applying the EMS algorithm to estimation of the NPMLE as described above, and denoting the parameter estimates at the k-th iteration by $\hat{\mathbf{f}}^{(k)} = \left(\hat{f}_1^{(k)} \ldots \hat{f}_m^{(k)}\right)$, the S-step uses a triangular kernel

$$\tilde{\hat{f}}_i^{(k)} = \frac{1}{4}\left[\hat{f}_{i-1}^{(k)} + 2\hat{f}_i^{(k)} + \hat{f}_{i+1}^{(k)}\right]; (2 \leq i < m\text{-}1).$$

To ensure that the smoothed estimator constitutes a proper distribution function, estimates at the two endpoints are given by

$$\tilde{\hat{f}}_1^{(k)} = \left(3\hat{f}_1^{(k)} + \hat{f}_2^{(k)}\right)/4 \ , \ \tilde{\hat{f}}_m^{(k)} = \left(3\hat{f}_m^{(k)} + \hat{f}_{m-1}^{(k)}\right)/4 .$$

If however the last interval is infinite due to right censoring, i.e. $u_{m+1}=\infty$, the last parameter is excluded from the smoothing step since it is determined by equation 3.1.

The addition of the S-step speeds up the convergence rate of the EM algorithm and the resulting smooth non-parametric MLE (SNP) from the EMS algorithm can be shown to be an extremum of a penalized likelihood. In a small simulation study where there is only one follow-up visit, Pan & Chappell compare the NPMLE and SNP estimators and conclude SNP performs satisfactorily in terms of bias and variation.

By definition, the EMS algorithm smoothes away any parameter estimates that would have been estimated as zero by an EM algorithm. An issue that arises in this approach is how to handle the elements of **f** that are zero *a priori* (section 3.2.4). Pan and Chappell

ignore these elements as recommended by Peto and Turnbull when fitting the NPMLE. The rationale for this is unclear since it is merely the arbitrary set of individual censoring intervals that determines whether a particular element of **f** is zero *a priori*. It would be expected that including these elements in the algorithm would enhance the smoothness of the estimator. Therefore, two smoothed estimators have been examined. Firstly, the *a priori* zero elements were excluded before estimation and the EMS algorithm fitted as described by Pan and Chappell to estimate the SNP as outlined above. Secondly, an alternative smooth estimator (SNP2) was fitted to the data by estimating a parameter for each of the intervals defined by the unique ordered values of $\{0,\{L_i\},\{R_i\}\}$ and NOT excluding the a priori elements.

3.3 SIMULATION STUDY

3.3.1 Methods

A simulation study, with similarities to that of section 2.4, was carried out to assess the accuracy of the various estimation methods described in section 3.2. Data were generated from a hypothetical follow-up study of duration D with v evenly spaced scheduled follow-up visits at $t=D/v, 2D/v, \ldots, D$. Each visit was missed at random with a fixed probability p. In addition, to allow for the fact that a patient may not attend on the scheduled date, the actual visit time was assumed to be normally distributed, centred at the scheduled visit time with standard deviation $D/(3*v)$. Survival times were generated from a Weibull distribution and the individual censoring intervals determined.

1,000 data sets were simulated for each combination of the following factors:

(a) Shape parameter of Weibull distribution, γ=0.6, 1.0, 1.5 representing a decreasing, constant, and increasing hazard function respectively. For each value of the shape parameter, the scale parameter, λ, was selected so that S(D)=0.05.
(b) Sample size, n=25, 50, 100, 200, 400, 800 individuals.
(c) Probability of missing a scheduled visit, p=0, 0.4.
(d) Number of evenly spaced scheduled follow up visits, v=3, 6, 12.

The survivor function was then estimated for each of the six methods described in section 3.2 (RKM, MKM, W, NPMLE, SNP, SNP2). Following Aalen [19], the accuracy of the estimated survivor function was defined as the absolute difference between the estimated cumulative hazard function, -log($\hat{S}(t)$), and the true cumulative

hazard $-\log(S(t)) = -\lambda t^{\gamma}$ integrated over (and standardised for) the duration of the study i.e.

$$\frac{1}{D}\int_0^D \left| -\log(\hat{S}(t)) - \lambda t^{\gamma} \right| dt \qquad (3.2)$$

A composite trapezoidal rule with 673 segments was used to calculate the metric (3.2). For each combination of γ, n, p and v, the metric (3.2) was averaged over the 1000 simulations, using a 10% trimmed mean.

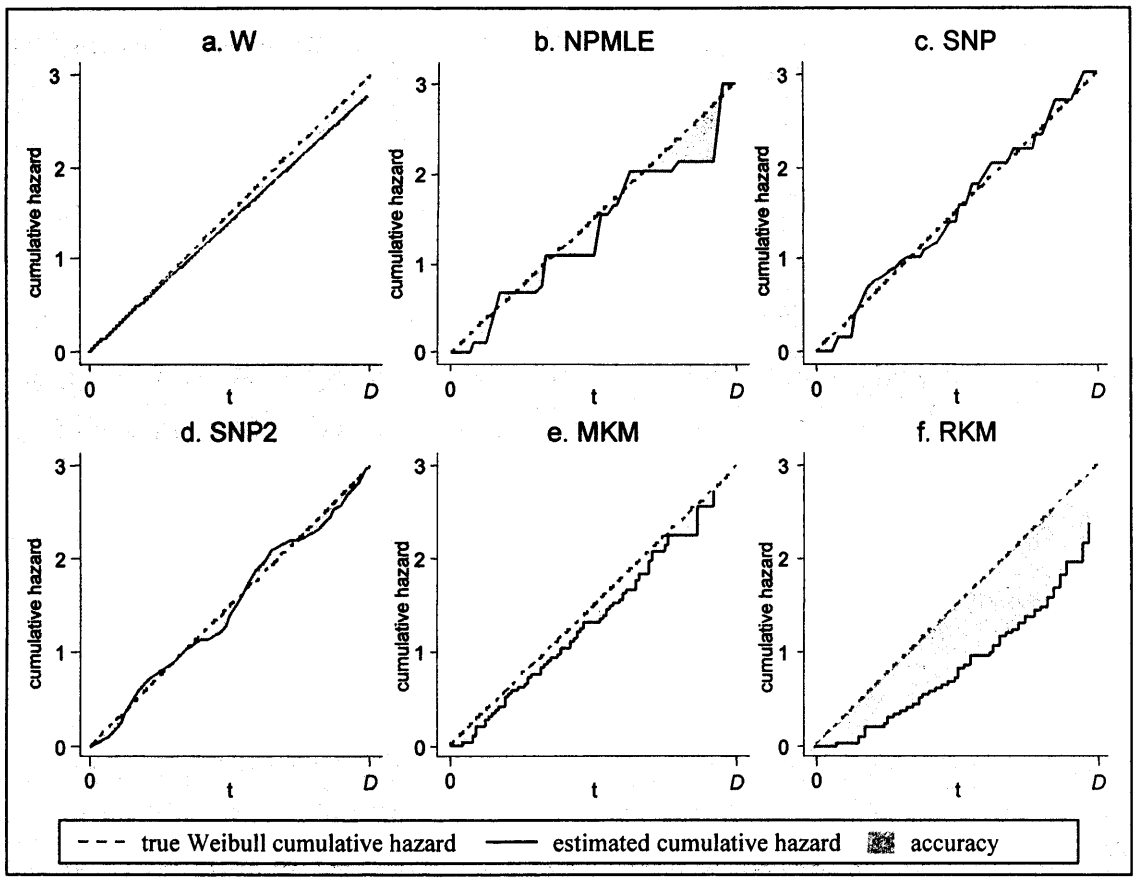

Figure 3.2: Illustration of accuracy metric for different methods of estimating the survivor function with interval-censored data

3.3.2 Illustration

Figure 3.2 shows the estimated cumulative hazard function for each of the six methods for a single simulation (γ=1, n=100, p=0.4, v=6), where the shaded area corresponds to the metric (3.2). The highly discrete nature of the NPMLE can be clearly seen in this example (Figure 3.2b). As was noted in section 3.2.4, $\hat{S}$ is undefined within intervals (u_j,u_{j+1}) for which $\hat{f}_j \neq 0$ for the NPMLE, SNP and SNP2 methods. Following convention $\hat{S}(t)$ was estimated by linear interpolation in these intervals, although the

survivor function is only strictly defined for times where the curve is horizontal.

Smoothing the NPMLE reduces the variation in step sizes and introduces additional steps by assigning mass to intervals where previously $\hat{f}_i = 0$. The resulting curves are therefore smoother than the NPMLE (Figure 3.2c, d), and the SNP2 estimator is smoother than the SNP estimator due to the increased number of parameters in the model. The severe underestimation of the cumulative hazard from the RKM method (Figure 3.2f) and the improved accuracy from fitting a correctly specified parametric model (Figure 3.2a) are apparent.

3.3.3 Implementation

All analyses were implemented using SAS Version 8.2 [20]. A program available from the SAS/IML sample library (ICE.sas) [21] was modified to fit the NPMLE using a quasi-Newton optimisation procedure and the SNP and SNP2 methods using an EMS algorithm (Appendices 3.A & 3.B). The Weibull model was fitted using PROC Lifereg and the MKM and RKM methods using PROC Lifetest.

3.3.4 Results

The mean accuracy metric for each method is given in Table 3.1. These reflect three characteristics of each estimator: bias, variance, and smoothness. It can be demonstrated that the exponentiated value can be roughly interpreted as the geometric mean of the ratio of the estimated survivor function to the true survivor function (inverted if <1) at any given time point. The ratio of the accuracy metric for selected comparisons of estimators are given in Table 3.1: the RKM method is reported relative to MKM, the SNP2 method relative to the Weibull, and the NPMLE, SNP and MKM methods relative to SNP2. These can be directly interpreted on a survivor function scale. For example, the value of 1.07 (= 0.554/0.520) for the NPMLE relative to the SNP2 estimator (Table 3.1, first row) implies that the error in estimating the survivor function at any time t is, on average, 7% higher for NPMLE than for SNP2. The standard errors of the ratios of mean accuracy metric were calculated using a Taylor series expansion.

As expected, general trends of increased accuracy as n or v increased and as p or s decreased were seen for all methods. Results for γ=1.0 and γ=1.5 were similar (Table 3.1, Figures 3.3-3.5).

Table 3.1: Simulation results: Accuracy metric

					Mean					Ratio of means [standard error]				
Shape	n	p	v	W	NP	SNP	SNP2	MKM	RKM	SNP2/W	NP/SNP2	SNP/SNP2	MKM/SNP2	RKM/MKM
0.6	25	0	3	0.447	0.554	0.551	0.520	0.486	0.749	1.16 [0.015]	1.07 [0.011]	1.06 [0.012]	0.94 [0.009]	1.54 [0.018]
0.6	25	0	6	0.369	0.448	0.441	0.419	0.425	0.502	1.14 [0.012]	1.07 [0.007]	1.05 [0.007]	1.02 [0.006]*	1.18 [0.013]
0.6	25	0	12	0.358	0.414	0.407	0.391	0.412	0.431	1.09 [0.012]	1.06 [0.004]	1.04 [0.004]	1.06 [0.005]	1.04 [0.008]
0.6	25	0.4	3	0.552	0.670	0.672	0.673	0.606	1.090	1.22 [0.024]	1.00 [0.013]*	1.00 [0.014]*	0.90 [0.014]	1.80 [0.017]
0.6	25	0.4	6	0.402	0.528	0.518	0.499	0.473	0.718	1.24 [0.019]	1.06 [0.013]	1.04 [0.012]	0.95 [0.012]	1.52 [0.017]
0.6	25	0.4	12	0.361	0.443	0.433	0.415	0.419	0.501	1.15 [0.012]	1.07 [0.006]	1.04 [0.006]	1.01 [0.007]*	1.20 [0.014]
1	25	0	3	0.327	0.475	0.447	0.399	0.356	0.598	1.22 [0.016]	1.19 [0.012]	1.12 [0.011]	0.89 [0.010]	1.68 [0.020]
1	25	0	6	0.269	0.363	0.346	0.318	0.319	0.384	1.18 [0.014]	1.14 [0.008]	1.09 [0.007]	1.00 [0.006]*	1.20 [0.015]
1	25	0	12	0.268	0.331	0.322	0.307	0.323	0.330	1.15 [0.012]	1.08 [0.004]	1.05 [0.004]	1.05 [0.004]	1.02 [0.010]*
1	25	0.4	3	0.404	0.596	0.567	0.527	0.437	0.860	1.30 [0.022]	1.13 [0.017]	1.08 [0.016]	0.83 [0.014]	1.97 [0.021]
1	25	0.4	6	0.311	0.449	0.420	0.384	0.356	0.566	1.23 [0.016]	1.17 [0.012]	1.09 [0.011]	0.93 [0.010]	1.59 [0.021]
1	25	0.4	12	0.288	0.375	0.361	0.339	0.332	0.395	1.18 [0.013]	1.11 [0.006]	1.06 [0.006]	0.98 [0.007]	1.19 [0.016]
1.5	25	0	3	0.241	0.408	0.374	0.324	0.280	0.525	1.34 [0.018]	1.26 [0.013]	1.15 [0.012]	0.86 [0.011]	1.87 [0.021]
1.5	25	0	6	0.229	0.317	0.302	0.275	0.272	0.334	1.20 [0.013]	1.15 [0.007]	1.10 [0.007]	0.99 [0.006]*	1.23 [0.017]
1.5	25	0	12	0.214	0.269	0.262	0.250	0.258	0.266	1.17 [0.013]	1.08 [0.005]	1.05 [0.004]	1.03 [0.005]	1.03 [0.012]
1.5	25	0.4	3	0.308	0.530	0.492	0.412	0.346	0.702	1.34 [0.021]	1.29 [0.017]	1.19 [0.016]	0.84 [0.012]	2.03 [0.025]
1.5	25	0.4	6	0.245	0.389	0.360	0.324	0.290	0.484	1.33 [0.019]	1.20 [0.012]	1.11 [0.012]	0.89 [0.010]	1.67 [0.020]
1.5	25	0.4	12	0.227	0.313	0.297	0.278	0.271	0.336	1.22 [0.015]	1.13 [0.007]	1.07 [0.007]	0.98 [0.007]	1.24 [0.016]
2	25	0	3	0.181	0.363	0.333	0.276	0.250	0.482	1.53 [0.023]	1.32 [0.013]	1.21 [0.012]	0.91 [0.010]	1.93 [0.020]
2	25	0	6	0.165	0.273	0.254	0.227	0.222	0.296	1.38 [0.017]	1.20 [0.008]	1.12 [0.007]	0.98 [0.006]	1.33 [0.018]
2	25	0	12	0.165	0.235	0.228	0.215	0.221	0.225	1.30 [0.015]	1.09 [0.005]	1.06 [0.005]	1.03 [0.005]	1.02 [0.012]*
2	25	0.4	3	0.214	0.471	0.436	0.342	0.292	0.598	1.60 [0.028]	1.38 [0.017]	1.28 [0.016]	0.85 [0.013]	2.05 [0.026]
2	25	0.4	6	0.185	0.344	0.316	0.275	0.251	0.427	1.48 [0.021]	1.25 [0.011]	1.15 [0.010]	0.91 [0.009]	1.70 [0.022]
2	25	0.4	12	0.170	0.274	0.259	0.237	0.230	0.300	1.39 [0.019]	1.16 [0.008]	1.09 [0.007]	0.97 [0.007]	1.31 [0.017]
0.6	50	0	3	0.285	0.389	0.385	0.360	0.374	0.711	1.27 [0.012]	1.08 [0.007]	1.07 [0.007]	1.04 [0.007]	1.90 [0.018]
0.6	50	0	6	0.257	0.321	0.314	0.298	0.308	0.419	1.16 [0.011]	1.08 [0.006]	1.05 [0.006]	1.03 [0.005]	1.36 [0.016]
0.6	50	0	12	0.248	0.299	0.291	0.285	0.296	0.332	1.15 [0.012]	1.05 [0.005]	1.02 [0.005]	1.04 [0.004]	1.12 [0.011]
0.6	50	0.4	3	0.337	0.486	0.484	0.456	0.509	1.095	1.35 [0.018]	1.06 [0.010]	1.06 [0.011]	1.12 [0.012]	2.15 [0.015]
0.6	50	0.4	6	0.275	0.380	0.356	0.343	0.368	0.675	1.25 [0.013]	1.11 [0.009]	1.04 [0.008]	1.07 [0.009]	1.84 [0.019]

* no evidence of a difference between the methods at P<0.01

					Mean					Ratio of means [standard error]				
Shape	n	p	v	W	NP	SNP	SNP2	MKM	RKM	SNP2/W	NP/SNP2	SNP/SNP2	MKM/SNP2	RKM/MKM
0.6	50	0.4	12	0.242	0.309	0.298	0.289	0.298	0.419	1.19 [0.012]	1.07 [0.006]	1.03 [0.005]	1.03 [0.006]	1.41 [0.016]
1	50	0	3	0.205	0.326	0.290	0.267	0.258	0.591	1.30 [0.015]	1.22 [0.008]	1.09 [0.007]	0.97 [0.008]	2.29 [0.024]
1	50	0	6	0.197	0.267	0.249	0.235	0.236	0.336	1.19 [0.012]	1.13 [0.006]	1.06 [0.005]	1.00 [0.005]*	1.43 [0.019]
1	50	0	12	0.188	0.239	0.230	0.224	0.230	0.252	1.19 [0.012]	1.07 [0.005]	1.03 [0.004]	1.03 [0.004]	1.10 [0.013]
1	50	0.4	3	0.246	0.415	0.374	0.337	0.335	0.872	1.37 [0.017]	1.23 [0.011]	1.11 [0.010]	0.99 [0.009]*	2.61 [0.022]
1	50	0.4	6	0.210	0.320	0.285	0.271	0.258	0.540	1.29 [0.016]	1.18 [0.008]	1.05 [0.008]	0.95 [0.008]	2.09 [0.025]
1	50	0.4	12	0.202	0.270	0.252	0.245	0.241	0.348	1.21 [0.013]	1.10 [0.006]	1.03 [0.006]	0.98 [0.006]*	1.44 [0.018]
1.5	50	0	3	0.169	0.285	0.246	0.223	0.211	0.531	1.32 [0.015]	1.28 [0.009]	1.10 [0.008]	0.94 [0.008]	2.52 [0.023]
1.5	50	0	6	0.156	0.224	0.200	0.191	0.187	0.289	1.23 [0.013]	1.17 [0.007]	1.05 [0.006]	0.98 [0.006]	1.55 [0.021]
1.5	50	0	12	0.145	0.192	0.184	0.177	0.181	0.210	1.22 [0.012]	1.09 [0.005]	1.04 [0.004]	1.02 [0.004]	1.16 [0.015]
1.5	50	0.4	3	0.207	0.377	0.332	0.291	0.249	0.708	1.40 [0.019]	1.30 [0.012]	1.14 [0.011]	0.86 [0.010]	2.84 [0.031]
1.5	50	0.4	6	0.158	0.267	0.231	0.215	0.195	0.473	1.36 [0.016]	1.24 [0.009]	1.07 [0.007]	0.90 [0.007]	2.43 [0.027]
1.5	50	0.4	12	0.160	0.223	0.204	0.197	0.191	0.287	1.23 [0.013]	1.13 [0.007]	1.03 [0.006]	0.97 [0.005]	1.50 [0.021]
2	50	0	3	0.131	0.253	0.217	0.193	0.185	0.489	1.48 [0.021]	1.31 [0.009]	1.12 [0.008]	0.96 [0.008]	2.64 [0.024]
2	50	0	6	0.121	0.199	0.175	0.167	0.162	0.278	1.38 [0.017]	1.19 [0.007]	1.05 [0.005]	0.97 [0.006]	1.71 [0.022]
2	50	0	12	0.118	0.169	0.160	0.152	0.157	0.178	1.29 [0.015]	1.11 [0.005]	1.05 [0.005]	1.03 [0.004]	1.13 [0.016]
2	50	0.4	3	0.152	0.327	0.289	0.238	0.207	0.616	1.57 [0.024]	1.37 [0.012]	1.21 [0.011]	0.87 [0.009]	2.97 [0.034]
2	50	0.4	6	0.136	0.253	0.220	0.203	0.179	0.422	1.49 [0.022]	1.25 [0.009]	1.09 [0.008]	0.88 [0.007]	2.36 [0.030]
2	50	0.4	12	0.128	0.197	0.179	0.171	0.166	0.266	1.34 [0.016]	1.15 [0.007]	1.05 [0.006]	0.97 [0.006]	1.60 [0.021]
0.6	100	0	3	0.184	0.281	0.274	0.257	0.295	0.708	1.40 [0.015]	1.09 [0.006]	1.07 [0.005]	1.15 [0.007]	2.40 [0.018]
0.6	100	0	6	0.172	0.228	0.218	0.209	0.222	0.372	1.22 [0.012]	1.09 [0.005]	1.04 [0.004]	1.06 [0.005]	1.67 [0.019]
0.6	100	0	12	0.162	0.202	0.197	0.192	0.200	0.249	1.18 [0.012]	1.05 [0.004]	1.03 [0.004]	1.04 [0.003]	1.25 [0.014]
0.6	100	0.4	3	0.220	0.348	0.335	0.312	0.441	1.101	1.42 [0.015]	1.11 [0.006]	1.07 [0.006]	1.41 [0.009]	2.50 [0.014]
0.6	100	0.4	6	0.184	0.266	0.242	0.235	0.288	0.651	1.27 [0.013]	1.13 [0.006]	1.03 [0.005]	1.23 [0.008]	2.26 [0.019]
0.6	100	0.4	12	0.169	0.220	0.208	0.203	0.219	0.365	1.20 [0.011]	1.08 [0.004]	1.03 [0.004]	1.08 [0.006]	1.67 [0.019]
1	100	0	3	0.142	0.240	0.199	0.192	0.200	0.599	1.35 [0.015]	1.25 [0.007]	1.04 [0.006]	1.04 [0.007]	2.99 [0.024]
1	100	0	6	0.133	0.189	0.167	0.163	0.165	0.310	1.22 [0.013]	1.16 [0.006]	1.03 [0.005]	1.02 [0.005]	1.88 [0.023]
1	100	0	12	0.128	0.164	0.156	0.153	0.156	0.199	1.19 [0.013]	1.08 [0.004]	1.02 [0.004]	1.02 [0.003]	1.27 [0.016]
1	100	0.4	3	0.171	0.298	0.256	0.235	0.271	0.864	1.37 [0.016]	1.27 [0.008]	1.09 [0.007]	1.15 [0.009]	3.19 [0.024]
1	100	0.4	6	0.144	0.227	0.191	0.185	0.192	0.539	1.29 [0.014]	1.23 [0.007]	1.03 [0.005]	1.04 [0.007]	2.81 [0.028]
1	100	0.4	12	0.135	0.186	0.168	0.166	0.164	0.308	1.23 [0.014]	1.12 [0.006]	1.01 [0.004]	0.99 [0.005]*	1.87 [0.024]

* no evidence of a difference between the methods at P<0.01

					Mean					Ratio of means [standard error]				
Shape	n	p	v	W	NP	SNP	SNP2	MKM	RKM	SNP2/W	NP/SNP2	SNP/SNP2	MKM/SNP2	RKM/MKM
1.5	100	0	3	0.114	0.205	0.162	0.156	0.160	0.536	1.37 [0.016]	1.31 [0.008]	1.04 [0.006]	1.03 [0.008]	3.34 [0.027]
1.5	100	0	6	0.106	0.159	0.136	0.134	0.133	0.286	1.27 [0.013]	1.19 [0.007]	1.01 [0.004]	1.00 [0.005]*	2.15 [0.025]
1.5	100	0	12	0.103	0.137	0.128	0.126	0.128	0.174	1.22 [0.012]	1.09 [0.005]	1.02 [0.003]	1.02 [0.004]	1.36 [0.018]
1.5	100	0.4	3	0.139	0.263	0.215	0.194	0.178	0.709	1.40 [0.016]	1.35 [0.009]	1.11 [0.008]	0.92 [0.008]	3.97 [0.042]
1.5	100	0.4	6	0.116	0.195	0.161	0.155	0.144	0.474	1.34 [0.016]	1.25 [0.007]	1.04 [0.005]	0.93 [0.006]	3.28 [0.035]
1.5	100	0.4	12	0.109	0.156	0.138	0.136	0.133	0.274	1.25 [0.013]	1.15 [0.006]	1.01 [0.004]	0.98 [0.005]	2.05 [0.027]
2	100	0	3	0.095	0.187	0.146	0.141	0.144	0.491	1.49 [0.019]	1.33 [0.008]	1.04 [0.006]	1.02 [0.009]*	3.42 [0.026]
2	100	0	6	0.085	0.140	0.115	0.114	0.114	0.274	1.34 [0.015]	1.23 [0.007]	1.01 [0.005]	1.00 [0.006]*	2.41 [0.026]
2	100	0	12	0.081	0.117	0.106	0.104	0.107	0.156	1.28 [0.014]	1.12 [0.005]	1.02 [0.004]	1.02 [0.003]	1.47 [0.019]
2	100	0.4	3	0.108	0.235	0.190	0.172	0.152	0.615	1.59 [0.022]	1.37 [0.009]	1.11 [0.008]	0.88 [0.009]	4.04 [0.045]
2	100	0.4	6	0.094	0.174	0.142	0.134	0.123	0.430	1.43 [0.017]	1.29 [0.008]	1.06 [0.006]	0.92 [0.006]	3.48 [0.039]
2	100	0.4	12	0.087	0.137	0.119	0.116	0.113	0.261	1.34 [0.015]	1.18 [0.006]	1.02 [0.004]	0.98 [0.005]	2.31 [0.028]
0.6	200	0	3	0.133	0.212	0.206	0.189	0.246	0.696	1.42 [0.014]	1.12 [0.006]	1.09 [0.004]	1.30 [0.008]	2.83 [0.016]
0.6	200	0	6	0.122	0.165	0.156	0.151	0.169	0.349	1.24 [0.011]	1.09 [0.005]	1.03 [0.003]	1.12 [0.005]	2.06 [0.020]
0.6	200	0	12	0.111	0.139	0.134	0.132	0.139	0.201	1.18 [0.012]	1.06 [0.004]	1.02 [0.003]	1.06 [0.003]	1.44 [0.018]
0.6	200	0.4	3	0.150	0.252	0.247	0.228	0.394	1.108	1.53 [0.016]	1.10 [0.006]	1.08 [0.005]	1.73 [0.010]	2.81 [0.011]
0.6	200	0.4	6	0.130	0.193	0.176	0.170	0.239	0.653	1.30 [0.013]	1.14 [0.006]	1.04 [0.004]	1.41 [0.009]	2.73 [0.018]
0.6	200	0.4	12	0.118	0.155	0.144	0.143	0.163	0.347	1.21 [0.012]	1.09 [0.004]	1.01 [0.003]	1.14 [0.007]	2.13 [0.021]
1	200	0	3	0.102	0.180	0.147	0.138	0.161	0.601	1.36 [0.015]	1.30 [0.007]	1.06 [0.005]	1.16 [0.008]	3.74 [0.023]
1	200	0	6	0.094	0.138	0.119	0.116	0.122	0.304	1.24 [0.012]	1.19 [0.006]	1.02 [0.003]	1.05 [0.005]	2.50 [0.026]
1	200	0	12	0.088	0.116	0.108	0.106	0.109	0.172	1.20 [0.011]	1.09 [0.004]	1.01 [0.003]	1.03 [0.003]	1.57 [0.020]
1	200	0.4	3	0.118	0.222	0.179	0.168	0.232	0.875	1.43 [0.016]	1.32 [0.008]	1.06 [0.005]	1.38 [0.010]	3.77 [0.022]
1	200	0.4	6	0.101	0.165	0.133	0.130	0.148	0.544	1.30 [0.013]	1.27 [0.007]	1.02 [0.004]	1.14 [0.008]	3.67 [0.032]
1	200	0.4	12	0.094	0.130	0.115	0.114	0.117	0.301	1.21 [0.011]	1.14 [0.005]	1.01 [0.003]	1.02 [0.005]	2.58 [0.029]
1.5	200	0	3	0.081	0.153	0.118	0.114	0.130	0.539	1.41 [0.016]	1.34 [0.008]	1.03 [0.005]	1.14 [0.009]	4.15 [0.025]
1.5	200	0	6	0.074	0.116	0.095	0.095	0.097	0.288	1.27 [0.013]	1.22 [0.006]	1.01 [0.003]*	1.02 [0.005]	2.97 [0.028]
1.5	200	0	12	0.071	0.095	0.086	0.086	0.088	0.151	1.21 [0.012]	1.11 [0.004]	1.01 [0.002]*	1.02 [0.003]	1.72 [0.022]
1.5	200	0.4	3	0.099	0.190	0.145	0.137	0.140	0.713	1.39 [0.016]	1.38 [0.009]	1.05 [0.006]	1.02 [0.010]*	5.11 [0.046]
1.5	200	0.4	6	0.080	0.142	0.110	0.109	0.102	0.480	1.36 [0.015]	1.29 [0.007]	1.01 [0.004]*	0.94 [0.006]	4.68 [0.044]
1.5	200	0.4	12	0.074	0.111	0.094	0.093	0.091	0.277	1.27 [0.012]	1.18 [0.006]	1.00 [0.003]*	0.97 [0.005]	3.04 [0.033]

* no evidence of a difference between the methods at P<0.01

					Mean					Ratio of means [standard error]				
Shape	n	p	v	W	NP	SNP	SNP2	MKM	RKM	SNP2/W	NP/SNP2	SNP/SNP2	MKM/SNP2	RKM/MKM
2	200	0	3	0.068	0.138	0.102	0.101	0.119	0.494	1.48 [0.018]	1.37 [0.008]	1.01 [0.004]	1.18 [0.009]	4.15 [0.023]
2	200	0	6	0.062	0.102	0.081	0.081	0.084	0.278	1.31 [0.014]	1.27 [0.007]	1.00 [0.004]*	1.05 [0.006]	3.30 [0.028]
2	200	0	12	0.060	0.084	0.075	0.074	0.077	0.146	1.23 [0.012]	1.13 [0.005]	1.00 [0.002]*	1.03 [0.003]	1.90 [0.024]
2	200	0.4	3	0.079	0.170	0.125	0.120	0.110	0.616	1.51 [0.019]	1.42 [0.010]	1.05 [0.006]	0.92 [0.010]	5.62 [0.063]
2	200	0.4	6	0.067	0.126	0.095	0.094	0.086	0.435	1.42 [0.017]	1.33 [0.008]	1.01 [0.004]	0.91 [0.007]	5.04 [0.050]
2	200	0.4	12	0.063	0.097	0.081	0.080	0.078	0.259	1.28 [0.013]	1.21 [0.006]	1.01 [0.003]*	0.98 [0.005]	3.30 [0.035]
0.6	400	0	3	0.091	0.159	0.155	0.140	0.216	0.705	1.53 [0.017]	1.14 [0.006]	1.11 [0.004]	1.54 [0.010]	3.27 [0.015]
0.6	400	0	6	0.086	0.121	0.114	0.109	0.132	0.340	1.27 [0.012]	1.11 [0.006]	1.04 [0.003]	1.20 [0.006]	2.58 [0.022]
0.6	400	0	12	0.082	0.103	0.099	0.098	0.106	0.182	1.20 [0.012]	1.05 [0.004]	1.01 [0.002]	1.08 [0.003]	1.71 [0.020]
0.6	400	0.4	3	0.110	0.196	0.192	0.172	0.372	1.109	1.56 [0.017]	1.14 [0.006]	1.11 [0.004]	2.16 [0.013]	2.98 [0.009]
0.6	400	0.4	6	0.092	0.140	0.128	0.122	0.204	0.654	1.33 [0.012]	1.15 [0.006]	1.05 [0.004]	1.67 [0.011]	3.20 [0.017]
0.6	400	0.4	12	0.082	0.110	0.101	0.100	0.124	0.336	1.22 [0.011]	1.09 [0.005]	1.01 [0.002]	1.24 [0.007]	2.71 [0.024]
1	400	0	3	0.073	0.137	0.109	0.100	0.138	0.602	1.37 [0.016]	1.37 [0.008]	1.09 [0.004]	1.38 [0.010]	4.36 [0.022]
1	400	0	6	0.067	0.102	0.086	0.083	0.092	0.304	1.24 [0.013]	1.23 [0.006]	1.03 [0.003]	1.11 [0.005]	3.29 [0.030]
1	400	0	12	0.064	0.084	0.078	0.077	0.080	0.155	1.20 [0.012]	1.10 [0.004]	1.01 [0.002]	1.05 [0.003]	1.93 [0.024]
1	400	0.4	3	0.083	0.167	0.131	0.120	0.204	0.874	1.46 [0.017]	1.39 [0.008]	1.09 [0.005]	1.69 [0.012]	4.29 [0.021]
1	400	0.4	6	0.072	0.122	0.097	0.094	0.119	0.546	1.29 [0.013]	1.30 [0.007]	1.03 [0.003]	1.27 [0.009]	4.60 [0.034]
1	400	0.4	12	0.066	0.094	0.081	0.080	0.085	0.299	1.22 [0.012]	1.17 [0.005]	1.01 [0.002]	1.06 [0.006]	3.50 [0.037]
1.5	400	0	3	0.056	0.116	0.085	0.081	0.111	0.539	1.44 [0.017]	1.43 [0.008]	1.04 [0.004]	1.37 [0.010]	4.86 [0.021]
1.5	400	0	6	0.053	0.087	0.069	0.068	0.074	0.288	1.29 [0.013]	1.27 [0.007]	1.01 [0.003]	1.08 [0.006]	3.89 [0.030]
1.5	400	0	12	0.051	0.069	0.062	0.062	0.065	0.146	1.21 [0.012]	1.12 [0.005]	1.00 [0.002]*	1.04 [0.003]	2.25 [0.028]
1.5	400	0.4	3	0.067	0.143	0.101	0.097	0.111	0.712	1.44 [0.017]	1.47 [0.009]	1.04 [0.005]	1.14 [0.011]	6.42 [0.049]
1.5	400	0.4	6	0.058	0.105	0.079	0.078	0.077	0.480	1.34 [0.014]	1.35 [0.008]	1.01 [0.003]	0.99 [0.007]*	6.26 [0.056]
1.5	400	0.4	12	0.052	0.078	0.065	0.065	0.064	0.277	1.25 [0.013]	1.20 [0.006]	1.00 [0.002]*	0.98 [0.004]	4.34 [0.043]
2	400	0	3	0.048	0.103	0.072	0.071	0.102	0.495	1.47 [0.017]	1.45 [0.009]	1.02 [0.003]	1.45 [0.012]	4.85 [0.022]
2	400	0	6	0.044	0.075	0.058	0.058	0.065	0.278	1.31 [0.014]	1.31 [0.008]	1.00 [0.003]*	1.12 [0.007]	4.29 [0.029]
2	400	0	12	0.042	0.059	0.051	0.051	0.054	0.144	1.22 [0.012]	1.15 [0.005]	1.00 [0.002]*	1.06 [0.004]	2.65 [0.030]
2	400	0.4	3	0.057	0.127	0.088	0.085	0.084	0.617	1.49 [0.017]	1.50 [0.010]	1.03 [0.005]	0.99 [0.011]*	7.32 [0.080]
2	400	0.4	6	0.048	0.093	0.067	0.066	0.062	0.435	1.38 [0.015]	1.40 [0.009]	1.00 [0.003]*	0.93 [0.008]	7.03 [0.067]
2	400	0.4	12	0.046	0.071	0.058	0.057	0.056	0.261	1.26 [0.013]	1.23 [0.007]	1.00 [0.002]*	0.98 [0.005]	4.66 [0.048]

* no evidence of a difference between the methods at P<0.01

					Mean					Ratio of means [standard error]				
Shape	n	p	v	W	NP	SNP	SNP2	MKM	RKM	SNP2/W	NP/SNP2	SNP/SNP2	MKM/SNP2	RKM/MKM
0.6	800	0	3	0.067	0.123	0.118	0.105	0.196	0.707	1.56 [0.016]	1.17 [0.006]	1.13 [0.004]	1.87 [0.011]	3.61 [0.011]
0.6	800	0	6	0.061	0.088	0.082	0.080	0.106	0.339	1.31 [0.012]	1.10 [0.005]	1.03 [0.003]	1.33 [0.007]	3.19 [0.022]
0.6	800	0	12	0.057	0.072	0.070	0.070	0.078	0.166	1.23 [0.012]	1.03 [0.003]	1.00 [0.001]*	1.12 [0.003]	2.13 [0.023]
0.6	800	0.4	3	0.080	0.151	0.147	0.129	0.352	1.109	1.61 [0.016]	1.17 [0.006]	1.14 [0.004]	2.73 [0.017]	3.15 [0.008]
0.6	800	0.4	6	0.066	0.105	0.095	0.091	0.185	0.656	1.38 [0.013]	1.15 [0.006]	1.05 [0.004]	2.04 [0.014]	3.55 [0.014]
0.6	800	0.4	12	0.061	0.081	0.076	0.075	0.102	0.332	1.24 [0.011]	1.08 [0.005]	1.00 [0.002]*	1.36 [0.009]	3.24 [0.026]
1	800	0	3	0.049	0.103	0.078	0.070	0.122	0.604	1.41 [0.016]	1.48 [0.009]	1.12 [0.004]	1.75 [0.012]	4.94 [0.016]
1	800	0	6	0.047	0.074	0.060	0.058	0.072	0.306	1.22 [0.013]	1.28 [0.007]	1.04 [0.002]	1.25 [0.007]	4.26 [0.030]
1	800	0	12	0.045	0.059	0.053	0.053	0.058	0.151	1.17 [0.012]	1.11 [0.005]	1.01 [0.002]	1.09 [0.004]	2.62 [0.029]
1	800	0.4	3	0.059	0.127	0.095	0.084	0.189	0.874	1.43 [0.016]	1.50 [0.009]	1.12 [0.005]	2.24 [0.017]	4.62 [0.017]
1	800	0.4	6	0.052	0.090	0.069	0.066	0.100	0.549	1.27 [0.013]	1.37 [0.008]	1.05 [0.003]	1.53 [0.011]	5.47 [0.033]
1	800	0.4	12	0.047	0.068	0.058	0.057	0.065	0.299	1.22 [0.012]	1.18 [0.005]	1.01 [0.002]	1.13 [0.007]	4.63 [0.040]
1.5	800	0	3	0.040	0.087	0.060	0.057	0.100	0.538	1.44 [0.016]	1.52 [0.010]	1.05 [0.004]	1.74 [0.014]	5.38 [0.019]
1.5	800	0	6	0.037	0.063	0.048	0.048	0.058	0.291	1.31 [0.014]	1.31 [0.007]	1.01 [0.002]	1.22 [0.007]	4.97 [0.028]
1.5	800	0	12	0.035	0.049	0.043	0.043	0.046	0.146	1.22 [0.012]	1.14 [0.004]	1.00 [0.002]*	1.09 [0.004]	3.15 [0.030]
1.5	800	0.4	3	0.048	0.109	0.075	0.071	0.099	0.715	1.46 [0.018]	1.55 [0.010]	1.06 [0.004]	1.40 [0.013]	7.24 [0.049]
1.5	800	0.4	6	0.041	0.077	0.055	0.054	0.057	0.478	1.33 [0.014]	1.42 [0.008]	1.01 [0.002]	1.06 [0.008]	8.39 [0.068]
1.5	800	0.4	12	0.038	0.058	0.047	0.047	0.048	0.278	1.26 [0.012]	1.22 [0.006]	1.00 [0.002]*	1.00 [0.005]*	5.85 [0.054]
2	800	0	3	0.034	0.078	0.052	0.051	0.096	0.495	1.48 [0.019]	1.54 [0.010]	1.02 [0.003]	1.88 [0.015]	5.17 [0.018]
2	800	0	6	0.031	0.055	0.041	0.041	0.053	0.280	1.30 [0.014]	1.35 [0.008]	1.00 [0.002]*	1.30 [0.008]	5.25 [0.026]
2	800	0	12	0.031	0.043	0.038	0.038	0.041	0.145	1.23 [0.012]	1.15 [0.005]	1.00 [0.002]*	1.10 [0.004]	3.53 [0.030]
2	800	0.4	3	0.040	0.097	0.063	0.061	0.071	0.618	1.51 [0.018]	1.58 [0.011]	1.03 [0.004]	1.17 [0.013]	8.67 [0.080]
2	800	0.4	6	0.034	0.069	0.047	0.047	0.046	0.436	1.38 [0.016]	1.46 [0.009]	1.01 [0.002]*	0.97 [0.009]	9.56 [0.085]
2	800	0.4	12	0.032	0.051	0.041	0.041	0.041	0.262	1.26 [0.012]	1.25 [0.006]	1.00 [0.001]*	1.01 [0.006]*	6.39 [0.058]

* no evidence of a difference between the methods at $P<0.01$

a) Right-censored Kaplan-Meier estimator

The RKM approach was less accurate than the MKM in all simulations (Table 3.1) due to the systematic underestimation of the cumulative hazard function with this method. The relative inaccuracy of RKM increased with increasing sample size, a larger shape parameter, and greater sparseness of data. The highest ratio of 8.39 was obtained for n=800, γ=1.5, v=6, p=0.4.

b) Effect of smoothing the non-parametric estimator

Figures 3.3a-c show the mean accuracy metric for the NPMLE and SNP models relative to the SNP2 method by v, p and n when γ=0.6, γ=1.0 and γ=1.5 respectively. In general, smoothing the non-parametric estimator gave improved accuracy, with SNP2 superior to SNP (Table 3.1).

When γ=0.6, the corresponding ratios were only slightly greater than one (NPMLE range 1.0 – 1.17, SNP range 1.0 – 1.14) implying all three non-parametric methods had similar accuracy (Figure 3.3a).

When γ=1.0 or 1.5, the benefit of the SNP2 method over the NPMLE increased as sample size increased and the number of visits decreased. In contrast, for the SNP estimator the effect of sample size reversed, with appreciable difference in accuracy compared to SNP2 being evident only for n<200 (Figure 3.3b&c). This effect was reduced as the number of follow-up visits increased. Both smooth estimators had similar accuracy when the number of visits and sample size were large. The effect of frequency of missing visits upon the relative accuracy of the methods was minimal.

c) Midpoint Kaplan-Meier estimator

Figures 3.4a-c show the accuracy of the MKM approach relative to the SNP2 method (the optimal non-parametric method) by v, p and n when γ=0.6, γ=1.0 and γ=1.5 respectively. This comparison was affected by a complex interaction between the shape parameter, sample size and sparseness of the data, which is discussed below.

When there were no missing visits, the relative accuracy of the MKM and SNP2 methods was not affected by the shape parameter (Table 3.1). When v=12, the relative accuracy was close to 1 for all sample sizes (range 1.02 – 1.12), but the MKM method

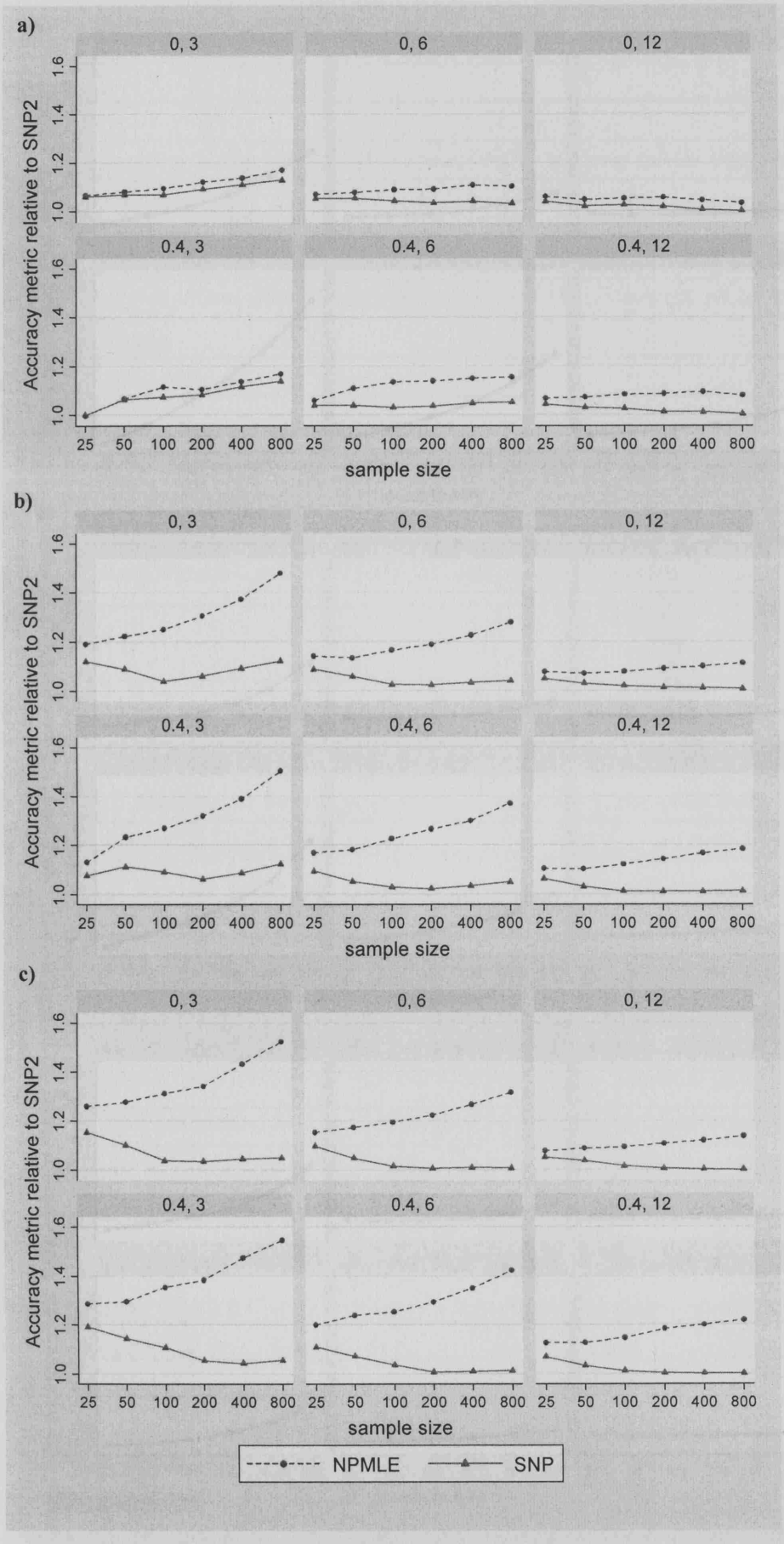

Figure 3.3: Comparing non-parametric methods: The effect of sample size and sparseness of data when a) γ=0.6, b) γ=1.0, c) γ=1.5.

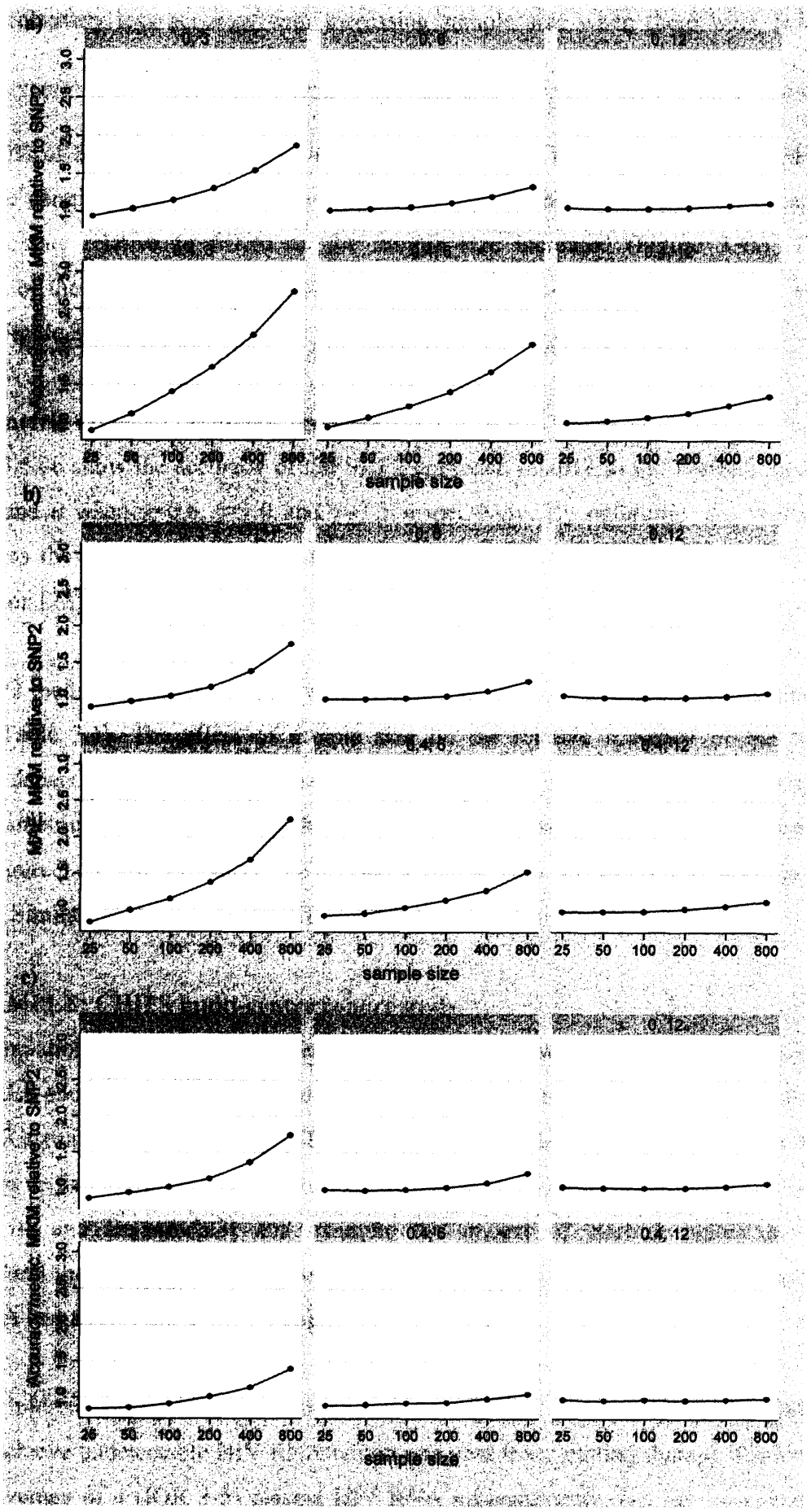

Figure 3.4: Comparing MKM and SNP2 models: The effect of sample size and sparseness of data when a) γ=0.6, b) γ =1.0, c) γ =1.5.

became relatively less accurate with increasing sample size when there were fewer visits.

When 40% of follow-up visits were missed the strength of these trends diminished as the shape parameter increased. For γ=0.6 and v=3 or 6, the SNP2 approach became more accurate relative to MKM as the sample size increased (Figure 3.4a). The accuracy of the two methods was comparable when γ=1.0, 1.5 except when v=3 and n>200 (Figure 3.4b&c).

d) Parametric vs. non-parametric methods

Figures 3.5a-c show the accuracy of the SNP2 approach relative to the Weibull model by v, p and n when γ=0.6, γ=1.0 and γ=1.5 respectively. By definition, the Weibull model was the most accurate method since the data were generated from a Weibull distribution [22].

The SNP2 method was less accurate by a factor of 1.09 – 1.61 relative to the Weibull model (Table 3.1). The effect of sample size on the relative accuracy of the two methods depended strongly on the shape parameter. For γ=0.6, the gain from fitting a fully parametric model increased with sample size, with a steeper gradient with fewer visits (Figure 3.5a). For γ=1.0, 1.5 the impact of sample size was much reduced (Figure 3.5b&c). Similar trends were seen for the NPMLE and SNP methods.

3.4 EXAMPLE: CHIPS multi-centre cohort study

In the simulation study described above the SNP2 estimator was consistently superior to the NPMLE and SNP estimators and the MKM method was found to perform as well as the SNP2 estimator in certain cases. To illustrate the MKM and SNP2 approaches we now focus on data from the CHIPS multi-centre cohort study of HIV-1 infected children in the UK and Ireland [23] and examine the impact of starting highly active antiretroviral therapy (HAART) on the time to achieving undetectable levels of plasma HIV RNA less than or equal to 500 copies/ml.

Of 599 children starting HAART with available HIV RNA measurements, 166 (28%) did not achieve undetectable HIV RNA within one year from starting therapy. Children had an average of 4 (IQR 3-5) plasma HIV RNA measurements over the year, and a median gap between measurements of 8 (IQR 5-13) weeks. The censoring intervals

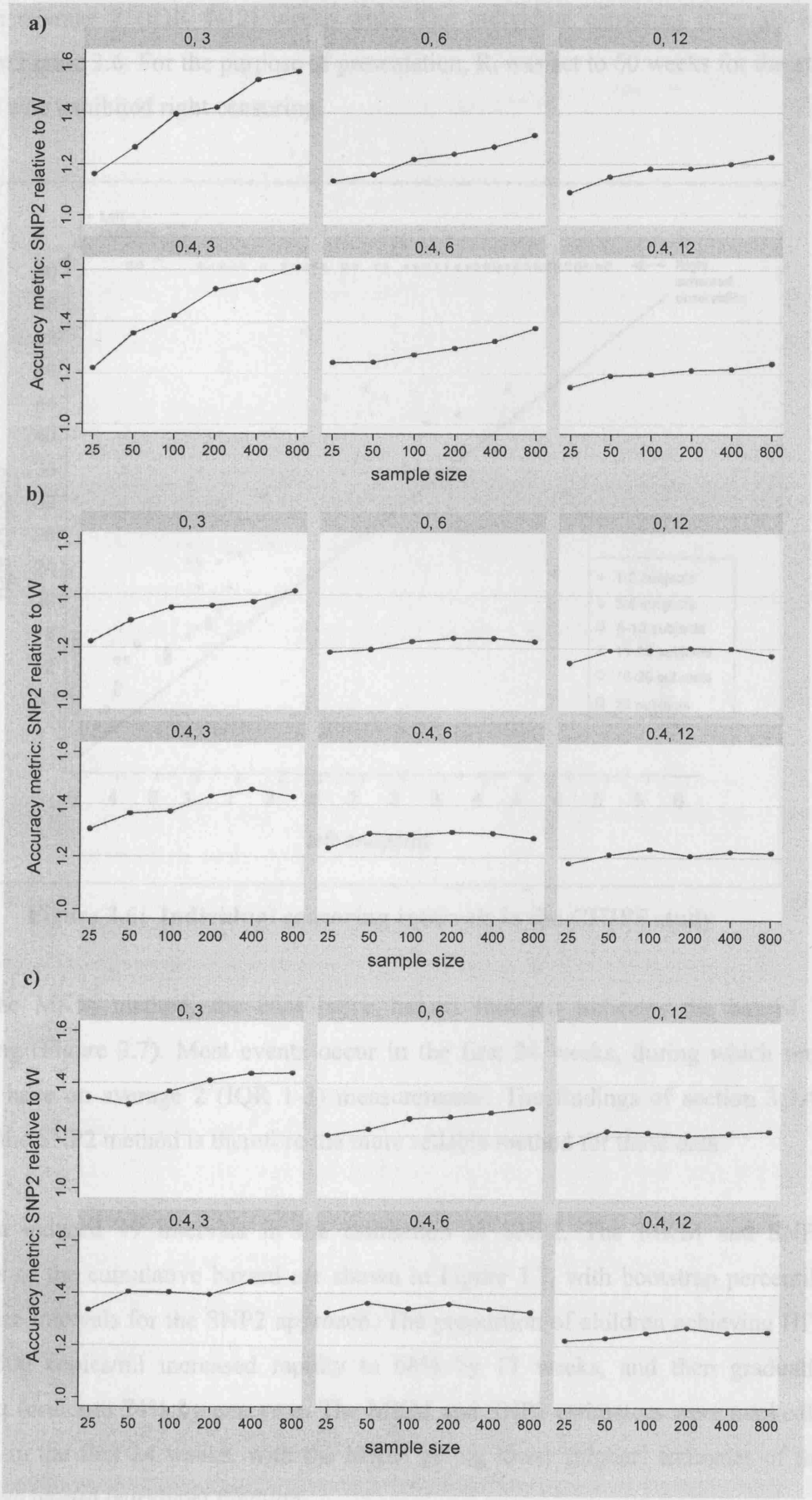

Figure 3.5: Comparing SNP2 and Weibull models: The effect of sample size and sparseness of data when a) γ=0.6, b) γ =1.0, c) γ =1.5.

were on average 8 (IQR 4-12) weeks wide. The individual censoring intervals are shown in Figure 3.6. For the purpose of presentation, R_i was set to 60 weeks for the 166 subjects who exhibited right censoring.

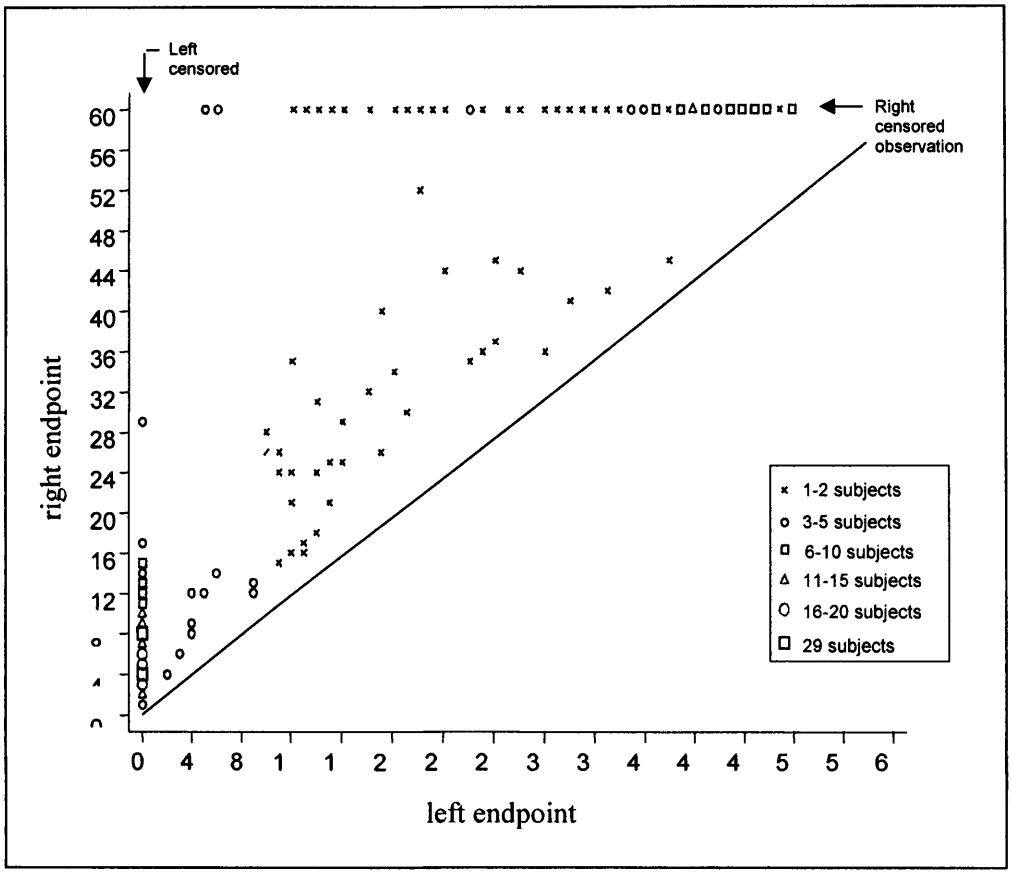

Figure 3.6: Individual censoring intervals in the CHIPS study

Using the MKM method, the cumulative hazard function indicates the hazard is decreasing (Figure 3.7). Most events occur in the first 24 weeks, during which time children have on average 2 (IQR 1-3) measurements. The findings of section 3.3.4c indicate the SNP2 method is therefore the more reliable method for these data.

The data induced 49 intervals in the estimation of SNP2. The MKM and SNP2 estimates of the cumulative hazard are shown in Figure 3.7, with bootstrap percentile confidence intervals for the SNP2 approach. The proportion of children achieving HIV RNA $\leq$500 copies/ml increased rapidly to 68% by 17 weeks, and then gradually increased further to 74% by one year. The MKM and SNP2 estimators were markedly different in the first 24 weeks, with the MKM giving lower (higher) estimates of the cumulative hazard (survivor) function.

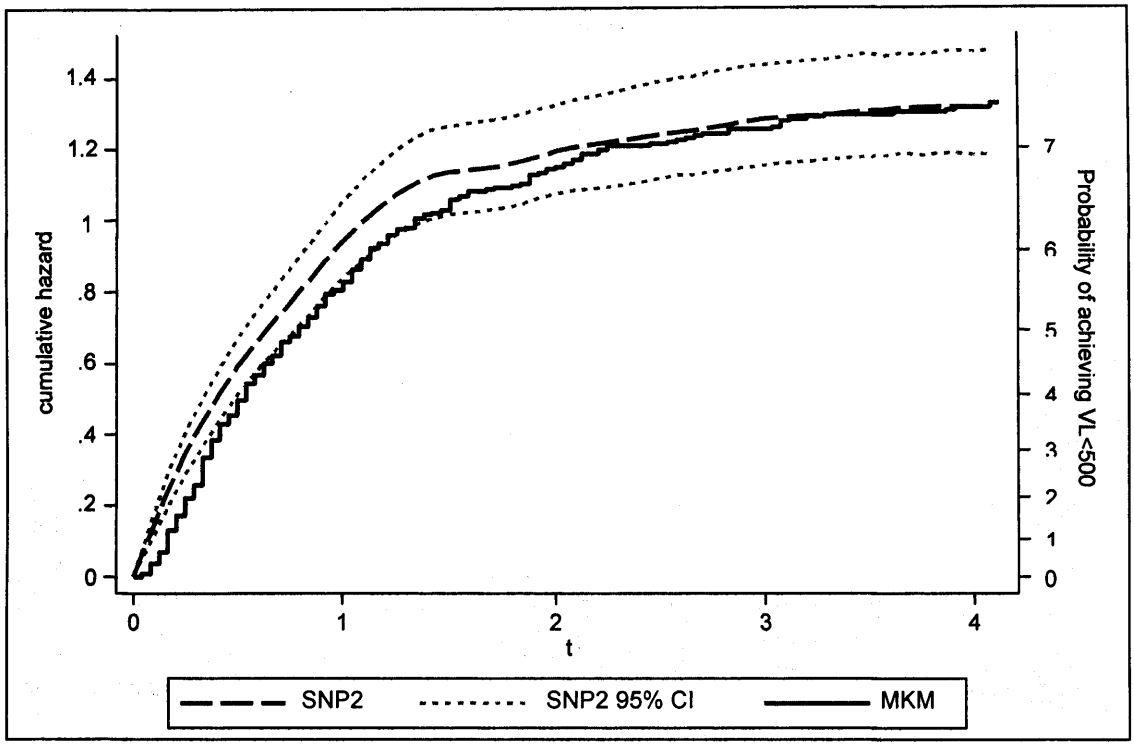

Figure 3.7: Cumulative hazard in the CHIPS study

3.5 DISCUSSION

Six methods for estimating the survivor function from interval-censored data have been described and assessed: a fully parametric Weibull model (W), the non-parametric maximum likelihood estimator (NPMLE), two closely related smoothed versions of this estimator (SNP & SNP2), a Kaplan-Meier estimator that assumes the event of interest occurs either at the middle of the observed censoring interval (MKM) or at the time it is observed (RKM).

The simulation study established that smoothing resulted in a significant increase in accuracy, with the SNP2 estimator consistently superior to the SNP and NPMLE estimators. The SNP2 method is therefore the recommended non-parametric approach among those examined. An extension of this work would be to compare SNP2 against other smoothing approaches developed for interval-censored survival data as reviewed in the Introduction, particularly the methods discussed by Pan [1] and Lesaffre et al [7].

The simulation study showed clearly that the RKM method results in significant bias, yet this method is still commonly applied in practice, perhaps because the main focus of an analysis is often a comparison of two or more groups. This is examined in the

following chapter. However, it remains desirable to obtain reasonably unbiased estimates of the survivor function for the individual groups. Since the MKM method is as easy to fit as the RKM approach, use of the latter cannot be justified.

In contrast, MKM performed well compared with SNP2 in certain situations. The results of the simulation study indicate that the SNP2 method is superior if n>200 and there are fewer than 5-6 visits per individual (during the period where most events occur), with this superiority being more marked when the hazard is decreasing. In all other scenarios considered, MKM was as accurate, or almost as accurate, as SNP2. A simple way of determining the shape of the hazard function is to plot the cumulative hazard -log((S(t)) against time, using the MKM method to estimate S(t).

Inspection of the results of individual simulations showed that the poor performance of the MKM when visits are infrequent was mainly due to over-estimation of the survivor function when t is small. The estimator fails to capture early changes in the underlying survival distribution as imputed event times are restricted by the distribution of the earliest visits. In particular, the first step of the estimator cannot occur earlier than midway to the first visit time. This phenomenon is stronger when the underlying hazard is decreasing as proportionately more events occur at earlier times.

As noted in section 3.3.1, the accuracy metric reflects both bias and variance of the estimators. For small samples sizes the accuracy metric is dominated by the variance component. The similarity of the metric for MKM and SNP2 for small sample sizes implies that the variance of these estimators is comparable. The superiority of the SNP2 method for larger samples sizes suggests that this approach is less biased. This finding has some theoretical justification in that the NPMLE method, to which the SNP2 method is closely related, has been shown to be asymptotically consistent [24,25].

The simulation study was based on data generated from a Weibull distribution, and the results confirmed the theoretical prediction that a correctly specified parametric model would be the most accurate method. The maximum benefit of the Weibull model compared to SNP2 was seen when the hazard was decreasing, sample size large and there were few follow-up visits. However, in practice the family of distributions is not known, and mis-specification leads to an asymptotically biased estimator [22]. The simulation study is limited by its restriction to data drawn only from a Weibull

distribution, however a range of hazard types was considered, although all were monotone.

Finally, there is little research into methods for estimating confidence intervals for the smoothed non-parametric estimator. Resampling methods, as demonstrated in the example described in this paper, offer a tractable solution. Also, confidence intervals for the MKM method by standard methods (e.g. Greenwood [26]) are likely to be too narrow since the method assumes the event time is known exactly. Again, simple non-parametric bootstrap confidence intervals are recommended.

References

1. Pan W. Smooth estimation of the survival function for interval-censored data. *Statistics in Medicine* 2000; **19**(19):2611-2624.

2. Kooperberg C, Stone CJ. Logspline density estimation for censored data. *Journal of computational and graphical statistics* 1992; **1**:301-328.

3. S-PLUS 2000 Guide to Statistics, Volume 2. 1999; (2).

4. Rosenberg PS. Hazard function estimation using B-splines. *Biometrics* 1995; **51**:874-887.

5. Betensky RA, Lindsey JC, Ryan LM, Wand MP. Local EM estimation of the hazard function for interval-censored data. *Biometrics* 1999; **55**:238-245.

6. Bebchuk JD, Betensky RA. Multiple imputation for simple estimation of the hazard function based on interval-censored data. *Stat.Med.* 2000; **19**:405-419.

7. Lesaffre E, Komarek A, Declerck D. An overview of methods for the analysis of interval-censored data with an emphasis on applications in dentistry. *Statistical Methods in Medical Research* 2005; **14**:539-552.

8. Pan W, Chappell R. Estimating survival curves with left truncated and interval-censored data via the EMS algorithm. *Communications in Statistics-Theory and Methods* 1998; **27**(4):777-793.

9. Silverman BW, Jones MC, Wilson JD. A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography. *Journal of the Royal Statistical Society B* 1990; **52**:271-324.

10. Odell PM, Anderson KM, D'Agostino RB. Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics* 1992; **48**:951-959.

11. Dorey FJ, Little RJA, Schenker N. Multiple imputation for threshold-crossing data with interval-censoring. *Statistics in Medicine* 1993; **12**:1589-1603.

12. Law CG, Brookmeyer R. Effects of mid-point imputation on the analysis of doubly censored data. *Stat.Med.* 1992; **11**(12):1569-1578.

13. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 1958; **53**:457-481.

14. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B* 1977; **39**:1-38.

15. Turnbull BW. Nonparametric estimation of a survivorship function with doubly censored data. *Journal of the American Statistical Association* 1974; **69**(345):169-173.

16. Alioum A, Commenges D. A proportional hazards model for arbitrarily censored and truncated data. *Biometrics* 1996; **52**(2):512-524.

17. Nychka N. Some properties of adding a smoothing step to the EM algorithm. *Statistics and Probability Letters* 1990; **9**:187-193.

18. Leung K-M, Elashoff RM. A three-state disease model with interval-censored data: estimation and application to AIDS and cancer. *Lifetime Data Analysis* 1996; **2**:175-194.

19. Aalen O. Non-parametric inference in connection with multiple decrement models. *Scandanavian Journal of Statistics* 1976; **3**:15-27.

20. SAS Institute Inc., Cary, North Carolina, USA. 1999; (8.2).

21. ICE.sas. *http://support.sas.com/ctx/samples/index.jsp?sid=477&tab=about* 2006. 6-17-2006.

22. Miller RG. What price Kaplan-Meier? *Biometrics* 1983; **39**:1077-1081.

23. Gibb DM, Duong T, Tookey PA, Sharland M, Tudor-Williams G, Novelli V, on behalf of the National Study of HIV in Pregnancy and Childhood (NSHPC) and of the Collaborative HIV Paediatric Study (CHIPS). Decline in mortality, AIDS and hospital admissions in perinatally HIV-1 infected children in the UK and Ireland. *BMJ* 2003; **327**:1019-1025.

24. Huang J, Wellner JA. Interval-censored survival data: a review of recent progress. Lin, D. Y. and Fleming, T. R. Lecture Notes in Statistics: Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis. 1997; 123-169. New York, Springer-Verlag.

25. Yu Q, Li L, Wong GWC. On consistency of the self-consistent estimator of survival functions with interval-censored data. *Scandanavian Journal of Statistics* 2000; **27**(1):35-44.

26. Collett D. Modelling Survival Data in Medical Research. 2003; **2nd**. London, Chapman and Hall.

CHAPTER 4: ESTIMATING THE EFFECT OF COVARIATES

4.1 INTRODUCTION

It is often desirable to quantifying the effect of different factors, for example age or treatment, on the survival time. A variety of methods have been proposed for estimation of covariate effects upon the survivor or hazard function when data are interval-censored (section 1.3). However, these are often computationally intensive, requiring complex programming.

In the case of categorical factors, the non-parametric maximum likelihood estimator (NPMLE) [1,2] (section 2.2.2) of the survivor function can be fitted to data from each level of the covariate and several significance tests have been proposed, mainly for the two sample problem [3]. Most of these are adaptations of non-parametric tests available for right-censored data but assume that the data are grouped, that is, have a finite number of possible follow-up times common to all individuals. This is often not the case.

Fang et al. [4] propose a test statistic for continuous survival times based on an integrated weighted survival difference that is implemented using simple bootstrap. This is a generalisation of the test statistic proposed by Pepe and Fleming [5] for right-censored survival data. The authors report that the method does not perform well when the difference between the survivor functions under examination is non-proportional, based on a simulation study (results not given).

Zhao and Sun [6] describe a generalised log-rank test, also for continuous survival times, that reduces to the usual log-rank test [7] when data are right-censored. The covariance matrix of the test statistic is calculated using multiple imputation, and the standardised test statistic is then referred to a χ^2-distribution on p-1 degrees of freedom, where p is the number of groups under study. This distributional assumption is shown to

work well in a simulation study (where survival times = discretized exponential, p=2, n=100 per group, right censoring = fixed or random; 10%, 25% or 40%).

Alternatively, Pan suggests a multiple imputation approach based on the Approximate Bayesian Bootstrap [8] that imputes survival times for interval-censored observations and then implements standard methods for right-censored data to compute a test statistic [9]. None of these methods are available in the standard statistical packages, although S-plus [10] code is available from Pan for the multiple imputation method.

In the case of continuous covariates when data are exact or right-censored, the proportional hazards model proposed by Cox [11] is often applied. The appeal of this method arises from its semi-parametric nature; the effect of covariates on survival can be investigated using a partial likelihood framework that does not require specification of the baseline hazard function. Instead only the rank order of event and censoring times are required. However, when the data are interval-censored these are not, in general, possible to derive, precluding the use of partial likelihood inference.

Several methods combine multiple imputation and the Cox proportional hazards model for right-censored data. Satten has proposed using either Gibbs sampling [12] or a parametric model for the baseline hazard [13] to determine the rank ordering of event times consistent with the observed censoring intervals. In contrast, Pan [14] uses the Breslow estimator of the survivor function to estimate the ranks whilst Goggins et al [15] use a Monte Carlo EM approach. Although these methods all reduce to the usual Cox proportional hazards model when the data are right censored the imputation approach is slightly ad hoc in each case. Alternatively, when the baseline hazard can be parametrically specified or modelled using splines the full likelihood can be maximised. Rosenberg et al. [16] and Kooperberg and Clarkson [17] demonstrate the use of regression splines, whilst Betensky et al. [18] extend their work on local likelihood estimation to allow estimation of covariate effects. These methods provide a smooth estimate of the baseline hazard/survivor function but do not reduce to the Cox proportional hazards model when the data are right censored. The fitting of each of these methods is complex and computationally intensive.

Finkelstein also describes a full likelihood approach where nuisance parameters associated with a non-parametric baseline hazard and covariate effects are estimated

simultaneously [19]. In this case, as for the NPMLE, the baseline hazard is only defined within a number of time intervals which are determined by the individual censoring intervals. This approach is similar to the piecewise exponential model [20,21], where the hazard is assumed to be constant across fewer, predefined time intervals. Both models require an iterative fitting procedure, such as Newton Rhapson, or the use of GLM or an EM approach has been suggested for the piecewise exponential model [22-24]. These methods are discussed in more detail in section 4.2. An approximation to the piecewise exponential model is also proposed that can be fitted using standard generalised linear modelling software without an iterative procedure (section 4.2.8).

As with the methods discussed in chapter 3, the above approaches developed specifically for interval-censored have been neglected in applied medical research. In practice, interval-censored data are usually analysed by fitting a Cox model assuming events occur at the right endpoint of the observed censoring interval for each individual. The following authors have discussed the use of simple imputation, where either the right endpoint or midpoint is imputed.

Lindsey and Ryan [22] compared a Cox model using the left, mid or right endpoint of the censoring interval to full parametric, piecewise exponential and Finkelstein models by estimating the hazard ratio between two groups using two illustrative datasets. For one dataset with light censoring (frequent visits and moderate proportion right-censored) parameter estimates of the hazard ratio from all methods of analysis were similar. For the other dataset, with heavier censoring, the Finkelstein model was unstable and estimates from the other models varied considerably.

Three simulation studies investigating right endpoint imputation have been reported. So [25] suggests that using the right endpoint in a Cox proportional hazards model may be satisfactory when ties are handled using an exact likelihood. However, just one scenario is considered with 100 observations and data that are grouped into 7 intervals. In contrast Williams et al. [26] consider more general interval-censoring and vary the number of follow-up visits and amount of right-censoring. They report moderate bias in covariate estimation as the number of follow-up visits decreases and significant bias when the patterns of follow-up visits are not identical in the two groups of subjects under comparison. However, each scenario had 5000 observations and the effect upon

their conclusions of reducing the number of observations to a size more common in practice is unclear.

The impact of different measurement frequencies on covariate estimation was also considered in a more extensive simulation study by Griffin et al. [27]. Here, right- or mid- point imputation in a Cox proportional hazards model was compared to right endpoint imputation in parametric models (Weibull, log-logistic, or cubic-spline baseline hazard models) and to estimation using the same parametric interval-censored models. Data were generated to mimic CD4 and HIV RNA measurements in patients infected by HIV. Two groups of 200 observations were simulated 2000 times, with the frequency of visits varying from equal schedules per group to one group having visits twice as frequently as the other. As the difference in visit frequency increased between the groups, imputation with the right endpoint, and to a lesser extent the midpoint, resulted in biased estimation of the covariate effect, due to underestimation in the group with fewer measurements. A difference in levels of detection bias in the two groups led to some bias when the interval-censored methods were used. The bias in estimated covariate effect led to poor coverage when there was a difference in measurement frequency between groups.

Finally, Odell et al [28] compared an accelerated failure time Weibull model based on the full interval-censored data (WIC) and a Weibull model based on midpoint imputation (WMID). The impact of sample size, percentage censored and width of censoring intervals on estimation of a single continuous covariate was assessed via a simulation study. When all three factors were small WMID was satisfactory. This was also the case when sample size was small and the censoring intervals relatively wide, or there was less than 25% censoring, since WIC estimates had high variance. When sample size was large the shape of the underlying hazard function affected the choice of method. As the hazard became steeper the bias in WMID increased. The paper did not consider any non-parametric approaches.

The two main objectives of this chapter are i) to assess a proposed approximation to the piecewise exponential model. This method is suggested as a more rigorous alternative to simple imputation methods whilst remaining easy to fit; and ii) to extend the work described above in a simulation study by considering a wider variety of generated datasets and regression models. The impact of the true size of covariate effect, baseline

hazard/survivor function, sample size, and the frequency and pattern of inspection times upon estimation of covariate effects were considered. Several different methods of analysis are assessed: a fully parametric model, the Finkelstein proportional hazards model, a piecewise exponential model, an approximation to the piecewise exponential model, and a Cox model using either the midpoint or right endpoint of the observed censoring interval. Situations under which the simpler methods (right- or mid- point imputation in the Cox model, approximation to the piecewise exponential model) achieve estimates of covariate effects that could be considered sufficiently accurate for practical purposes are determined. The extent to which the simple imputation methods are affected by the method of treating ties in the data is also examined.

After describing the approximate piecewise exponential model and other methods (section 4.2), the results of the simulation study are reported (section 4.3). Data from the Delta Trial Virology Substudy [29] and CHIPS study [30] are used to illustrate the different methods (section 4.4). A summary of the main findings and their practical implications is given (section 4.5).

4.2 NOTATION AND THEORY

4.2.1 Observed data

The same notation as in section 2.2.1 is used. In addition, let $M_i = (L_i+R_i)/2$ denote the midpoint of the observed censoring interval, and let x_i (i=1,...,n) be a fixed covariate for the i-th individual. The censoring mechanism is assumed to be independent of both the true failure time and the covariate.

4.2.2 Proportional hazards model

The hazard function for a given covariate vector x,

$$h(t \mid x) = \frac{\lim_{\Delta \to 0+} P(t \le T < t + \Delta \mid t \le T, x)}{\Delta t}$$

represents the instantaneous death rate for an individual surviving to time t. In a proportional hazards model, the hazard function is expressed as the product of an unspecified baseline hazard, $h_0(t)$ and relative risk function, usually the exponential of a linear function of covariates such that

$$h(t \mid x) = h_0(t) \exp(\beta' x) \qquad (4.1)$$

The survivor function under the assumption of proportional hazards can be expressed as

$$S(t \mid x) = [S_0(t)]^{\exp(\beta' x)} \tag{4.2}$$

where the baseline survivor function is given by

$$S_0(t) = \exp\left(- \int_0^t h_0(u)du \right).$$

4.2.3 Partial likelihood

In a Cox proportional hazards model, the regression coefficients β in (4.1) are estimated without the need to specify a form for the baseline hazard, which is itself not estimated. Using a partial likelihood that treats the baseline hazard as a nuisance parameter, only information on the rank of the censoring and event times is used rather than the actual times themselves [11,31].

Assuming there are j=1,...,m events, $x_{(j)}$ are the covariates for the individual experiencing the j-th event and $R_{(j)}$ denotes the set of all individuals at risk at the j-th event, the partial likelihood is given by

$$L(\beta) = \prod_{j=1}^{m} \left[\frac{\exp(\beta' x_{(j)})}{\sum_{i \in R_{(j)}} \exp(\beta' x_i)} \right] \tag{4.3}$$

When ties exist in the data, all terms from the partial likelihood (4.3) that are consistent with the observed data can be summed. For example, if two individuals with covariates $x_{(j1)}$ and $x_{(j2)}$ are tied at the j-th event time, and the set $R_{(j)}$ contains these two individuals plus one other with covariates $x_{(j3)}$, the contribution from the j-th event time to the likelihood would be

$$\frac{\exp(\beta' x_{(j1)})}{[\exp(\beta' x_{(j1)}) + \exp(\beta' x_{(j2)}) + \exp(\beta' x_{(j3)})]} \frac{\exp(\beta' x_{(j2)})}{[\exp(\beta' x_{(j2)}) + \exp(\beta' x_{(j3)})]}$$
$$+ \frac{\exp(\beta' x_{(j2)})}{[\exp(\beta' x_{(j1)}) + \exp(\beta' x_{(j2)}) + \exp(\beta' x_{(j3)})]} \frac{\exp(\beta' x_{(j1)})}{[\exp(\beta' x_{(j1)}) + \exp(\beta' x_{(j3)})]}$$

This sum does not simplify and the corresponding log likelihood is difficult to compute if there are a large number of tied observations. Instead it is common to use an approximation to the partial likelihood. Let d_j be the number of tied observations, and

$D_{(j)}$ the set of individuals with tied observations, at the j-th event time. In the simplest approximation proposed by Breslow [32] the contribution for each tied event is given by

$$\frac{\prod_{i \in D_{(j)}} \exp(\beta' x_i)}{\left[\sum_{i \in R_{(j)}} \exp(\beta' x_i)\right]^{d_j}}$$

4.2.4 Imputing event times

As discussed in section 3.2.2, it is possible to implement standard methods available for the analysis of right-censored data by imputing an exact failure time within each observed censoring interval. Estimates of covariate effects can be obtained by fitting a Cox proportional hazards model as described above using i) the censoring interval right endpoint, R_i (RCox), or ii) the midpoint, M_i (MCox). This can give rise to a large number of tied observations if there is a limited set of possible visit times.

4.2.5 Parametric Weibull model

By assuming a parametric model for the baseline hazard a proportional hazards model can be easily fitted by maximising the full likelihood. Assuming the observed event times follow a Weibull distribution (section 3.2.3), the baseline hazard is given by

$$h_0(t) = \lambda\gamma t^{\gamma-1}$$

where λ is a scale parameter and γ the shape parameter. Under an assumption of proportional hazards (4.1), the effect of any covariate will be to modify the scale parameter, leaving the shape parameter unchanged. The event times of all individuals will therefore have a Weibull distribution with shape parameter γ, and the Weibull distribution is said to have the proportional hazards property.

4.2.6 Adapting the Cox proportional hazards model

Finkelstein [19] describes a proportional hazards model fitted to interval-censored data that maximises the full likelihood. Following the notation in section 2.2.2, let $\{u_j; j=1,...,m+1\}$ denote the unique ordered values of $\{0,\{L_i\},\{R_i\}\}$ and define an indicator variable $\alpha_{ij}=1$ if $(u_j, u_{j+1}) \subseteq I_i$ and 0 otherwise. Under the assumption of proportional hazards, the survivor function is given by (4.2) and the likelihood (2.1) is written

$$L = \sum_{i=1}^{n} \log\left(\sum_{j=1}^{m} \alpha_{ij} \left[S_0(u_{j-1})^{\exp(\beta' x_i)} - S_0(u_j)^{\exp(\beta' x_i)} \right] \right).$$

To remove range restrictions on the parameters for the underlying survival curve, the likelihood is parameterised in terms of the log cumulative hazard such that $\gamma_j=\log[-\log(S_0(u_j))]$. The likelihood can then be rewritten as

$$L(\gamma,\beta)=\sum_{i=1}^{N}\log\sum_{j=1}^{m}\alpha_{ij}[\exp(-\exp(\beta' x_i+\gamma_{j-1}))-\exp(-\exp(\beta' x_i+\gamma_j))]$$

where $\gamma_0=-\infty$ ($S_0(u_1)=1$) and $\gamma_m=\infty$ ($S_0(u_m)=0$).

As for the NPMLE of the survivor function, the baseline survivor function is constant outside all distinct intervals with endpoints $u_j \in \{L_i\}$ and $u_{j+1} \in \{R_i\}$ (section 2.2.2) [33] and so the number of parameters in the model can be reduced. An iterative procedure is required to maximise the likelihood and Finkelstein proposes a Newton-Raphson approach. Estimates of the standard errors of the parameters are obtained from the second derivative of the log-likelihood. Setting $\beta=0$ in the score equation gives the NPMLE described by Peto and Turnbull.

This approach is closely related to a piecewise exponential model (section 4.2.7), where the time scale is divided into intervals over which the hazard is assumed to be constant (the piecewise exponential model parameterises the hazard rather than the log-cumulative hazard). However, in contrast with the piecewise exponential model, the intervals in Finkelstein's model are data dependent, and tend to increase with sample size. As mentioned in the discussion of the NPMLE (section 2.2.3) the asymptotic results may not therefore be justified. Finkelstein notes that continuous data can be grouped and therefore the method can be appropriately applied. However, the method can be numerically unstable and computationally intensive for some datasets (section 4.3.4).

4.2.7 Piecewise exponential model (PE)

The piecewise exponential (PE) model imposes the relatively weak assumption of constant hazards within a fixed number of predefined time intervals. For example, when analysing data from a clinical trial, follow-up can be divided into intervals centred at the scheduled visit times. Let follow-up time be divided into p pre-specified intervals $I_j=(t_{j-1}\ t_j)$, $j=1...p$, over which failure rate λ_j, $j=1...p$, is considered constant. Multiplicative covariate effects (i.e. proportional hazards) can be easily incorporated into the model using equation (4.1). Let δ_i be an indicator of whether person i has an

event. If all event times are known exactly, the likelihood contribution for a person with follow-up time $T_i \in I_j$ is

$$\lambda_j^{\delta_i} \exp\left[-\sum_{k=1}^{j-1}\lambda_k(t_k - t_{k-1}) - \lambda_j(T_i - t_{j-1})\right]$$

and the log-likelihood can be written as

$$L = \sum_{j=1}^{m}\{D_j \ln \lambda_j - \lambda_j Y_j\} \tag{4.4}$$

where D_j and Y_j are the number of events and person-time at risk in interval j respectively. However the exact times of failure are not known. An EM algorithm [34] can be used to obtain parameter estimates where the complete-data are defined as the exact failure times for people experiencing an event and the censoring times of those right-censored.

Lindsey and Ryan give expressions for an individual's conditional probability that an event occurred in interval j, and the conditional expected time at risk in interval j, given the observed data and current estimates of $\hat{\lambda}_j$ [22]. These are calculated in the E-step and the most recent estimates of D_j and Y_j then used in the M-step where the likelihood (4.4) is maximised to update the estimates of $\hat{\lambda}_j$. The likelihood (4.4) is proportional to the likelihood obtained by assuming the D_j are Poisson distributed with mean parameter $\lambda_j Y_j$, and hence a Poisson regression package can be used to estimate the λ_j by treating the outcome as D_j and $\log(Y_j)$ as offset. Standard errors can be estimated using the methods of Meng and Rubin [35]. Survival curves can then be calculated as

$$\hat{S}(t) = \exp\left[-\sum_{m=1}^{j-1}\hat{\lambda}_m[t_m - t_{m-1}] - \hat{\lambda}_j(t - t_{j-1})\right];\ t_{j-1} < t \le t_j \tag{4.5}$$

where $\hat{\lambda}_j$ is the estimated rate in interval j. These rates can be modelled as log linear combinations of covariates, which may be dependent on interval.

Seaman and Bird [21] extend the PE model to allow for time-varying coefficients and left truncation, in an EM framework. Other authors [23] have suggested a generalised linear model approach [24,36]that extends Efron's proposal for fitting logistic models to right censored survival data using an associated Bernoulli model [37]. The fitting procedure requires user-written macros to iterate between estimation of the baseline hazard parameters and covariates effects using standard GLM software. However, there are often problems in estimation as the rate parameters should be restricted to positive

values. In addition, standard errors must be obtained from the observed information matrix, which is not always the default package output.

Deciding the number or breakpoints of intervals is often based on natural divisions of time e.g. six-monthly, or yearly. However the number of intervals that can be reliably estimated is dependant on the number of events in the data. Friedman recommends starting with 5-7 intervals and suggests choosing the interval breakpoints such that the expected numbers of events in each interval are equal [20]. Two methods have been described in the literature for refining this choice, which are based on the model with no covariates. Estimated rate parameters and their standard errors can be investigated to show whether adjacent intervals have similar hazard rates and can be combined, or there are sharp changes in underlying hazard rates and intervals should be further divided (assuming there are enough events in the new intervals for estimation) [20]. Alternatively the model can be progressively simplified by assigning the same indicator variable to adjacent time points (starting with I_p) as long as a likelihood ratio test shows no significant difference as a result of combining the intervals [38]. The number and breakpoints of the intervals has been investigated using real-life data and the method found to be fairly robust to their choice [20,22].

4.2.8 Approximate piecewise exponential model (APE)

An approximate piecewise exponential (APE) method that can be fitted in standard packages without an iterative procedure is proposed. As for the PE model, constant rates λ_j, j=1...p are assumed within p pre-defined time intervals. First, each interval is allocated total number of events D_j and person-time at risk Y_j accounting for the interval-censoring, as described below. A standard Poisson regression model can then be fitted, with $D_j \sim \text{Po}(Y_j\lambda_j)$.

This approach therefore maximises the same log likelihood (4.4) as that for the complete data when fitting a PE model using an EM algorithm[34]. However, instead of computing the conditional expectation of D_j and Y_j in the E-step at each iteration as for the PE model, the APE approach calculates the conditional expectation of D_j and Y_j once assuming the event time t is uniformly distributed. So the APE approach is equivalent to the first iteration of the PE model using an EM algorithm given that starting values are calculated assuming that an individual experiences the event halfway through the interval. This method is explained graphically in Figure 4.1.

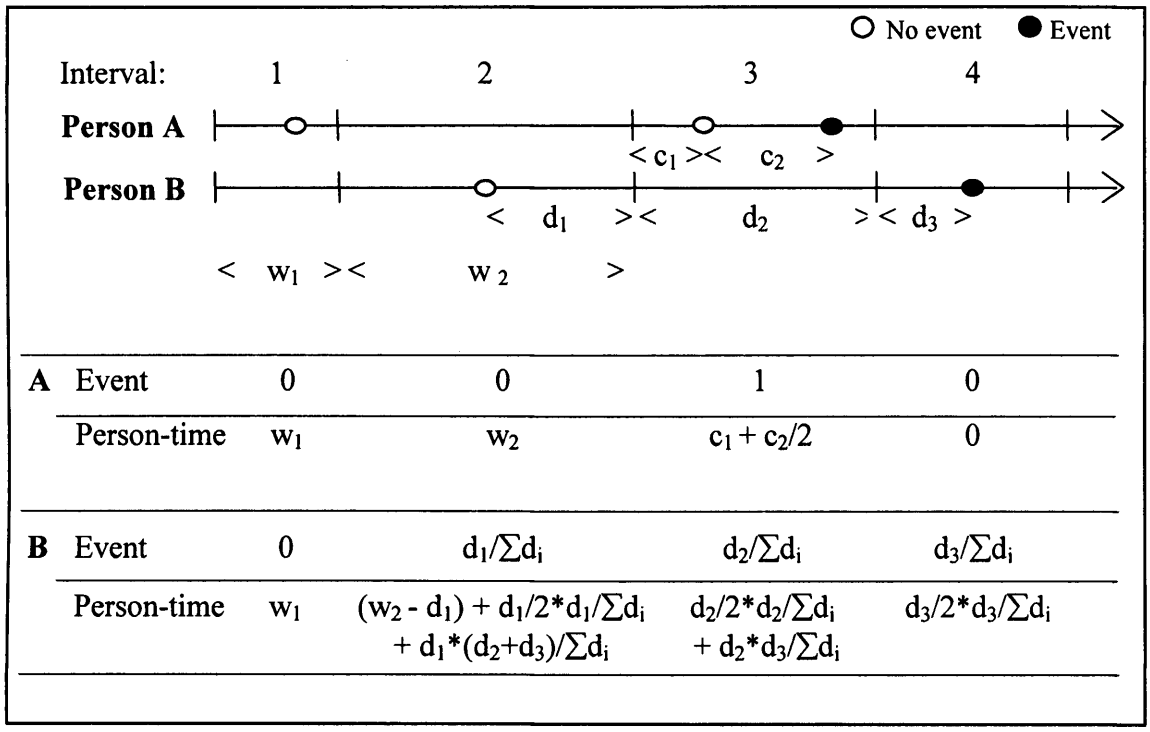

Figure 4.1: Dividing events and follow-up into intervals for the approximate piecewise exponential method

If the censoring interval I_i, for the *i*-th individual lies wholly within the *j*-th pre-specified interval, as for Person A, this individual will contribute one event to D_j and person years are calculated assuming the event occurred at the midpoint of the censoring interval, M_i. If, however, one or more intervals are overlapped by the observed censoring interval, the event and the person time at risk for that individual are divided between all intervals during which the event may have occurred.

Consider person B, whose interval overlaps intervals 2, 3 and 4. Their event is divided between these intervals in proportion to the length of time common to each interval and their censoring interval; namely $d_1/\sum d_i$, $d_2/\sum d_i$ and $d_3/\sum d_i$ for intervals 2, 3 and 4 respectively. The person-time at risk is then calculated for each proportion of an event, as if the event occurred in the middle of the person-time at risk within the relevant interval. For example, the person-time at risk in interval 2 is calculated as follows: if the event occurred in interval 2 this corresponds to $((w_2-d_1)+d_1/2)*(d_1/\sum d_i)$ person-time at risk; if the event occurred in interval 3 this corresponds to $w_2*(d_2/\sum d_i)$ person-time at risk; and if the event occurred in interval 4 this corresponds to $w_2*(d_3/\sum d_i)$ person-time at risk. The person-time attributed to interval 2 is therefore $(w_2-d_1) + d_1/2*(d_1/\sum d_i) + d_1*(d_2+d_3)/\sum d_i)$.

4.3 SIMULATION STUDY

A simulation study was carried out to assess the performance of each of the following estimation methods as described in section 4.2:

- the Finkelstein proportional hazards model (F).
- a piecewise exponential model (PE) with time divided into 5 periods with constant hazard within periods. Breakpoints were chosen so that the expected number of events using an MKM model (section 3.2.2) with no covariates were equal within each interval.
- an approximate PE model (APE) with 5 intervals and breakpoints as for the PE model.
- MCox and RCox models, using both an exact likelihood and a Breslow approximation for the treatment of ties.
- a parametric Weibull survival model (W).

4.3.1 Methods

Data were generated from a hypothetical follow-up study of duration D as described in section 3.3.1. Survival times were generated from a Weibull distribution with a single binary covariate corresponding to a log hazard ratio β. In each dataset $n/2$ observations were generated for each covariate value. The individual censoring intervals were then determined.

1,000 data sets were simulated for each combination of the following factors:

(a) Shape parameter of Weibull distribution in individuals, γ=0.6, 1.0, 1.5 representing a decreasing, constant, and increasing hazard function respectively. For each value of the shape parameter, the scale parameter, λ, was selected so that $S(D)$=0.05 when x_i=0.

(b) Total sample size, n=50, 100, 200, 400, 800 individuals.

(c) Probability of missing a scheduled visit, p=0, 0.4.

(d) Number of evenly spaced scheduled follow up visits, v=3, 6, 12.

(e) Log hazard ratio, β=0, 0.4, 0.8 corresponding to a hazard ratio of 1, 1.49 and 2.23 respectively.

Figure 4.2 shows the survivor functions for each shape parameter γ, and log hazard ratio, β. For each combination of β, γ, n, p and v, the mean log hazard ratio, mean squared error (MSE), mean estimated standard error of the log hazard ratio, empirical

standard error of the log hazard ratio and coverage probability were estimated. For 1000 repetitions the empirical variance is estimated to within 7% of the true value (or 3% for the standard deviation).

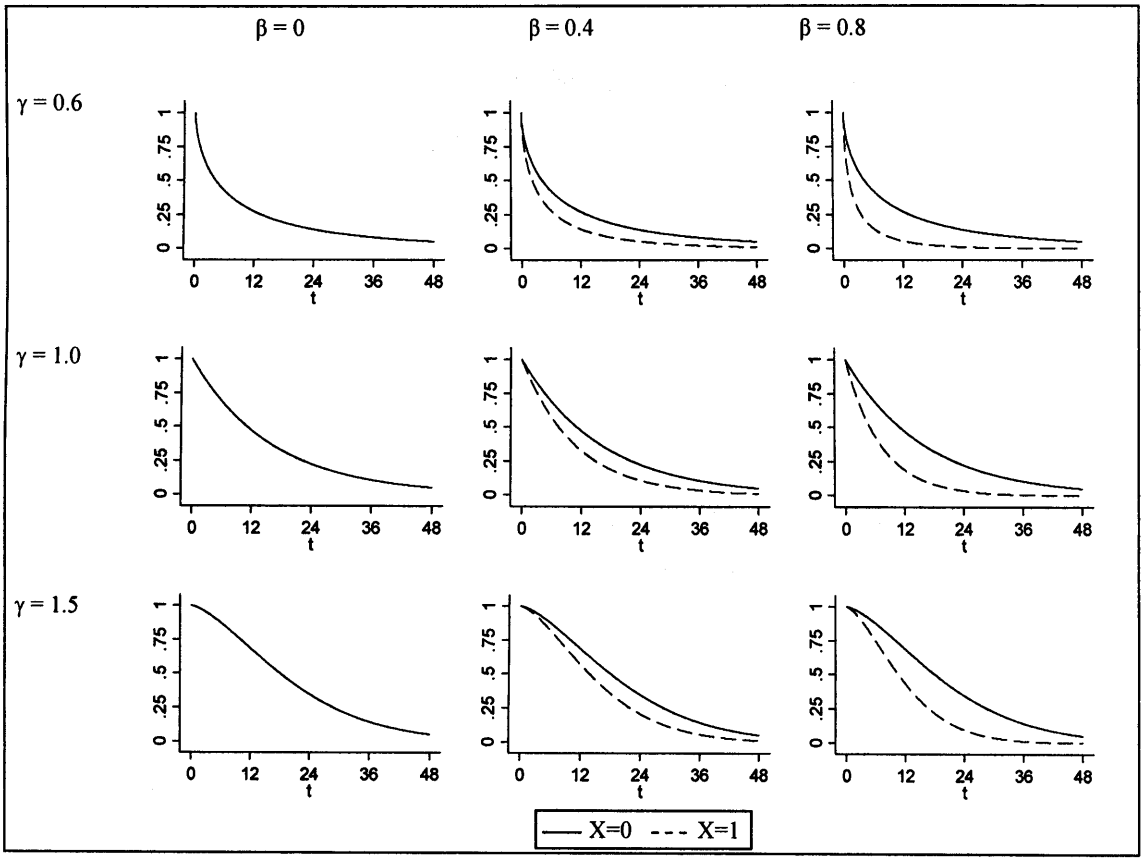

Figure 4.2: Survivor distributions in simulated data

4.3.2 Implementation

All analyses were implemented using SAS Version 9.1 [39]. Data were rounded from days to fortnightly units prior to analysis. Finkelstein models were fitted using the (Dual) Quasi-Newton or Newton-Raphson Ridge non-linear optimisation routines in IML (Appendix 4.A). PE models were fitted using (Dual) Quasi-Newton non-linear optimisation routines in IML (Appendix 4.B), Weibull models were fitted using PROC LIFEREG, APE models were fitted using PROC GENMOD after creation of the necessary data sets (Appendix 4.C), and MCox and RCox methods fitted using PROC PHREG.

4.3.3 Convergence

The default convergence criterion for PROC LIFEREG was used, namely that the maximum change in the parameter estimates between Newton-Raphson steps was less than 0.001. The non-linear optimisation routines in IML stop the iteration process when at least one of the default termination criterion are met. This was usually relative gradient convergence for the Finkelstein model and absolute function convergence for the piecewise exponential model. Initially the Finkelstein model failed to converge for around 10% of all simulations. This was determined to be due to the maximisation procedure attempting to take the logarithm of zero and resolved by setting the contribution from such terms to be a large negative number.

The Weibull model failed to converge in 45 (0.02%) simulations overall, from six scenarios. The Finkelstein model failed to converge in 88 (0.03%) simulations overall, spread between 14 different scenarios. The piecewise exponential model failed to converge in 149 (0.06%) simulations overall, spread between 33 different scenarios. All failed runs for the Weibull model and the majority of failed runs for the piecewise exponential and Finkelstein models were from scenarios with n=50 and v=3.

To investigate the reason for lack of convergence, the data from scenarios with n=50 and v=3 that had not converged were examined and a separate Weibull model fitted to each group (where X=0 or X=1). Problems occurred when the censoring intervals in a sample were wide and the left endpoint for most individuals was zero. A Weibull model could not be fitted to data from the single group in this case due to the lack of information provided by the censoring intervals. Covariate estimates were unstable in the proportional hazards model since all the information is being driven by one sample.

4.3.4 Results

a) Treatment of ties for Cox models

For β=0.4 or 0.8, both the MCox and RCox methods under-estimated β when using either an exact likelihood or the Breslow approximation (Figure 4.3). However, the bias was more severe using the Breslow approximation, at least 4% higher for RCox and 2% higher for MCox. The results of RCox and MCox analyses in the rest of this chapter use an exact likelihood approach.

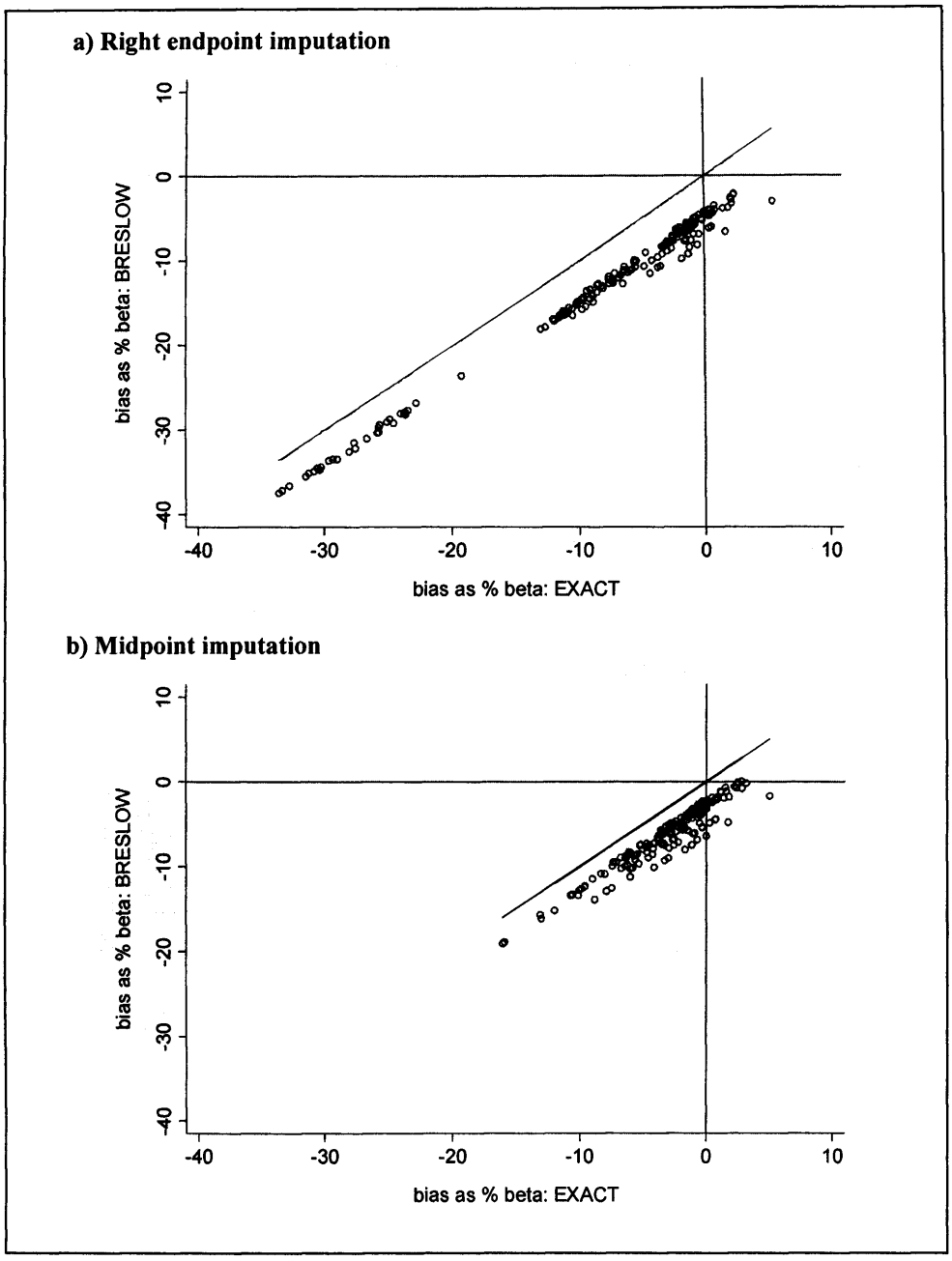

Figure 4.3: Effect of method of treating ties in Cox model on bias in estimation of covariate effect

Table 4.1: Results from simulation study

					Mean Bias						Coverage (%)						MSE					
β	γ	n	P	v	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox
0	0.6	50	0	3	0.002	0.002	0.004	-0.004	-0.004	-0.005	93.7	94.7	93.2	94.8	93.2	93.8	0.130	0.149	0.138	0.086	0.095	0.097
0	0.6	50	0	6	-0.015	-0.018	-0.017	-0.016	-0.014	-0.015	93.3	93.7	93.1	93.6	93.5	93.5	0.109	0.117	0.113	0.099	0.097	0.097
0	0.6	50	0	12	0.002	-0.000	0.002	0.002	0.001	0.001	95.7	95.2	94.8	95.0	95.9	95.9	0.090	0.094	0.093	0.091	0.086	0.085
0	0.6	50	0.4	3	-0.008	-0.016	-0.012	-0.003	-0.005	-0.005	95.7	98.5	96.2	97.6	94.5	94.2	0.161	0.197	0.174	0.074	0.106	0.109
0	0.6	50	0.4	6	0.003	0.003	0.004	0.005	0.006	0.005	94.4	95.1	94.0	95.9	94.9	94.4	0.126	0.144	0.132	0.085	0.096	0.100
0	0.6	50	0.4	12	0.012	0.014	0.013	0.012	0.014	0.014	95.1	95.1	94.7	95.0	95.1	94.6	0.102	0.108	0.107	0.094	0.094	0.097
0	0.6	100	0	3	0.001	-0.001	0.001	0.001	-0.002	-0.003	95.6	96.1	95.0	96.6	95.8	95.3	0.055	0.060	0.057	0.039	0.044	0.044
0	0.6	100	0	6	-0.001	0.001	-0.000	-0.001	-0.000	-0.000	93.7	93.6	93.2	93.4	94.3	93.7	0.050	0.052	0.051	0.046	0.045	0.045
0	0.6	100	0	12	0.004	0.004	0.004	0.004	0.003	0.003	94.5	93.8	93.8	93.9	94.4	94.3	0.049	0.050	0.051	0.049	0.047	0.048
0	0.6	100	0.4	3	0.000	0.004	0.001	0.001	0.001	0.004	94.6	95.8	94.7	97.4	93.3	95.1	0.072	0.083	0.076	0.037	0.053	0.053
0	0.6	100	0.4	6	-0.008	-0.009	-0.008	-0.007	-0.007	-0.008	94.7	95.6	94.6	94.0	92.4	89.7	0.055	0.059	0.056	0.049	0.058	0.066
0	0.6	100	0.4	12	0.012	0.013	0.013	0.012	0.013	0.013	94.9	95.3	94.9	95.1	94.3	94.1	0.048	0.049	0.049	0.044	0.045	0.046
0	0.6	200	0	3	-0.002	-0.002	-0.002	-0.001	0.000	0.001	93.8	94.3	94.0	95.4	94.4	94.1	0.028	0.030	0.029	0.020	0.023	0.022
0	0.6	200	0	6	-0.006	-0.007	-0.007	-0.006	-0.006	-0.006	95.0	95.3	95.0	95.7	96.2	95.9	0.022	0.023	0.023	0.021	0.020	0.021
0	0.6	200	0	12	-0.010	-0.010	-0.010	-0.010	-0.011	-0.010	95.3	95.2	94.9	95.1	95.5	95.5	0.021	0.021	0.021	0.020	0.020	0.020
0	0.6	200	0.4	3	-0.003	-0.004	-0.003	-0.001	-0.001	-0.001	93.0	93.2	92.5	96.9	93.8	95.9	0.037	0.040	0.038	0.019	0.026	0.025
0	0.6	200	0.4	6	0.002	0.002	0.002	0.003	0.003	0.004	95.3	95.3	95.2	95.8	94.7	94.4	0.026	0.026	0.026	0.020	0.023	0.024
0	0.6	200	0.4	12	-0.002	-0.002	-0.002	-0.003	-0.004	-0.006	94.1	94.4	93.9	93.9	92.5	91.1	0.024	0.025	0.024	0.023	0.025	0.027
0	0.6	400	0	3	-0.004	-0.004	-0.004	-0.004	-0.005	-0.004	96.2	96.6	96.1	97.4	96.5	96.3	0.012	0.012	0.012	0.009	0.010	0.010
0	0.6	400	0	6	-0.006	-0.007	-0.006	-0.006	-0.006	-0.005	94.5	95.1	94.3	94.6	94.8	95.3	0.012	0.012	0.012	0.011	0.011	0.011
0	0.6	400	0	12	0.005	0.011	0.005	0.005	0.005	0.005	95.4	95.0	94.9	95.2	95.5	95.2	0.011	0.017	0.011	0.011	0.011	0.011
0	0.6	400	0.4	3	-0.003	-0.003	-0.003	-0.001	-0.001	0.002	94.4	95.2	94.2	97.8	94.7	95.3	0.016	0.017	0.016	0.008	0.012	0.012
0	0.6	400	0.4	6	0.005	0.005	0.005	0.004	0.004	0.003	94.9	95.2	95.1	95.9	94.1	94.0	0.013	0.013	0.013	0.010	0.011	0.012
0	0.6	400	0.4	12	0.001	0.001	0.001	0.001	0.001	0.001	95.2	95.8	95.3	95.8	95.6	95.3	0.011	0.012	0.012	0.010	0.011	0.011
0	0.6	800	0	3	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002	94.9	96.1	94.7	96.6	94.7	94.7	0.006	0.007	0.007	0.005	0.006	0.006
0	0.6	800	0	6	0.002	0.002	0.002	0.002	0.002	0.002	93.6	94.6	93.4	93.6	93.1	93.6	0.006	0.006	0.006	0.005	0.005	0.006
0	0.6	800	0	12	-0.001	0.013	-0.001	-0.001	-0.001	-0.001	95.7	95.9	95.6	96.0	96.0	96.1	0.005	0.018	0.005	0.005	0.005	0.005
0	0.6	800	0.4	3	0.003	0.003	0.003	0.002	0.002	0.001	95.6	96.1	95.7	98.2	95.6	94.8	0.008	0.008	0.008	0.004	0.006	0.006
0	0.6	800	0.4	6	-0.001	-0.000	-0.000	-0.000	-0.000	0.001	95.1	95.5	95.0	94.4	92.0	90.0	0.006	0.006	0.006	0.006	0.007	0.007
0	0.6	800	0.4	12	0.003	0.006	0.003	0.003	0.003	0.003	94.1	95.9	95.1	95.9	95.3	94.7	0.006	0.009	0.006	0.005	0.006	0.006
0	1.0	50	0	3	0.002	0.007	0.001	0.002	0.004	0.011	93.9	95.9	93.9	96.0	94.3	94.4	0.113	0.130	0.118	0.082	0.094	0.096
0	1.0	50	0	6	-0.007	-0.009	-0.008	-0.007	-0.010	-0.009	95.6	95.6	95.3	95.7	95.7	95.5	0.095	0.105	0.100	0.090	0.090	0.089
0	1.0	50	0	12	-0.008	-0.009	-0.010	-0.010	-0.009	-0.010	93.9	93.4	93.9	94.1	94.4	94.5	0.100	0.107	0.105	0.101	0.098	0.099
0	1.0	50	0.4	3	0.012	0.017	0.017	0.008	0.008	0.006	94.8	96.5	95.7	98.2	95.8	94.6	0.162	0.202	0.172	0.073	0.105	0.105
0	1.0	50	0.4	6	0.011	0.006	0.009	0.008	0.009	0.005	93.3	94.2	93.4	94.8	93.6	93.0	0.123	0.138	0.128	0.095	0.107	0.117
0	1.0	50	0.4	12	-0.005	-0.004	-0.004	-0.005	-0.004	-0.006	94.6	94.9	94.2	94.7	94.9	95.0	0.097	0.104	0.102	0.091	0.090	0.093

					Mean Bias						Coverage (%)						MSE					
β	γ	n	P	v	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox
0	1.0	100	0	3	0.003	0.003	0.003	0.003	0.003	0.002	96.1	96.6	95.8	97.1	96.6	96.0	0.048	0.053	0.049	0.036	0.041	0.042
0	1.0	100	0	6	0.002	0.002	0.002	0.002	0.002	0.001	94.5	95.1	94.6	95.1	94.8	94.2	0.044	0.046	0.045	0.041	0.042	0.043
0	1.0	100	0	12	0.002	0.002	0.001	0.002	0.002	0.002	94.5	94.8	94.7	95.0	94.6	94.9	0.042	0.044	0.043	0.042	0.042	0.042
0	1.0	100	0.4	3	0.001	-0.001	0.001	0.002	0.001	0.000	95.4	96.2	95.4	98.2	94.9	95.5	0.067	0.078	0.070	0.038	0.051	0.049
0	1.0	100	0.4	6	-0.010	-0.011	-0.010	-0.009	-0.009	-0.008	94.6	94.6	94.7	95.9	95.1	95.0	0.052	0.057	0.053	0.041	0.046	0.046
0	1.0	100	0.4	12	-0.003	-0.004	-0.003	-0.003	-0.003	-0.003	94.1	94.9	94.2	94.7	94.4	94.5	0.051	0.052	0.051	0.047	0.048	0.048
0	1.0	200	0	3	0.008	0.007	0.008	0.007	0.007	0.006	95.7	95.8	95.5	96.7	95.9	96.1	0.024	0.025	0.024	0.018	0.021	0.021
0	1.0	200	0	6	0.003	0.003	0.003	0.003	0.003	0.003	93.6	94.3	93.2	94.0	93.6	94.1	0.024	0.025	0.024	0.023	0.023	0.023
0	1.0	200	0	12	0.005	0.005	0.005	0.005	0.005	0.005	95.4	95.7	95.5	95.8	95.9	95.3	0.022	0.022	0.022	0.022	0.022	0.022
0	1.0	200	0.4	3	-0.002	-0.003	-0.002	-0.003	-0.003	-0.004	94.6	95.0	94.5	97.9	95.6	95.7	0.032	0.035	0.033	0.018	0.024	0.024
0	1.0	200	0.4	6	-0.003	-0.003	-0.003	-0.002	-0.003	-0.003	94.5	94.5	94.3	95.4	95.1	93.9	0.025	0.026	0.025	0.020	0.023	0.025
0	1.0	200	0.4	12	0.003	0.003	0.002	0.002	0.002	0.001	94.9	95.1	94.9	95.3	94.3	92.8	0.023	0.023	0.023	0.022	0.023	0.027
0	1.0	400	0	3	-0.003	-0.003	-0.004	-0.003	-0.004	-0.004	95.1	95.3	95.2	96.3	95.1	95.8	0.012	0.013	0.012	0.009	0.011	0.011
0	1.0	400	0	6	0.001	0.001	0.001	0.001	0.001	0.001	95.0	94.8	94.8	95.2	95.0	94.6	0.011	0.011	0.011	0.010	0.010	0.011
0	1.0	400	0	12	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	94.6	94.9	94.8	94.8	95.0	94.6	0.011	0.011	0.011	0.011	0.011	0.011
0	1.0	400	0.4	3	0.005	0.006	0.005	0.005	0.005	0.006	95.0	95.5	95.3	96.8	95.0	94.8	0.015	0.016	0.015	0.009	0.013	0.013
0	1.0	400	0.4	6	-0.005	-0.005	-0.005	-0.004	-0.004	-0.003	94.7	94.7	94.6	96.1	95.3	95.2	0.012	0.012	0.012	0.010	0.011	0.011
0	1.0	400	0.4	12	-0.003	-0.003	-0.003	-0.003	-0.003	-0.004	94.6	94.9	94.6	94.9	94.7	92.8	0.011	0.012	0.011	0.011	0.011	0.013
0	1.0	800	0	3	0.002	0.002	0.002	0.002	0.001	0.001	95.8	95.5	95.9	97.0	95.8	95.2	0.006	0.006	0.006	0.004	0.005	0.005
0	1.0	800	0	6	-0.003	-0.002	-0.003	-0.002	-0.002	-0.002	94.0	94.4	94.3	94.6	94.3	93.4	0.006	0.007	0.006	0.005	0.005	0.005
0	1.0	800	0	12	0.002	0.003	0.001	0.001	0.001	0.001	95.0	94.8	95.0	95.1	94.5	94.1	0.005	0.007	0.005	0.005	0.005	0.005
0	1.0	800	0.4	3	0.000	0.001	0.000	0.001	0.002	0.005	95.3	95.3	95.3	97.1	94.9	93.7	0.007	0.008	0.007	0.004	0.006	0.007
0	1.0	800	0.4	6	0.000	0.000	0.000	-0.000	0.000	-0.000	94.6	94.6	94.5	95.5	94.4	94.0	0.006	0.006	0.006	0.005	0.005	0.006
0	1.0	800	0.4	12	0.000	0.001	0.000	0.000	0.000	0.000	94.9	94.6	95.0	95.4	94.8	94.8	0.005	0.006	0.005	0.005	0.005	0.005
0	1.5	50	0	3	-0.021	-0.024	-0.020	-0.018	-0.016	-0.020	93.9	94.8	93.9	96.6	94.1	94.1	0.112	0.140	0.121	0.077	0.092	0.094
0	1.5	50	0	6	-0.006	-0.007	-0.007	-0.007	-0.007	-0.008	94.0	93.8	93.7	94.3	93.3	94.0	0.099	0.109	0.103	0.090	0.095	0.095
0	1.5	50	0	12	-0.010	-0.008	-0.009	-0.008	-0.008	-0.007	93.3	92.8	92.9	93.4	94.0	93.8	0.098	0.106	0.102	0.097	0.097	0.098
0	1.5	50	0.4	3	-0.007	-0.008	-0.005	-0.006	-0.005	0.001	93.0	95.0	93.5	97.2	93.7	95.3	0.181	0.259	0.209	0.083	0.120	0.113
0	1.5	50	0.4	6	-0.007	-0.006	-0.005	-0.004	-0.005	0.003	95.9	96.6	95.5	97.5	96.2	96.0	0.103	0.122	0.109	0.077	0.091	0.094
0	1.5	50	0.4	12	0.003	0.003	0.002	0.003	0.002	0.003	93.4	94.0	93.4	94.4	94.3	94.4	0.103	0.111	0.107	0.094	0.098	0.099
0	1.5	100	0	3	-0.007	-0.005	-0.007	-0.005	-0.005	-0.003	94.0	94.8	93.9	96.5	94.6	94.8	0.051	0.057	0.052	0.037	0.044	0.045
0	1.5	100	0	6	0.012	0.013	0.012	0.012	0.011	0.011	93.7	93.7	93.6	94.7	94.3	94.5	0.047	0.050	0.048	0.043	0.045	0.044
0	1.5	100	0	12	-0.007	-0.007	-0.008	-0.007	-0.007	-0.007	93.7	93.4	93.9	93.9	94.3	94.0	0.047	0.049	0.048	0.046	0.047	0.047
0	1.5	100	0.4	3	0.010	0.012	0.012	0.007	0.007	0.008	95.1	95.9	95.0	98.2	95.3	97.0	0.070	0.083	0.074	0.037	0.051	0.045
0	1.5	100	0.4	6	0.012	0.013	0.013	0.011	0.011	0.009	95.1	95.8	95.4	96.5	95.5	95.4	0.052	0.057	0.053	0.040	0.046	0.046
0	1.5	100	0.4	12	0.005	0.005	0.005	0.005	0.005	0.003	94.6	94.4	94.2	94.9	94.7	94.5	0.048	0.050	0.049	0.044	0.046	0.047

					Mean Bias						Coverage (%)						MSE					
β	γ	n	P	v	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox
0	1.5	200	0	3	-0.002	-0.003	-0.002	-0.002	-0.002	-0.002	94.2	94.8	94.4	95.5	94.8	94.7	0.025	0.026	0.025	0.019	0.022	0.023
0	1.5	200	0	6	-0.002	-0.001	-0.001	-0.001	-0.002	-0.003	94.9	94.9	95.1	95.6	95.1	95.3	0.022	0.023	0.023	0.021	0.022	0.022
0	1.5	200	0	12	0.004	0.003	0.003	0.003	0.003	0.003	95.3	95.6	95.3	95.4	95.9	95.2	0.023	0.023	0.022	0.022	0.022	0.022
0	1.5	200	0.4	3	0.000	0.001	0.001	0.001	0.000	0.003	93.8	94.4	94.0	96.6	93.9	94.6	0.038	0.042	0.038	0.021	0.029	0.027
0	1.5	200	0.4	6	-0.006	-0.007	-0.006	-0.006	-0.005	-0.006	95.0	94.9	95.5	96.9	95.2	95.1	0.026	0.028	0.026	0.020	0.023	0.023
0	1.5	200	0.4	12	-0.004	-0.004	-0.004	-0.004	-0.004	-0.005	94.1	94.3	94.2	94.7	94.1	91.0	0.024	0.024	0.024	0.022	0.023	0.029
0	1.5	400	0	3	0.003	0.003	0.003	0.002	0.003	0.001	94.9	94.6	95.1	96.3	94.7	94.2	0.012	0.012	0.012	0.009	0.011	0.011
0	1.5	400	0	6	0.005	0.005	0.005	0.005	0.005	0.005	94.6	94.5	94.5	95.1	94.8	94.2	0.011	0.011	0.011	0.010	0.011	0.011
0	1.5	400	0	12	-0.003	-0.003	-0.003	-0.003	-0.003	-0.004	94.9	94.9	95.1	95.3	94.9	94.7	0.012	0.012	0.011	0.011	0.011	0.011
0	1.5	400	0.4	3	0.001	0.000	0.001	-0.000	-0.001	-0.003	94.7	95.1	94.3	97.0	94.6	94.0	0.016	0.017	0.016	0.010	0.013	0.014
0	1.5	400	0.4	6	-0.003	-0.003	-0.003	-0.003	-0.003	-0.001	94.7	95.1	95.0	96.3	94.8	94.6	0.013	0.014	0.013	0.011	0.012	0.012
0	1.5	400	0.4	12	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	94.9	94.9	94.8	95.8	94.7	94.8	0.012	0.012	0.012	0.011	0.011	0.011
0	1.5	800	0	3	-0.002	-0.002	-0.002	-0.002	-0.002	-0.003	93.7	93.4	93.9	96.1	93.6	94.1	0.006	0.006	0.006	0.005	0.005	0.006
0	1.5	800	0	6	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	94.9	94.8	95.0	95.5	95.0	94.9	0.005	0.005	0.005	0.005	0.005	0.005
0	1.5	800	0	12	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	93.4	93.5	93.8	94.1	93.8	93.4	0.006	0.006	0.006	0.006	0.006	0.006
0	1.5	800	0.4	3	-0.001	-0.001	-0.001	-0.001	-0.001	0.001	93.9	93.8	93.9	97.2	94.5	94.6	0.008	0.008	0.008	0.005	0.006	0.007
0	1.5	800	0.4	6	-0.002	-0.002	-0.002	-0.001	-0.001	-0.001	95.1	95.1	95.2	95.9	95.1	90.4	0.006	0.006	0.006	0.005	0.006	0.008
0	1.5	800	0.4	12	0.002	0.003	0.002	0.002	0.002	0.004	94.7	94.6	94.7	95.1	94.5	92.1	0.006	0.006	0.006	0.005	0.006	0.006
0.4	0.6	50	0	3	0.031	0.048	0.035	-0.021	-0.024	-0.026	94.8	95.1	94.3	95.5	94.3	94.1	0.127	0.145	0.131	0.093	0.101	0.103
0.4	0.6	50	0	6	0.022	0.030	0.028	0.014	0.000	-0.002	93.4	93.3	93.3	93.5	94.2	94.4	0.117	0.125	0.122	0.111	0.104	0.104
0.4	0.6	50	0	12	0.033	0.040	0.039	0.041	0.020	0.022	92.1	92.4	91.9	91.9	93.5	93.6	0.114	0.119	0.119	0.119	0.108	0.108
0.4	0.6	50	0.4	3	0.040	0.061	0.043	-0.074	-0.041	-0.111	95.2	97.6	96.1	96.5	94.8	93.4	0.144	0.166	0.147	0.084	0.104	0.123
0.4	0.6	50	0.4	6	0.038	0.052	0.044	-0.010	-0.007	-0.027	95.2	95.9	95.0	95.9	94.5	94.5	0.119	0.134	0.126	0.093	0.099	0.101
0.4	0.6	50	0.4	12	0.013	0.024	0.022	0.007	-0.002	-0.006	93.5	94.2	93.7	94.0	94.1	93.7	0.108	0.115	0.113	0.103	0.101	0.101
0.4	0.6	100	0	3	0.007	0.015	0.010	-0.039	-0.035	-0.042	94.6	95.1	94.9	95.3	94.5	94.8	0.055	0.060	0.056	0.044	0.048	0.048
0.4	0.6	100	0	6	0.001	0.006	0.005	-0.009	-0.017	-0.017	95.4	95.6	94.8	94.7	95.8	96.0	0.049	0.050	0.050	0.047	0.046	0.046
0.4	0.6	100	0	12	0.014	0.018	0.019	0.018	0.007	0.007	94.3	93.8	93.7	93.5	94.4	94.3	0.049	0.051	0.051	0.051	0.048	0.048
0.4	0.6	100	0.4	3	0.022	0.038	0.027	-0.075	-0.048	-0.119	95.4	96.6	95.2	94.1	92.1	87.5	0.067	0.076	0.070	0.052	0.063	0.087
0.4	0.6	100	0.4	6	0.008	0.016	0.010	-0.034	-0.026	-0.046	95.2	96.5	94.8	96.0	95.9	95.0	0.052	0.054	0.052	0.042	0.045	0.046
0.4	0.6	100	0.4	12	0.002	0.005	0.005	-0.009	-0.013	-0.019	95.4	94.9	94.8	95.1	94.7	95.2	0.047	0.049	0.048	0.045	0.045	0.047
0.4	0.6	200	0	3	0.009	0.014	0.011	-0.035	-0.030	-0.036	94.4	94.4	94.2	95.6	94.8	94.7	0.026	0.028	0.026	0.022	0.023	0.024
0.4	0.6	200	0	6	0.012	0.016	0.015	0.000	-0.003	-0.005	93.4	94.0	93.4	94.0	94.5	94.1	0.026	0.027	0.026	0.024	0.024	0.025
0.4	0.6	200	0	12	0.001	0.004	0.005	0.003	-0.004	-0.005	93.6	94.0	93.4	93.6	93.9	94.0	0.024	0.025	0.024	0.024	0.024	0.024
0.4	0.6	200	0.4	3	0.014	0.025	0.018	-0.079	-0.052	-0.125	96.5	96.6	96.4	95.4	95.1	90.2	0.029	0.031	0.030	0.024	0.024	0.037
0.4	0.6	200	0.4	6	0.004	0.009	0.006	-0.036	-0.025	-0.046	94.1	95.0	94.3	94.6	93.7	93.2	0.028	0.029	0.028	0.024	0.025	0.027
0.4	0.6	200	0.4	12	-0.000	0.003	0.002	-0.010	-0.011	-0.015	93.9	93.7	93.7	94.0	94.1	94.2	0.026	0.027	0.026	0.025	0.025	0.025

					Mean Bias						Coverage (%)						MSE					
β	γ	n	P	v	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox
0.4	0.6	400	0	3	0.004	0.006	0.005	-0.037	-0.032	-0.039	94.4	94.4	94.4	93.5	93.6	92.8	0.014	0.014	0.014	0.012	0.013	0.013
0.4	0.6	400	0	6	0.001	0.003	0.003	-0.012	-0.012	-0.014	95.0	95.1	94.9	95.4	94.9	94.7	0.012	0.012	0.012	0.012	0.012	0.012
0.4	0.6	400	0	12	-0.003	0.002	-0.000	-0.001	-0.007	-0.007	93.9	93.7	94.0	93.6	93.8	93.6	0.012	0.016	0.012	0.012	0.012	0.012
0.4	0.6	400	0.4	3	-0.002	0.003	-0.000	-0.090	-0.064	-0.133	93.8	94.3	93.8	88.1	88.5	75.3	0.016	0.017	0.016	0.020	0.020	0.035
0.4	0.6	400	0.4	6	0.007	0.010	0.009	-0.034	-0.023	-0.045	95.1	95.1	95.0	94.6	93.8	92.3	0.013	0.013	0.013	0.012	0.012	0.014
0.4	0.6	400	0.4	12	0.006	0.007	0.007	-0.006	-0.006	-0.012	95.1	95.4	95.0	94.6	93.9	92.7	0.012	0.012	0.012	0.012	0.012	0.013
0.4	0.6	800	0	3	0.004	0.004	0.005	-0.037	-0.032	-0.038	95.3	95.1	95.1	92.7	93.2	91.6	0.006	0.006	0.006	0.006	0.006	0.007
0.4	0.6	800	0	6	-0.001	-0.000	0.000	-0.014	-0.013	-0.015	96.5	96.9	96.3	96.1	96.1	95.8	0.006	0.006	0.006	0.006	0.006	0.006
0.4	0.6	800	0	12	-0.001	0.013	0.001	0.000	-0.005	-0.005	95.2	94.5	95.1	95.0	95.2	95.3	0.006	0.015	0.006	0.006	0.006	0.006
0.4	0.6	800	0.4	3	-0.003	-0.001	-0.002	-0.090	-0.064	-0.134	96.3	96.5	96.3	82.0	88.2	60.0	0.007	0.008	0.007	0.013	0.010	0.024
0.4	0.6	800	0.4	6	0.001	0.003	0.002	-0.038	-0.027	-0.049	95.3	95.4	95.1	93.5	93.9	90.7	0.006	0.007	0.006	0.007	0.007	0.008
0.4	0.6	800	0.4	12	0.001	0.004	0.002	-0.011	-0.010	-0.017	94.9	95.2	95.1	95.3	95.2	94.5	0.006	0.008	0.006	0.006	0.006	0.006
0.4	1.0	50	0	3	0.028	0.049	0.034	-0.015	-0.007	-0.023	95.7	96.4	95.5	96.8	96.0	95.9	0.109	0.126	0.113	0.084	0.092	0.098
0.4	1.0	50	0	6	0.003	0.013	0.008	-0.006	-0.011	-0.013	94.0	94.1	93.4	93.6	93.8	93.9	0.113	0.121	0.117	0.109	0.108	0.109
0.4	1.0	50	0	12	0.011	0.021	0.018	0.014	0.006	0.006	93.6	94.0	93.7	94.1	95.0	94.5	0.101	0.106	0.105	0.102	0.097	0.098
0.4	1.0	50	0.4	3	0.012	0.052	0.025	-0.080	-0.052	-0.131	94.7	96.3	94.5	95.1	93.4	92.0	0.164	0.209	0.175	0.103	0.124	0.138
0.4	1.0	50	0.4	6	0.028	0.047	0.037	-0.008	0.000	-0.024	94.4	95.7	94.3	95.7	95.1	94.6	0.111	0.130	0.120	0.090	0.098	0.100
0.4	1.0	50	0.4	12	0.023	0.033	0.028	0.015	0.011	0.009	95.0	94.8	94.5	95.0	95.3	95.0	0.100	0.110	0.106	0.096	0.095	0.099
0.4	1.0	100	0	3	0.012	0.025	0.014	-0.026	-0.014	-0.029	94.6	95.7	94.6	96.1	95.6	95.3	0.050	0.054	0.051	0.041	0.045	0.047
0.4	1.0	100	0	6	0.020	0.025	0.022	0.011	0.011	0.009	94.3	95.2	94.7	95.2	95.0	94.7	0.052	0.054	0.053	0.049	0.050	0.050
0.4	1.0	100	0	12	-0.003	-0.000	-0.001	-0.004	-0.008	-0.009	96.3	95.9	96.1	96.1	96.1	96.2	0.044	0.045	0.044	0.044	0.043	0.043
0.4	1.0	100	0.4	3	0.009	0.031	0.014	-0.073	-0.042	-0.122	94.4	95.3	94.4	95.8	94.2	92.4	0.066	0.077	0.069	0.047	0.054	0.069
0.4	1.0	100	0.4	6	0.013	0.022	0.014	-0.019	-0.007	-0.026	95.0	95.5	94.5	95.8	94.8	93.8	0.051	0.055	0.052	0.044	0.048	0.056
0.4	1.0	100	0.4	12	0.004	0.008	0.006	-0.006	-0.006	-0.012	93.9	94.3	94.5	94.3	93.9	92.8	0.052	0.054	0.053	0.050	0.051	0.054
0.4	1.0	200	0	3	0.002	0.008	0.004	-0.032	-0.019	-0.034	95.7	95.4	95.0	95.6	95.2	93.8	0.024	0.026	0.025	0.021	0.023	0.025
0.4	1.0	200	0	6	0.005	0.007	0.005	-0.005	-0.003	-0.005	94.6	94.8	94.7	94.8	94.9	94.7	0.023	0.024	0.023	0.022	0.022	0.023
0.4	1.0	200	0	12	0.001	0.002	0.002	-0.001	-0.002	-0.003	95.2	94.6	94.7	94.9	95.1	95.2	0.024	0.024	0.024	0.024	0.024	0.024
0.4	1.0	200	0.4	3	0.006	0.017	0.007	-0.070	-0.040	-0.121	95.3	95.8	95.8	95.0	95.0	88.2	0.031	0.034	0.032	0.026	0.027	0.041
0.4	1.0	200	0.4	6	0.005	0.010	0.006	-0.026	-0.012	-0.039	94.1	94.6	94.2	93.4	93.0	87.0	0.028	0.030	0.029	0.027	0.029	0.038
0.4	1.0	200	0.4	12	0.003	0.005	0.003	-0.007	-0.005	-0.010	93.9	94.3	93.9	94.3	94.1	92.6	0.025	0.026	0.026	0.025	0.025	0.027
0.4	1.0	400	0	3	0.007	0.009	0.007	-0.028	-0.014	-0.029	94.1	94.2	94.2	94.6	94.5	94.2	0.013	0.013	0.013	0.011	0.012	0.013
0.4	1.0	400	0	6	-0.001	0.000	-0.001	-0.010	-0.007	-0.010	95.3	95.1	95.1	95.4	95.1	95.3	0.012	0.012	0.012	0.011	0.012	0.012
0.4	1.0	400	0	12	-0.003	-0.002	-0.002	-0.005	-0.005	-0.005	94.1	93.7	94.0	93.8	93.9	93.7	0.013	0.013	0.013	0.013	0.013	0.013
0.4	1.0	400	0.4	3	0.002	0.009	0.004	-0.070	-0.038	-0.117	93.9	93.9	93.7	90.0	90.4	77.3	0.016	0.017	0.017	0.018	0.017	0.034
0.4	1.0	400	0.4	6	0.003	0.006	0.004	-0.028	-0.014	-0.041	94.3	94.4	94.3	94.8	94.5	92.9	0.014	0.014	0.014	0.013	0.013	0.015
0.4	1.0	400	0.4	12	0.002	0.003	0.002	-0.008	-0.004	-0.010	94.2	94.7	94.6	94.8	94.3	92.9	0.012	0.012	0.012	0.012	0.012	0.013

					Mean Bias						Coverage (%)						MSE					
β	γ	n	P	v	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox
0.4	1.0	800	0	3	0.002	0.003	0.002	-0.031	-0.017	-0.033	94.7	95.0	95.1	93.8	94.3	91.6	0.006	0.006	0.006	0.006	0.006	0.007
0.4	1.0	800	0	6	0.003	0.004	0.003	-0.006	-0.002	-0.006	95.1	95.4	95.6	95.3	95.8	95.1	0.006	0.006	0.006	0.005	0.006	0.006
0.4	1.0	800	0	12	0.001	0.002	0.002	-0.001	-0.000	-0.001	96.0	96.4	96.2	96.3	96.2	96.2	0.005	0.006	0.005	0.005	0.005	0.005
0.4	1.0	800	0.4	3	-0.003	-0.001	-0.003	-0.075	-0.043	-0.123	95.5	95.5	95.6	87.4	93.3	66.3	0.007	0.008	0.007	0.011	0.008	0.021
0.4	1.0	800	0.4	6	-0.003	-0.002	-0.003	-0.033	-0.019	-0.045	94.4	94.4	94.3	92.1	92.4	85.5	0.007	0.007	0.007	0.007	0.007	0.010
0.4	1.0	800	0.4	12	0.002	0.003	0.003	-0.008	-0.003	-0.009	94.5	94.7	94.7	94.2	94.2	93.1	0.006	0.006	0.006	0.006	0.006	0.007
0.4	1.5	50	0	3	0.007	0.034	0.014	-0.039	-0.023	-0.042	94.1	95.6	94.0	95.5	93.9	94.4	0.114	0.137	0.120	0.088	0.100	0.106
0.4	1.5	50	0	6	0.023	0.036	0.030	0.010	0.013	0.008	94.5	94.8	94.5	95.0	94.6	95.0	0.100	0.112	0.106	0.094	0.096	0.098
0.4	1.5	50	0	12	0.011	0.019	0.016	0.008	0.004	0.003	94.0	94.2	93.7	94.2	94.4	94.8	0.101	0.109	0.107	0.101	0.101	0.100
0.4	1.5	50	0.4	3	0.035	0.089	0.049	-0.066	-0.030	-0.112	95.0	96.6	95.4	96.6	95.2	94.5	0.170	0.232	0.190	0.096	0.121	0.127
0.4	1.5	50	0.4	6	0.012	0.037	0.020	-0.029	-0.015	-0.047	94.5	95.9	94.6	95.8	94.7	93.5	0.116	0.135	0.123	0.091	0.102	0.109
0.4	1.5	50	0.4	12	0.014	0.025	0.018	-0.001	0.001	-0.011	95.5	95.7	96.0	96.3	96.4	95.2	0.099	0.106	0.103	0.092	0.094	0.096
0.4	1.5	100	0	3	0.002	0.016	0.004	-0.040	-0.023	-0.052	94.8	95.7	95.2	95.8	95.5	93.3	0.054	0.059	0.055	0.044	0.049	0.053
0.4	1.5	100	0	6	0.007	0.014	0.009	-0.006	-0.002	-0.005	95.2	94.6	94.6	95.5	95.0	95.6	0.045	0.048	0.046	0.042	0.044	0.046
0.4	1.5	100	0	12	-0.000	0.004	0.001	-0.005	-0.004	-0.005	94.4	94.9	94.4	94.8	94.9	94.8	0.049	0.051	0.050	0.048	0.049	0.049
0.4	1.5	100	0.4	3	0.021	0.050	0.027	-0.061	-0.025	-0.110	93.2	94.2	92.8	95.1	93.0	92.5	0.080	0.096	0.084	0.054	0.065	0.072
0.4	1.5	100	0.4	6	0.015	0.028	0.019	-0.022	-0.005	-0.037	95.0	95.5	94.5	96.1	95.3	94.3	0.056	0.061	0.057	0.046	0.051	0.057
0.4	1.5	100	0.4	12	0.014	0.019	0.016	0.000	0.005	-0.003	94.7	94.6	94.5	95.2	94.8	94.2	0.050	0.052	0.051	0.047	0.048	0.051
0.4	1.5	200	0	3	0.010	0.019	0.011	-0.030	-0.011	-0.036	93.6	94.0	93.8	94.8	93.9	92.4	0.026	0.028	0.027	0.022	0.024	0.027
0.4	1.5	200	0	6	0.007	0.010	0.007	-0.007	-0.000	-0.006	94.6	95.0	94.6	95.3	94.8	95.2	0.025	0.026	0.025	0.023	0.024	0.024
0.4	1.5	200	0	12	0.004	0.006	0.003	-0.002	0.001	0.002	94.9	94.8	94.9	95.1	95.0	95.1	0.022	0.023	0.023	0.022	0.022	0.023
0.4	1.5	200	0.4	3	0.006	0.021	0.008	-0.069	-0.036	-0.126	95.6	96.6	95.5	95.3	94.8	89.9	0.031	0.034	0.031	0.025	0.027	0.040
0.4	1.5	200	0.4	6	0.007	0.013	0.008	-0.027	-0.010	-0.044	94.8	95.2	94.8	95.1	94.8	92.3	0.027	0.028	0.027	0.024	0.025	0.030
0.4	1.5	200	0.4	12	0.000	0.003	0.001	-0.013	-0.007	-0.013	94.8	94.9	94.9	94.9	94.5	94.4	0.024	0.025	0.025	0.023	0.024	0.025
0.4	1.5	400	0	3	-0.003	0.001	-0.003	-0.041	-0.022	-0.051	94.7	94.2	94.6	94.8	94.9	92.4	0.012	0.013	0.013	0.012	0.012	0.014
0.4	1.5	400	0	6	0.005	0.007	0.004	-0.009	-0.001	-0.006	94.0	94.2	94.1	94.1	93.9	93.8	0.012	0.012	0.012	0.011	0.012	0.012
0.4	1.5	400	0	12	0.002	0.003	0.000	-0.005	-0.000	-0.001	93.5	93.5	93.5	94.2	93.5	93.8	0.012	0.012	0.012	0.012	0.012	0.012
0.4	1.5	400	0.4	3	0.012	0.021	0.013	-0.061	-0.026	-0.116	94.8	95.0	94.8	93.5	93.8	82.8	0.017	0.019	0.017	0.015	0.015	0.028
0.4	1.5	400	0.4	6	0.001	0.005	0.002	-0.032	-0.014	-0.048	95.5	95.4	95.7	94.8	94.7	91.0	0.012	0.013	0.013	0.012	0.012	0.016
0.4	1.5	400	0.4	12	-0.000	0.000	-0.002	-0.015	-0.007	-0.013	94.1	94.1	94.1	94.9	94.2	93.6	0.012	0.012	0.012	0.012	0.012	0.013
0.4	1.5	800	0	3	0.002	0.004	0.002	-0.036	-0.017	-0.045	95.5	95.4	95.6	94.3	94.4	89.7	0.006	0.006	0.006	0.006	0.006	0.008
0.4	1.5	800	0	6	-0.001	0.000	-0.001	-0.014	-0.006	-0.012	95.0	95.1	95.0	94.8	95.2	95.0	0.006	0.006	0.006	0.006	0.006	0.006
0.4	1.5	800	0	12	0.005	0.005	0.003	-0.002	0.003	0.002	95.2	94.6	94.5	94.9	94.8	95.0	0.006	0.006	0.006	0.006	0.006	0.006
0.4	1.5	800	0.4	3	0.002	0.006	0.002	-0.069	-0.033	-0.121	94.7	94.3	94.8	87.6	92.4	64.5	0.008	0.009	0.008	0.011	0.008	0.025
0.4	1.5	800	0.4	6	0.005	0.007	0.005	-0.028	-0.010	-0.045	93.0	92.7	93.0	91.8	93.1	83.9	0.007	0.007	0.007	0.007	0.007	0.011
0.4	1.5	800	0.4	12	-0.000	0.000	-0.001	-0.014	-0.005	-0.013	94.9	94.9	94.9	94.3	94.4	93.4	0.006	0.006	0.006	0.006	0.006	0.006

					Mean Bias						Coverage (%)						MSE					
β	γ	n	P	v	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox
0.8	0.6	50	0	3	0.024	0.052	0.034	-0.037	-0.048	-0.061	94.9	95.8	94.6	95.2	95.0	94.4	0.128	0.146	0.134	0.106	0.110	0.115
0.8	0.6	50	0	6	0.032	0.045	0.041	0.027	-0.003	-0.007	94.4	94.5	94.1	94.1	94.5	94.4	0.115	0.123	0.119	0.114	0.105	0.108
0.8	0.6	50	0	12	0.023	0.032	0.034	0.040	0.002	0.003	94.4	94.8	94.6	94.0	95.3	95.3	0.116	0.121	0.121	0.122	0.111	0.111
0.8	0.6	50	0.4	3	0.057	0.103	0.071	-0.100	-0.064	-0.183	96.2	97.7	96.1	95.3	94.3	91.7	0.156	0.194	0.167	0.103	0.118	0.145
0.8	0.6	50	0.4	6	0.035	0.056	0.043	-0.031	-0.033	-0.068	95.1	96.8	95.6	95.5	95.3	94.1	0.126	0.138	0.127	0.101	0.107	0.113
0.8	0.6	50	0.4	12	0.049	0.061	0.055	0.038	0.014	0.003	95.1	95.8	94.8	95.4	95.8	95.8	0.120	0.127	0.122	0.114	0.110	0.111
0.8	0.6	100	0	3	0.027	0.041	0.032	-0.024	-0.034	-0.045	94.2	94.5	94.1	94.4	94.4	93.7	0.063	0.069	0.064	0.055	0.057	0.059
0.8	0.6	100	0	6	0.004	0.011	0.008	-0.005	-0.021	-0.024	94.6	95.0	94.7	94.5	94.7	95.0	0.055	0.057	0.057	0.055	0.054	0.055
0.8	0.6	100	0	12	0.016	0.023	0.025	0.028	0.006	0.005	94.0	94.0	94.0	93.4	94.5	94.4	0.055	0.058	0.058	0.059	0.055	0.055
0.8	0.6	100	0.4	3	0.034	0.063	0.043	-0.113	-0.077	-0.206	94.1	96.0	93.8	93.2	92.9	84.3	0.082	0.094	0.086	0.065	0.068	0.103
0.8	0.6	100	0.4	6	0.012	0.024	0.017	-0.045	-0.038	-0.076	94.8	95.4	94.2	94.6	93.9	92.6	0.061	0.064	0.063	0.055	0.056	0.062
0.8	0.6	100	0.4	12	0.003	0.008	0.008	-0.007	-0.019	-0.027	94.3	94.5	94.3	94.0	94.3	93.2	0.056	0.058	0.057	0.055	0.054	0.056
0.8	0.6	200	0	3	0.018	0.026	0.022	-0.027	-0.037	-0.049	94.7	94.9	94.5	94.0	94.4	93.5	0.029	0.031	0.029	0.027	0.027	0.028
0.8	0.6	200	0	6	0.012	0.016	0.015	0.001	-0.009	-0.012	95.1	94.8	94.3	94.5	94.9	94.3	0.028	0.028	0.028	0.028	0.027	0.027
0.8	0.6	200	0	12	0.013	0.017	0.019	0.020	0.005	0.005	93.8	93.8	93.5	93.1	94.2	94.3	0.028	0.030	0.029	0.030	0.028	0.028
0.8	0.6	200	0.4	3	0.014	0.027	0.019	-0.117	-0.079	-0.199	94.3	95.0	94.2	88.4	88.9	73.0	0.036	0.039	0.037	0.043	0.042	0.081
0.8	0.6	200	0.4	6	0.012	0.018	0.014	-0.042	-0.030	-0.061	94.9	94.9	94.5	94.4	94.3	93.2	0.029	0.030	0.030	0.027	0.028	0.030
0.8	0.6	200	0.4	12	0.014	0.018	0.019	0.004	-0.003	-0.012	95.4	95.6	95.1	94.5	94.7	92.9	0.027	0.027	0.027	0.028	0.028	0.030
0.8	0.6	400	0	3	0.008	0.010	0.011	-0.036	-0.043	-0.055	94.8	94.4	94.9	92.8	92.1	89.8	0.014	0.015	0.014	0.014	0.015	0.016
0.8	0.6	400	0	6	-0.004	-0.002	-0.001	-0.014	-0.024	-0.027	93.7	93.8	93.7	93.2	93.2	92.9	0.014	0.014	0.014	0.014	0.014	0.014
0.8	0.6	400	0	12	-0.001	0.007	0.005	0.005	-0.008	-0.008	94.1	93.2	94.0	93.5	93.8	93.7	0.013	0.017	0.014	0.014	0.013	0.013
0.8	0.6	400	0.4	3	0.010	0.016	0.013	-0.119	-0.080	-0.201	95.0	95.1	95.0	83.7	87.6	59.1	0.016	0.018	0.017	0.027	0.022	0.056
0.8	0.6	400	0.4	6	0.001	0.005	0.003	-0.053	-0.040	-0.074	96.1	96.2	95.9	94.2	94.6	90.8	0.013	0.013	0.013	0.014	0.013	0.016
0.8	0.6	400	0.4	12	0.001	0.003	0.005	-0.010	-0.015	-0.024	95.5	95.9	95.4	95.1	95.2	94.4	0.012	0.012	0.012	0.012	0.012	0.013
0.8	0.6	800	0	3	-0.000	-0.001	0.002	-0.040	-0.047	-0.059	96.1	95.5	96.0	92.7	92.3	89.2	0.006	0.007	0.007	0.008	0.008	0.010
0.8	0.6	800	0	6	-0.001	0.001	0.002	-0.010	-0.018	-0.021	95.3	95.4	95.0	94.6	94.9	94.7	0.006	0.006	0.006	0.006	0.006	0.006
0.8	0.6	800	0	12	0.003	0.014	0.008	0.007	-0.003	-0.003	95.2	93.8	95.4	94.8	95.3	95.0	0.006	0.010	0.006	0.006	0.006	0.006
0.8	0.6	800	0.4	3	0.003	0.004	0.005	-0.124	-0.085	-0.206	93.5	93.6	93.6	69.5	81.3	31.9	0.009	0.010	0.009	0.022	0.015	0.050
0.8	0.6	800	0.4	6	0.002	0.004	0.003	-0.051	-0.036	-0.069	95.6	95.8	95.7	89.7	90.0	82.3	0.007	0.007	0.007	0.010	0.009	0.014
0.8	0.6	800	0.4	12	-0.001	0.003	0.003	-0.013	-0.015	-0.025	95.1	95.3	95.1	94.4	94.5	93.8	0.006	0.007	0.006	0.006	0.006	0.007
0.8	1.0	50	0	3	0.027	0.063	0.038	-0.031	-0.023	-0.045	94.0	95.4	94.4	95.0	94.3	94.7	0.125	0.147	0.131	0.101	0.108	0.112
0.8	1.0	50	0	6	0.041	0.060	0.051	0.030	0.019	0.017	95.3	95.6	95.2	95.6	96.1	96.4	0.111	0.122	0.116	0.106	0.106	0.106
0.8	1.0	50	0	12	0.031	0.044	0.041	0.036	0.018	0.019	95.0	94.5	94.5	94.7	95.0	95.2	0.110	0.118	0.117	0.113	0.108	0.108
0.8	1.0	50	0.4	3	0.079	0.141	0.098	-0.063	-0.021	-0.155	94.7	97.6	94.9	95.8	94.6	91.6	0.179	0.232	0.202	0.116	0.134	0.168
0.8	1.0	50	0.4	6	0.040	0.071	0.055	-0.011	-0.001	-0.045	94.2	95.3	94.1	94.4	93.9	92.4	0.131	0.156	0.144	0.111	0.120	0.126
0.8	1.0	50	0.4	12	0.031	0.044	0.038	0.017	0.009	-0.001	95.4	96.0	95.5	95.9	96.3	96.3	0.108	0.114	0.110	0.099	0.100	0.101

					Mean Bias						Coverage (%)						MSE					
β	γ	n	P	ν	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox
0.8	1.0	100	0	3	0.018	0.039	0.023	-0.033	-0.018	-0.045	93.4	94.1	93.5	94.2	93.7	92.7	0.061	0.069	0.064	0.054	0.056	0.059
0.8	1.0	100	0	6	0.020	0.028	0.022	0.006	0.004	-0.000	94.8	95.1	94.6	94.5	95.0	94.8	0.055	0.057	0.056	0.053	0.053	0.053
0.8	1.0	100	0	12	0.012	0.017	0.016	0.011	0.003	0.003	93.4	93.4	92.8	93.5	93.8	93.6	0.057	0.060	0.059	0.058	0.057	0.057
0.8	1.0	100	0.4	3	0.018	0.053	0.027	-0.100	-0.057	-0.193	95.3	96.5	95.0	94.6	95.1	88.2	0.070	0.082	0.073	0.057	0.058	0.089
0.8	1.0	100	0.4	6	0.039	0.054	0.044	-0.011	0.004	-0.038	94.2	94.8	94.1	95.6	95.1	93.5	0.064	0.069	0.065	0.053	0.057	0.060
0.8	1.0	100	0.4	12	0.014	0.023	0.020	0.002	0.000	-0.007	95.5	95.9	95.6	95.7	95.6	95.7	0.054	0.056	0.055	0.052	0.052	0.054
0.8	1.0	200	0	3	0.014	0.025	0.017	-0.035	-0.016	-0.044	93.7	93.8	93.9	94.5	93.6	93.1	0.029	0.031	0.029	0.026	0.027	0.030
0.8	1.0	200	0	6	0.009	0.014	0.010	-0.004	-0.001	-0.007	95.4	95.9	95.5	95.3	95.6	95.7	0.025	0.025	0.025	0.024	0.024	0.024
0.8	1.0	200	0	12	0.011	0.015	0.014	0.010	0.006	0.006	94.6	94.9	94.3	94.3	94.8	95.3	0.026	0.027	0.026	0.026	0.026	0.026
0.8	1.0	200	0.4	3	0.018	0.039	0.023	-0.094	-0.048	-0.190	94.9	95.6	94.2	92.4	93.2	79.9	0.035	0.040	0.037	0.034	0.032	0.064
0.8	1.0	200	0.4	6	0.016	0.025	0.019	-0.029	-0.011	-0.052	94.4	95.3	94.9	94.8	94.8	93.2	0.029	0.031	0.030	0.026	0.027	0.030
0.8	1.0	200	0.4	12	0.002	0.005	0.003	-0.013	-0.009	-0.018	95.4	95.3	95.5	95.5	95.6	94.5	0.026	0.026	0.026	0.025	0.025	0.026
0.8	1.0	400	0	3	-0.004	0.001	-0.002	-0.051	-0.033	-0.062	95.6	95.1	95.7	94.2	95.2	91.6	0.013	0.014	0.013	0.014	0.013	0.016
0.8	1.0	400	0	6	0.006	0.009	0.007	-0.007	-0.003	-0.008	93.1	94.0	93.7	93.9	94.0	94.0	0.014	0.014	0.014	0.013	0.013	0.014
0.8	1.0	400	0	12	-0.002	0.000	-0.000	-0.004	-0.005	-0.005	94.5	94.0	94.5	94.3	94.4	94.1	0.013	0.013	0.013	0.013	0.013	0.013
0.8	1.0	400	0.4	3	0.010	0.020	0.013	-0.096	-0.050	-0.188	94.6	95.0	94.9	88.5	92.6	62.7	0.018	0.019	0.018	0.023	0.019	0.053
0.8	1.0	400	0.4	6	0.008	0.013	0.009	-0.036	-0.016	-0.058	95.0	95.1	95.2	94.1	93.8	91.2	0.014	0.015	0.014	0.014	0.014	0.018
0.8	1.0	400	0.4	12	0.001	0.003	0.002	-0.013	-0.008	-0.016	95.7	95.7	95.5	95.6	95.6	95.3	0.013	0.013	0.013	0.013	0.013	0.014
0.8	1.0	800	0	3	0.005	0.005	0.006	-0.041	-0.022	-0.052	94.6	94.3	94.6	93.1	94.0	90.1	0.007	0.007	0.007	0.008	0.007	0.009
0.8	1.0	800	0	6	0.003	0.005	0.003	-0.009	-0.004	-0.010	96.2	96.1	96.1	96.2	96.0	95.7	0.006	0.006	0.006	0.006	0.006	0.006
0.8	1.0	800	0	12	0.003	0.004	0.004	0.001	0.001	0.001	95.6	95.7	95.6	95.9	95.7	95.9	0.006	0.006	0.006	0.006	0.006	0.006
0.8	1.0	800	0.4	3	0.008	0.011	0.009	-0.098	-0.051	-0.191	94.8	94.5	94.9	79.4	90.4	39.5	0.009	0.010	0.009	0.016	0.010	0.044
0.8	1.0	800	0.4	6	0.003	0.006	0.004	-0.041	-0.020	-0.062	94.6	94.4	94.7	90.1	92.6	81.0	0.007	0.007	0.007	0.009	0.008	0.013
0.8	1.0	800	0.4	12	0.005	0.007	0.006	-0.009	-0.002	-0.011	95.1	94.7	94.6	94.5	94.8	94.1	0.006	0.007	0.006	0.006	0.006	0.007
0.8	1.5	50	0	3	0.023	0.074	0.038	-0.050	-0.027	-0.072	96.1	97.0	95.7	95.3	95.6	94.5	0.119	0.152	0.132	0.096	0.105	0.111
0.8	1.5	50	0	6	0.036	0.058	0.047	0.012	0.013	0.007	93.8	93.0	93.3	94.0	93.6	94.5	0.124	0.142	0.135	0.117	0.122	0.120
0.8	1.5	50	0	12	0.032	0.048	0.044	0.028	0.020	0.019	94.4	95.2	94.7	95.0	95.5	95.5	0.109	0.118	0.116	0.109	0.108	0.109
0.8	1.5	50	0.4	3	0.056	0.136	0.082	-0.090	-0.042	-0.197	95.0	96.9	94.4	95.9	95.2	90.9	0.178	0.257	0.206	0.121	0.137	0.180
0.8	1.5	50	0.4	6	0.019	0.052	0.032	-0.043	-0.024	-0.078	94.5	95.5	94.4	95.8	95.2	94.8	0.133	0.159	0.143	0.109	0.120	0.122
0.8	1.5	50	0.4	12	0.033	0.052	0.043	0.012	0.012	0.004	94.2	94.6	93.9	95.0	94.7	95.1	0.123	0.138	0.133	0.116	0.119	0.117
0.8	1.5	100	0	3	0.033	0.060	0.038	-0.035	-0.009	-0.059	94.9	94.9	94.7	94.7	94.3	92.7	0.059	0.068	0.061	0.048	0.052	0.058
0.8	1.5	100	0	6	0.015	0.029	0.019	-0.006	0.001	-0.008	95.2	95.4	94.9	95.3	94.9	94.6	0.052	0.056	0.054	0.049	0.051	0.051
0.8	1.5	100	0	12	0.002	0.010	0.006	-0.006	-0.004	-0.005	94.7	94.2	94.2	94.7	94.5	94.6	0.054	0.057	0.056	0.054	0.055	0.055
0.8	1.5	100	0.4	3	0.032	0.078	0.045	-0.090	-0.041	-0.190	94.6	95.6	94.5	94.3	94.2	88.5	0.081	0.099	0.086	0.061	0.065	0.099
0.8	1.5	100	0.4	6	0.021	0.042	0.026	-0.032	-0.009	-0.066	95.2	95.6	95.4	95.5	95.4	95.6	0.062	0.070	0.065	0.054	0.057	0.066
0.8	1.5	100	0.4	12	0.011	0.021	0.016	-0.010	-0.002	-0.012	94.4	93.8	93.8	95.2	94.5	94.9	0.055	0.058	0.056	0.052	0.054	0.055

					Mean Bias						Coverage (%)						MSE					
β	γ	n	P	v	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox	W	F	PE	APE	MCox	RCox
0.8	1.5	200	0	3	0.002	0.016	0.004	-0.058	-0.031	-0.082	95.6	96.2	95.7	95.0	95.2	93.0	0.027	0.029	0.027	0.025	0.025	0.031
0.8	1.5	200	0	6	-0.001	0.005	0.000	-0.022	-0.012	-0.020	94.0	94.5	94.3	93.7	94.4	93.7	0.027	0.028	0.027	0.026	0.026	0.027
0.8	1.5	200	0	12	0.000	0.003	-0.002	-0.011	-0.005	-0.007	94.8	94.5	94.6	94.8	94.6	94.1	0.026	0.026	0.026	0.025	0.026	0.026
0.8	1.5	200	0.4	3	0.002	0.027	0.006	-0.108	-0.059	-0.214	96.0	96.5	96.0	92.3	94.5	76.8	0.034	0.038	0.035	0.036	0.032	0.076
0.8	1.5	200	0.4	6	0.005	0.014	0.007	-0.047	-0.023	-0.079	95.7	95.9	95.7	95.9	96.1	93.1	0.027	0.028	0.027	0.025	0.026	0.032
0.8	1.5	200	0.4	12	0.006	0.011	0.006	-0.017	-0.005	-0.017	94.8	94.9	95.0	95.4	94.9	95.2	0.027	0.028	0.028	0.026	0.027	0.027
0.8	1.5	400	0	3	0.007	0.014	0.008	-0.052	-0.023	-0.077	95.3	95.1	95.1	93.8	94.6	89.2	0.013	0.014	0.014	0.014	0.013	0.019
0.8	1.5	400	0	6	0.003	0.006	0.002	-0.020	-0.007	-0.017	93.7	94.1	93.6	93.7	93.7	93.6	0.014	0.014	0.014	0.013	0.014	0.014
0.8	1.5	400	0	12	0.005	0.006	0.002	-0.008	0.001	-0.000	95.5	95.3	95.5	95.6	95.4	94.7	0.013	0.013	0.012	0.012	0.012	0.013
0.8	1.5	400	0.4	3	0.001	0.014	0.003	-0.104	-0.054	-0.207	95.7	95.5	95.7	88.3	93.0	61.7	0.017	0.019	0.017	0.023	0.017	0.056
0.8	1.5	400	0.4	6	-0.005	0.001	-0.004	-0.056	-0.029	-0.087	94.2	94.7	94.6	91.9	93.0	87.2	0.016	0.016	0.016	0.016	0.015	0.022
0.8	1.5	400	0.4	12	0.008	0.009	0.006	-0.016	-0.002	-0.014	95.2	95.6	95.6	96.1	95.7	94.7	0.013	0.013	0.013	0.012	0.013	0.014
0.8	1.5	800	0	3	0.000	0.003	0.001	-0.057	-0.027	-0.082	95.6	95.2	95.7	90.6	94.8	81.0	0.007	0.007	0.007	0.009	0.007	0.013
0.8	1.5	800	0	6	0.000	0.003	0.000	-0.021	-0.008	-0.017	94.6	94.5	94.2	94.3	94.3	94.2	0.006	0.007	0.007	0.007	0.007	0.007
0.8	1.5	800	0	12	0.000	0.001	-0.003	-0.012	-0.002	-0.003	94.7	94.6	94.6	94.7	94.4	94.4	0.006	0.006	0.006	0.006	0.006	0.006
0.8	1.5	800	0.4	3	0.005	0.011	0.006	-0.100	-0.049	-0.206	95.2	95.4	95.3	80.2	91.5	34.5	0.009	0.010	0.009	0.017	0.010	0.049
0.8	1.5	800	0.4	6	-0.003	0.001	-0.002	-0.053	-0.026	-0.081	94.9	94.1	94.1	91.2	93.2	83.3	0.007	0.007	0.007	0.009	0.007	0.013
0.8	1.5	800	0.4	12	0.005	0.006	0.003	-0.019	-0.004	-0.016	94.3	94.0	94.4	94.5	94.8	93.9	0.007	0.007	0.007	0.007	0.007	0.007

b) Comparison of methods

The mean bias, coverage and mean squared error (MSE) for the Weibull, Finkelstein, PE, APE, RCox and MCox methods are presented in Table 4.1.

i) Bias

For β=0 all the methods had mean bias less than -0.0005 (Table 4.1). This is as predicted since the two groups defined by the covariate are "exchangeable" in terms of the failure times. Note that inference under the MCox and RCox methods depends on the rank order of the imputed failure times. The imputation results in a re-ordering of the unobserved true failures times, but this process is independent of the covariate and will not therefore induce any bias.

When β=0.4 or 0.8, general trends of increased bias in the estimation of beta as v decreased and p increased were seen for all methods. Figure 4.4 shows the mean bias as a percentage of the true value of beta for each method by v, p and n when β=0.8. Results were similar for β=0.4 (Table 4.1). The shape of the underlying hazard had little effect on the relative performance of the different methods.

The sparseness of the data, determined by the number of scheduled visits v and the proportion of missing visits p, had the greatest impact upon the performance of the different methods (Figure 4.4). When the number of follow-up visits was relatively high (p=0 & v=12 or 6, p=0.4 & v=12), all the methods estimated β with negligible bias. With a smaller number of visits (p=0 & v=3, p=0.4 & v=6 or 3), the Weibull, Finkelstein and PE models estimated β with minimal bias for n$\geq$200, but with a positive bias, overestimating β, for smaller sample sizes. This illustrates that these methods are only *asymptotically* unbiased.

In contrast, the APE, MCox and RCox methods all estimated β with a negative bias, which was greatest for RCox, followed by APE. In each case the difference between the distributions of failure times for each value of x is diluted when the failure times are approximated by the method. The extent of the bias, and differences between the methods, were clearest when the data were most sparse (p=0.4 & v=3) and was similar for all sample sizes.

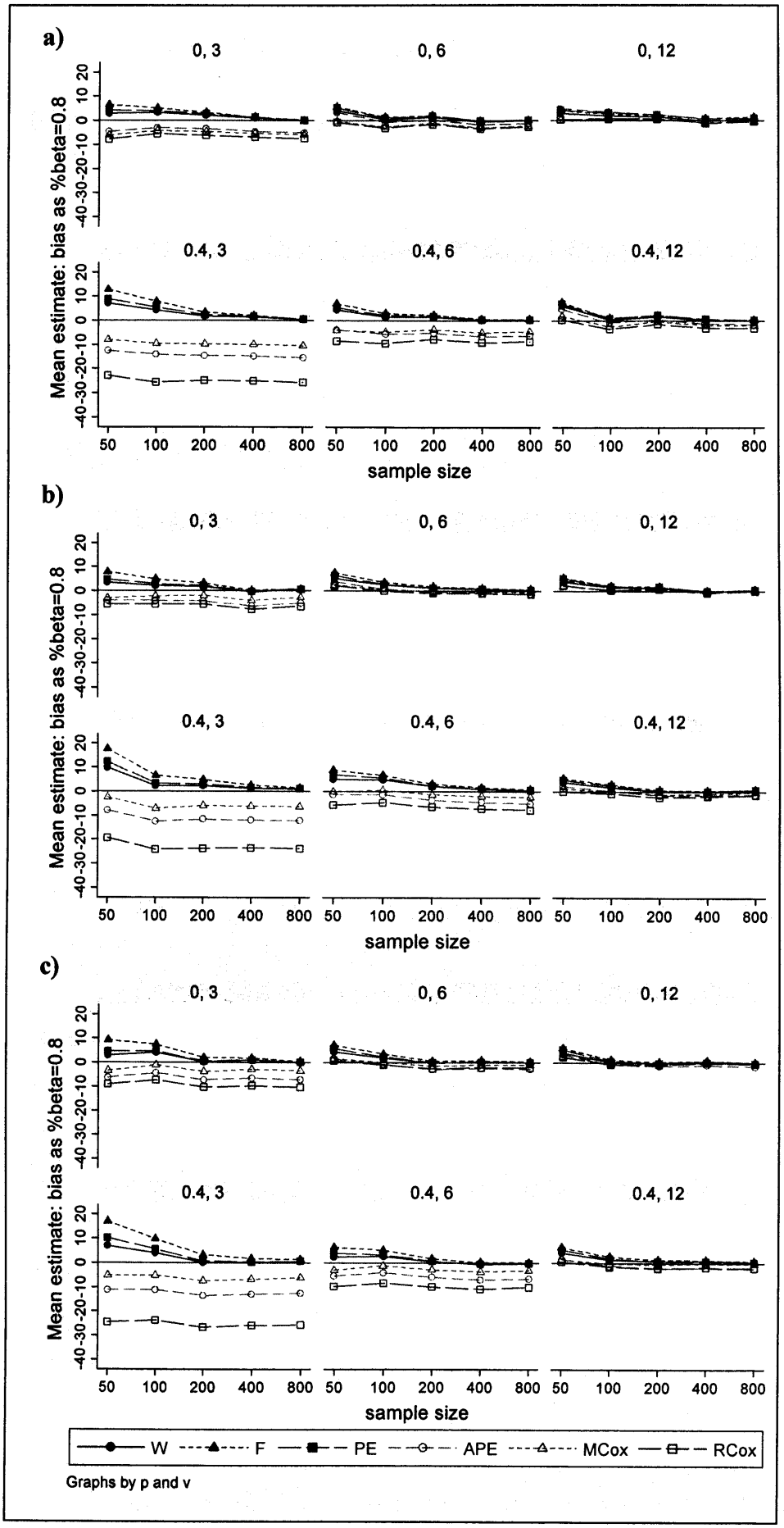

Figure 4.4: Comparing models: the effect of sample size and sparseness of data on bias when β=0.8 and a) γ=0.6, b) γ=1.0, c) γ=1.5.

A surprising finding was the large difference between the PE and APE methods when the data were sparse: APE underestimated β whereas PE was either unbiased or slightly overestimated β for small n. The difference between the estimates was greatest when n=50 or there were few follow-up visits, indicating the assumption of uniform event times is inadequate when the width of intervals is large and the amount of information available is small.

ii) MSE

Figure 4.5 shows the MSE for each method by v, p and n when β=0.8. Results were similar for β=0 & β=0.4 (Table 4.1).

As expected, the MSE decreased for all methods as sample size increased. As for the bias, clear differences between most of the methods were only apparent for infrequent visits and small sample sizes (Figure 4.5), namely:

- the Weibull, Finkelstein and PE models had greater MSE for n≤100 than the other methods;
- the Finkelstein model had the largest MSE followed by the PE model;
- the APE and MCox methods had the smallest MSEs;

The RCox method had larger MSE than all other methods at large n, reflecting the large bias with this method for sparse data.

iii) Coverage

Figure 4.6 shows the concordance between the estimated standard error of β and the empirical standard error of β for each method.

The mean estimated standard errors from all models were in close agreement with the empirical standard error, with only slight underestimation when sample size was small. This result was surprising for the RCox, MCox and APE approaches since they theoretically underestimate the variability in the data by assuming event times are known.

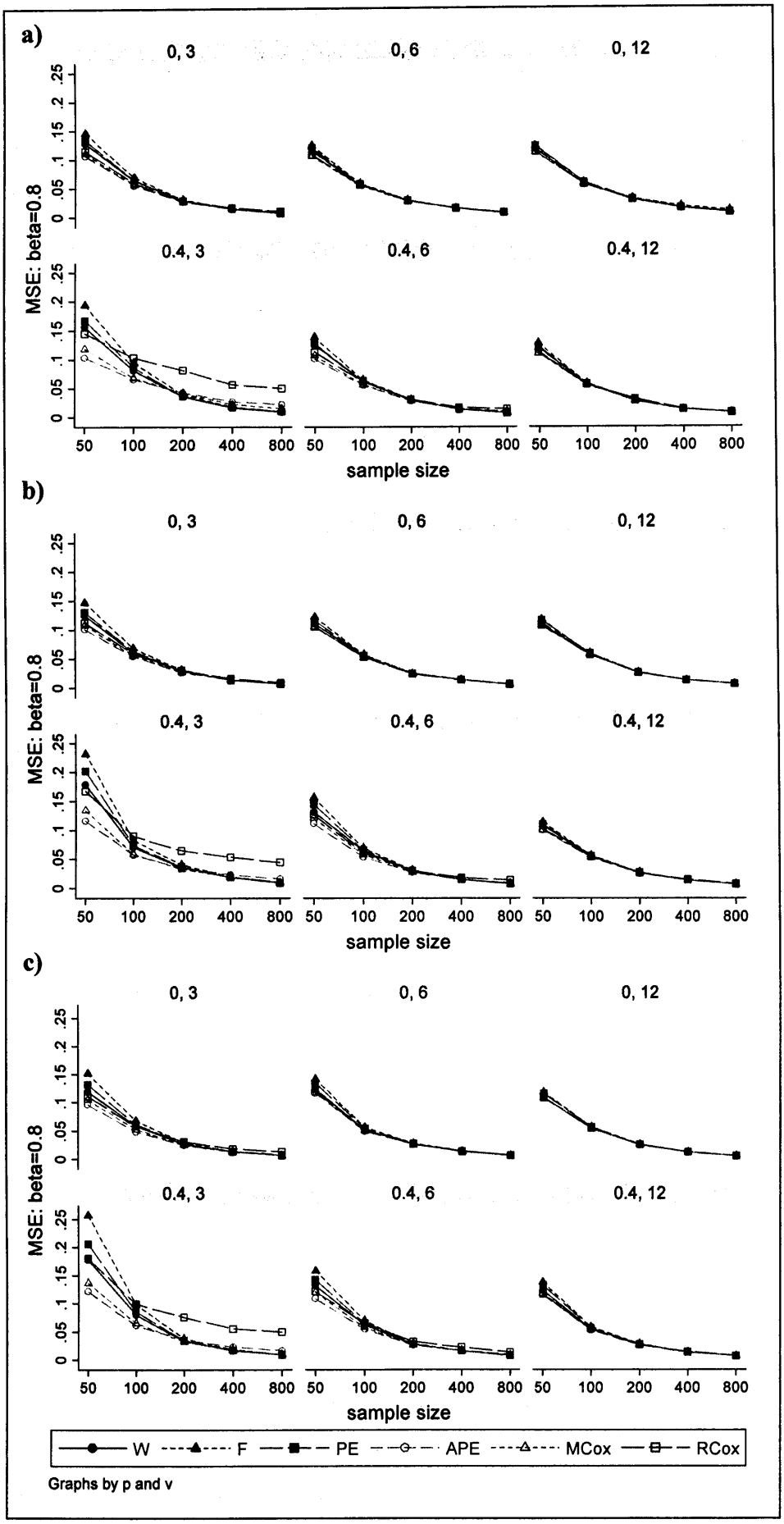

Figure 4.5: Comparing models: the effect of sample size and sparseness of data on MSE when β=0.8 and a) γ=0.6, b) γ=1.0, c) γ=1.5.

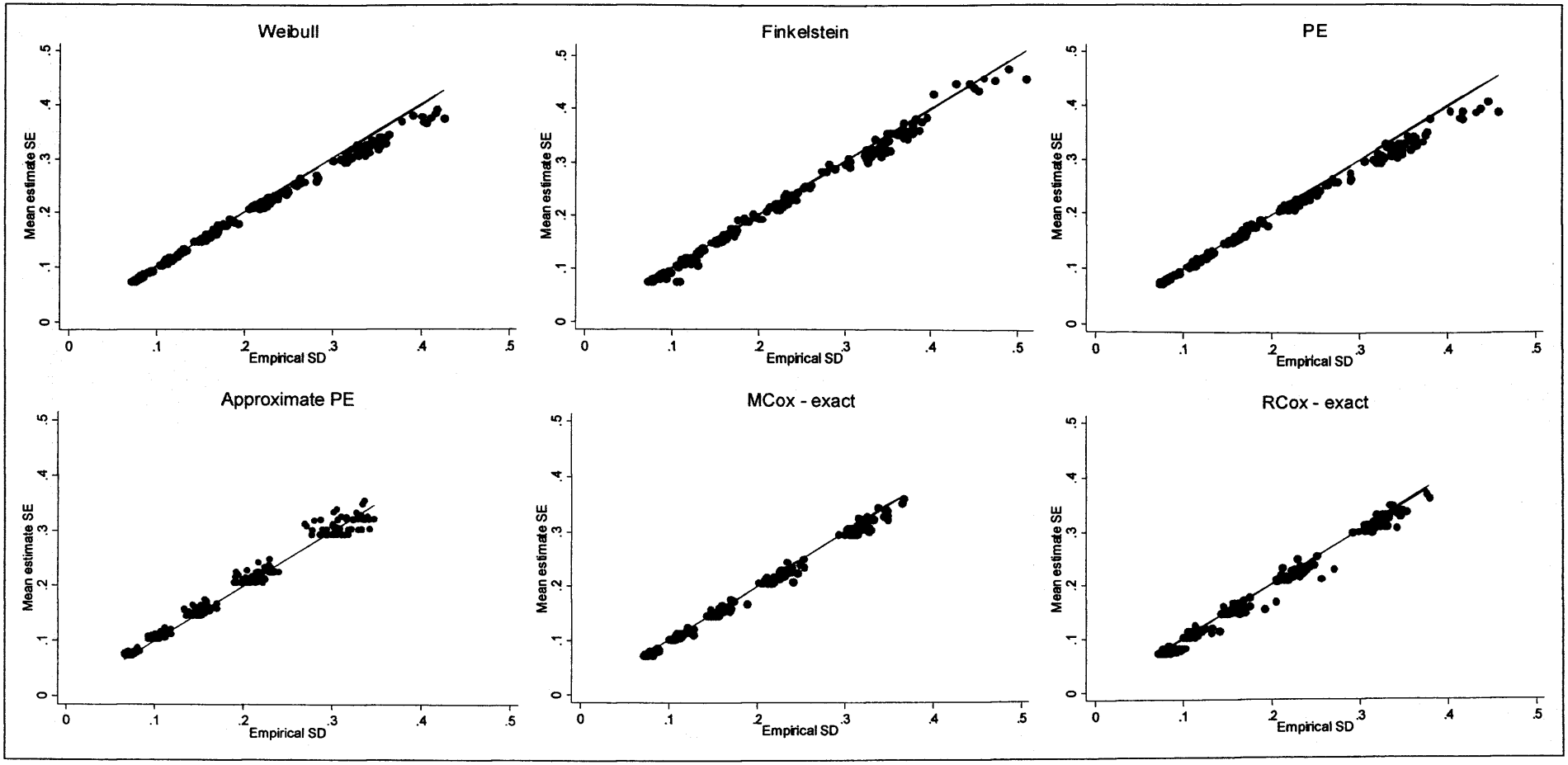

Figure 4.6: Comparison of mean estimated standard error and empirical standard error by method

Each dot represents one scenario.

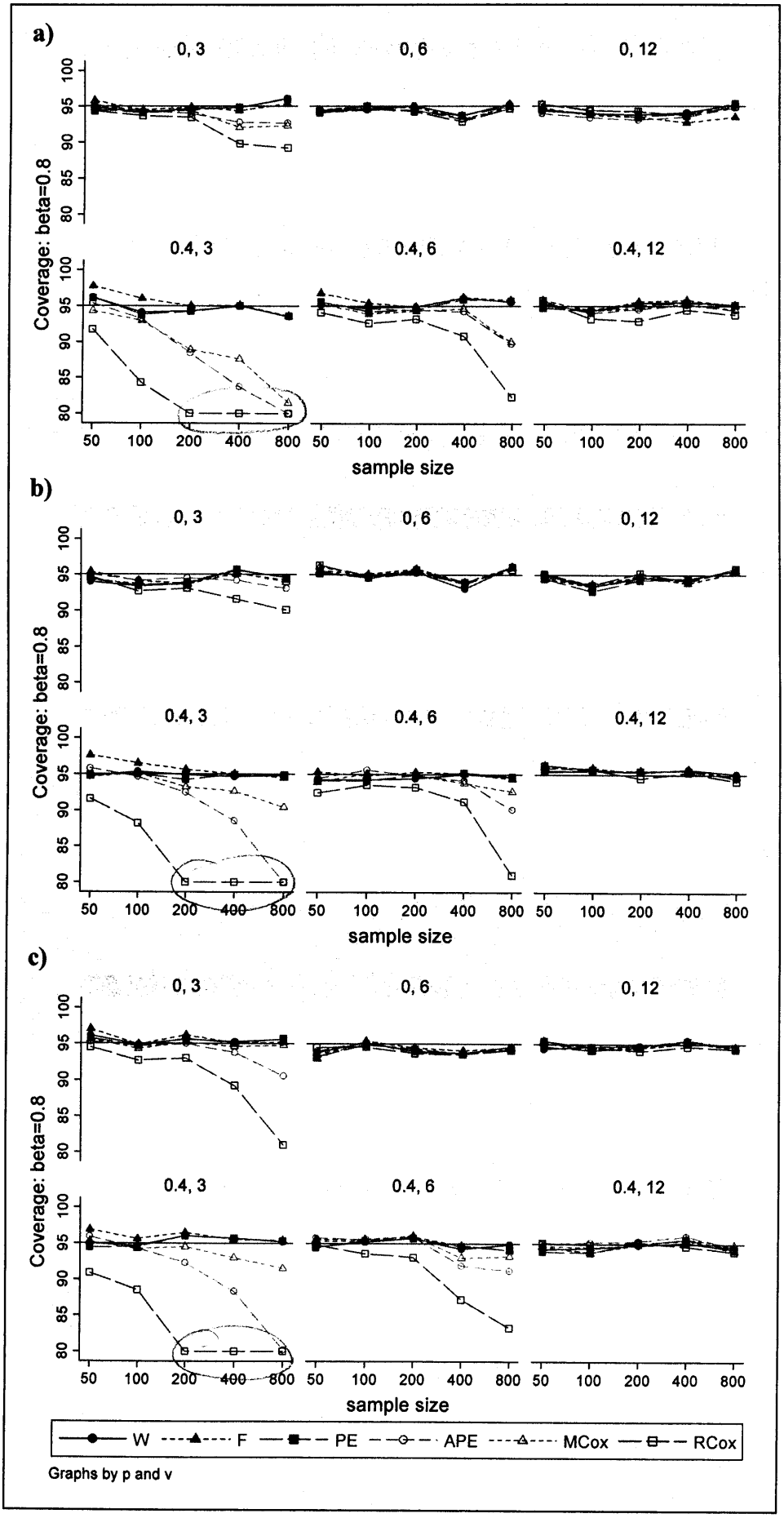

Figure 4.7: Comparing models: the effect of sample size and sparseness of data on coverage when β=0.8 and a) γ=0.6, b) γ=1.0, c) γ=1.5.

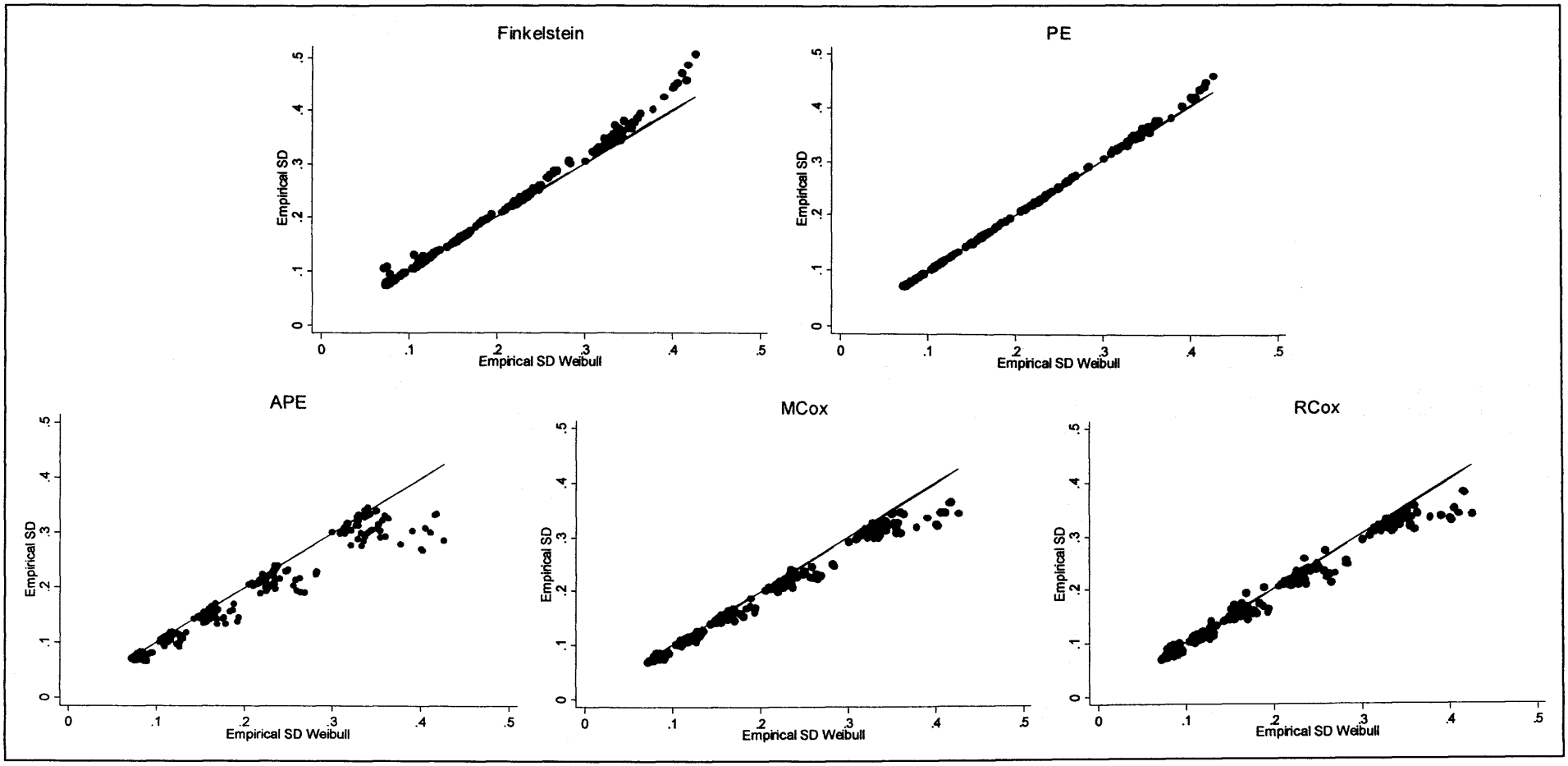

Figure 4.8: Comparison of empirical standard error for each method to the empirical standard error for the Weibull model.

Each dot represents one scenario.

When β=0, the coverage for all methods was close to the nominal value of 95% reflecting the lack of bias and correct estimation of the standard error of β for these scenarios (Table 4.1). Figure 4.7 shows the coverage, as a percentage, for each method by v, p and n when β=0.8. Any coverage less than 80% are plotted as 80% exactly. Results were similar for β=0.4 (Table 4.1). The shape of the underlying hazard had little effect on the coverage.

When β>0, coverage for the Weibull, Finkelstein and PE models were close to the nominal value of 95% for all scenarios (Figure 4.7). The APE, MCox and RCox methods had coverage close to the nominal value of 95% when there were frequent visits, but when the data were sparse the coverage became under-conservative. The RCox method had the lowest coverage due to having the greatest bias, followed by the APE approach. Coverage was particularly low for large n, when the estimated standard errors were small.

iv) Comparison of empirical standard errors

Figure 4.8 compares the empirical standard error for the Finkelstein, PE, APE, MCox and RCox methods to that for the Weibull model.

The empirical standard errors for the Finkelstein and PE models were in close agreement to those from the Weibull model, except for a few scenarios with the largest variability i.e. those with small sample size. The empirical standard errors were smaller for APE, MCox and RCox than the Weibull model as sample size decreased, most markedly when p=0.4 & v=3. The APE approach showed greater "underestimation" compared to the Weibull model than either Cox model.

4.4 EXAMPLES

The selected methods are illustrated using two datasets. The first is from the Delta trial [40] (section 2.3). This example focuses on the 1280 subjects who participated in the virology substudy [29]. One aim of the study was to determine how rapidly the minimum HIV RNA level (limit of detection 800 copies/ml) according to treatment regimen was achieved. This was to inform clinicians at what time a change of therapy should be considered if the initial response was unsatisfactory.

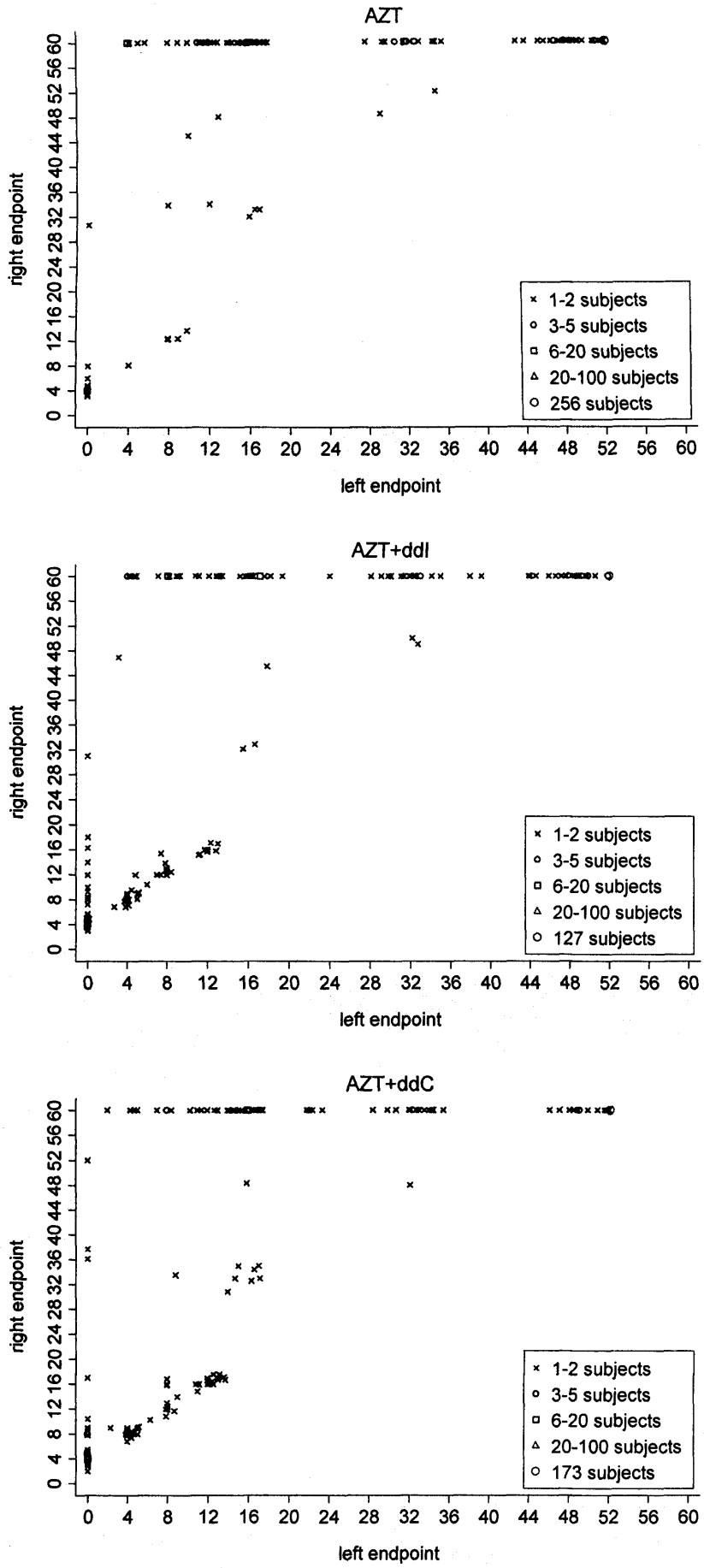

Figure 4.9: Individuals censoring intervals in the DELTA Virology Substudy by randomised treatment

The 1243 trial participants with a viral load greater than 800 copies/ml at randomisation and at least one subsequent result within 48 weeks are included in the analysis. Of these, 400, 423 and 420 individuals were randomised to receive AZT, AZT+ddI and AZT+ddC respectively. 369 (92%), 213 (50%) and 247 (59%) did not achieve undetectable HIV RNA within 48 weeks from randomisation. Nine follow-up visits were scheduled during the first year of therapy at weeks 4, 8, 12, 16, 20, 24, 32, 40 and 48. An HIV RNA measurement was unavailable at 39% of these visits. The number of plasma HIV RNA measurements over the year [median 6, IQR 4-7 measurements] and during the first 24 weeks [median 5, IQR 4-5 measurements] was similar in the three treatment groups. The individual censoring intervals are shown in Figure 4.9. For the purpose of presentation, R_i was set to 60 weeks for the 414 subjects who exhibited right censoring.

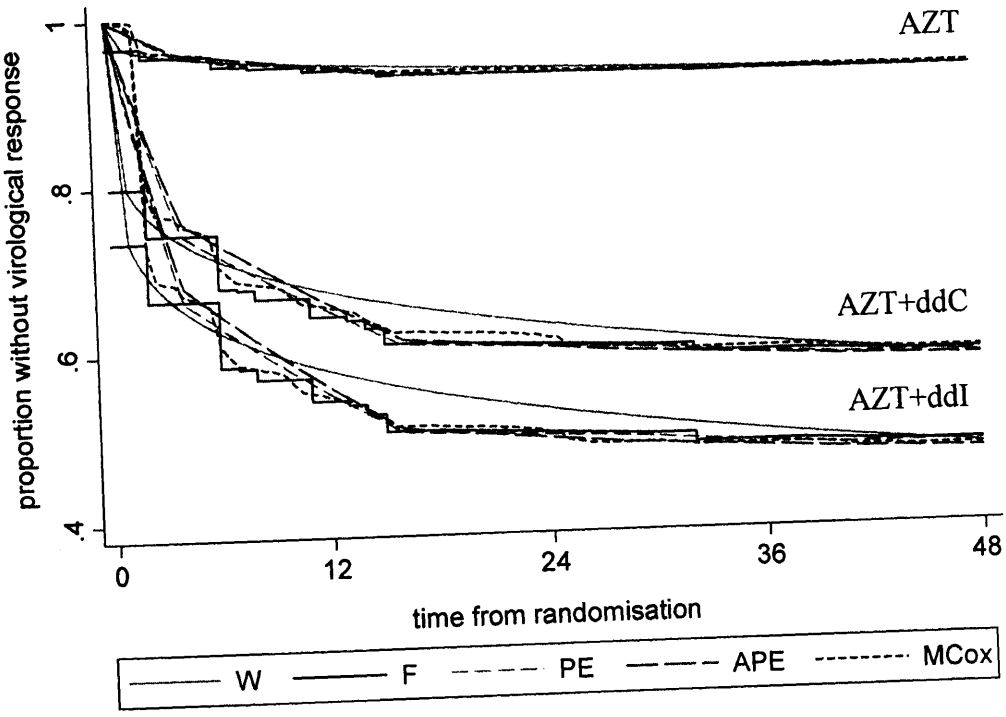

Figure 4.10: Time to achieving viral load ≤800 copies/ml by randomised treatment

Time to achieving a virological response was much faster for individuals in both combination therapy arms compared to those starting treatment with AZT monotherapy, and the majority of responses on combination therapy had occurred by 24 weeks. This implies that patients without a virological response by 24 weeks should be considered as virological failures and their therapy altered. The best response was seen in individuals

starting therapy with AZT+ddI (Figure 4.10). There was close agreement between the hazard ratios and confidence intervals estimated using all the different analytical approaches (Table 4.2).

Table 4.2: Parameter estimates and hazard ratios for the Delta example

DELTA			
ddI+AZT vs. AZT			
Method	**β**	**(SE β)**	**HR (95% CI)**
Weibull (W)†	2.198	(0.193)	9.00 (6.17, 13.14)
Finkelstein (F)	2.189	(0.193)	8.93 (6.11, 13.04)
Piecewise exponential:	2.194	(0.193)	8.97 (6.15, 13.09)
Approximate piecewise exponential:	2.195	(0.193)	8.98 (6.15, 13.11)
Cox model:			
Right endpoint (RCox): Breslow	2.113	(0.193)	8.28 (5.67, 12.07)
exact	2.191	(0.193)	8.95 (6.13, 13.05)
Mid point (MCox): Breslow	2.114	(0.193)	8.28 (5.67, 12.08)
Exact	2.189	(0.193)	8.92 (6.12, 13.02)
ddC+AZT vs. AZT			
Method	**β**	**(SE β)**	**HR (95% CI)**
Weibull (W)†	1.880	(0.195)	6.55 (4.47, 9.61)
Finkelstein (F)	1.869	(0.195)	6.48 (4.42, 9.51)
Piecewise exponential:	1.875	(0.195)	6.52 (4.45, 9.56)
Approximate piecewise exponential:	1.878	(0.195)	6.54 (4.46, 9.59)
Cox model:			
Right endpoint (RCox): Breslow	1.833	(0.195)	6.25 (4.26, 9.16)
exact	1.874	(0.195)	6.52 (4.45, 9.56)
Mid point (MCox): Breslow	1.835	(0.195)	6.27 (4.27, 9.19)
exact	1.873	(0.195)	6.51 (4.44, 9.54)

† Estimated Weibull Shape parameter = 0.227

The number of intervals in the piecewise exponential model was initially taken as 5, the breakpoints determined by the expected number of events estimated using a Kaplan-Meier survivor function where events are assumed to occur at the midpoint of the censoring interval. This model was then simplified in a stepwise fashion by joining adjacent intervals where a likelihood ratio test indicated that the rate could be assumed to be constant across the combined interval. This resulted in a final model with 3 time intervals: 0-4, 4-16 and 16-52 weeks. The effect upon the treatment covariates was minimal – results between the 5 and 3 interval models agreed to 3 d.p. The final model for the APE method, determined in a similar fashion, had 4 time intervals: 0-4, 4-16, 16-28 and 28-52 weeks. Changing the time intervals again had minimal effect upon the treatment covariates.

The second example is from the CHIPS multi-centre cohort study of HIV-1 infected children in the UK and Ireland [30] (section 3.4). Treatment guidelines recommend that initial antiretroviral therapy in children should be a 3-drug/2-class regimen of two nucleoside reverse transcriptase inhibitors (NRTI) plus either a non-nucleoside reverse transcriptase inhibitor (NNRTI) or a protease inhibitor (PI) [41]. In the absence of evidence from paediatric trials there is still wide variation in the use of these two classes of drugs. Data from the CHIPS study was used to examine the impact of starting highly active antiretroviral therapy (HAART) with either an NNRTI or PI based regimen on the time to achieving undetectable levels of plasma HIV RNA less than or equal to 500 copies/ml.

Of 599 children starting HAART with available HIV RNA measurements, 219 started a PI-based regimen and 308 an NNRTI-based regimen. The remaining 72 children started other HAART regimens and are excluded from this analysis. Seventy-nine (36%) and 54 (18%) failed to achieve undetectable HIV RNA within one year from starting therapy with a PI or NNRTI respectively. The number of plasma HIV RNA measurements over the year [median 4, IQR 3-5 measurements] and during the first 24

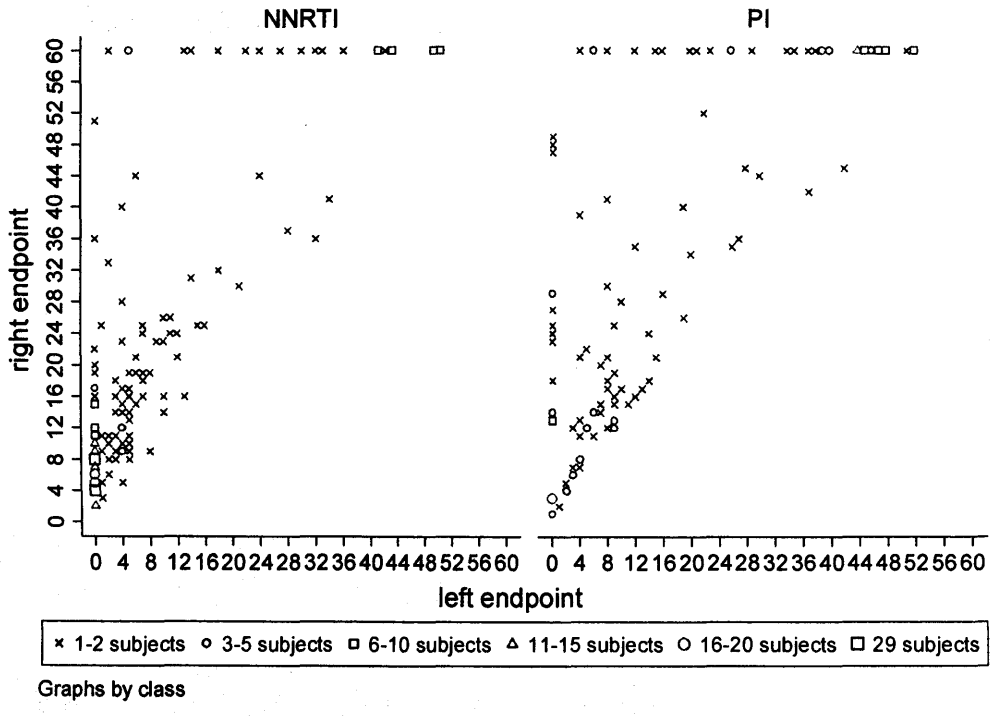

Figure 4.11: Individual censoring intervals in the CHIPS study by drug class of initial regimen

weeks [median 2, IQR 1-3 measurements] was similar in the two groups of children. The individual censoring intervals are shown in Figure 4.11. For the purpose of presentation, R_i was set to 60 weeks for the 133 subjects who exhibited right censoring.

Time to achieving a virological response was significantly shorter in those starting a NNRTI-based regimen for each methods of analysis (Figure 4.12), with close agreement between the estimated hazard ratios and confidence intervals (Table 4.3). The likelihood of achieving a virological response was reduced by 46-48% in children starting therapy with a PI compared to those starting an NNRTI containing regimen, depending upon the choice of analysis. This result is likely to be partly explained by the confounding effect of improved response in CHIPS children seen with calendar year [42] and the fact that NNRTI-based regimens have increased in popularity over time [30],

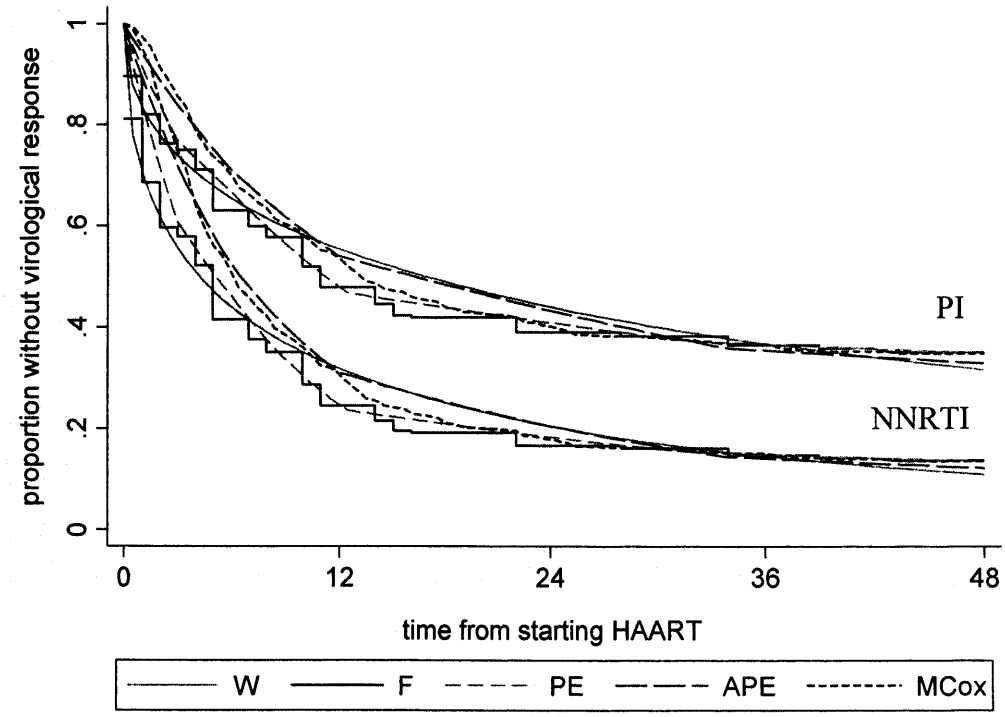

Figure 4.12: Time to achieving virological response according to drug class of initial regimen in the CHIPS study

The initial piecewise exponential model had 5 time intervals determined by the expected number of events as for the Delta example. This model was then simplified to a final model with 4 time intervals: 0-3, 3-13, 13-32 and 32-52 weeks. The final model for the APE method, determined in a similar fashion, had 3 time intervals: 0-13, 13-32 and 32-

52 weeks. Changing the time intervals again had minimal effect upon the treatment covariate for either method.

Comparing the survival curves to the non-parametric Finkelstein model suggests that the Weibull model does not fit the data as well as the PE model, and that the APE and MCox methods underestimate the event rate at the early time points. Nevertheless, the log hazard ratios for the different methods were all within 2% of the Finkelstein

Table 4.3: Parameter estimates and hazard ratios for the CHIPS example

CHIPS: PI vs NNRTI				
Method		**β**	**(SE β)**	**HR (95% CI)**
Weibull (W)*		-0.659	(0.108)	0.52 (0.42, 0.64)
Finkelstein (F)		-0.643	(0.111)	0.53 (0.42, 0.65)
Piecewise exponential:		-0.641	(0.108)	0.53 (0.43, 0.65)
Approximate piecewise exponential:		-0.651	(0.106)	0.53 (0.43, 0.65)
Cox model:				
Right endpoint (RCox):	Breslow	-0.617	(0.107)	0.54 (0.44, 0.67)
	exact	-0.632	(0.107)	0.53 (0.43, 0.66)
Mid point (MCox):	Breslow	-0.624	(0.106)	0.54 (0.44, 0.66)
	exact	-0.636	(0.106)	0.53 (0.42, 0.64)

*Estimated Weibull Shape parameter = 0.480

estimate. This is closer agreement between the methods than expected given the simulation results (although the results are the mean of 1000 simulations and there were examples of individual simulations where the variation in the estimates between the different methods was consistent with that in the CHIPS example). Examination of Figure 4.9 and Figure 4.11 shows a difference in the pattern of censoring for the CHIPS and Delta data. In Figure 4.9 most data points form a diagonal line but in Figure 4.11 there are many "off diagonal" points. This reflects the wider variation in interval endpoints and width for the CHIPS data (median width 9 weeks, IQR 5-15 weeks) compared to the Delta data (median width 8 weeks, IQR 5-8 weeks). This is because there is less regularity in the timing of clinic visits for children in the CHIPS study than there are for the scheduled follow-up visits for individuals in the Delta trial. The number of baseline hazard parameters for the Finkelstein model was higher for the CHIPS data (46 compared to 31 for Delta) reflecting this.

4.5 DISCUSSION

Six methods for estimating the effect of a covariate upon failure time have been described and assessed: a parametric Weibull model (W), the full likelihood proportional hazards model proposed by Finkelstein (F), a piecewise exponential model (PE), an approximation to the PE model (APE), a Cox proportional hazards model that assumes the event of interest occurs at the middle of the observed censoring interval (MCox), and finally a Cox proportional hazards model where the event is assumed to occur at the time it is observed (RCox).

The number of scheduled visits and the amount of missing visits had the greatest impact upon the performance of the different methods. There were no material differences between all six methods when follow-up visits were frequent in terms of bias, MSE and coverage. In addition, all methods estimated the standard error of the log hazard ratio accurately, including the APE, MCox and RCox methods which by imputing event times ignore the uncertainty in the observed data.

The picture was more complicated however when the data were more sparse. In this situation, the RCox model was the poorest of all the methods with substantial bias and low coverage at all sample sizes and is therefore discounted as an appropriate method of analysis. This concurs with the findings of Williams et al, who conducted a very limited simulation study [26].

As in previous chapters, the simulation study was based on data generated from a Weibull distribution, and even the correctly specified parametric model exhibited some positive bias at small sample sizes when the data were sparse, as did the F and PE methods. This reflects the fact that these methods are only asymptotically unbiased. In contrast the APE, and to a lesser extent the MCox method were consistently biased for *all* sample sizes. The ranking of failure times assuming events occurred at the midpoint of the censoring interval are plotted against the true rank of the known event time for one scenario from the simulation study in Figure 4.13. The effect of the imputation on the ranking of event times will be more extreme as the number of ties in the data increases, resulting in poor performance of MCox when the data was sparse. As for the estimation of the survivor curve (section 3.5), the distribution of imputed event times is determined by a function of the distribution of the visit times. This problem becomes more marked as the number of visits decreases since the number of ties in the data then

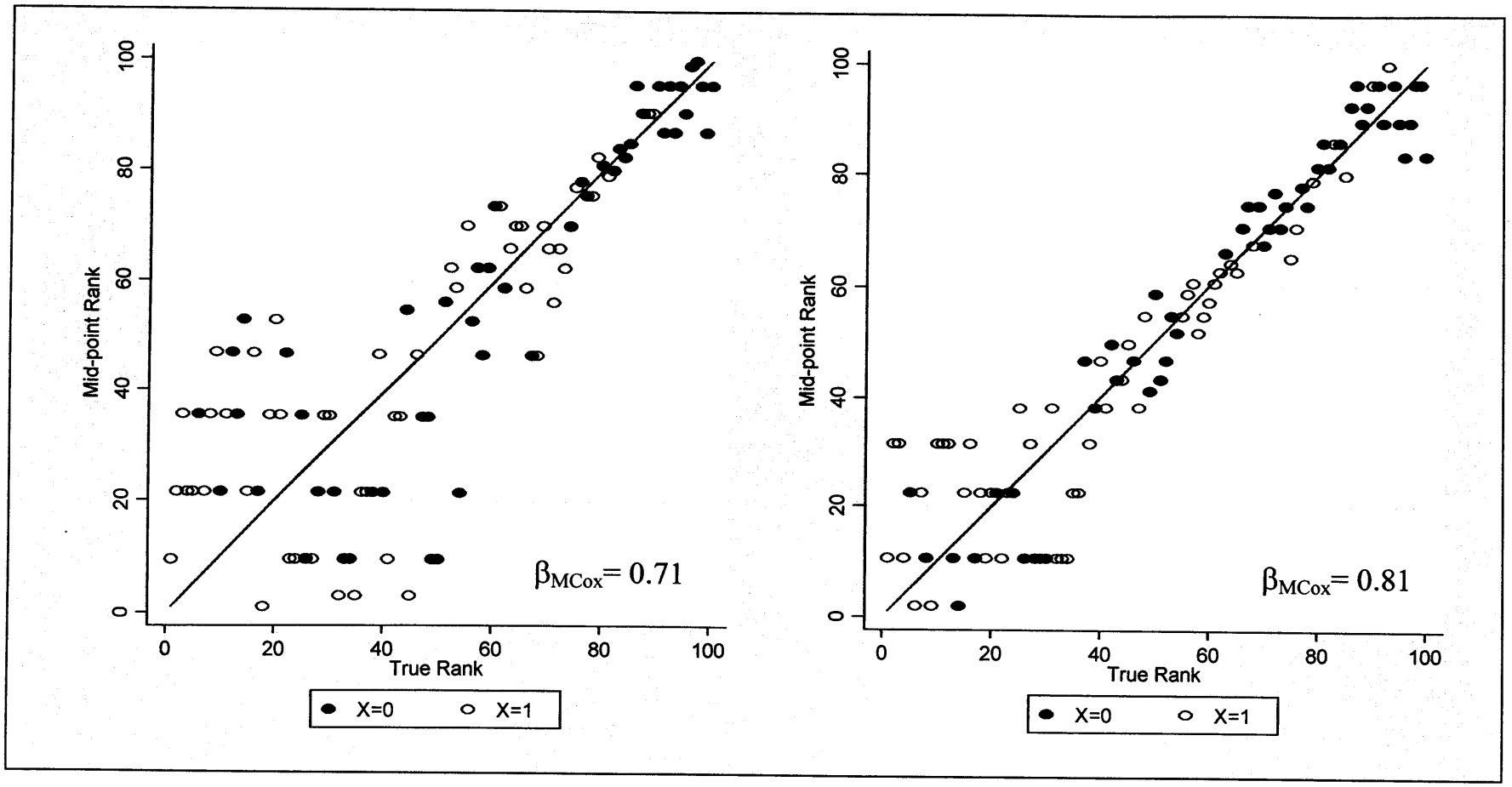

Figure 4.13: Comparing imputed and true ranks in the MCox model (β=0.8, γ=0.6, n=100, p=0, v=6)

increases. Estimation is unbiased when $\beta=0$ since the rank of the true failure times is a random process in this case. A similar argument can be used to explain the bias in the RCox approach.

Given this relationship between the ranking of the failure times and the performance of the method, the mechanism for the treatment of ties is important. Results from the simulation study indicate that the bias is increased if a Breslow approximation is used rather than the exact likelihood. This agrees with the work of So who performed a small simulation study on grouped data [25].

In contrast, a major finding was that the MSE for all methods were similar at large sample sizes and smaller for the APE and MCox methods when sample size was small, due to more precise estimation of beta rather than smaller bias. This is surprising given the theoretical prediction that a correctly specified parametric model should be the most efficient. In practice the family of distributions is not known, and although mis-specification of the baseline hazard may have little effect upon the estimation of the relative importance of covariates, it may lead to a reduction in the size or power of significance tests [43].

The coverage for the W, F and PE methods was good for all scenarios when the data were sparse, but the APE and MCox methods had lower than nominal coverage at large sample sizes.

The satisfactory estimation of the estimates and their standard errors for the MCox approach when the number of visits was large suggests the use of this method is reasonable unless the data are sparse, with fewer than 5-6 visits per individual (during the period where most events occur). When the data are sparse the sample size affects the choice of method. At small sample sizes, the low MSE and good coverage for the MCox method mean this is still the method of choice. For larger sample sizes the PE model or Finkelstein model should be used. The PE model has the advantage for allowing estimation of time-varying covariates.

The simulation study is limited by its restriction to data drawn only from a Weibull distribution, although constant, decreasing, and increasing hazards were considered. The ability of the piecewise exponential or Finkelstein models to model flexibly a wide

variety of hazard functions implies this is a robust approach when the underlying hazard may be non-monotone. The simulation study considered a single binary covariate. However, it seems likely that the findings should generalise to a continuous covariate and to multivariate analysis.

References

1. Peto R. Experimental survival curves for interval-censored data. *Applied Statistics* 1973; **22**:86-91.

2. Turnbull BW. The empirical distribution function with arbitrary grouped, censored and truncated data. *Journal of the Royal Statistical Society B* 1976; **38**:290-295.

3. Lesaffre E, Komarek A, Declerck D. An overview of methods for the analysis of interval censored data with an emphasis on applications in dentistry. *Statistical Methods in Medical Research* 2005; **14**:539-552.

4. Fang H-B, Sun J, Lee M-LT. Non-parametric survival comparisons for interval-censored continuous data. *Statistics Sinica* 2002; **12**:1073-1083.

5. Pepe MS, Fleming TR. Weighted Kaplan-Meier statistics: a class of distance tests for censored survival data. *Biometrics* 1989; **45**:497-507.

6. Zhao Q, Sun J. Generalized log-rank test for mixed interval-censored failure time data. *Stat.Med.* 2004; **23**:1621-1629.

7. Mantel N. Evaluation of survival data and two new rank order sstatistics arising in its consideration. *Cancer Chemotherapy Reports* 1966; **50**:163-170.

8. Rubin DB. The Bayes bootstrap. *Annals of Statistics* 1981; **9**:130-134.

9. Pan W. A two-sample test with interval-censored data via multiple imputation. *Stat.Med.* 2000; **19**:1-11.

10. S-PLUS 2000 Guide to Statistics, Volume 2. 1999; (2).

11. Cox DR. Regression models with lifetables (with discussion). *Journal of the Royal Statistical Society B* 1972; **34**:187-220.

12. Satten GA. Rank-based inference in the proportional hazards model for interval-censored data. *Biometrika* 1996; **83**:355-370.

13. Satten GA, Datta S, Williamson JM. Inference based on imputed failure times for the proportional hazards model with interval-censored data. *Journal of the American Statistical Association* 1998; **93**:318-327.

14. Pan W. A multiple imputation approach to Cox regression with interval-censored data. *Biometrics* 2000; **56**:199-203.

15. Goggins WB, Finkelstein DM, Schoenfeld DA, Zaslavsky AM. A Markov Chain Monte Carlo EM algorithm for analysing interval-censored data under the Cox proportional hazard model. *Biometrics* 1998; **54**:1498-1507.

16. Rosenberg PS, Goedert JJ, Giggar RJ. Effect of age at seroconversion on the natural AIDS incubation distribution. *AIDS* 1994; **8**:803-810.

17. Kooperberg C, Clarkson DB. Kooperberg C, Hazard regression with interval-censored data. *Biometrics* 1997; **53**:1485-1494.

18. Betensky RA, Lindsey JC, Ryan LM, Wand MP. A local likelihood proportional hazards model for interval-censored data. *Stat.Med.* 2002; **21**:263-275.

19. Finkelstein DM. A proportional hazards model for interval-censored failure time data. *Biometrics* 1986; **42**(4):845-854.

20. Friedman M. Piecewise exponential models for survival data with covariates. *Annals of Statistics* 1982; **10**:101-109.

21. Seaman SR, Bird SM. Proportional hazards model for interval-censored failure times and time-dependent covariates: application to hazard of HIV infection of injecting drug users in prison. *Stat.Med.* 2001; **20**:1855-1870.

22. Lindsey JC, Ryan LM. Tutorial in biostatistics - Methods for interval-censored data. *Statistics in Medicine* 1998; **17**(2):219-238.

23. Carsenten B. Regression models for interval-censored survival data: application to HIV infection in Danish homosexual men. *Stat.Med.* 1996; **15**:2177-2189.

24. Farrington CP. Interval censored survival data: a generalized linear modelling approach. *Statistics in Medicine* 1996; **15**(3):283-292.

25. So Y. Interval-censored survival data. *Proceedings of the 19th Annual SAS Users Group, SAS Institute, NC, USA.* 1994; (1107):1113.

26. Williams G, He K, Chen G, Chi G, Pazdur R. Operational bias in assessing time to progression (TTP). *Proc.Am.Soc.Clin.Oncol.* 2002; **21**(244a):(abstr 975).

27. Griffin JT, Fraser C, Gras L, deWolf F, Ghani AC. The effect on treatment comparisons of different measurement frequencies in Human Immunodeficiency Virus observaional databases. *Am J Epidemiol.* 2006; **163**:676-683.

28. Odell PM, Anderson KM, D'Agostino RB. Maximum likelihood estimation for interval censored data using a Weibull-based accelerated failure time model. *Biometrics* 1992; **48**:951-959.

29. Delta Coordinating Committee and Delta Virology Committee. HIV-1 RNA response to antiretroviral treatment in 1280 participants in the Delta Trial: an extended virology study. *AIDS* 1999; **13**(1):57-65.

30. Gibb DM, Duong T, Tookey PA, Sharland M, Tudor-Williams G, Novelli V, on behalf of the National Study of HIV in Pregnancy and Childhood (NSHPC) and of the Collaborative HIV Paediatric Study (CHIPS). Decline in mortality, AIDS and hospital admissions in perinatally HIV-1 infected children in the UK and Ireland. *BMJ* 2003; **327**:1019-1025.

31. Cox DR. Partial Likelihood. *Biometrika* 1975; **62**:269-276.

32. Breslow N. Covariance analysis of censored survival data. *Biometrika* 1974; **30**:89-99.

33. Alioum A, Commenges D. A proportional hazards model for arbitrarily censored and truncated data. *Biometrics* 1996; **52**(2):512-524.

34. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B* 1977; **39**:1-38.

35. Meng XL, Rubin DB. Using EM to obtain asymptotic variance-covariance matrices: the SEM algorithm. *Journal of the American Statistical Association* 1991; **86**:899-909.

36. Smith PJ, Thompson TJ, Jereb JA. A model for interval-censored tuberculosis outbreak data. *Stat.Med.* 2006; **16**:485-496.

37. Efron B. Logistic regression, survival analysis and the Kaplain-Meier curve. *Journal of the American Statistical Association* 1988; **83**:414-425.

38. Clark DE, Ryan LM. Concurrent prediction of hospital mortality and length of stay from rish factors on admission. *Health Services Report* 2002; **37**:631-645.

39. SAS Institute Inc., Cary, North Carolina, USA. 1999; (8.2).

40. Delta Coordinating Committee. Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. *Lancet* 1996; **348**:283-291.

41. Sharland M, Blanche S, Castelli G, Ramos J, Gibb D, on behalf of the PENTA Steering Committee. PENTA guidelines for the use of antiretroviral therapy, 2004. *HIV Medicine* 2004; **5**:61-86.

42. Judd A, Lee KJ, Duong T, Walker AS, Butler Keta. Exposure and response to highly active antiretroviral therapy (HAART) in ART naive children in the UK and Ireland. *XVI International AIDS Conference, Toronto 2006* 2006; **(abstr.THPE0106)**.

43. Li Y-H, Klein JP, Moeschberger ML. Effects of model mis-specification in estimating covariate effects in survival analysis for small sample sizes. *Computational Stats.* 1996; **22**:177-192.

CHAPTER 5: DISCUSSION AND FUTURE WORK

The aim of this research was to build upon previously published methods that are accessible via self-written macros in the standard packages - with the goal of increasing the use of appropriate methods in applied medical research, and to determine the impact of simple imputation upon both the estimation of the survivor function and the effect of covariates on the survival time. The key findings of this research are summarised below.

In the first part of this thesis, three methods for calculating pointwise confidence intervals for the non-parametric survivor function estimated from interval-censored data were described and assessed: the first based on the full (observed) information matrix (Wald-1), the second a modification of this approach involving deletion of rows and columns of the information matrix corresponding to zero estimates prior to inversion (Wald-2), and the third based on likelihood ratio inference.

The simulation study showed clearly that the Wald-1 method substantially and consistently over-estimated the standard error (of the estimated survivor function at fixed time points), and therefore resulted in confidence intervals that were too wide. In contrast, the Wald-2 method produced accurate standard errors, using the Monte Carlo estimates as a benchmark, and for most of the combinations of factors considered, the coverage was, for practical purposes, acceptably close to the nominal value. However, for some combinations, the coverage was unacceptably low, understating the true uncertainty in the point estimates. The likelihood ratio method gave the most accurate confidence intervals with coverage consistently close to the nominal level. This reassuring result was not predictable from a theoretical standpoint since a high proportion of the estimates lay on the boundary of the parameter space [1].

In Chapter 3, six methods for estimating the survivor function from interval-censored data were described and assessed: a fully parametric Weibull model (W), the non-parametric maximum likelihood estimator (NPMLE), two closely related smoothed

versions of this estimator (SNP & SNP2), a Kaplan-Meier estimator that assumes the event of interest occurs at the middle of the observed censoring interval (MKM), and finally a Kaplan-Meier estimator where the event is assumed to occur at the time it is observed (RKM).

The simulation study established that smoothing resulted in a significant increase in accuracy, with the SNP2 estimator consistently superior to the SNP and NPMLE estimators. It could also be clearly seen that the RKM method results in significant underestimation of the survival probabilities. In contrast, the MKM method performed well, comparable to the SNP2 estimator when the sample size was small, except when the underlying hazard function was decreasing and the data were sparse.

This work was then extended in chapter 4 to determine which methods could be successfully employed to obtain the most precise estimates of covariate effects. A fully parametric Weibull model (W), the full likelihood proportional hazards model proposed by Finkelstein (F), a piecewise exponential model (PE), an approximation to the PE model (APE), and Cox proportional hazards models where event times were imputed as the right- (RCox) or mid- (MCox) point of the censoring interval, were compared in a simulation study.

The number of scheduled visits and the proportion of missing visits had the greatest impact upon the performance of the different methods. There were no material differences between all six methods when follow-up visits were frequent in terms of bias, MSE or coverage. In addition, all methods estimated the standard error of the log hazard ratio accurately, including the APE, MCox and RCox methods. This was unexpected since by imputing event times these methods ignore some of the uncertainty in the observed data.

When the data were more sparse the RCox model estimated the log hazard ratio with substantial bias and low coverage at all sample sizes. The W, F and PE models estimated the covariate effect with a small positive bias when sample size was small, and had good coverage for all scenarios. In contrast, the MCox and APE models were consistently biased for all sample sizes and had lower than nominal coverage at large sample sizes. An important finding was that the MSE for the W, F and PE models was

larger than for the APE and MCox methods when sample size was small, due to less precise estimation of the hazard ratio rather than smaller bias.

Results from the simulation studies carried out in chapters 3 & 4 show clearly that the commonly employed method of imputing the event time as the right endpoint of the observed censoring interval for the analysis of interval-censored data can be substantially biased, both in terms of obtaining reliable estimates of the effect of covariates upon survival and the survivor function itself. Unbiased estimates of covariate effects can be obtained when the numbers of follow up visits are frequent but as the number of visits decreases the bias becomes considerable. In contrast, survival probabilities are always underestimated.

However, in certain situations simply replacing the right endpoint by the midpoint of the censoring interval results in unbiased estimators. In chapter 3, accurate estimates of the survivor function were obtained using the MKM method except when the sample size was large, the underlying hazard was decreasing and the number of follow-up visits was few. Inspection of the results of individual simulations showed that the poor performance of the MKM when visits are infrequent was mainly due to over-estimation of the survivor function when t is small. The estimator fails to capture early changes in the underlying survival distribution as imputed event times are restricted by the distribution of the earliest visits. In particular, the first step of the estimator cannot occur earlier than midway to the first visit time. This phenomenon is stronger when the underlying hazard is decreasing as proportionately more events occur at earlier times.

Midpoint imputation has also been shown, in chapter 4, to reliably estimate the hazard ratio except when there are few follow-up visits and the sample size was large. The shape of the underlying hazard was no longer influential. When there are few visits the method fails due to an inability to approximate the correct ranking of event and failure times. This problem is heightened as the number of ties in the data increases, which is synonymous with a decrease in the possible visits times. It is therefore important to use an exact likelihood for the treatment of ties rather than an approximation, such as proposed by Breslow, when fitting the MCox method.

The results for the APE model were disappointing. The method was no better, and often more biased, with lower coverage, than the MCox approach. Given the additional

complexity of reshaping the data into the required format before fitting the APE model, this method would be a second choice to the simple MCox approach in situations where both performed adequately.

Although circumstances have been identified under which the straightforward midpoint imputation approach may be successfully applied it should be noted that when the intellectual or computational capacity are available more complex models may be preferred. The choice of interval-censoring method, either when midpoint imputation is not recommended or when wishing to apply the appropriate methodology, will depend largely on the type of data being analysed.

The NPMLE can be used to estimate a non-parametric survivor function from interval-censored data in standard packages with the aid of user written macros. This method is an important addition to the fully parametric models which are readily available. Even if a parametric approach is preferred the NPMLE can be used to determine which family of distributions to use or assess goodness of fit. From the work in chapter 2 we can conclude that the likelihood ratio method gives the most consistent and accurate coverage for point estimates of the survivor function estimated using the NPMLE. The Wald-2 method is easier to compute than the likelihood ratio method and also achieved coverage close to the nominal value in most situations, particularly for large sample sizes, but should be used with care given coverage was sometime underestimated.

The results of chapter 3 imply the SNP2 approach is in fact preferential to the NPMLE since smoothing the estimator resulted in a significant increase in accuracy. Estimating a smooth function also gives more desirable estimates. In particular, estimation of median (or other percentile) survival times is now possible since the function is fully defined. However, there has been little research into methods of calculating confidence intervals for the SNP2 estimator although re-sampling methods can be used as in the example in section 3.4.

A large number of methods have been proposed in the literature for the estimation of the effect of covariates upon survival or more commonly the hazard distribution. Many are complex and computationally intensive and the methods included in this thesis were chosen because they could be easily fitted in the more common statistical packages once a generic macro program has been written. The full likelihood Cox model proposed by

Finkelstein extends the NPMLE and is attractive since no parametric assumptions about the baseline hazard are necessary. The model was sometimes unstable however at large sample sizes in the simulation study in chapter 4. The PE model is a versatile alternative which becomes more non-parametric in nature as the number of baseline hazard parameters increases. This method can also be easily extended to include time varying covariates. Either method would be recommended in situations where the MCox model was inappropriate. A fully parametric model may also be a good alternative and should not be overlooked, providing the distributional assumptions are checked, possibly by fitting the NPMLE.

A number of areas have been identified in which the work in this thesis can be extended. Although a range of hazard types were considered in the simulation study, all were monotone and consideration of bathtub or uni-modal hazard functions would be a useful extension of the simulation work presented. The simulation study could also be extended by considering parametric models other than the Weibull to investigate the robustness of midpoint imputation when the family of distributions for the hazard is misspecified. In addition, more precise identification of situations in which the Wald-2 method underestimates the coverage, and a comparison of this and likelihood ratio methods to the resampling methods proposed by Sun would compliment the work in chapter 2. Further work on the estimation of confidence intervals for the SNP2 estimator would facilitate the use of this method in practice.

For biological measurements such as HIV RNA levels, a value exists at each visit time which is summarised by the binary outcome of success or failure as considered here. It would therefore be possible to refine the imputation of event times by taking these measurements into account e.g. using linear interpolation, rather than assuming the events occur at the midpoint [2]. In related work Hsu et al. [3] describe imputation methods that incorporate information from auxiliary variables e.g. CD4 counts, rather than the outcome variable itself. Finally, the methods applied in this thesis would not be appropriate if the timing of visits was related to an individual's outcome or visits were informatively missing [4-6], for example if low levels of a biological marker were correlated to poor health and therefore more frequent clinic attendance or conversely non-attendance. As such, consideration of informative censoring or visit times would also be an important extension of this work.

References

1. Self SG, Liang KY. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non-standard conditions. *Journal of the American Statistical Association* 1987; **82**(398):605-610.

2. Dorey FJ, Little RJA, Schenker N. Multiple imputation for threshold-crossing data with interval censoring. *Statistics in Medicine* 1993; **12**:1589-1603.

3. Hsu C-H, Taylor JMG, Murray S, Commenges D. Multiple imputation for interval censored data with auxiliary variables. *Stat.Med.* 2007; **26**:769-781.

4. Farrington CP, Gay NJ. Interval-censored survival data with informative examination times: parametric models and approximate inference. *Statistics in Medicine* 1999; **18**(10):1235-1248.

5. Finkelstein DM, Goggins WB, Schoenfeld DA. Analysis of failure time data with dependent interval censoring. *Biometrics* 2002; **58**:298-304.

6. Shardell M, Scharfstein DO, Bozzette SA. Survival curve estimation for informatively coarsened discrete event-time data. *Stat.Med.* 2007; **26**:2184-2202.

APPENDICES

APPENDIX 1: Proof that the NPMLE is constant outside all distinct intervals with endpoints $\mathbf{u_j \in \{L_i\}}$ and $\mathbf{u_{j+1} \in \{R_i\}}$

Following the notation in Section 2.2.2, define the distinct intervals with endpoints $u_j \in \{L_i\}$ and $u_{j+1} \in \{R_i\}$ in order as $(q_1\, p_1)(q_2\, p_2)\ldots(q_m\, p_m)$, and define the set of points t_j , $j=i,\ldots,m-1$ where t_j is some value greater than all the right and less than all the left endpoints in $(p_j\, q_{j+1})$.

Define a function $\bar{S}$ that decreases outside the distinct intervals $(q_j\, p_j)$; $j=i,\ldots,m$. There exists at least one t_j such that either

$$\text{A} \qquad \bar{S}(p_j) > \bar{S}(t_j) \geq \bar{S}(q_{j+1})$$

or

$$\text{B} \qquad \bar{S}(p_j) \geq \bar{S}(t_j) > \bar{S}(q_{j+1})$$

Define a function S* to be constant outside the distinct intervals $(q_j\, p_j)$; $j=i,\ldots,m$ with $S^*(p_j) = S^*(q_{j+1}) = \bar{S}(t_j)$.

For case A, by construction of the distinct intervals, there exists at least one individual i such that $p_j = R_i$, implying $S^*(R_i) < \bar{S}(R_i)$. Therefore, denoting the contribution of the i-th individual to the log-likelihood ℓ_i,

$$\ell_i(S^*) > \ell_i(\bar{S})$$

and it follows that

$$\ell(S^*) > \ell(\bar{S})$$

and therefore $\bar{S}$ is not the MLE. The same follows for case B.

APPENDIX 2.A: Confidence intervals for the NPMLE in SAS

The simulation study (Section 2.4) was implemented using the SAS statistical software [1]. The ICE.sas program available in the SAS/IML sample library (under 'Interval Censored Estimation Macro') fits the non-parametric maximum likelihood estimator (NPMLE) to interval-censored data but gives confidence intervals using the Wald-1 method only [2]. This SAS code was adapted to allow estimation of Wald-2 and likelihood ratio confidence limits.

To calculate the Wald-2 confidence intervals, contributions to the Information matrix by parameters estimated as zero are removed before inverting to calculate the covariance matrix. Using code compatible with ICE.sas, for information matrix *h* and *m* parameters estimated by array *rx*, the code to calculate the covariance matrix *sigma2* by the Wald-2 method is given below:

```
top=m;
j=1;
do i=1 to m-2 ;
      if rx[i]<10**-8 then do;
        top=top-1;
        I1=h[1:i-j,1:i-j];
        I2=h[1:i-j,i-j+2:top];
        I3=h[i-j+2:top,1:i-j];
        I4=h[i-j+2:top,i-j+2:top];
        I=( I1 || I2) // (I3 || I4);
        j=j+1;
      end;
end;
top=top-1;
if rx[m-1]<10**-8 then do;
        top=top-1;
        I=I[1:top,1:top];
end;
sigma2=inv(I);
```

This can be directly inserted into the IML library program at line 469 after the section of code titled '/* covariance matrix of the first mm parameters */ ' . The ICE.sas program will then automatically output Wald-2 standard errors.

The likelihood ratio confidence limits are calculated for a specified time *t*, where *S(t)* corresponds to the sum of the first *i* parameters. A user written procedure then calculates the lower (upper) confidence limit by using interval bisection to find the root of

$$\hat{S}(t) - \tilde{S}(t) - \frac{1}{2}\chi_{\alpha}^{2} = 0 \qquad (1)$$

where $\hat{S}(t)$ is the survival probability calculated at the maximum likelihood estimate $\hat{f}$, and $\tilde{S}(t)$ is the desired lower (upper) limit. This is achieved as follows:

1. An initial interval for the lower (upper) limit is set as the MLE of S(t) at time t, $\hat{S}(t)$, and 0 (1).
2. The initial estimate of the confidence limit $\tilde{S}(t)$ is set as the lower (upper) confidence limit estimate calculated using the Wald-1 method. (In light of our findings, using the Wald-2 confidence limit would be more efficient).
3. The program then loops through the following process until the root is found:
    - constrain the sum of the last (*m-i*-1) parameters to be equal to the current estimate C
    - set initial values of the probability densities such that this new constraint holds and call the NLPQN optimisation subroutine to perform quasi-Newton optimisation
    - calculate the current value of the likelihood function, $\tilde{S}(t)$
    - calculate the value of equation (1) with the current estimate of $\tilde{S}(t)$
    - replace one end of the current estimation interval by the current estimate of $\tilde{S}(t)$ such that the interval still contains the root of the equation
    - set the new current estimate of $\tilde{S}(t)$ as the midpoint of the estimation interval

Systematically subdividing the interval in this way produces a final interval containing the zero. This final interval has a length bounded by the user's specified error requirements which are set as termination criterion for the above loop.

References

1. SAS Institute Inc., Cary, North Carolina, USA. 1999; (8.2).
2. ICE.sas. *http://support.sas.com/ctx/samples/index.jsp?sid=477&tab=about* 2006. 6-17-2006.

APPENDIX 2.B: SAS program to fit the NPMLE with Wald-2 confidence intervals

The left and right interval endpoints of the observed censoring intervals for each individual are recorded in variables LEFT and RIGHT of data set DATASET, where LEFT = 0 if the observation is left-censored, LEFT = RIGHT if the observation is exact, and RIGHT = an arbitrary fixed value beyond the last examination time if the observation is right-censored.

```
-----------------------------------------------------------------------

proc iml;

% fit the NPMLE using the quasi-Newton non-linear programming routine;
start interval(l,r) global(_x,nobs,nparm,end,lll,rrr);

   lll=l;
   rrr=r;
   nobs= nrow(l);

   /* GENERATE NON-OVERLAPPING INTERVALS */
    p=0;
    q=0;
    call nolap(nparm, p, q, l, r);

   /* GENERATE THE ALPHA-MATRIX */
    _x= j(nobs, nparm, 0);
    do j= 1 to nparm;
       _x[,j]= choose(l <= q[j]  & p[j] <= r, 1, 0);
       end;

   /* USING NLP TO MAXIMIZE LIKELIHOOD FUNCTION */
   /* options */
    optn= {1 0};

   /* constraints */
    con= j(3, nparm + 2, .);
    con[1, 1:nparm]=1.e-10;
    con[2:3, 1:nparm]= 1;
    con[3,nparm + 1]=0;
    con[3,nparm + 2]=1;

   /* initial estimates */
    x0= j(1, nparm, 1/nparm);

   * **** termination criterion *****;
    tc=j(1,13,0);
    tc[1,1]=4000;
    tc[1,2]=8000;
```

```
 tc[1,3]=.;
 tc[1,9]=1e-12;

/* call the optimization routine */
 call nlpqn(rc,rx,"LL",x0,optn,con,tc,,,"GRAD");
 optx=rx;
 optf=LL(optx);

 q= q`;
 p= p`;
 theta= rx`;
 do i=1 to nparm ;
   if rx[i]<10**-8 then rx[i]=0;
 end;

/* COMPUTE THE SURVIVAL DISTRIBUTION FUNCTION (NPMLE) */
 tmp1= cusum(rx[nparm:1]);
 sdf= tmp1[nparm-1:1];

/* COMPUTE THE WALD-2 CONFIDENCE LIMITS OF THE NPMLE */
 mm= nparm -1;

/* calculate full covariance matrix of the first mm parameters */
 _xx= _x - _x[,nparm] * (j(1, mm, 1) || {0});
 h= j(mm, mm, 0);
 ixtheta= 1 / (_xx * ((rx[,1:mm]) || {1})`);
 do i= 1 to nobs;
    rowtmp= ixtheta[i] # _xx[i,1:mm];
    h= h + (rowtmp` * rowtmp);
 end;

/* remove rows and columns of parameters estimated as zero */
A=h;
top=nparm;
j=1;
do i=1 to mm-1 ;
      if rx[i]<10**-8 then do;
        top=top-1;
        A1=A[1:i-j,1:i-j];
        A2=A[1:i-j,i-j+2:top];
        A3=A[i-j+2:top,1:i-j];
        A4=A[i-j+2:top,i-j+2:top];
        A5=A1 || A2;
        A6=A3 || A4;
        A=A5 // A6;
        j=j+1;
      end;
end;
top=top-1;
if rx[mm]<10**-8 then do;
        top=top-1;
        A=A[1:top,1:top];
```

```
   end;

   /* estimated variance of the NPMLE */
   sigma2p= inv(A);
   ind=rx>10**-8;
   sigma3p= j(mm, 1, 0);
   tmp1= j(top, 1, 0);
   j=1;
   do i= 1 to mm;
      tmp1[j]= 1;
      sigma3p[i]= tmp1` * sigma2p * tmp1;
      j=j+ind[1,i+1];
   end;
   sep=sqrt(sigma3p);

   /* confidence limits */
    tmp1= probit(1 - .5 * 0.05);
    tmp1= tmp1 *sqrt(sigma3p);
    lcl= choose(sdf>tmp1,sdf - tmp1,0);
    ucl= sdf + tmp1;

   /* PRINTOUT #3*/
    left= {0} // p;
    right= q // end;
    sdf= {1} // sdf // {0};
    se= {.} // se // {.};
    sep= {.} // sep // {.};
    lcl= {.} // lcl // {.};
    ucl= {.} // ucl // {.};

   /* print distinct intervals and parameter estimates */
    print q p theta;
   /* print NPMLE and Wald-2 SEs/CLs of survivor function */
    print left right sdf sep lcl ucl;

finish  interval;

 %* loglikelihood function *;
 start LL(theta) global(_x,nparm,lll,rrr);
   tmp=_x * theta`;
   xlt= log(_x * theta`);
   f= xlt[+];
   return(f);
 finish LL;

 %* gradient vector *;
 start GRAD(theta) global(_x,nparm);
   g= j(1,nparm,0);
   tmp= _x # (1 /  (_x * theta`) );
   g= tmp[+,];
   return(g);
 finish GRAD;
```

```
%****** CONSTRUCT THE NON-OVERLAPPING TIME ******;
%****** INTERVALS FOR THE TURNBULL METHOD *******;
start nolap(nq,p,q,l,r) global(end,le,ri);
   le=l;ri=r;
   pp= unique(r); npp= ncol(pp);
   qq= unique(l); nqq= ncol(qq);
   end=pp[npp];
   q= j(1,npp, .);
   do i= 1 to npp;
      do j= 1 to nqq;
         if ( qq[j] <= pp[i] ) then q[i]= qq[j];
      end;
      if q[i] = qq[nqq] then goto lab1;
   end;

lab1:
   if i > npp then nq= npp;
   else            nq= i;
   q= unique(q[1:nq]);
   nq= ncol(q);
   p= j(1,nq, .);
   do i= 1 to nq;
     do j= npp to 1 by -1;
        if ( pp[j] >= q[i] ) then p[i]= pp[j];
     end;
   end;
finish nolap;


   * ******** MAIN PROGRAM *********;

   use DATASET;
   read all var{left right};
   call interval(left,right);

   quit;
```

APPENDIX 2.C: SAS IML macro to calculate profile likelihood confidence limits for the NPMLE

```
Macro arguments:
Input:
i = The profile likelihood confidence limits are calculated for a
specified time t, where S(t) corresponds to the sum of the first i
parameters.

Global arguments (from macro INTERVAL above):
nparm = number of estimated parameters
con = constraint matrix from estimation of NPMLE
sdf = NPMLE of survivor function
ucl = upper Wald-2 confidence limit of NPMLE
lcl = lower Wald-2 confidence limit of NPMLE

Output arguments:
lower = upper profile likelihood confidence limit
upper = lower profile likelihood confidence limit

-----------------------------------------------------------------------

* ****** PROFILE LIKELIHOOD CI ********;
start plci(i,lower,upper) global(nparm,con,sdf,lcl,us,optf);

   * **** termination criterion *****;
    tc=j(1,13,0);
    tc[1,1]=20000;
    tc[1,2]=18000;
    tc[1,3]=.;
    tc[1,9]=1e-12;

   /* options */
    optn= {1 0};

   * ***** add constraint *******;
    conadd=j(1, nparm + 2, .);
    conadd[1,nparm+1]=0;
    con2=con//conadd;
    do k=i+1 to nparm;
        con2[4,k]=1;
    end;
    chiprob=cinv(1-0.05,1);

       * lower limit ***;
        lest=0;
        rest=sdf[i+1];
        currentl=lcl[i+1];
        if currentl=0 then currentl=10**-8;
        diff=optf;
```

```
        do while ( abs(diff)>10**-2);
            con2[4,nparm+2]=currentl;
            x0=j(1,nparm,(1-currentl)/i);
            do m = i+1 to nparm;
                x0[1,m]=currentl/(nparm-i);
            end;
            * call optimisation routine and calc constrained ests ;
            call nlpqn(rc,rx,"LL",x0,optn,con2,tc,,,"GRAD");
            estlike=LL(rx);

            diff=optf-estlike-chiprob*0.5;
            if diff<0 then do;
                 rest=currentl;
                 currentl=lest+0.5*(currentl-lest);
            end;
            if diff>0 then do;
                 lest=currentl;
                 currentl=currentl+0.5*(rest-currentl);
            end;
        end;
        lower=currentl;

        * upper limit ;
        rest=1;
        lest=sdf[i+1];
        currentu=us[i+1];
        if currentu>=1 then currentu=1-10**-8;
        diff=optf;
        do while ( abs(diff)>10**-2);
            con2[4,nparm+2]=currentu;
            x0=j(1,nparm,(1-currentu)/i);
            do m = 1 to i;
                x0[1,m]=currentu/(nparm-i);
            end;
            * call optimisation routine and calc constrained ests ;
            call nlpqn(rc,rx,"LL",x0,optn,con2,tc,,,"GRAD");
            estlike=LL(rx);

            diff=optf-estlike-chiprob*0.5;
            if diff>0 then do;
                 rest=currentu;
                 currentu=lest+0.5*(currentu-lest);
            end;
            if diff<0 then do;
                 lest=currentu;
                 currentu=currentu+0.5*(rest-currentu);
            end;
        end;
        upper=currentu;

finish plci;
```

APPENDIX 3.A: Fitting the SNP and SNP2 methods in SAS

The only difference between estimation of the SNP and SNP2 methods is in the estimation of the time intervals during which the survivor function jumps. The SNP method uses the same routine as the NPMLE to calculate intervals with endpoints that are in $\{L_i\}$ and $\{R_i\}$ respectively and contain no other endpoints. The SNP2 method uses time intervals defined by all unique values of $\{L_i\}$ and $\{R_i\}$.

APPENDIX 3.B: SAS program to fit SNP or SNP2 models

```
The left and right interval endpoints of the observed censoring
intervals for each individual are recorded in variables LEFT and RIGHT
of data set DATASET, where LEFT = 0 if the observation is left-
censored, LEFT = RIGHT if the observation is exact, and RIGHT = an
arbitrary fixed value beyond the last examination time if the
observation is right-censored.

-----------------------------------------------------------------------

start interval(l,r)
global(_x,nobs,nparm,ipar,lstar,end,vars3,con,sdf,ls,us,optf,lll,rrr);

lll=l;
rrr=r;
nobs= nrow(l);

-----------------------------------------------------------------------
SNP2:                                   SNP:
   /* GENERATE UNIQUE VALUES OF         /* GENERATE NON-OVERLAPPING
               {Li} and {Ri} */                          INTERVALS
                                        */
   timepoints=unique(lll||rrr);
   nparm=ncol(timepoints);                  p=0;
   nparm1=nparm-1;                          q=0;
   q= timepoints[1:nparm1]`;                call nolap(nparm, p, q, l,
   p= timepoints[2:nparm]`;             r);
   nparm=nparm-1;
-----------------------------------------------------------------------

/* GENERATE THE ALPHA-MATRIX */
_x= j(nobs, nparm, 0);
do j= 1 to nparm;
   _x[,j]= choose(l <= q[j]  & p[j] <= r, 1, 0);
end;

/* initial estimates */
x0= j(1, nparm, 1/nparm);
```

```
emconv=1e-12;

/* CALL EM MACRO */
rx=em(x0, emconv);

q= q`;
p= p`;
theta= rx`;
do i=1 to nparm ;
   if rx[i]<10**-8 then rx[i]=0;
end;

/* COMPUTE THE SURVIVAL DISTRIBUTION FUNCTION */
tmp1= cusum(rx[nparm:1]);
sdf= tmp1[nparm-1:1];
/* PRINTOUT #3*/
left= {0} // p;
right= q // end;
sdf= {1} // sdf // {0};
print nparm q p theta;
print left right sdf;

finish interval;

/* CONSTRUCT THE NON-OVERLAPPING TIME INTERVALS */
start nolap(nq,p,q,l,r) global(end,le,ri);
le=l;ri=r;
pp= unique(r); npp= ncol(pp);
qq= unique(l); nqq= ncol(qq);
end=pp[npp];
q= j(1,npp, .);
do i= 1 to npp;
   do j= 1 to nqq;
      if ( qq[j] < pp[i] ) then q[i]= qq[j];
   end;
   if q[i] = qq[nqq] then goto lab1;
end;

lab1:
if i > npp then nq= npp;
else            nq= i;
q= unique(q[1:nq]);
nq= ncol(q);
p= j(1,nq, .);
do i= 1 to nq;
   do j= npp to 1 by -1;
      if ( pp[j] > q[i] ) then p[i]= pp[j];
   end;
end;
finish nolap;
```

```
%****** Self-Consistency Algorithm *******;
start em(theta0, conv) global(_x,nobs,nparm,_zfreq,_freq);
   iter=0;
   u= _x # theta0;
   xt= u[,+];
   lxt= log(xt);
   u= u # (1 / xt);
   ntot= nobs;
   ll0= lxt[+];
   theta0= u[+,] / ntot;
if nparm=3 then do ;
      s0= {3} || {3};
      s1= theta0[,2] || {0};
      s2= {0} || theta0[,1];
end;
if nparm>3 then do;
      s0= {3} || j(1,nparm-3,2) || {3};
      s1= theta0[,2:nparm-1] || {0};
      s2= {0} || theta0[,1:nparm-2];
end;
    s= (s0 # theta0[,1:nparm-1]) + s1 + s2;
    s = (0.25 # s ) || theta0[,nparm];
    theta0=s;
    difcrit= 1;

    do while ( difcrit > conv );
       iter= iter + 1;
       u= _x # theta0;
       xt= u[,+];
       lxt= log(xt);
       u= u # (1 / xt);
       ll= lxt[+];
       theta0= u[+,] / ntot;
if nparm=3 then do;
      s0= {3} || {3};
      s1= theta0[,2] || {0};
      s2= {0} || theta0[,1];
end;
if nparm>3 then do;
      s0= {3} || j(1,nparm-3,2) || {3};
      s1= theta0[,2:nparm-1] || {0};
      s2= {0} || theta0[,1:nparm-2];
end;
s= (s0 # theta0[,1:nparm-1]) + s1 + s2;
s = (0.25 # s ) || theta0[,nparm];
theta0=s;

       difcrit= ll - ll0;
       ll0= ll;
    end;
    return(theta0);
 finish em;
```

```
* ******** MAIN PROGRAM *********;

use DATASET;
read all var{left right};
call interval(left,right);
quit;
```

APPENDIX 4.A: Fitting the Finkelstein, PE AND APE models in SAS

The left and right interval endpoints of the observed censoring intervals for each individual are recorded in variables LEFT and RIGHT of data set DATASET, where LEFT = 0 if the observation is left-censored, LEFT = RIGHT if the observation is exact, and RIGHT = an arbitrary fixed value beyond the last examination time if the observation is right-censored. DATASET contains a single covariate X.

APPENDIX 4.B: SAS program to fit the Finkelstein model

```
proc iml;

/* CONSTRUCT THE NON-OVERLAPPING TIME INTERVALS */
start nolap(nq,p,q,l,r);

pp= unique(r); npp= ncol(pp);
qq= unique(l); nqq= ncol(qq);
q= j(1,npp, .);
do i= 1 to npp;
   do j= 1 to nqq;
      if ( qq[j] < pp[i] ) then q[i]= qq[j];
   end;
   if q[i] = qq[nqq] then goto lab1;
end;

lab1:
if i > npp then nq= npp;
else            nq= i;
q= unique(q[1:nq]);
nq= ncol(q);
p= j(1,nq, .);
do i= 1 to nq;
   do j= npp to 1 by -1;
      if ( pp[j] > q[i] ) then p[i]= pp[j];
   end;
end;
finish nolap;

/* ************************************ */
use DATASET;
read all var{left right x};
l=left; r=right;
nobs= nrow(l);

/* GENERATE NON-OVERLAPPING INTERVALS */
p=0;
q=0;
call nolap(nparm, p, q, l, r);
nparm1=nparm-1;
```

```
/* GENERATE THE ALPHA-MATRIX */
alpha1= j(nobs, 1, 0);
_x= j(nobs, nparm1, 0);
do i= 1 to nobs;
   /* sort out first col of alpha */
   if l[i]<=q[1] then alpha1[i]=1;
   do j= 2 to nparm;
      if ( l[i]>=p[j-1] & l[i]<=q[j] ) then _x[i,j-1]=1;
      if ( r[i]>=p[j-1] & r[i]<=q[j] ) then _x[i,j-1]=-1;
   end;
end;

/* no. parameters */
ncoef=1;

start LL(theta) global(nobs,nparm1,ncoef,alpha1,_x,x);
   u=nparm1+1;
   v=nparm1+ncoef;
   ll=0;
   do i=1 to nobs;
      xbeta= x[i,]*theta[u:v] ;
      l0=alpha1[i];
      do j=1 to nparm1;
         tmp1=-exp(theta[j]+xbeta);
         tmp2=exp(tmp1);
         l0=l0+_x[i,j]*tmp2;
      end;
   ll=ll+log(l0);
   end;
   return(ll);
finish;

start GRAD(theta)  global(nobs,nparm1,ncoef,alpha1,_x,x);
   u=nparm1+1;
   v=nparm1+ncoef;
   g1=j(1,nparm1,0);
   g2=j(1,ncoef,0);
   h=g1;
   do i = 1 to nobs;
      xbeta= x[i,]*theta[u:v] ;
      lli=alpha1[i];
      l0=0;
      do j= 1 to nparm1;
         tmp1=-exp(theta[j]+xbeta);
         tmp2=exp(tmp1)*_x[i,j];
         lli=lli+tmp2;
         tmp3=tmp2*tmp1;
         l0=l0+tmp3;
         h[j]=tmp3;
      end;
      g1=g1+(h/lli);
      g2=g2+(x[i,]*l0/lli);
```

```
   end;
   g = g1||g2;
   return (g);
finish;
/* estimate the regression parameters */

/* options */
opt= {1,0};

/* termination criteria */
tc=j(1,10,.);
tc[1,1]=500;
tc[1,2]=1000;
tc[1,6]=0;

* constraints;
nparm=nparm1+ncoef;
nparm2=nparm1-1;
con=j(nparm2,nparm+2,0);
do i=1 to nparm2;
   con[i,i]=-1;
   con[i,i+1]=1;
   con[i,nparm+1]=1;
end;
con=j(1,nparm+2,.)//j(1,nparm+2,.)//con;

/* initial estimates */
beta=j(1,ncoef,0);
gamma=j(1,nparm1,0);
do j=1 to nparm1;
   gamma[j]=log(-log(1-j/nparm));
end;
z0=gamma || beta;
b={.};
seb={.};
rcode={.};
method=2;

call nlpqn(rc,est,"ll",z0,opt,con,tc,,,"grad");

rcode=rc;
if rcode <15 & rcode>0 then do ;
      b=est[nparm];
      call nlpfdd(f, g, h, "ll", est,,"grad");
      help=-1*h;
      v=inv(help);
      betav=v[nparm,nparm];
      seb=sqrt(betav);
end;
finkres= method || rcode || b || seb || nparm1 || nrcens || nlcens;

quit;
```

APPENDIX 4.C: SAS program to fit the PE model

```
proc lifetest data=DATASET outsurv=surv noprint;
time midpt*event(0);run;

data a (KEEP= cutpt rep); set surv; where sdf_lcl ne .;
cutpt=0;
rep=0;
if survival<0.2 & lag(survival)>0.2 then do; cutpt=midpt; rep=rep+1;
end;
if survival<0.4 & lag(survival)>0.4 then do; cutpt=midpt; rep=rep+1;
end;
if survival<0.6 & lag(survival)>0.6 then do; cutpt=midpt; rep=rep+1;
end;
if survival<0.8 & lag(survival)>0.8 then do; cutpt=midpt; rep=rep+1;
end;
run;
data pcent5; set a; where rep ne 0; run;

proc iml;

start FITPE(l,r,covar,event)
global(nparm,nobs,nint,w,tj_1,tj,vmat,j1,i1,xx,resp,A,B,ests,optx);

resp=event;
xx=covar;
nobs=nrow(l);
nint=nrow(w);
nparm=nint+1;
i1=j(nobs,1,1);
j1=j(1,nint,1);
A=j(nobs,nint,0);
B=j(nobs,nint,0);
  do i= 1 to nobs;
  do j= 1 to nint;
    if (l[i] > tj_1[j] & l[i]<=tj[j]) then A[i,j]=l[i]-tj_1[j];
    else if (l[i] > tj[j]) then A[i,j]=w[j];
    else A[i,j]=0;

    if (r[i] >= tj[j] & l[i]<tj[j] & l[i]>tj_1[j] & event[i]=1) then
      B[i,j]=tj[j]-l[i];
    else if (r[i] >= tj[j] & l[i]<=tj_1[j] & event[i]=1) then
      B[i,j]=w[j];
    else if (r[i] < tj[j] & l[i]>=tj_1[j] & event[i]=1) then
      B[i,j]=r[i]-l[i];
    else if (r[i] < tj[j] & r[i]>tj_1[j] & l[i]<tj_1[j] & event[i]=1)
    then B[i,j]=r[i]-tj_1[j];
    else B[i,j]=0;
  end;
  end;
```

```
/* USING NLP TO MAXIMIZE LIKELIHOOD FUNCTION */
/* options */
optn= {1 0};

/* initial estimates */
startt=j(nobs,1,0);
startt=choose(event=0,l,l+0.5*(r-l));
pd0=i1*w`;
pevent=j(nobs,nint,-9);
do j=1 to nint;
   pevent[,j]=choose(startt>tj_1[j] & startt<=tj[j],1,0);
   pd0[,j]=choose(startt<=tj_1[j],0,pd0[,j]);
   pd0[,j]=choose(startt>tj_1[j] & pevent[,j]=1,startt-
      tj_1[j],pd0[,j]);
end;
totevent=pevent[+,];
totpd0=pd0[+,];
x0= totevent/totpd0 || {0};

* **** termination criterion *****;
tc=j(1,13,0);
tc[1,1]=12000;
tc[1,2]=10000;
tc[1,3]=.;
tc[1,9]=1e-10;

/* constraint */
conl=j(1,nparm,-20);
conu=j(1,nparm,.);
con=conl//conu;

/* call the optimization routine */
call nlpqn(rc,rx,"LL",x0,optn,con,tc,,,"GRAD");
optx=rx;

call nlpfdd(f, g, h, "ll", rx,,"grad");
negh=-1*h;
help=det(negh);
if help ^= 0 then do;
      vmat=inv(negh);
      seb=sqrt(vmat[nparm,nparm]);
      ests=optx[1,nparm] || seb;
end;
if help = 0 then do;
      seb={.};
      ests=optx[1,nparm] || seb;
end;

finish FITPE;
```

```
start PCT5;
use pcent5;
read all var{cutpt rep};
nint=nrow(cutpt);
total=rep[+];
cutpt=cutpt // {364};
rep[nint]=rep[nint]+(4-total);
tmp=cutpt;
do j = 1 to nint;
   do k = 1 to rep[j];
      new= cutpt[j] + k*(cutpt[j+1]-cutpt[j])/rep[j];
        tmp=union(tmp,new);
   end;
end;
tj=tmp`;
tj_1 = {0} // tj[1:4];
w = tj - tj_1;
tmp2=tmp[1:4];
output= {1} || tmp2`;
varnames={z p20 p40 p60 p80};
create tmp5 from output [colname=varnames];
append from output;
close tmp5;

finish PCT5;

/* loglikelihood function */
 start LL(theta) global(xx,A,B,resp,i1,j1,nparm,nint,nobs);
   LPIij=i1*theta[1,1:nint];
   LPXij=xx*theta[1,nparm];
   tmp=LPXij*j1;
   LPij=LPIij + tmp;
   sbrate=(-exp(LPij)#B)[,+];
   mmm=choose(1-exp(sbrate)=0,1,1-exp(sbrate));
   respbit=(resp # log(mmm))[+];
   abit=(-exp(LPij) # A )[+];
   f= abit+respbit;
   return(f);
 finish LL;

/* gradient vector */
 start GRAD(theta) global(xx,A,B,resp,i1,j1,nparm,nint,nobs);
   LPIij=i1*theta[1,1:nint];
   LPXij=xx*theta[1,nparm];
   tmp=LPXij*j1;
   tmp2=xx*j1;
   LPij=LPIij + tmp;
   sbrate=(-exp(LPij)#B)[,+];
   abit= -exp(LPij) # A ;
   bbit= (exp(sbrate)*j1) # exp(LPij) # B;
   cbit= choose(1-exp(sbrate)=0,1,1-exp(sbrate));
   Gij=abit + bbit/(cbit*j1);
```

```
  intgrad=Gij[+,];
  lpgrad=(tmp2 # Gij)[+];
  g= intgrad || lpgrad;
  return(g);
finish GRAD;

%* -ve of hessian matrix *;
start HESS(theta) global(xx,A,B,resp,i1,j1,nparm,nint,nobs);
  LPIij=i1*theta[1,1:nint];
  LPXij=xx*theta[1,nparm];
  tmp=LPXij*j1;
  tmp2=xx*j1;
  LPij=LPIij + tmp;
  sbrate=(-exp(LPij)#B)[,+];
  abit= -exp(LPij) # A ;
  bbit= (exp(sbrate)*j1) # exp(LPij) # B;
  cbit= choose(1-exp(sbrate)=0,1,1-exp(sbrate));
  cbit2=(cbit##2 )*j1;
  dbit= j(nobs,nint,1) - (exp(LPij)#B) - (exp(sbrate)*j1);
  Hij=abit + (bbit#dbit)/cbit2;
  int=Hij[+,];
  lp=  ((tmp2##2) # Hij)[+];
  intlp=(tmp2 # Hij)[+,];
  h=diag( int || lp );
  h[nparm,1:nint]=intlp;
  h[1:nint,nparm]=intlp`;
  return(h);
finish HESS;

/* MAIN PROGRAM */

use DATASET;
read all var{left right x event};
call pct5;
call fitpe(left,right,x,event);

quit;
```

APPENDIX 4.D: SAS program to fit the APE model

```
data DATASET; set DATASET;
if event=0 then right=.; z=1;
run;
data use5; merge DATASET tmp5; by z; run;

data use5;set use5;
keep left right t  event x id simnum i li ri wi;
i=1; li=0; ri=p20; wi=p20; output;
i=2; li=p20; ri=p40; wi=p40-p20; output;
i=3; li=p40; ri=p60; wi=p60-p40; output;
```

```
i=4; li=p60; ri=p80; wi=p80-p60; output;
i=5; li=p80; ri=52*7; wi=52*7-p80; output;
run;

%MACRO FITMP(data=,est=);

data &data; set &data;
if (left>=ri) then pd=wi;
if (left>li & left<=ri) then pd=left-li;
eflag=0;
if ( (pd=. or pd<wi) and right>li) then eflag=1;
if eflag=1 then atrisk=wi;
if eflag=1 & pd~=. then atrisk=wi-pd;
if (right>li & right<=ri) then atrisk=right-li;
if (left>=li & right<=ri) then atrisk=right-left;
if atrisk=. then atrisk=0;
cens=1-event;
if i ne 1 then cens=0;
run;
proc means data=&data noprint;
   var atrisk; class id;
   output out=tmp sum=tatrisk ;
run;
proc sort data=tmp; by id; run;
proc sort data=&data; by id descending i; run;
data &data; merge tmp &data; by id; where id ne .; run;
data &data; retain sum; set &data;
   by id descending i;
   if first.id then sum=0;
   sum=sum+atrisk;
run;
data &data; set &data;
   if eflag=1 then e=atrisk/tatrisk;
   else e=0;
   if pd=. then pd=0;
   pdbit=e*(sum-0.5*atrisk);
   totpd=pd+pdbit;
   if totpd=0 then delete;
   ltotpd=log(totpd);
run;

proc genmod data=&data;
class i;
model e = i x / noint dist=p offset=ltotpd;
make 'parameterestimates' out=parm ;
run;

%MEND;

%fitmp(data=use5,est=est5);
```

APPENDIX 4.A: Fitting the Finkelstein, PE AND APE models in SAS

The left and right interval endpoints of the observed censoring intervals for each individual are recorded in variables LEFT and RIGHT of data set DATASET, where LEFT = 0 if the observation is left-censored, LEFT = RIGHT if the observation is exact, and RIGHT = an arbitrary fixed value beyond the last examination time if the observation is right-censored. DATASET contains a single covariate X.

APPENDIX 4.B: SAS program to fit the Finkelstein model

```
proc iml;

/* CONSTRUCT THE NON-OVERLAPPING TIME INTERVALS */
start nolap(nq,p,q,l,r);

pp= unique(r); npp= ncol(pp);
qq= unique(l); nqq= ncol(qq);
q= j(1,npp, .);
do i= 1 to npp;
   do j= 1 to nqq;
      if ( qq[j] < pp[i] ) then q[i]= qq[j];
   end;
   if q[i] = qq[nqq] then goto lab1;
end;

lab1:
if i > npp then nq= npp;
else            nq= i;
q= unique(q[1:nq]);
nq= ncol(q);
p= j(1,nq, .);
do i= 1 to nq;
   do j= npp to 1 by -1;
      if ( pp[j] > q[i] ) then p[i]= pp[j];
   end;
end;
finish nolap;

/* ************************************ */
use DATASET;
read all var{left right x};
l=left; r=right;
nobs= nrow(l);

/* GENERATE NON-OVERLAPPING INTERVALS */
p=0;
q=0;
call nolap(nparm, p, q, l, r);
nparm1=nparm-1;
```

```
/* GENERATE THE ALPHA-MATRIX */
alpha1= j(nobs, 1, 0);
_x= j(nobs, nparm1, 0);
do i= 1 to nobs;
   /* sort out first col of alpha */
   if l[i]<=q[1] then alpha1[i]=1;
   do j= 2 to nparm;
      if ( l[i]>=p[j-1] & l[i]<=q[j] ) then _x[i,j-1]=1;
      if ( r[i]>=p[j-1] & r[i]<=q[j] ) then _x[i,j-1]=-1;
   end;
end;

/* no. parameters */
ncoef=1;

start LL(theta) global(nobs,nparm1,ncoef,alpha1,_x,x);
   u=nparm1+1;
   v=nparm1+ncoef;
   ll=0;
   do i=1 to nobs;
      xbeta= x[i,]*theta[u:v] ;
      l0=alpha1[i];
      do j=1 to nparm1;
         tmp1=-exp(theta[j]+xbeta);
         tmp2=exp(tmp1);
         l0=l0+_x[i,j]*tmp2;
      end;
   ll=ll+log(l0);
   end;
   return(ll);
finish;

start GRAD(theta)  global(nobs,nparm1,ncoef,alpha1,_x,x);
   u=nparm1+1;
   v=nparm1+ncoef;
   g1=j(1,nparm1,0);
   g2=j(1,ncoef,0);
   h=g1;
   do i = 1 to nobs;
      xbeta= x[i,]*theta[u:v] ;
      lli=alpha1[i];
      l0=0;
      do j= 1 to nparm1;
         tmp1=-exp(theta[j]+xbeta);
         tmp2=exp(tmp1)*_x[i,j];
         lli=lli+tmp2;
         tmp3=tmp2*tmp1;
         l0=l0+tmp3;
         h[j]=tmp3;
      end;
      g1=g1+(h/lli);
      g2=g2+(x[i,]*l0/lli);
```

```
   end;
   g = g1||g2;
   return (g);
finish;
/* estimate the regression parameters */

/* options */
opt= {1,0};

/* termination criteria */
tc=j(1,10,.);
tc[1,1]=500;
tc[1,2]=1000;
tc[1,6]=0;

* constraints;
nparm=nparm1+ncoef;
nparm2=nparm1-1;
con=j(nparm2,nparm+2,0);
do i=1 to nparm2;
   con[i,i]=-1;
   con[i,i+1]=1;
   con[i,nparm+1]=1;
end;
con=j(1,nparm+2,.)//j(1,nparm+2,.)//con;

/* initial estimates */
beta=j(1,ncoef,0);
gamma=j(1,nparm1,0);
do j=1 to nparm1;
   gamma[j]=log(-log(1-j/nparm));
end;
z0=gamma || beta;
b={.};
seb={.};
rcode={.};
method=2;

call nlpqn(rc,est,"ll",z0,opt,con,tc,,,"grad");

rcode=rc;
if rcode <15 & rcode>0 then do ;
      b=est[nparm];
      call nlpfdd(f, g, h, "ll", est,,"grad");
      help=-1*h;
      v=inv(help);
      betav=v[nparm,nparm];
      seb=sqrt(betav);
end;
finkres= method || rcode || b || seb || nparm1 || nrcens || nlcens;

quit;
```

APPENDIX 4.C: SAS program to fit the PE model

```
proc lifetest data=DATASET outsurv=surv noprint;
time midpt*event(0);run;

data a (KEEP= cutpt rep); set surv; where sdf_lcl ne .;
cutpt=0;
rep=0;
if survival<0.2 & lag(survival)>0.2 then do; cutpt=midpt; rep=rep+1;
end;
if survival<0.4 & lag(survival)>0.4 then do; cutpt=midpt; rep=rep+1;
end;
if survival<0.6 & lag(survival)>0.6 then do; cutpt=midpt; rep=rep+1;
end;
if survival<0.8 & lag(survival)>0.8 then do; cutpt=midpt; rep=rep+1;
end;
run;
data pcent5; set a; where rep ne 0; run;

proc iml;

start FITPE(l,r,covar,event)
global(nparm,nobs,nint,w,tj_1,tj,vmat,j1,i1,xx,resp,A,B,ests,optx);

resp=event;
xx=covar;
nobs=nrow(l);
nint=nrow(w);
nparm=nint+1;
i1=j(nobs,1,1);
j1=j(1,nint,1);
A=j(nobs,nint,0);
B=j(nobs,nint,0);
  do i= 1 to nobs;
  do j= 1 to nint;
    if (l[i] > tj_1[j] & l[i]<=tj[j]) then A[i,j]=l[i]-tj_1[j];
    else if (l[i] > tj[j]) then A[i,j]=w[j];
    else A[i,j]=0;

    if (r[i] >= tj[j] & l[i]<tj[j] & l[i]>tj_1[j] & event[i]=1) then
      B[i,j]=tj[j]-l[i];
    else if (r[i] >= tj[j] & l[i]<=tj_1[j] & event[i]=1) then
      B[i,j]=w[j];
    else if (r[i] < tj[j] & l[i]>=tj_1[j] & event[i]=1) then
      B[i,j]=r[i]-l[i];
    else if (r[i] < tj[j] & r[i]>tj_1[j] & l[i]<tj_1[j] & event[i]=1)
    then B[i,j]=r[i]-tj_1[j];
    else B[i,j]=0;
  end;
  end;
```

```
/* USING NLP TO MAXIMIZE LIKELIHOOD FUNCTION */
/* options */
optn= {1 0};

/* initial estimates */
startt=j(nobs,1,0);
startt=choose(event=0,1,1+0.5*(r-1));
pd0=i1*w`;
pevent=j(nobs,nint,-9);
do j=1 to nint;
   pevent[,j]=choose(startt>tj_1[j] & startt<=tj[j],1,0);
   pd0[,j]=choose(startt<=tj_1[j],0,pd0[,j]);
   pd0[,j]=choose(startt>tj_1[j] & pevent[,j]=1,startt-
      tj_1[j],pd0[,j]);
end;
totevent=pevent[+,];
totpd0=pd0[+,];
x0= totevent/totpd0 || {0};

* **** termination criterion *****;
tc=j(1,13,0);
tc[1,1]=12000;
tc[1,2]=10000;
tc[1,3]=.;
tc[1,9]=1e-10;

/* constraint */
conl=j(1,nparm,-20);
conu=j(1,nparm,.);
con=conl//conu;

/* call the optimization routine */
call nlpqn(rc,rx,"LL",x0,optn,con,tc,,,"GRAD");
optx=rx;

call nlpfdd(f, g, h, "ll", rx,,"grad");
negh=-1*h;
help=det(negh);
if help ^= 0 then do;
      vmat=inv(negh);
      seb=sqrt(vmat[nparm,nparm]);
      ests=optx[1,nparm] || seb;
end;
if help = 0 then do;
      seb={.};
      ests=optx[1,nparm] || seb;
end;

finish FITPE;
```

```
start PCT5;
use pcent5;
read all var{cutpt rep};
nint=nrow(cutpt);
total=rep[+];
cutpt=cutpt // {364};
rep[nint]=rep[nint]+(4-total);
tmp=cutpt;
do j = 1 to nint;
   do k = 1 to rep[j];
      new= cutpt[j] + k*(cutpt[j+1]-cutpt[j])/rep[j];
        tmp=union(tmp,new);
   end;
end;
tj=tmp`;
tj_1 = {0} // tj[1:4];
w = tj - tj_1;
tmp2=tmp[1:4];
output= {1} || tmp2`;
varnames={z p20 p40 p60 p80};
create tmp5 from output [colname=varnames];
append from output;
close tmp5;

finish PCT5;

/* loglikelihood function */
 start LL(theta) global(xx,A,B,resp,i1,j1,nparm,nint,nobs);
   LPIij=i1*theta[1,1:nint];
   LPXij=xx*theta[1,nparm];
   tmp=LPXij*j1;
   LPij=LPIij + tmp;
   sbrate=(-exp(LPij)#B)[,+];
   mmm=choose(1-exp(sbrate)=0,1,1-exp(sbrate));
   respbit=(resp # log(mmm))[+];
   abit=(-exp(LPij) # A )[+];
   f= abit+respbit;
   return(f);
 finish LL;

/* gradient vector */
 start GRAD(theta) global(xx,A,B,resp,i1,j1,nparm,nint,nobs);
   LPIij=i1*theta[1,1:nint];
   LPXij=xx*theta[1,nparm];
   tmp=LPXij*j1;
   tmp2=xx*j1;
   LPij=LPIij + tmp;
   sbrate=(-exp(LPij)#B)[,+];
   abit= -exp(LPij) # A ;
   bbit= (exp(sbrate)*j1) # exp(LPij) # B;
   cbit= choose(1-exp(sbrate)=0,1,1-exp(sbrate));
   Gij=abit + bbit/(cbit*j1);
```

```
    intgrad=Gij[+,];
    lpgrad=(tmp2 # Gij)[+];
    g= intgrad || lpgrad;
    return(g);
  finish GRAD;

  %* -ve of hessian matrix *;
  start HESS(theta) global(xx,A,B,resp,i1,j1,nparm,nint,nobs);
    LPIij=i1*theta[1,1:nint];
    LPXij=xx*theta[1,nparm];
    tmp=LPXij*j1;
    tmp2=xx*j1;
    LPij=LPIij + tmp;
    sbrate=(-exp(LPij)#B)[,+];
    abit= -exp(LPij) # A ;
    bbit= (exp(sbrate)*j1) # exp(LPij) # B;
    cbit= choose(1-exp(sbrate)=0,1,1-exp(sbrate));
    cbit2=(cbit##2 )*j1;
    dbit= j(nobs,nint,1) - (exp(LPij)#B) - (exp(sbrate)*j1);
    Hij=abit + (bbit#dbit)/cbit2;
    int=Hij[+,];
    lp=  ((tmp2##2) # Hij)[+];
    intlp=(tmp2 # Hij)[+,];
    h=diag( int || lp );
    h[nparm,1:nint]=intlp;
    h[1:nint,nparm]=intlp`;
    return(h);
  finish HESS;

/* MAIN PROGRAM */

use DATASET;
read all var{left right x event};
call pct5;
call fitpe(left,right,x,event);

quit;
```

APPENDIX 4.D: SAS program to fit the APE model

```
data DATASET; set DATASET;
if event=0 then right=.; z=1;
run;
data use5; merge DATASET tmp5; by z; run;

data use5;set use5;
keep left right t  event x id simnum i li ri wi;
i=1; li=0; ri=p20; wi=p20; output;
i=2; li=p20; ri=p40; wi=p40-p20; output;
i=3; li=p40; ri=p60; wi=p60-p40; output;
```

```
i=4; li=p60; ri=p80; wi=p80-p60; output;
i=5; li=p80; ri=52*7; wi=52*7-p80; output;
run;

%MACRO FITMP(data=,est=);

data &data; set &data;
if (left>=ri) then pd=wi;
if (left>li & left<=ri) then pd=left-li;
eflag=0;
if ( (pd=. or pd<wi) and right>li) then eflag=1;
if eflag=1 then atrisk=wi;
if eflag=1 & pd~=. then atrisk=wi-pd;
if (right>li & right<=ri) then atrisk=right-li;
if (left>=li & right<=ri) then atrisk=right-left;
if atrisk=. then atrisk=0;
cens=1-event;
if i ne 1 then cens=0;
run;
proc means data=&data noprint;
   var atrisk; class id;
   output out=tmp sum=tatrisk ;
run;
proc sort data=tmp; by id; run;
proc sort data=&data; by id descending i; run;
data &data; merge tmp &data; by id; where id ne .; run;
data &data; retain sum; set &data;
   by id descending i;
   if first.id then sum=0;
   sum=sum+atrisk;
run;
data &data; set &data;
   if eflag=1 then e=atrisk/tatrisk;
   else e=0;
   if pd=. then pd=0;
   pdbit=e*(sum-0.5*atrisk);
   totpd=pd+pdbit;
   if totpd=0 then delete;
   ltotpd=log(totpd);
run;

proc genmod data=&data;
class i;
model e = i x / noint dist=p offset=ltotpd;
make 'parameterestimates' out=parm ;
run;

%MEND;

%fitmp(data=use5,est=est5);
```