uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Li Chen

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Epidemiology)

GRADE / DEGREE

Department of Epidemiology and Community Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

A Comparison of Methods in the Presence of Censored Cost Data Under Different Censoring Mechanisms

TITRE DE LA THÈSE / TITLE OF THESIS

G. Wells

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

D. Coyle

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

A. Dabrowski

R. Nair

Gary W. Slater

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

A COMPARISON OF METHODS IN THE PRESENCE OF CENSORED COST DATA UNDER DIFFERENT CENSORING MECHANISMS

LI CHEN

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the MSc degree in Epidemiology (Biostatistics Specialization)

Department of Epidemiology and Community Medicine
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 0-494-11235-2
Our file Notre référence
ISBN: 0-494-11235-2

ABSTRACT

Several approaches have recently been proposed in order to derive an accurate estimate of mean costs given censoring. The aim of this study was to compare methods for estimating mean costs given censoring across different censoring mechanisms and censoring levels. 736 "complete" cases from the CHART study were used to form a "complete" set where the mean cost was known. This "complete" cohort was used to generate simulated data sets. The accuracy of methods was measured by comparing the difference between estimates and the "true" cost. The Uncensored cases method, Cox's PH model and the Weighted method CHU consistently gave better estimates of mean costs across different censoring mechanisms and censoring levels. Estimates of mean costs from all methods deteriorated as the censoring level increased. The Uncensored cases method, Cox's PH model and the Weighted method CHU may be appropriate methods for estimating mean costs given censoring in short-term studies.

ACKNOWLEDGEMENTS

I want to sincerely thank my supervisor, Dr. George Wells for providing guidance in statistical matters and encouraging advice. You have been a source of inspiration not only for this thesis, but for all my life.

I would also like to thank my supervisor, Dr. Douglas Coyle for his availability, his help and useful discussion. Thank you for all you taught me.

I am also very grateful to my husband William Guo for his love and support to my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

There is increasing necessity to establish cost-effectiveness as well as clinical effectiveness when evaluating health care interventions (1). In clinical trials, therefore, there has been a corresponding move towards the collection of resource use data from which the total costs associated with the alternative treatments can be estimated (2, 3, 4). Proper analysis of treatment cost data may lead to the adoption of more cost-effective therapies, resulting in potential savings in resource utilization and incentives to develop a more standardized and efficient medical intervention. For instance, when evaluating the cost-effectiveness of a cancer screening program, the potential savings in treatment costs due to earlier diagnosis through screening are of interest. As another example, comparisons of the average costs associated with the alternative therapies may lead to substantial cost reduction. Clinical trial data as well as administrative databases from medical centers, disease registries and insurance companies present excellent opportunities to evaluate the cost of medical care in contrast to the clinical benefits.

Despite the tremendous interest in the analysis of medical costs, there remains a host of challenging issues, both practical and methodological. From a practical point of view, the availability of and access to reliable and representative data are critical. Although a number of resources are now available to researchers, each is subject to limitations. In the case of administrative databases, these include poor sensitivity of diagnosis codes (5), lack of direct information about comorbidity, and other confounding factors (6), and the potential lack of representative-ness from one data set to other populations (7). An additional difficulty is the finite nature of follow-up, which leads to the incompleteness of the available data. In long-term clinical studies to collect cost data, it is inevitable that

some patients are not followed until the endpoint of interest so that their medical costs are not fully observed. This phenomenon is referred to as censoring, which is well known for survival time data. Censored economic data are common, because the obstacles to the collection of economic data still are substantial. The censoring of cost histories poses two important problems for the analysis of long-term cumulative costs. First, analyses are difficult to perform when the interval of observation needed for the study exceeds the length of the follow-up period in the data for most or all of the study subjects. Second, the analysis of censored data is not straightforward. Cost data are difficult to analyze by standard means because of their typically nonstandard statistical distribution. Medical cost distributions often exhibit a mass at zero representing nonusers of medical resources, and relatively small numbers of extremely high users which tend to highly skew nonzero costs (8). These features of medical cost data have been recognized for some time, but historically, the methods developed to deal with them were not designed to accommodate censored cost data.

Until recently, the mean total cost for a group of patients has commonly been estimated by the sample mean of the observed costs from all study subjects or from only the uncensored cases. The former method, to be referred to as the Available sample method is always biased downward because the costs incurred after censoring times are not accounted for. The latter method, called the Uncensored cases method, is also always biased toward the costs of the patients with shorter survival times, because longer survival times are more likely to be censored. It is now a well-recognized fact that censoring occurs in costs as well as survival time data. Statistical methods for handling censoring in survival time data have been well developed, such as the Kaplan-Meier

method or Cox's proportional hazard regression model to adjust for censored data within context of the analysis. However, these methods may result in substantial bias to the problem of cost evaluation under censoring, because the assumption of independence between the censoring mechanism and costs is violated (9, 10). With the increasing prominence of medical cost as an outcome, several methods have been developed to adjust for censored cost data in order to derive a more accurate estimate of mean costs over recent years (10, 11, 12, 13). Some methods make use of the detailed patient cost histories recorded throughout studies and the remainder makes use of each patient's total study costs. Several studies have shown that these methods produced an unbiased estimate of mean total study costs providing that censoring occurs in a random fashion.

It is important to better understand the potential impact of censored cost data by considering the mechanisms leading to the censoring. Censoring can occur for many reasons, for instance, patients may drop out of a study due to moving out of the study area, the study may end before all patients experience the event of interest, or patients may be withdrawn from the study due to the reasons that are related to the event of interest.

The aim of this present study was to compare various methods for estimating the mean total costs given censoring. The methods were compared under three censoring mechanisms (random, end-of-study and informative censoring) respectively, and the analysis identified whether any particular method consistently performed well. In addition, various methods given different levels of censoring were also compared under each of the three censoring mechanisms. Finally, the consistency of the above two objectives was determined through the Monte Carlo simulation analysis. A cohort of

patients from the UK clinical trials comparing conventional radiotherapy and continuous hyper-fractionated accelerated radiotherapy (CHART) were used throughout this paper to illustrate the methods.

The outline of the thesis is as follows. Chapter 2 provides the background on the problem of censoring in the analysis of cost. We also explain various censoring mechanisms and give a brief summary of different methods to deal with censored cost data. We finally provide an overview of the Monte Carlo simulation research process. Chapter 3 is the core of the thesis. The reader will find twelve methods for estimating the mean total costs in the presence of censoring in detail. Description of the study design is illustrated on data from the cost-effectiveness study of CHART conducted at the ten medical centers in the UK (14). Chapter 4 presents results from simulated data sets generated from the actual CHART trial data using twelve methods described in the Chapter 3. Discussions are presented in Chapter 5 where some concluding remarks and extensions are given.

2. BACKGROUND

2.1 Issues with the economic evaluation of costs

The rising costs of health care have created serious national concern in Canada and other industrialized countries. This concern has prompted tremendous recent interest in the economic evaluation of medical care (2, 3, 4). For instance, many clinical trials include data collection relating to the costs of treatments. Similar data are routinely collected by hospitals, insurance companies and disease registries. Proper understanding of medical costs plays a crucial role in ascertaining the costs of treating patients with a particular disease, in searching for cost-effective intervention/prevention strategies, and in identifying determinants of medical costs.

There are several complications with these available data sources. First, subjects may not survive beyond the time period of interest, and survival time is related to cost accumulation. Secondly, both survival time and cost accumulation process are subject to right censoring. Censoring is caused by the limited study duration; in other studies, loss to follow-up is also a major source of censoring. Thirdly, the cost data are normally recorded in broad time intervals, such as monthly or yearly intervals, and no information is available on how the data is accumulated within an interval. Finally, the costs in different time intervals tend to be correlated. These complications pose major challenges in the statistical analysis of cost data.

The literature on evaluating the cost-effectiveness of treatments indicates that methods of assessing costs remain fairly primitive. It is critical to note how rarely the issue of costs is addressed in clinical trials. In a recent survey of randomized controlled trials literature dating from 1966 to 1988, Adams et al find that only 0.2% of clinical reports

included economic analysis (15). Furthermore, when costs are considered, a number of problems often arise: costs are not identified appropriately; they are not analyzed rigorously; and they tend to be collected separately from the clinical trials or only be collected on a convenient subset of trial patients, which hamper not only their identification, but also their interpretation and generalization (16). The findings of the Adams survey are informative. Based on a random sample of the studies in their survey which included an economic assessment, 28% aggregated costs and benefits in some form. Of those trials that planned for an economic assessment in their design (74% of the sample), 68% did not appropriately identify the costs involved. Of those which collected cost data separately or only on a convenient subset of trial patients, none correctly measured costs. Eighty-four percent of all studies attempting some economic evaluation did not conduct sensitivity analysis of the cost data.

While identification of costs is important and evaluation at time of trial helps to guarantee this takes place correctly, an important question that has not, to date, been addressed is the appropriate method for analyzing these costs. Costs often are presented as some average, such as the average cost per case or the average cost per treatment, but it often is unclear what this "average" points to. Furthermore, attempts rarely are made to relate this figure either to the underlying trial population or to the sample from which it is drawn. Accuracy and efficiency suggest there is necessity to combine the collection of cost data into a clinical trial itself and to establish an appropriate and efficient methodology through which the underlying sample is made explicitly and the data can be subject to inference. In particular, this methodology needs to address the issue of censoring, which often exists in randomized controlled trials, in relation to cost data.

2.2 The problem of censoring

Censoring types

Censoring comes in many forms and occurs for many different reasons (17). The most basic distinction is between left censoring and right censoring. An observation on a variable T is right censored if all you know about T is that it is greater than some value c. In survival analysis, T is typically the time of occurrence for some event, and cases are right censored because observation is terminated before the event occurs. This notion of censoring is not restricted to event times. For instance, if the censored variable is costs from the start of treatment until death, you may know only that costs are greater than $20,000, in which case the person's costs are right censored at $20,000. Symmetrically, left censoring, which is not common in clinical trials, occurs when all you know about an observation on a variable T is that it is less than some value c. In the context of survival data, left censoring is most likely to occur when you begin observing a sample at a time when some of the individuals may have already experienced the event. For example, in biomedical research one may know that a patient entered the hospital at a particular date and that he/she survived for a certain amount of time thereafter; however, the researcher does not know when exactly the symptoms of the disease first occurred or were diagnosed. Unless you can obtain information on the starting date of the disease for those patients, the time of disease is left censored at the admission date. Because lifetime data generally occur over chronological time, a variety of schemes are used to obtain data according to prevailing time and resource constraints. This can produce other forms of incompleteness besides right censoring and left censoring. Another well-known form of

censoring in survival analysis is interval censoring. Here the event time is never observed exactly. Only a random interval containing it is available. For instance, in many longitudinal studies on humans it is feasible to see individuals only at rather widely spaced intervals, say monthly or yearly. So the exact timing of some types of events is interval censored.

Censoring mechanisms

The appropriate strategy for analyzing particular censored costs will be dependent on the mechanism that leads to censoring (18). There are many reasons why censoring might occur and these may be classified into different types of censoring mechanisms (19).

- Random censoring

Random censoring occurs when patients are lost to follow-up for reasons that are independent of the event of interest. For example, patients may drop out of a study due to moving out of the study area.

- End-of-study censoring

In most clinical studies, patients are not recruited into a study at the same time point but over a period of time. However, studies may end at a pre-specified time point. Thus it is possible that a percentage of patients may have incomplete follow-up data over a pre-specified period.

- Informative censoring

Informative censoring occurs when patients drop out of a study due to the reasons that are related to the event of interest. For instance, patients may be withdrawn from a study because their condition improves and this improvement is related to the treatment they

are receiving. Informative censoring may be neglected because it is not always possible to determine the reason a patient has dropped out of a study.

In both the survival time and cost accumulation process, right censoring is far more common than left censoring and interval censoring (17). In human clinical trials, censored data typically arise when the course of treatment extends beyond the end of the trial period. A common example of this is when treatment is administered to chronically ill patients up to the time of their death, an event likely to postdate the end of the trial. Censoring also is likely to occur when the trial design is such that sequential admission implies variable periods of follow-up, and when patients are likely to be withdrawn from the trial for administrative reasons unconnected with treatment objectives. Besides, patients may refuse to continue the trial due to the improvement or deterioration, so informative censoring arises. Due to the prominence of the right censoring in cost evaluation, most of the existing methods for estimating the mean total study costs are developed to adjust for right censoring. Thus, in this present study we aimed to assess the performance of different methods for estimating the mean total costs in the presence of right censoring (random, end-of-study and informative censoring).

2.3 Estimating mean total costs in the presence of censoring

Censoring is an issue in estimating the mean total cost for treating a particular disease (or similarly the cost until cure, or the cost in a 12-month period). Most often, the complete costs of some patients are not fully observed because the patients are lost to follow-up or because they are still alive (or not cured or discharged or have not been enrolled for 12 months) at the time of data analysis. In the terminology of survival

analysis, the lifetime costs for such patients are defined as censored. Mean cost is frequently estimated using the average total costs from all study patients, even though some of these costs are incomplete, or from only the uncensored patients. The estimated costs will be too low if one analyzes all the available cost data. Alternatively, calculating costs for only the uncensored patients is likely to give too much emphasis to patients who died early, because people who survive for a longer time are more likely to be censored. Therefore each method leads to false inference.

Statistical methods for handling censoring in survival data have been well developed. Due to the similarity between censored medical costs and censored survival times, some investigators have applied standard survival analysis methods such as the Kaplan-Meier method, and Cox's proportional hazard regression model to the problem of cost evaluation under censoring, by analyzing censored costs as though they were censored survival times (20, 21, 22, 23). However, Lin et al. (10) indicate that this approach is invalid because the cumulative cost at censoring time is positively correlated with the cumulative cost at the endpoint of interest, even if the underlying censoring mechanism is independent. This phenomenon of dependent censoring is caused by the fact that the patients are heterogeneous such that those who accumulate costs at higher rates over time tend to generate higher cumulative costs at all time points as compared with those with lower accumulation rates.

To overcome the problem of independent censoring, Lipscomb et al. (24) advocate the Stratified Cox model for estimating the mean total costs. The study period is partitioned into smaller time intervals (strata), then a Cox's proportional hazard model is fitted within each strata. The censoring mechanism differs for this method as patient costs are

censored within strata if they are incomplete for the full strata, reasons for censoring include death or incomplete cost collection due to lost to follow-up, patients with complete costs within each strata are not censored. The estimated mean total costs are then summed across strata.

To minimize the bias induced by censoring, Lin et al (10) propose partitioning the entire time period of interest into a number of small intervals and then estimating the mean total costs either by the sum of the Kaplan-Meier estimator for the probability of surviving to the start of each interval multiplied by an appropriate estimator for the average cost over the interval conditional on surviving to the start of the interval or by the sum of the Kaplan-Meier estimator for the probability of dying in each time interval multiplied by the sample mean of the total costs from those who are observed to die in that interval.

The other method proposed by Zhao and Tsiatis (25) for estimating the quality-adjusted survival function, can be further used to estimate the mean costs. This method computes the weighted sum of the number of patients with costs exceeding a specific value y, where the weights are inversely proportional to the probability of non-censoring with respect to y. More efficient estimators can make use of the cost histories for the censored observations.

With increasing demand to evaluate the health care costs, studies of the costs of illness and the factors affecting the costs will be important to clinicians and policy-makers. Several studies have indicated that patients characteristics (such as age), as well as more specific measures of the clinical severity of a particular illness affect clinical outcome (26, 27). It is reasonable hypothesis that these factors may also affect the cost of treating a

patient with a particular illness. Thus, the regression methodology would be particularly valuable in identifying cost-effective intervention/prevention programs. Lin (11) develops the further methods using linear regression analysis to estimate the mean total costs to adjust for censored cost data. The familiar normal equations for the least squares estimation are modified in several ways to properly account for the incompleteness of the data. A weighted linear regression model is fitted to the total study costs of patients with complete follow-up, and also adjusted for the factors known to influence costs. The weighting is calculated by the inverse probability of the Kaplan-Meier survival estimate, where the censoring indicator is reversed. More efficient estimators are available when the cost data are recorded in multiple time intervals.

The other regression-based method for estimating the mean treatment costs in the presence of right-censoring is proposed by Carides (12). This estimator for mean costs exploits the underlying relationship between total treatment costs and survival time. The proposed method utilizes either parametric or nonparametric regression to estimate this relationship and is consistent when this relationship is consistently estimated.

2.4 Overview of Monte Carlo simulation

What is a Monte Carlo study? According to the definition on Webster's dictionary, Monte Carlo relates to or involves "the use of random sampling techniques and often the use of computer simulation to obtain approximate solutions to mathematical or physical problems especially in terms of a range of values, each of which has a calculated probability of being the solution" (28). This definition provides a brief and accurate description for Monte Carlo studies. For example, suppose that we are interested in

finding what are the chances of obtaining three as the sum from rolling a die twice (assuming a fair die). There are basically three ways of obtaining an answer to our question. The first is to do it the hard way by rolling a die twice tens of thousands of times so that you could reasonably estimate the chances of obtaining three as the sum of rolling a die twice. Alternatively, you could estimate the chances for this event by relying on theoretical probability theory. Instead of these two methods, we can also take an empirical approach to obtain the answer to our question without physically rolling a die. This approach is called the Monte Carlo simulation (MCS) in which the outcomes of rolling a die twice are simulated, rather than physically rolling a die twice. Monte Carlo simulation offers researchers an alternative to the theoretical approach (29). There are many situations where the theoretical approach is difficult to implement, much less to find an exact solution. MCS approach is only possible with a computer and some appropriate software, such as EXCEL. As a matter of fact, with computing power becoming increasingly cheap and with powerful computers more widely available than ever, this computing-intensive approach is becoming more and more popular with researchers. In short, MCS simulates the sampling process from a defined population repeatedly by using a computer instead of physically drawing multiple samples to estimate the sampling distributions of the events of interest. This approach can be applied to a variety of situations in different disciplines, especially in economic cost evaluation.

Why is Monte Carlo simulation needed or necessary? There are many situations where MCS is needed, or where MCS is the only feasible approach to providing analytic solutions to some quantitative questions (29). First, as we know, the validity of any statistical theory is typically based on some theoretical assumptions. However, when the

assumptions of a theory are violated in the data that we have, the validity of the estimates about certain sampling distribution characteristics based on the theory is often compromised and uncertain; consequently, we often ignore how much we can trust the theoretical estimates, or about how erroneous our conclusion might be if we blindly rely on the theory, even if some assumptions of the theory have been violated. In this situation, MCS becomes very useful to researchers, because this approach relies on empirical estimation of sampling distribution characteristics, rather than on theoretical expectations of those characteristics. With a large number of replications, the empirical results should asymptotically approach the theoretical results, and this can be testified when the theoretical results can be obtained. In addition, there are some other situations where statistical theories are either so weak that they can not be fully relied upon or statistical theories simply do not exist. In these situations, MCS may be the only feasible approach to giving answers to a variety of questions.

There are some typical situations where a Monte Carlo study is needed (29). First, MCS can be used to assess the consequence of assumption violations. As we know, most popular statistical techniques are parametric statistics which need assumptions about the distribution of the data. If the assumptions are not maintained, the validity of the results produced from applying these techniques may be in question. However, statistical theory itself does not usually provide any indication about what the reality would be if the conditions are not met by the data. Again, MCS becomes, in many situations, the only feasible way to obtain answers to these questions. Secondly, MCS can help to estimate the sampling distribution of a statistic that has no theoretical distribution. Due to the complexity of a particular statistic, a theoretical sampling distribution of the statistic may

not be available in some situations. In such situations, if we are interested in how the statistic will vary from sample to sample, that is, the sampling distribution of the statistic, MCS becomes one realistic approach to obtaining such information.

Why we select the EXCEL system for conducting Monte Carlo studies in this present study? Because MCS is computation-intensive, it is obvious that MCS research typically requires programming capabilities. Furthermore, because many MC studies involve some type of statistical techniques and/or mathematical functions, statistical/mathematical capabilities are also required. The EXCEL system has the combination of a powerful variety of built-in statistical procedures, mathematical functions, and the versatile programming capabilities associated with Visual Basic. This combination makes the EXCEL system ideal for conducting the Monte Carlo simulation analysis in our study, especially the study related to statistical techniques. Such a combination of built-in statistical procedures and versatile programming capabilities makes it much more convenient for MCS researchers. Without such a combination of statistical capabilities and programming capabilities within the same system, we may have to deal with different systems, and consequently worry about the interface among different systems.

2.5 Objectives

There were three related objectives within this study:

The primary objective was to compare various methods for estimating the mean total costs given censoring. The methods were compared under each of the three censoring mechanisms (random, end of study and informative censoring) and analysis identified whether any particular method consistently performed well.

The secondary objective was to compare various methods given different levels of censoring. The methods again were compared under each of the three censoring mechanisms.

The third objective was to determine the consistency of the above two objectives through the Monte Carlo simulation analysis.

3. METHODOLOGY

3.1 Methods for estimating mean total costs in the presence of censoring

3.1.1 Available sample method

The computationally simplest strategy is the Available sample method, in which the analysis ignores censoring and calculates the sample mean of the observed costs from all study patients. The mean total costs will be biased downward because of failure to account for costs incurred after censoring. However, this method is still a frequently used method in cost-effectiveness studies. The estimator of mean total costs is given as follow,

$$\overline{C} = \sum_{i=1}^{N} C_i / N \qquad i = 1,\ldots,N . \tag{3.1}$$

3.1.2 Uncensored cases method

This method discards data from censored cases and estimates the mean total costs from complete cases only. This method has obvious advantages. It is very easy to describe, and since the data structure is as would have resulted from a complete experiment, standard statistical software can be used. However, the method suffers from severe drawbacks. First, the mean costs will be biased toward the costs of the patients with shorter survival times, because longer survival times are more likely to be censored. Second, there is nearly always substantial loss of information. Moreover, investigators generally wish to make inferences about the entire target population, and deleting patients may restrict generalizablity. The estimator of mean total costs with only complete cases can be written as follow,

$$\overline{C} = \sum_{i=1}^{n} C_i / n \qquad i = 1,\ldots n . \tag{3.2}$$

3.1.3 Kaplan-Meier cost method

Fenn et al. (21) propose a novel survival analysis technique to the problem of cost evaluation by thinking of cumulative costs as a measure of "time". They have applied the standard Kaplan-Meier (Kaplan-Meier, 1958) survival analysis technique to the cost estimation by treating costs as potentially right-censored survival times (i.e., attaching the censoring indicator to the observed total costs). Mean cost is estimated by calculating the area under the Kaplan-Meier survival curve, using a cost rather than time scale. This method, however, is biased toward inflated cost estimates, because the inherent patient heterogeneity with respect to cost accumulation implies that the total cost at the survival time is positively correlated with the total cost at the censoring time. The assumption of independence between the censoring mechanism and costs has been violated (9, 10).

3.1.3.1 Estimator

If there is no censoring, the average time of survival time T is simply given as the sum of the survival times, saying, T_i, divided by the number of patients N.

$$\overline{T} = \sum_{i=1}^{N} T_i / N \tag{3.3}$$

If the entire time period is partitioned into a number of equal and small intervals, saying Δ_{t0}, Δ_{t1} and $\Delta_{t2}, \ldots$, then each T_i can be denoted as $T_i = \Delta_{t0} + \Delta_{t1} + \cdots + \Delta_{t[T_i/\Delta]}$. So the equation (3.3) can be rewritten as follow,

$$\overline{T} = \sum_{i=0}^{\infty} S_i \Delta_{ti} \tag{3.4}$$

where S_i is the proportion of patients surviving to the time interval i.

Similarly, the mean cost of survival is shown as the sum of the costs, saying, C_i, divided by the number of patients N.

$$\overline{C} = \sum_{i=1}^{N} C_i / N \tag{3.5}$$

The equation (3.5) can be rewritten in one way similar to that just done for the average time of survival.

$$\overline{C} = \sum_{i=0}^{\infty} S_i \overline{C}_i \tag{3.6}$$

where S_i is the proportion of patients who have accrued costs to the cost interval i and $\overline{C}_i$ is the mean cost that the patients who experienced the time interval i incurred during that time interval.

For the censored data, an unbiased estimator $\hat{T}$ of $\overline{T}$ can be calculated if we can find an unbiased estimator $\hat{S}_i$ of S_i. The Kaplan-Meier survival method can be used to calculate an unbiased estimator $\hat{S}_i$, under the assumption that the patients censored have the same probability of survival as the patients who are not censored; that is, "censoring is independent of failure." Therefore, the Kaplan-Meier method can also be applied to cost data to get an unbiased estimator $\hat{S}_i$ of S_i, and so an unbiased estimate of $\overline{C}$. To obtain an unbiased estimator of $\overline{C}$, we would also require an unbiased estimate of $\overline{C}_i$. If the mean costs in the Δ_{ti} of those who are censored before Δ_{ti} is the same as the mean costs of those observed in Δ_{ti}, then the observed mean costs will be an unbiased estimator. That is, if the assumptions of "censoring is independent of costs" and

"censoring is independent of failure" are maintained, then the equation (3.6) does give an unbiased estimator of mean costs.

3.1.3.2 Application

In practice, the problem is that the low-cost patients, those who accrue a small cost per unit of time, will be censored sooner than the high-cost patients on the cost scale. Because low-cost patients are censored sooner, the estimator proposed by Fenn et al. (21) will overestimate the true mean costs. Because the assumption of independence between the censoring mechanism and costs has been violated, the Kaplan-Meier cost method always results in bias. However, this approach does provide a basis for using more sophisticated survival techniques to estimate the economic cost data. It is also reasonable to further suppose that the same technique such as Cox's proportional hazard regression model could be used to investigate the relationship between various factors and costs.

3.1.4 Cox's proportional hazard regression model

A natural extension to Kaplan-Meier survival analysis technique is to apply Cox's proportional hazard regression model to the analysis of costs (20). The survival time is replaced by the costs of treatment in Cox's model as the dependent variable, and clinical factors can be analyzed as predictors of costs within a regression analysis framework. Particularly, Cox's proportional hazard model needs no assumptions about the distribution of the dependent variable (costs) and the residuals from the regression. The influence of extreme values is avoided by using the ranks of the dependent variable in Cox's model. These characteristics of Cox's model offer the particular advantages in the

analysis of cost data. Except for the independence assumption between the censoring mechanism and costs, the proportional hazards between different levels of fitted factors should be checked in order to obtain unbiased estimates.

3.1.4.1 Estimator

A Cox's semi-parametric proportional hazards model is shown as,

$$S(y) = [S_0(y)]^{\exp(\sum a_i x_i)} \tag{3.7}$$

where $S_0(y)$ is the underlying survivorship function. And $\sum a_i x_i$ is the effect of a linear combination of independent variables, where a_i are regression coefficients and $\{x_i\}$ are a set of independent variables. Thus the predicted mean cost is then calculated by numerical integration of the area under the estimated survival curve for Cox's model.

$$\hat{y} = \int_0^\infty S_0(y)^{\exp(\sum a_i x_i)} dy \tag{3.8}$$

3.1.4.2 Application

Cox's proportional hazard regression technique appears to be particularly helpful in the analysis of skewed, censored data often encountered in the evaluation of hospital costs, due to the few assumptions made concerning the distribution of costs. And Cox's model depend only on the ranks of the costs, so the influence of extreme values is reduced. Cox's model also has a potential ability to handle censored observations which often occur in longitudinal studies. Ability to handle censored data may be useful in the analysis of such longitudinal data. Furthermore, accurate predictions of the costs of care based on clinical characteristics have the potential to assist clinicians in their attempts to

provide cost-effective care to patients, as well as facilitating a fair method of adjusting payments to providers based on clinical indicators of severity of illness.

3.1.5 Stratified Cox model

To overcome the problem of independent censoring, Lipscomb et al. (24) advocate the Stratified Cox model for estimating the mean total costs, where there is no assumption about independent censoring. The study period is divided into smaller time intervals (strata), then a Cox's proportional hazard model is fitted within each strata. The censoring mechanism differs for this method as patient costs are censored within strata if they are incomplete for the full strata, reasons for censoring include death or incomplete cost collection due to patient censoring, patients with complete costs within each strata are not censored. The estimated costs are then summed over strata to obtain mean total study costs.

3.1.5.1 Estimator

The cost model is denoted as a linear function determining the impacts of explanatory variables on costs and thus allowing conditional predictions of costs:

$$G(\beta, X_{it}) = \beta_0 + \sum_j \beta_j X_{ijt} \tag{3.9}$$

where X_{it} is a vector of explanatory variables for patient i in interval t whose j th element is X_{ijt}, and β is a vector of regression coefficients whose j th element is β_j. Because it makes no specific assumption about an error term distribution, Cox's proportional hazard model is particularly useful for modeling costs whose probability distributions may be complex. The time period for cost analysis is divided into mutually

exclusive smaller intervals. Then we adopt the following stratified variant of Cox's model:

$$\Pr(C_{it} \geq c | t, X_{it}) = K_t(c)^{\exp[G(\beta, X_{it})]} \tag{3.10}$$

where $K_t(c)$ is the baseline cumulative cost function defined for each interval t, describing the level of cost expected when all explanatory variables are set to zero. Then $K_t(c) = 1 - F_t(c)$, where $F_t(c)$ is the baseline cumulative distribution function of costs for interval t. Thus, the equation (3.10) is analogous to the stratified survival function of Cox's model. The predicted mean cost for person i in interval t is then calculated as

$$\hat{C}_{it} = \int_0^\infty K_t(c)^{\exp[G(\hat{\beta}, X_{it})]} dc \tag{3.11}$$

For computational purposes, we define that $K_t(c) = \exp[-\Lambda_t(c)]$, where $\Lambda_t(c)$ is the baseline cumulative hazard function.

3.1.5.2 Application

A well-known and attractive feature of Cox's model is the flexibility provided by the fact that no assumptions about the baseline distribution of costs are necessary. In the Stratified Cox model for estimating the mean total costs, we fit a proportional-hazard model within each stratum, so that the random variable of interest is the total costs incurred within a pre-specified time interval. In this way, we overcome the "one-to-one" link between the flow of costs and the flow of time, and so there is no assumption about independent censoring. Thus we believe that an analog of Cox's proportional hazard regression for relating costs to factors, such as stratifying by fixed time intervals, or by

other phenomena whose operational definitions do not rely on the proportional-hazards event of interest, would be an emerging area for future research.

3.1.6 Lin's method where cost histories are known (CHK)

To minimize the bias induced by using the simple methods and the survival analysis methods, Lin et al. (10) propose two approaches to estimate mean total costs in the presence of censoring. The first approach is called Lin's method where cost histories are known (CHK). This method partitions the entire time period of interest into a number of equally spaced small intervals, and then estimates mean total costs by the sum of the Kaplan-Meier estimator for the probability of surviving to the start of each interval multiplied by an appropriate estimator for mean costs over the interval conditional on surviving to the start of the interval. This approach makes use of the cost information.

3.1.6.1 Estimator

Notations and assumptions

Let the random variable C denote the total costs for a patient during the specified time period $[0,\tau)$. Let T be the survival time which is assumed continuous, and U be the censoring time which is assumed to be either continuous or discrete. If $T < \tau$, then C becomes the total cost up to T. The aim of this approach is to derive an estimate of mean total costs $\mu = E(C)$ where E denotes expectation. Because no patient is followed beyond τ, we exclude the costs incurred after τ in the definition of mean total costs.

The entire time period $[0,\tau)$ is partitioned into K subintervals $[a_k, a_{k+1})$ $(k = 1,\cdots,K)$, where $a_1 = 0, a_{K+1} = \tau$. If a patient can survive to the start of the interval a_k, a patient can

accrue costs over $[a_k, a_{k+1})$. The random variable C_k is denoted as the costs incurred during interval $[a_k, a_{k+1})$.

The observables from a study in the presence of censoring are $X = \min(T, U)$, i.e. the last contact date; $\delta = I(T \leq U)$, where $I(\cdot)$ is the indicator function taking the value of 1 if the observation is uncensored and zero otherwise. The observable data for n patients are then the independent and identically distributed random vectors $(X, \delta, \widetilde{C})$, where $\widetilde{C}$ is the observed total costs accrued from the start of the follow-up to the last contact date X. If cost histories are collected, $\widetilde{C}$ is divided as $(\widetilde{C}_1, \cdots, \widetilde{C}_k)$, where $\widetilde{C}_k$ is the observed costs during $[a_k, a_{k+1})$. The subscript i identifies a patient. If $\delta_i = 1$ or $X_i = \tau$, $\widetilde{C}_i = C_i$. If $X_i < a_k$, $\widetilde{C}_{ki}$ is either zero or missing dependent on whether $\delta_i = 1$ or 0. If $X_i \geq a_k$, $\widetilde{C}_{ki} = C_{ki}$ if the i th patient is uncensored before a_{k+1} and $\widetilde{C}_{ki}$ equals the costs accrued over $[a_k, U_i)$ otherwise. $\widetilde{C}_{ki}$ can be missing before X_i in a completely random fashion. Except for the independence assumption between time to failure and censoring time, we also need to ensure that patients are not censored because they accrue unusually high or low costs at any follow-up time t.

Mathematical description

From $C = \sum_{k=1}^{K} C_k$, we can get that $\mu = \sum_{k=1}^{K} E(C_k)$. According to the conditional expectation law, μ can be rewritten as

$$\mu = \sum_{k=1}^{K} E\{E(C_k | T \geq a_k)\} = \sum_{k=1}^{K} \Pr(T \geq a_k) E(C_k | T \geq a_k).$$

Therefore, μ can be given as

$$\mu = \sum_{k=1}^{K} S_k E_k \tag{3.12}$$

where $S_k = \Pr(T \geq a_k)$ and $E_k = E(C_k | T \geq a_k)$. S_k are the probabilities of surviving to a_k, and are consistently estimated by the Kaplan-Meier method as

$$\hat{S}_k = \prod_{j;t_j<a_k} \frac{n_j - d_j}{n_j} \tag{3.13}$$

where $t_1 < t_2 < \cdots < t_j$ are the observed survival times, and d_j counts the number of observed deaths at t_j, and n_j counts the number of patients under risk at t_j (j=1,…,J).

Based on the assumption of independent censoring, it implies that $E_k = E(C_k | T \geq a_k) = E(C_k | X \geq a_k)$. That is, E_k is an estimator for mean costs in interval $[a_k, a_{k+1})$ and is derived from those who are under observation at the start of the interval. Thus, E_k's are estimated by

$$\hat{E}_k = \frac{\sum_{i=1}^{n} Y_{ki} \tilde{C}_{ki}}{\sum_{i=1}^{n} Y_{ki}}, k = 1, \cdots, K \tag{3.14}$$

where Y_{ki} are defined as $Y_{ki} = I(X_i \geq a_k)$ taking the value of 1 if $\tilde{C}_{ki}$ is included in the estimation of $\hat{E}_k$ and zero otherwise. Note that $\hat{E}_k$ is a consistent estimator for E_k, if censoring occurs only at the boundaries of the intervals, and also consistent for small time interval regardless of the censoring pattern.

Therefore, μ can be estimated by $\hat{\mu} = \sum_{k=1}^{K} \hat{S}_k \hat{E}_k$.

For large samples, the estimator $\hat{\mu}$ is shown to be approximately normal and its variance estimator $\hat{V}$ is derived using the martingale version of the central limit theorem (30, 31, 32) and is given as

$$\hat{V} = \sum_{i=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{K} W_{ki} W_{li}$$

where

$$W_{ki} = \frac{\hat{S}_k Y_{ki} (\tilde{C}_{ki} - \hat{E}_k)}{\sum_{j=1}^{n} Y_{ki}} - \hat{S}_k \hat{E}_k \left\{ \frac{I(X_i \le a_k)\delta_i}{R_i} - \sum_{j:X_j \le \min(a_k, X_i)} \frac{\delta_j}{R_j^2} \right\}$$

$$R_i = \sum_{l=1}^{n} I(X_l \ge X_i).$$

3.1.6.2 Application

It is obvious that $\hat{E}_k$ is an unbiased estimator of E_k, when censoring occurs only at the boundaries of the interval. This censoring pattern can occur when patients enter the study at discrete time points and are withdrawn from the study prematurely at a limited set of time points. Therefore, the bias of $\hat{E}_k$ depends on the amount and timing of censoring and diminishes as the intervals shrink. This method requires that the cost histories be recorded on patients. If the cost histories are available, then this approach is usually preferable, especially when there is heavy censoring and / or truncation, as it makes fuller use of the cost information and requires smaller sample sizes.

3.1.7 Lin's method where cost histories are unknown (CHU)

The other approach proposed by Lin et al. (10) is called Lin's method where cost histories are unknown (CHU). Similarly, the entire time period of interest is divided into

a number of equally spaced small intervals. Then the mean total cost can be estimated by the sum of the Kaplan-Meier estimator for the probability of dying in each time interval multiplied by the sample mean of total costs from those who are observed to die in that interval. The costs for patients alive at the end of the study are included in mean total costs for the final interval. This approach does not require the detailed information on the patient cost histories, as only the patient total study costs are required.

3.1.7.1 Estimator

According to the argument of conditional expectation,

$\mu = \sum_{k=1}^{K} E(C|a_k \le T < a_{k+1})\Pr(a_k \le T < a_{k+1}) + E(C|T \ge \tau)\Pr(T \ge \tau)$ or

$\mu = \sum_{k=1}^{K+1} E(C|a_k \le T < a_{k+1})\Pr(a_k \le T < a_{k+1})$ with $a_{k+2} = \infty$. Thus,

$$\mu = \sum_{k=1}^{K+1} A_k (S_k - S_{k+1}) \tag{3.15}$$

where the survival probabilities S_k are consistently estimated by the Kaplan-Meier method with $(S_k - S_{k+1})$ being the estimated Kaplan-Meier probability of death over the interval $[a_k, a_{k+1})$, and $A_k = E(C|a_k \le T < a_{k+1})$.

Under the assumption of independent censoring,

$$A_k = E(C|a_k \le T < a_{k+1}, U \ge a_k) = E(C|X \ge a_k, T < a_{k+1}) \tag{3.16}$$

If censoring occurs only at the boundaries of the interval, then

$I(X \ge a_k, T < a_{k+1}) = I(a_k \le X < a_{k+1}, \delta = 1)$. That is, A_k is an unbiased estimator for mean total costs over the interval $[a_k, a_{k+1})$ and can be derived from the patients who are observed to die in the interval. Thus, $\hat{A}_k$ is given by

$$\hat{A}_k = \frac{\sum_{i=1}^{n} Y_{ki}\widetilde{C}_{ki}}{\sum_{i=1}^{n} Y_{ki}}, k = 1, \cdots, K \tag{3.17}$$

where $Y_{ki} = I(a_k \le X_i < a_{k+1}, \delta_i = 1)$.

Note that the estimation of interval costs $\hat{A}_k$ does not require cost information on those patients who are censored before τ, therefore need not to be recorded. The costs for those who are observed to die or whose censoring times equal to τ are allowed to be missing in a completely random fashion.

Given the $\hat{S}_k$'s and $\hat{A}_k$'s, we can estimate μ by

$$\hat{\mu} = \sum_{k=1}^{K+1} \hat{A}_k(\hat{S}_k - \hat{S}_{k+1}) \tag{3.18}$$

As long as the $\hat{A}_k$'s are consistently estimated, the estimator $\hat{\mu}$ will be consistently estimated. The consistency of the $\hat{A}_k$'s can be enhanced by selecting an appropriate cut point of the time interval so that the bias can be minimized. However, the reliable estimation of A_k requires a reasonable number of observed deaths in each interval (say 5 or more).

For large samples, the estimator $\hat{\mu}$ is shown to be approximately normal and its variance is derived using the same theoretical framework as for the previous estimator of Lin's method CHK and is given as

$$\hat{V} = \sum_{i=1}^{n} \sum_{k=1}^{K+1} \sum_{l=1}^{K+1} W_{ki} W_{li}$$

where

$$W_{ki} = \frac{(\hat{S}_k - \hat{S}_{k+1})Y_{ki}(C_i - \hat{A}_k)}{\sum_{j=1}^{n} Y_{kj}} + \hat{E}_k(\hat{S}_{k+1}D_{k+1,i} - \hat{S}_k D_{ki})$$

$$D_{ki} = \frac{I(X_i \le a_k)\delta_i}{R_i} - \sum_{j:X_j \le \min(a_k, X_i)} \frac{\delta_j}{R_j^2}, \text{ and } R_i = \sum_{l=1}^{n} I(X_l \ge X_i).$$

3.1.7.2 Application

In this approach it is unnecessary to collect detailed information on patient cost histories. It requires only the total costs at the last contact dates among those who are observed to die before the largest observed time τ or who are still alive at τ. We recommend that this method be used when censoring occurs solely at the boundaries of the intervals given that there are a few (say, 5 or more) observed deaths in each interval, but the results need to be interpreted with caution when there is heavy censoring in the interior of an interval.

3.1.8 Lin's regression estimate where cost histories are unknown (CHU)

Until now, no valid regression method was presented for evaluating the effects of covariates, such as therapies and patients characteristics on medical costs based on censored data. Therefore, the main motivation for developing regression methods is to adjust for a large number of continuous and discrete covariates. Lin (11) develops two further methods using linear regression analysis to estimate mean total costs to adjust for censored cost data. The first strategy is called Lin's regression estimate CHU. A weighted linear regression model is fitted to the total study costs of patients who die during the study period or have complete costs to the end of the study period, and also adjusted for the factors known to influence costs. The weighting is calculated by the

inverse probability of the Kaplan-Meier survival estimate, where the censoring indicator is reversed. That is, the deaths are indicated as 0 (rather than 1) and censored cases are indicated as 1 (rather than 0). Then the mean total study costs can be estimated by applying the model to the entire study cohort including the censored cases. The above idea of weighting the complete cases by their inverse probabilities of inclusion was originated by Horvitz and Thompson (33) in the context of sample surveys. The adaptation of this idea to the setting of censored survival data was initially considered by Koul et al. (34), and later on by Robins and Rotnitzky (35) and Lin and Ying (36). Recently, Zhao and Tsiatis (25) also applied this idea to the problem of quality adjusted survival time.

3.1.8.1 Estimator

Notations and assumptions

The aim in which we are interested is to derive an estimate of mean total costs $\mu = E(Y)$ over a specified period when the data is right censored, where Y denotes the total medical costs for a patient during the time period $[0,\tau]$. Generally, τ is no larger than the overall length of study. Because no patient is followed beyond τ, we exclude the costs incurred after τ in the definition of mean total costs. No patient will accumulate medical costs after death, thus the total costs incurred in $[0,\tau]$ is the same as the cumulative costs at $T^* = \min(T,\tau)$, where T is the survival time assumed continuous over $[0,\tau]$. In most cases, however, medical costs are incomplete due to right censoring. Defining therefore a potential time to censoring denoted as C, the observables from a study in the presence of censoring are $X = \min(T,C)$, $\delta = I(C \geq T)$, and $\delta^* = I(C \geq T^*)$,

where $I(\cdot)$ is the indicator function taking the value of 1 if the observation is uncensored and zero otherwise. So Y is known if and only if $\delta^* = 1$, whereas T is known if and only if $\delta = 1$. Thus a patient whose survival time is censored at or after τ has a complete observation on medical costs during $[0, \tau]$, while a patient whose survival time is censored before τ has an incomplete observation. The observable data includes n patients, with i th patient identified by the subscript i. The assumption underlying this approach is that censoring occurs in a purely random fashion so that C is independent of all other random variables. $G(t)$ is the probability that patient i has survived to T_i without being censored and is calculated by $G(t) = \Pr(C \geq t)$. The unknown function $G(t)$ is estimated by the Kaplan-Meier estimator based on the data $(X_i, 1-\delta_i)(i = 1, \cdots, n)$ as $\hat{G}(t) = \prod_{j; t_j < a_k} \frac{n_j - c_j}{n_j}$ where c_j counts the number of censored cases at t_j, and n_j counts the number of patients under risk at t_j (j=1,…, J). So the censoring indicator is reversed. It means that the deaths are indicated as 0 (rather than 1) whereas censored cases are indicated as 1.

Mathematical description

Let Y be the cumulative costs at τ or T^*, and let Z be a $p \times 1$ vector of covariates whose effects on Y are of interest. The relationship between the cumulative costs Y and covariates Z can be related through the linear regression model

$$Y = \beta' Z + \varepsilon \tag{3.19}$$

where β is a $p \times 1$ vector of unknown regression parameters, and ε is an error term with zero mean. The first component of Z should be 1 so that the first component of β corresponds to the intercept.

For uncensored data, β can be estimated by the least-squares normal equation as:

$$\sum_{i=1}^{n}(Y_i - \beta' Z_i)Z_i = 0 \tag{3.20}$$

In practice, medical costs always subject to right censoring. As noted, Y_i is known if and only if $\delta^* = 1$. Because $E\{\delta_i^* / G(T_i^*)\} = 1$, we can rewrite (3.20) as

$$\sum_{i=1}^{n}\frac{\delta_i^*}{G(T_i^*)}(Y_i - \beta' Z_i)Z_i = 0 \tag{3.21}$$

Thus, the left-hand sides of (3.20) and (3.21) have the same expectation, which is zero. Note that only the patients with complete cost data are included in (3.21), their costs are weighted inversely by their probabilities G which can be consistently estimated by the Kaplan-Meier survival estimator $\hat{G}$, where the censoring indicator is reversed. So G in (3.21) can be replaced with $\hat{G}$ as follows

$$\sum_{i=1}^{n}\frac{\delta_i^*}{\hat{G}(T_i^*)}(Y_i - \beta' Z_i)Z_i = 0 \tag{3.22}$$

Then we can get β as,

$$\hat{\beta} = \left\{\sum_{i=1}^{n}\frac{\delta_i^*}{\hat{G}(T_i^*)}Z_i^{\otimes 2}\right\}^{-1}\sum_{i=1}^{n}\frac{\delta_i^*}{\hat{G}(T_i^*)}Y_i Z_i \tag{3.23}$$

where $a^{\otimes 0} = 1, a^{\otimes 1} = a$, and $a^{\otimes 2} = aa'$. If no censoring occurs, (3.23) becomes to the ordinary least squares estimator for β.

$n^{\frac{1}{2}}(\hat{\beta}-\beta)$ is proved to converge in distribution to a p-variate zero-mean normal random vector with a covariance matrix which can be consistently estimated by $\hat{A}^{-1}\hat{B}\hat{A}^{-1}$ (11), where

$$\hat{A} = n^{-1}\sum_{i=1}^{n} Z_i^{\otimes 2} \tag{3.24}$$

$$\hat{B} = n^{-1}\sum_{i=1}^{n}\left[\frac{\delta_i^*(Y_i-\hat{\beta}'Z_i)Z_i}{\hat{G}(T_i^*)} + \bar{\delta}_i Q(X_i) - \sum_{j=1}^{n}\frac{\bar{\delta}_j I(X_j \le X_i)Q(X_j)}{\sum_{l=1}^{n} I(X_l \ge X_j)}\right]^{\otimes 2} \tag{3.25}$$

and

$$Q(t) = \sum_{i=1}^{n}\frac{I(T_i^* > t)\delta_i^*(Y_i-\hat{\beta}'Z_i)Z_i}{\hat{G}(T_i^*)} \Big/ \sum_{j=1}^{n} I(X_j \ge t) \tag{3.26}$$

Note that $\hat{\beta}$ and its covariance matrix are all estimated based on the uncensored cases (patients who die during the study or have complete costs to the end of the study period).

3.1.8.2 Application

This approach can apply to arbitrary censoring patterns, whereas Lin's two non-parametric methods require censoring to occur only at the boundaries of the interval. Through assessing the effects of covariates, this regression methodology proposed by Lin will be greatly helpful in identifying cost-effective intervention or prevention programs. A more specific application is to devise risk-adjusted payment systems for Medicare or insurance companies which would reduce the incentives for hospitals to use unnecessarily expensive therapies and at the same time would avoid penalizing the hospitals that serve patients requiring more intensive care.

3.1.9 Lin's regression estimate where cost histories are known (CHK)

The estimator of Lin's regression estimate CHU may be inefficient when there is heavy censoring because the cost histories from the censored cases are excluded at all. In order to make use of cost information, Lin (11) proposes a further method, called Lin's regression estimate CHK. The study period is divided into smaller intervals and a linear regression model is fitted to the study costs incurred in each interval. The same factors known to influence the costs are adjusted in all models. The regression coefficients are then summed across the intervals, and the resultant model is then applied to the entire study cohort to obtain an estimate of mean total costs. The idea of partitioning the entire study period into smaller time intervals to improve the efficiency was originated by Lin et al. (10) in the one-sample study, however they handled the censoring with a very different approach.

3.1.9.1 Estimator

The entire time period of interest $[0,\tau]$ is partitioned into K subintervals $(t_{k-1},t_k]$, $k=1,\ldots,K$. Defining Y_{ki} is the cost of i th patient incurred during the time interval $(t_{k-1},t_k]$. The initial cost at t=0 is included in the calculation of the first time interval. Then we fit a linear regression model in each of the K intervals as

$$Y_{ki}=\beta_k' Z_i+\varepsilon_{ki}, \qquad k=1,\cdots,K; \qquad i=1,\cdots,n, \tag{3.27}$$

which is a semi-parametric marginal model for repeated measures in that only the marginal mean structure is modelled, where β_k $(k=1,\ldots,K)$ are $p\times 1$ vectors of unknown regression coefficients, and the error terms ε_{ki}'s are assumed to be independent among different patients but allowed to be correlated within the same patient. By

summing both sides of (3.27) over k intervals, the resultant regression equation becomes to

$$Y_i = \beta' Z_i + \varepsilon_i, \qquad i = 1, \cdots, n, \tag{3.28}$$

where $Y_i = \sum_{k=1}^{K} Y_{ki}, \beta = \sum_{k=1}^{K} \beta_k$ and $\varepsilon_i = \sum_{k=1}^{K} \varepsilon_{ki}$. The model is analogous to the model (3.19). Note that both models (3.19) and (3.28) don't need specification of the relationship between survival time and costs.

Let $T_{ki}^* = \min(T_i, t_k)$, and $\delta_{ki}^* = I(C_i \geq T_{ki}^*)$. Thus, Y_{ki} is known if and only if $\delta_{ki}^* = 1$. Similarly, the estimating equation for β_k is

$$\sum_{i=1}^{n} \frac{\delta_{ki}^*}{\hat{G}(T_{ki}^*)} (Y_{ki} - \beta_k' Z_i) Z_i = 0 \tag{3.29}$$

which has a solution

$$\hat{\beta}_k = \left\{ \sum_{i=1}^{n} \frac{\delta_{ki}^*}{\hat{G}(T_{ki}^*)} Z_i^{\otimes 2} \right\}^{-1} \sum_{i=1}^{n} \frac{\delta_{ki}^*}{\hat{G}(T_{ki}^*)} Y_{ki} Z_i. \tag{3.30}$$

Then the estimator of β is

$$\hat{\beta} = \sum_{k=1}^{K} \left[\left\{ \sum_{i=1}^{n} \frac{\delta_{ki}^*}{\hat{G}(T_{ki}^*)} Z_i^{\otimes 2} \right\}^{-1} \sum_{i=1}^{n} \frac{\delta_{ki}^*}{\hat{G}(T_{ki}^*)} Y_{ki} Z_i \right]. \tag{3.31}$$

This estimator follows the same spirit of the generalized estimating equation for repeated measures (37).

$n^{\frac{1}{2}}(\hat{\beta} - \beta)$ is proved to be asymptotically normal with zero mean and a covariance matrix can be consistently estimated by $\hat{A}^{-1} \hat{B} \hat{A}^{-1}$ (11), where $\hat{B} = \sum_{k=1}^{K} \sum_{l=1}^{K} \hat{B}_{kl}$. $\hat{A}$ is given in (3.24), $\hat{B}_{kl} = n^{-1} \sum_{i=1}^{n} \hat{\xi}_{ki} \hat{\xi}_{li}'$, and

$$\hat{\xi}_{ki} = \frac{\delta_{ki}^{*}(Y_{ki} - \hat{\beta}_k' Z_i)Z_i}{\hat{G}(T_{ki}^{*})} + (1-\delta_i)Q_k(X_i) - \sum_{j=1}^{n} \frac{(1-\delta_i)I(X_j \le X_i)Q_k(X_j)}{\sum_{l=1}^{n} I(X_l \ge X_j)} \tag{3.32}$$

where

$$Q_k(t) = \sum_{i=1}^{n} \frac{I(T_{ki}^{*} > t)\delta_{ki}^{*}(Y_{ki} - \hat{\beta}_k' Z_i)Z_i}{\hat{G}(T_{ki}^{*})} \Big/ \sum_{j=1}^{n} I(X_j \ge t) \tag{3.33}$$

3.1.9.2 Application

This method will be particularly valuable in estimating mean total costs with heavy censoring, because the patient cost information can be included in the regression method. A patient whose survival time is censored in the $(k+1)$ th interval still contributes the cost data from the first k time intervals to the estimation of mean total costs, whereas a patient whose survival time is censored before τ does not contribute any cost information to the estimator of Lin's regression estimate CHU. Thus, Lin's regression estimate CHK will produce a more efficient estimator for mean total costs than Lin's regression estimate CHU. In practice, the costs are generally collected in certain broad time intervals, such as 3 monthly intervals in CHART study, in which case it is possible to obtain more efficient estimators. Due to the availability of the cost data in multiple time intervals, we can also evaluate how the effects of covariates change over time.

3.1.10 Weighted method where cost histories are unknown (CHU)

Zhao and Tsiatis (25) modified the simple weighted complete-case estimator and used them to estimate the distribution of quality-adjusted life with censored data. Based on

their results, Bang and Tsiatis (13) applied these efficient estimators to estimate mean costs with censored cost data. The first method is called the Weighted method CHU. This method calculates mean total costs using only uncensored cases (the cases that die during the study or those whose complete study costs are recorded). For each uncensored case, their total cost is divided by the individual Kaplan-Meier survival probability. The Kaplan-Meier survival estimators are calculated with a reversed censoring indicator (death indicated as 0, censoring indicated as 1). Then mean total costs to adjust for censored cost data can be estimated by summing these results across time and divided by the study sample size. Note that the sample size should include the censored cases.

3.1.10.1 Estimator

Notations and assumptions

The purpose of this approach is to derive an estimate of mean total costs $\mu = E(M)$ and its variance over a specified period when the data is right censoring, where the random variable M denotes the total medical costs for a patient during some specified time T and E denotes expectation. The distribution of the random variable T is assumed continuous over $(0, L]$ where L denotes the upper bound of T, that is, the maximum time for which each patient is observed. In that case, M is considered as the total costs incurred by a patient over the time period $(0, L]$. If all the patients were observed over the time period $(0, L]$, then complete information on M would be available and mean total costs would be estimated by the average of the costs for each patient. However, the costs for all patients are generally not completely observed due to censoring. Defining therefore a potential time to censoring denoted by C, with survivor function

$K(u) = pr(C > u)$. We assume that censoring occurs completely at random, and C is assumed continuous. The observables from a study in the presence of censoring are $X = \min(T, C)$ and $\Delta = I(T \leq C)$, where $I(\cdot)$ is the indicator function taking the value of 1 if the observation is uncensored and zero otherwise. The costs M accrues from the start of the treatment to the last contact date X and other intermediate cost history for each patient. i.e. $M^H(t) = \{M(u), u \leq t\}$, where $M^H(t)$ indicates the cost history up to time t, $M = M(T)$, with $M(u)$ being the known accumulated costs up to time u and u denoting points in time at which cost information becomes available. The observable data for n individuals are then the independent and identically distributed random vectors $\{X_i = \min(T_i, C_i), \Delta_i = I(T_i \leq C_i), M_i^H(X_i)\}$, $i = 1, \ldots, n$ where i identifies an individual. In order to ensure that a number of patients are still under observation at L to enable calculation of the costs over the defined period $(0, L]$, we also assume that $pr(C_i \geq L) > 0$.

Mathematical description

All the proposed simple weighted estimators are based on the weighted complete-case estimator. If complete costs are followed up for each patient, then the mean total cost is simply estimated by $n^{-1} \sum M_i$. Under the conditions of independent censoring, at the time of death T_i, $K(T_i) = pr(U > T_i)$ is the probability that individual i has survived to T_i without being censored. Therefore, if individual i is observed to die at T_i, then he represents, on average, $1/K(T_i)$ individuals who might have been censored. Hence a weighted estimator using uncensored individuals only is given by

$$\frac{1}{n}\sum \frac{\Delta_i M_i}{K(T_i)} \tag{3.34}$$

The unknown survival function $K(\cdot)$ is estimated by the Kaplan-Meier method using a reverse censoring indicator, that is, the survival time T_i censors the censoring time C_i.

$$\hat{K}(t) = \prod_{u \le t} \left\{ 1 - \frac{dN^c(u)}{Y(u)} \right\} \tag{3.35}$$

where $N^c(u) = \sum I(X_i \le u, \Delta_i = 0)$ counting the number of individuals censored over time and $Y(u) = \sum I(X_i \ge u)$ counting the number of individuals at risk over time. Therefore the simple weighted complete-case estimator is defined as

$$\hat{\mu}_{WT} = \frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i M_i}{\hat{K}(T_i)} \tag{3.36}$$

For large samples, $\hat{\mu}_{WT}$ is consistent and asymptotically normal, and its asymptotic variance estimator is derived using the martingale version of the central limit theorem described by Fleming & Harrington (31) and is given as

$$\text{var}(\hat{u}_{WT}) = \frac{1}{n}\left[\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i (M_i - \hat{\mu}_{WT})^2}{\hat{K}(T_i)} + \frac{1}{n}\int_0^L \frac{dN^c(u)}{\hat{K}^2(u)} \left\{ \hat{G}(M^2, u) - \hat{G}^2(M, u) \right\} \right]$$

where

$$\hat{G}(M, u) = \frac{1}{n}\frac{1}{\hat{S}(u)}\sum_{i=1}^{n} \frac{\Delta_i M_i I(T_i \ge u)}{\hat{K}(T_i)}.$$

and $\hat{S}(u)$ corresponds to the Kaplan-Meier estimator for $S(u) = pr(T > u)$.

3.1.11 Weighted method where cost histories are known (CHK)

Bang and Tsiatis (13) also suggest a further method making use of the cost histories for the censored cases which are not used by the Weighted method CHU. The approach is similar to Lin's method CHK. However, in contrast to their estimator, the consistency and asymptotic normality of this method estimator do not depend on the selection of the partition or the discreteness of the censoring times. In order to make fuller use of the cost information of patients, the entire study period is partitioned into a number of smaller intervals. The estimated costs incurred in each interval of patients who die during this interval or those who have complete costs for the interval is weighted by the Kaplan-Meier survival estimate with a reversed censoring indicator. These estimates are then summed across intervals and over time and divided by the total sample size to obtain mean total costs.

3.1.11.1 Estimator

The duration of analysis $(0, L]$ is partitioned into K intervals $(t_j, t_{j+1}]$, $j = 0, \ldots, K-1$, the simple weighted estimator proposed in section 3.1.10 is then used to calculate the estimated costs incurred in each of K intervals and the final estimate of mean total costs is derived by summing across these intervals. By this way, patient i is defined as uncensored in j th interval whenever the censoring time C_i larger than the minimum of T_i and t_j. Therefore, there is an increase in the cost information being used by this estimator, as individuals who were treated as censored in the estimator of the Weighted method CHU having $C_i < T_i$ and whose cost information was thus not used in the estimation process will be now uncensored in some of the intervals of the partition in which their costs will contribute to the estimates.

Our partitioned estimator is given as

$$\hat{\mu}_p = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K}\frac{\Delta_i^j\{M_i(t_j)-M_i(t_{j-1})\}}{\hat{K}_j(T_i^j)} \tag{3.37}$$

where for the individual i: $\Delta_i^j = I\{\min(T_i,t_j)\le C_i\}$, $M_i(t_j)$ is the cumulative costs up to time t_j, and $\hat{K}_j(T_i^j)$ is the Kaplan-Meier estimator for the probability of not being censored based on the dataset $\{X_i^j,\Delta_i^j,i=1,\cdots,n\}$ where $X_i^j = \min(T_i^{t_j},C_i)$ and $T_i^{t_j} = \min(T_i,t_j)$.

Its consistency follows the similar fashion used for the simple weighted estimator. For large samples, the estimator $\hat{\mu}_P$ is asymptotic normal with true mean costs of zero and its variance is derived based on the theory of counting processes and the associated martingale theory (31) and is given as

$$\text{var}(\hat{\mu}_p) = \frac{1}{n}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i(M_i-\hat{\mu}_P)^2}{\hat{K}(T_i)} + \int_0^t\sum_{j=1}^{K}\sum_{l=1}^{K}\hat{S}_{j\wedge l}(u)\{\hat{G}_{j\wedge l}(M_jM_l,u)-\hat{G}_{j\wedge l}(M_j,u)\hat{G}_{j\wedge l}(M_l,u)\}\frac{dN^c(u)}{Y(u)\hat{K}(u)}\right]$$

where

$$\hat{G}_{j\wedge l}(M_l,u) = \frac{1}{n}\frac{1}{\hat{S}_{j\wedge l}(u)}\sum_{i=1}^{n}\frac{\Delta_i^{j\vee l}M_{il}I(T_i^{j\wedge l}\ge u)}{\hat{K}_{j\vee l}(T_i^{j\vee l})},$$

$$\hat{G}_{j\wedge l}(M_jM_l,u) = \frac{1}{n}\frac{1}{\hat{S}_{j\wedge l}(u)}\sum_{i=1}^{n}\frac{\Delta_i^{j\vee l}M_{ij}M_{il}I(T_i^{j\wedge l}\ge u)}{\hat{K}_{j\vee l}(T_i^{j\vee l})},$$

where $j\vee l = \max(j,l), j\wedge l = \min(j,l), T_i^{t_j} = T_i^j$, $M_{ij} = M_i(t_j)-M_i(t_{j-1})$, and $\hat{S}_j(u)$ is the Kaplan-Meier estimator of $pr\{\min(T_i,t_j)\ge u\}$.

3.1.11.2 Application

The two weighted methods on estimating mean total costs do not rely on the pattern of the censoring distribution and will all apply to the problems subject to right censoring. Furthermore, the cost history for the *ith* patient, $M_i^H(T_i) = \{M_i(u), u \le T_i\}$, can be expanded to include additional factors for which are related to costs or to a patient's health status, such as baseline covariates, individual components of medical costs, or other relevant information. Note that the Weighted method CHK estimator is not always more efficient than the Weighted method CHU estimator. However, the Weighted method CHK estimator does provide substantial improvement, especially when costs over subintervals are positively correlated. These situations are common in most real clinical studies.

3.1.12 Carides' regression method

Carides (12) proposes an estimator for mean total costs, which exploits the underlying relationship between total treatment costs and survival time. The proposed method involves two stages. In the first stage, an appropriate parametric or nonparametric regression model is selected to fit to patients who die during the study or those whose complete study costs are available to predict total costs from survival time, and the model selected should best fit the data. In the second stage, the study period is partitioned into smaller intervals. For each interval, the mean costs calculated from the regression model are weighted by the Kaplan-Meier probability of dying for those who die during that interval. Those with complete costs are included in the final interval's calculation. Then the mean costs are summed over intervals to obtain an estimate of mean total study costs.

3.1.12.1 Estimator

Notations and assumptions

For right censored data, $T_{ij}, i = 1,\cdots,N_j, j = 1,\cdots,k$, are indicated as the survival times of the i th patient in stratum j, and are assumed to be identically and independently distributed with survivor function $S_T(t) = P(T > t)$. Defining therefore a potential time to censoring denoted as $U_i, i = 1,\cdots,N$, which are assumed to be identically and independently distributed with survival function $S_U(u) = P(U > u)$, and independent of the survival time T_i. Note that one stratum is considered now for simplification. Thus, the observables from a study given censoring are $V_i = \min(T_i, U_i)$ and $\delta_i = I(T_i \le U_i)$, where $I(\cdot)$ is the indicator function taking the value of 1 if the observation is uncensored and zero otherwise. Therefore we only have information on costs accumulated up to the minimum of the survival and censoring time. Defining Y_i be the total treatment costs for patient i, and Y_i are assumed to be independent across patients. The total costs are supposed to include two components — a deterministic function of survival time and a random component. Thus,

$$Y_i = h\{g(T_i), \varepsilon_i\} \tag{3.38}$$

The function g is generally nondecreasing over T_i. And the function h can be written in multiplicative and additive models respectively as follow,

$$Y_i = g(T_i)Z_i \quad \text{where } Z_i = e^{\varepsilon_i}, \qquad \text{and } Y_i = g(T_i) + \varepsilon_i. \tag{3.39}$$

Generally, the multiplicative model is applied for the conditions where patients with longer survival time tend to have greater potential for events such as hospitalizations,

whereas the additive model is more appropriate for milder medical conditions without hospitalizations.

Let $g(t) = E(Y|T = t)$ be the expected costs for a patient with survival time t. Thus the mean total costs are given as

$$\mu = \int_0^\infty g(t)|dS_T(t)| \tag{3.40}$$

Mathematical description

The proposed method involves two stages. In the first stage, the deterministic component $g(t) = E(Y|T = t)$, which should be fitted to patients whose complete study costs are available or those who die during the study, is chosen to predict total costs from survival time, treating survival time as fixed. The model chosen should be the one that best fits the data. If a parametric model which is linear in the coefficients is selected, then ordinary linear regression can be used. However, if there is much uncertainty as to the functional form of $g(t)$, nonparametric regression is often a better choice (38). Note that only the uncensored patients can be used to estimate the regression function $g(t)$ due to the difference between the average costs of uncensored and censored cases.

In the second stage, we weight the estimate of the regression $\hat{g}(t)$ function by the Kaplan-Meier probability of dying at time t. The two-stage estimator is written as

$$\hat{\mu}_{TS} = \int_0^{V_{\max}} \hat{g}(t)|d\hat{S}_T(t)| \tag{3.41}$$

where $\hat{S}_T$ is the Kaplan -Meier estimator of the survival function, and $V_{\max}$ is the observed largest follow-up time. If $g(t)$ is consistently estimated by $\hat{g}(t)$ and the largest follow-up time is a death time rather than a censoring time, this estimator is consistent with

unrestricted mean estimator (3.40). However, if the largest follow-up time is a censoring time, (3.41) will underestimate the true mean cost due to the undefined Kaplan-Meier estimator beyond time $V_{\max}$. Until now, the problem of how to estimate the regression and survival probability beyond the largest follow-up time has not been solved. Because the aim of this approach is to estimate mean total costs restricted to the time interval $[0, L]$, where $L \leq V_{\max}$, the two-stage estimator of mean costs incurred during $[0, L]$ is defined as

$$\hat{\mu}_{TS} = \int_0^L \hat{g}(t)\left|d\hat{S}_T(t)\right| + \bar{Y}_{V \geq L}\hat{S}(L) \tag{3.42}$$

If $g(t)$ is monotonically nondecreasing, the estimator can be rewritten by applying the Kaplan-Meier estimator directly to $g(t)$ as follows,

$$\hat{\mu}_{TS} = \int_0^{\hat{g}(V_{\max})} \hat{S}_{g(T)}\{\hat{g}(t)\}\left|d\hat{g}(t)\right| \tag{3.43}$$

where $\hat{S}_{g(T)}$ is the Kaplan-Meier estimator for $\hat{g}$. If the mean costs are only calculated over the time interval $[0, L]$, the estimator assuming monotonicity becomes

$$\hat{\mu}_{TS} = \left[\int_0^{\hat{g}(L)} \hat{S}_{g(T)}\{\hat{g}(t)\}\left|d\hat{g}(t)\right| - \hat{g}(L)\hat{S}_T(L)\right] + \bar{Y}_{V \geq L}\hat{S}_T(L) \tag{3.44}$$

The four forms of proposed two-stage estimators in (3.41)-(3.44) are consistent as long as the function $g(t)$ is consistently estimated. If a parametric regression method is used, the estimator (3.43) will be asymptotically normal provided that the coefficients are consistently estimated. For instance, for the model $Y_i = \beta_0 + \beta_1 T_i + \varepsilon_i$, where ε_i are iid with zero mean and finite variance, the mean costs are estimated by $\hat{\mu}_\gamma = \hat{\beta}_0 + \hat{\beta}_1\hat{\mu}_T$, where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the least squares estimators of the intercept and slope respectively, and $\hat{\mu}_T$ is the Kaplan-Meier estimator of mean survival. Asymptotic normality follows

from the consistency of the least squares estimator (39), the asymptotic normality of the Kaplan-Meier estimator, and Slutsky's theorem. The asymptotic variance is shown to be $\beta_1^2 VAR(\hat{\mu}_T)$, where the variance expression for $\hat{\mu}_T$ is proved by Lawless (17). For the nonparametric regression method, the two-stage estimator is also asymptotically normal where the costs are nonlinear in known coefficients. For instance, for the model $Y_i = aT_i^r Z_i$, where the errors Z_i are iid with mean 1 and $r > 0$ is known, the mean costs are estimated by $\hat{\mu}_Y = \hat{a}\hat{\mu}_X$, where $\hat{a}$ is a consistent estimator of $a > 0$ and $\hat{\mu}_X$ is the Kaplan-Meier estimator of the mean of $X = T^r$. Asymptotic normality also follows by Slutsky's theorem and the asymptotic normality of the Kaplan-Meier estimator. The asymptotic variance is shown as $a^2 VAR(\hat{\mu}_x)$, where the variance expression for $\hat{\mu}_x$ follows the same expression for the parametric method.

3.1.12.2 Application

The proposed method exploits the underlying relationship between total treatment costs and survival time. A particular advantage of this approach is the gain in efficiency over purely nonparametric methods which result from the estimation and use of this relationship. The two-stage estimators for mean costs do not need any knowledge about the costs over time for individual patients. It would be particularly useful in some studies where these cost histories may not be recorded, such as retrospective studies. The two-stage method can also utilize the cost history information if we define $g(t)$ to be the expected costs incurred at follow-up time t. This approach can further include the effects of both categorical and continuous covariates on mean costs. Furthermore, this method

can also be applied to the estimation of the mean for other variables which are related to right-censored failure times, such as quality-adjusted survival time.

3.2 Case study

The present study used the cost data collected alongside two multi-centre clinical trials comparing conventional radiotherapy and continuous hyperfractionated accelerated radiotherapy (CHART) in patients with head and neck cancer or carcinoma of the bronchus. Patients were enrolled in the trials at the ten clinical centers in the UK and three centers elsewhere in Europe between October 1990 and December 1993. Details of the design of the clinical studies have already been published (40). The cost-effectiveness study was conducted on patients who were enrolled in the CHART trials at the ten clinical centers in the UK (14). A total of 970 patients enrolled for radiotherapy treatments at the ten centers participating in the study were followed from the start of treatment to 24 months post treatment. Detailed information on resource use at the clinical center was collected during the study, as was information on patient survival and clinical information. Resource use comprised four components: hospital resources, radiotherapy resources, community resources and travel. The collection of resource-use data was planned to coincide with clinical follow-up assessments at 3, 6, 9, 12, 18, 24 months after treatment started. Quality of life assessments were measured using the Hospital Anxiety and Depression Scale (41) and the Rotterdam Symptom Checklist (42) which were administered as a self-reported questionnaire from prior to treatment, 3, 4, 6 weeks after the start of treatment, three and six months after the start of treatment, and thereafter at six monthly intervals.

3.3 Sample

The data analysis here only used the hospital and radiotherapy cost data, and concentrated on the period up to 12 months from the start of treatment. Thus the cohort of patients analyzed here only consisted of 736 patients who were not censored within 12 months from recruitment. The other 234 patients were excluded because they did not have 12 months worth of follow up data.

3.4 Study design

3.4.1 Implementation of objective 1

In order to address objective 1, two steps were done. The initial step was to generate the simulated data sets according to the three censoring mechanisms. One simulated data set was generated for each of the three censoring mechanisms respectively. The second step was to do the statistical analysis. It compared the twelve different methods to adjust for censored cost data across different censoring mechanisms in order to assess whether any particular method consistently out performed the other methods.

3.4.1.1 Step1: Simulating the data sets

736 "complete" cases from the CHART study were used to form a "complete" data set where the mean total cost was known. This cohort of patients was used to generate the simulated data sets which encompassed 736 cases with a proportion of the cases randomly censored at a specific time point (either 3, 6, 9 or 12 months after the start of treatment).

Random censoring

For randomly censored data, one simulated data set was generated from the CHART cohort using the Monte Carlo simulation method. Random number generators were used to select which patients would be censored and at what time points the censoring would occur. Patient study costs were censored at the randomly determined censoring time points. A censoring level of 25% was selected corresponding to the actual level of censoring within the CHART trial.

End-of-study censoring

For end-of-study censoring, a date was chosen which would mark an artificial end to data collection. This would simulate a data set with 25% cases censored.

Informative censoring

Metastasis (MET) scores denote the spread of cancer from one part of the body to another by way of the lymph system or bloodstream, so the results from the MET scores in the CHART study were used to generate a data set that simulated the informative censoring. A simulated data set was produced: it assumed that patients with MET scores of 1 would drop out of the study due to ill health. In this simulated data set, patients whose MET scores changed from 0 to 1 at any time point (3, 6, 9, or 12 months) during the study were censored at the time point their score became 1, this resulted in 24% of the data being censored.

3.4.1.2 Step2: Statistical analysis

Data description

The demographic characteristics for the "complete" data set of 736 patients during the study period were summarized using the frequency counts and the percentage to describe the categorical variables and the mean and the standard error to describe the continuous variables.

Cost estimation

The mean total costs and standard error for the "complete" cohort were calculated first. The estimated mean total costs and standard errors in the presence of censoring for the twelve different methods under three censoring mechanisms were calculated and shown in tables and figures respectively. The accuracy of methods was evaluated by comparing the difference between estimates and the "true" cost for the cohort which means bias and using sampling standard errors.

3.4.2 Implementation of objective 2

In order to address objective 2, two steps were done. The initial step was to generate the simulated data sets using the different levels of censoring. The second step was to do the statistical analysis. It compared the twelve different methods to adjust for censored cost data under different levels of censoring in order to assess whether any particular method consistently out performed the other methods.

3.4.2.1 Step1: Simulating the data sets

The simulation methods for random censoring and end-of-study censoring described for objective 1 in section 3.4.1.1 were repeated using the different levels of censoring (10%, 25%, and 50%) respectively. For informative censoring, results from MET scores in the CHART study were still used to generate the data sets that simulated the different levels of informative censoring. Three simulated data sets were created: the first assumed that patients whose MET scores equaled to 1 at 3 months during the study were censored at 6 months; this resulted in 8% of the data being censored. And the second assumed that patients whose MET scores equaled to 1 at 6 months during the study were censored at 9 months; this resulted in 15% of the data being censored. And the third assumed that patients whose MET scores equaled to 1 at 9 months during the study were censored at 12 months; this resulted in 21% of the data being censored. Note that for each simulated data set patient study costs were set to zero from the time point of censoring.

3.4.2.2 Step2: Statistical analysis

The estimated mean total costs and standard errors in the presence of censoring for the twelve different methods under the different levels of censoring were calculated and tabulated respectively. The accuracy of methods was again evaluated by comparing the difference between estimates and the "true" cost for the cohort and using sampling standard errors.

3.4.3 Implementation of objective 3

To reach objective 3, six steps were done to implement a Monte Carlo (MC) study. The following were the basic steps necessary for a Monte Carlo study:

- Ask questions that can be examined through a Monte Carlo study
- Design a Monte Carlo study to provide answers to the questions
- Generate data
- Implement the quantitative technique
- Obtain and accumulate the statistic of interest from each replication
- Statistical analysis of the accumulated statistic of interest

3.4.3.1 Step1: Asking questions suitable for a Monte Carlo study

It may be obvious, but unless we ask the right question, it may not be possible or necessary to conduct a MC study in the first place. As discussed in Section 2.4, a MC study is essentially concerned about how a statistic of interest may vary from sample to sample. In other words, a MC study is about obtaining the sampling distribution of a statistic of interest by repeatedly drawing random samples from a specified population. In this sense, the question suitable for a MC study is typically related to some aspects of the sampling distribution of a statistic. Because we were interested in evaluating the uncertainty of the censoring of the simulated data sets in this study, our question was easily translated into a question about the sampling distribution of the estimated mean total costs from the different methods to adjust for censored cost data under the different combinations of censoring mechanism and censoring level conditions.

3.4.3.2 Step 2: Designing a Monte Carlo study

Once we identified the question suitable for a MC study, we need to figure out how we can answer our question by designing an appropriate MC study. To do this, we have to

consider the major factors that may influence the variability of sample mean total costs. According to the previous investigation, the sample mean total cost was influenced by the different censoring mechanisms and different levels of censoring. Then we need to consider another important issue: under each combination of censoring mechanism and censoring level conditions, how many random samples were we going to draw from a specified statistical population that represented the null hypothesis? The decision must be made carefully so that reasonably accurate answers to our question can be obtained. Because we were trying to obtain the sampling distribution of the estimated mean total costs from the different methods to adjust for censored cost data under the true null hypothesis, the number of samples drawn under each combination of censoring mechanism and censoring level conditions would greatly influence the accuracy of the simulated sampling distribution of the estimated mean total costs. If too few samples were drawn under each combination of censoring mechanism and censoring level conditions, our answers might be too crude to be useful. According to the findings from previously published reports (10, 12, 43), we decided that 5000 samples were the minimum number we can live with, and the sampling distribution of the estimated mean total costs from the different methods across the 5000 samples under each combination of censoring mechanism and censoring level conditions should be accurate enough for our purpose.

3.4.3.3 Step 3: Generating sample data

Once the study design was worked out, we need to generate the sample data to be used in the MC study. It is worth pointing out that data generation is probably the most

important step in any MC study. This is so because MC study results are based on the data generated in this process. If the data generated in this process are not what we think they should be, the validity of the MC study results will obviously be in serious question. From this perspective, the importance of data generation in a MC study can never be overemphasized. Depending on the aim of this present MC study, we need to generate 5000 simulated data sets under each combination of censoring mechanism and censoring level conditions. We can use EXCEL random number generator RAND () to accomplish this.

3.4.3.4 Step 4: Implementing the statistical technique in question

In many MC studies, some types of statistical technique are involved. For the present study, the mean total costs and standard errors from the different methods to adjust for censored cost data need to be computed for each of the 5000 samples under each combination of censoring mechanism and censoring level conditions.

3.4.3.5 Step 5: Obtaining and accumulating the statistic of interest

Once the statistical technique was implemented and the statistic of interest was computed, the statistic of interest from each random sample must be obtained, and it must be accumulated across samples. In our study, we need to obtain each of the 5000 sample mean total costs from the different methods under each combination of censoring mechanism and censoring level conditions and to accumulate them for later analyses.

3.4.3.6 Step 6: Statistical analysis of the accumulated statistic of interest

By the time the statistic of interest from all the samples had been obtained and accumulated, the simulation process of the MC study was complete. The estimated mean total costs from the different methods across the 5000 samples under each combination of censoring mechanism and censoring level conditions were shown in tables and frequency figures respectively, in order to assess which method was the most accurate through the MC simulation analysis.

In summary, data analyses for this study were very complex. Analyses were conducted using the SAS statistical software version 8.0 and EXCEL for windows 2000. The initial simulated data sets were generated using EXCEL. Data then were transported to SAS and the initial analyses using the twelve identified methods were conducted. Complex coding was required within EXCEL to reproduce the results from the SAS analyses – this would confirm that the complex statistical calculations can be reproduced. This was necessary, as the Monte Carlo simulation study required for objective 3 was conducted within EXCEL.

4. RESULTS

4.1 Data description

4.1.1 Eligibility

Patients were only eligible for the present analysis if the complete 12 months of hospital cost data were available. Thus the "complete" cohort of patients identified here consisted of 736 patients who were not censored within 12 months from recruitment. The other 234 patients were excluded because they did not have 12 months worth of follow up data. The mean total cost for the "complete" cohort was £3609.33 and the standard error was 101.89. The median total cost was £2944.23 with costs ranging from 0 to £23345.82 across the samples.

4.1.2 Patient characteristics

The demographic characteristics for the "complete" set of 736 patients having the complete 12 months hospital cost data are presented in Table 1. The "complete" cohort comprised 463(62.91%) head and neck cancer patients and 273(37.09%) carcinoma of bronchus patients, who were recruited during period October, 1990 to December, 1993. Of the total cohort, 451(61.28%) patients were given continuous hyperfractionated accelerated radiotherapy (CHART) treatment, and 285(38.72%) patients received conventional radiotherapy treatment. 230(31.25%) patients died during 12 months study period.

Table 1: Demographic patient characteristics (N=736)

	Number (%)
Gender	
Males	561 (76.22%)
Treatment	
CHART	451 (61.28%)
Conventional	285 (38.72%)
Status	
Dead	230 (31.25%)
Alive	506 (68.75%)
Center	
1	50 (6.79%)
2	106 (14.40%)
3	106 (14.40%)
4	89 (12.09%)
5	112 (15.22%)
6	57 (7.74%)
7	67 (9.10%)
8	31 (4.21%)
9	91 (12.36%)
10	27 (3.67%)
Site of cancer	
Head and neck	463 (62.91%)
Bronchus	273 (37.09%)

4.2 Results of statistical analysis

4.2.1 Analysis of primary objective

The purpose of this analysis was to screen for methods to be further evaluated in the Monto Carlo study. It compared the twelve different methods to adjust for censored cost data across three censoring mechanisms (random, end-of-study and informative censoring) to assess whether any particular method consistently out

performed the other methods. The estimated mean total costs and standard errors in the presence of censoring for the twelve different methods across three censoring mechanisms were calculated and shown in tables and figures respectively. The accuracy of methods was evaluated by comparing the difference between estimates and the "true" cost for the cohort which means bias and using sampling standard errors.

Random censoring

Table 2 presents the estimated mean total costs and standard errors in the presence of censoring for the twelve different methods for estimating the mean total study costs for random censoring. Cox's proportional hazard regression model gave the most accurate estimate (£3601.74) of the "true" cost of £3609.33. The method that gave the least accurate estimate of mean total costs was the Stratified Cox model where the estimate of mean total costs (£5708.82) was over 1.6 times greater than the "true" cost (£3609.33), and the prediction of standard error (442.40) was 4.3 times higher than the "true" standard error (101.89).

Table 2: The estimated mean total costs and standard errors for the twelve methods for random censoring

	Mean Total Cost* £	Standard** Error	Difference Rank
True cost	3609.33	101.89	
Available sample method	3101.08	99.70	8
Uncensored cases method	3567.42	119.56	3
Kaplan-Meier cost method	4038.70	137.80	6
Cox's PH regression model	3601.74	120.07	1
Stratified Cox model	5708.82	442.40	11
Lin's method,CHK	2830.77	122.47	9
Lin's method,CHU	2636.53	191.71	10
Lin's regression estimate CHU	3160.84	22.38	7
Lin's regression estimate CHK	3376.29	12.75	4
Weighted method, CHU	3574.90	145.33	2
Weighted method, CHK	3256.55	107.09	5
Carides' regression method	2636.53	Not estimable	10

*estimates of true mean total costs ** estimates of true standard error

End-of-study censoring

Table 3 presents the estimated mean total costs and standard errors in the presence of censoring for the twelve different methods for estimating the mean total study costs for end-of-study censoring. The Uncensored cases method gave the most accurate estimate (£3645.14) corresponding to the "true" cost of £3609.33. The method that gave the least accurate estimate of mean total costs was the Stratified Cox model where the estimate of mean total costs (£5793.08) was over 1.6 times greater than the "true" cost (£3609.33), and the prediction of standard error (423.89) was 4.2 times higher than the "true" standard error (101.89).

Table 3: The estimated mean total costs and standard errors for the twelve methods for end-of-study censoring

	Mean Total Cost* £	Standard** Error	Difference Rank
True cost	3609.33	101.89	
Available sample method	3360.29	100.89	7
Uncensored cases method	3645.14	119.08	1
Kaplan-Meier cost method	4234.40	144.10	9
Cox's PH regression model	3668.74	119.38	3
Stratified Cox model	5793.08	423.89	11
Lin's method,CHK	3026.46	121.15	8
Lin's method,CHU	2703.86	181.39	10
Lin's regression estimate CHU	3398.95	21.95	6
Lin's regression estimate CHK	3520.12	11.86	4
Weighted method, CHU	3651.38	143.71	2
Weighted method, CHK	3471.19	106.61	5
Carides' regression method	2703.86	Not estimable	10

*estimates of true mean total costs **estimates of true standard error

Informative censoring

Table 4 presents the estimated mean total costs and standard errors in the presence of censoring for the twelve different methods for estimating the mean total study costs for informative censoring. The Kaplan-Meier cost method gave the most accurate estimate (£3661.60) of the "true" cost of £3609.33. The method that gave the least accurate estimate of mean total costs was again the Stratified Cox model where the estimate of mean total costs (£5079.11) was over 1.4 times greater than the "true" cost (£3609.33), and the prediction of standard error (430.41) was 4.2 times higher than the "true" standard error (101.89).

Table 4: The estimated mean total costs and standard errors for the twelve methods for informative censoring

	Mean Total Cost* £	Standard** Error	Difference Rank
True cost	3609.33	101.89	
Available sample method	2921.78	88.89	8
Uncensored cases method	3221.06	101.27	4
Kaplan-Meier cost method	3661.60	122.90	1
Cox's PH regression model	3253.12	101.48	2
Stratified Cox model	5079.11	430.41	11
Lin's method,CHK	2729.48	111.76	9
Lin's method,CHU	2614.82	189.54	10
Lin's regression estimate CHU	2965.04	17.14	7
Lin's regression estimate CHK	3131.76	9.90	5
Weighted method, CHU	3222.85	120.99	3
Weighted method, CHK	3032.40	93.31	6
Carides' regression method	2614.82	Not estimable	10

*estimates of true mean total costs **estimates of true standard error

In order to visually compare the accuracy of the twelve different methods for estimating the mean total costs given censoring, Figures 1 to 3 show the estimated mean total study costs in the presence of censoring for the twelve different methods for estimating the mean study costs for random, end-of-study, informative censoring respectively.

Figure 1: The estimated mean total study costs in the presence of censoring for the twelve different methods for estimating the mean study costs for 25% random censoring

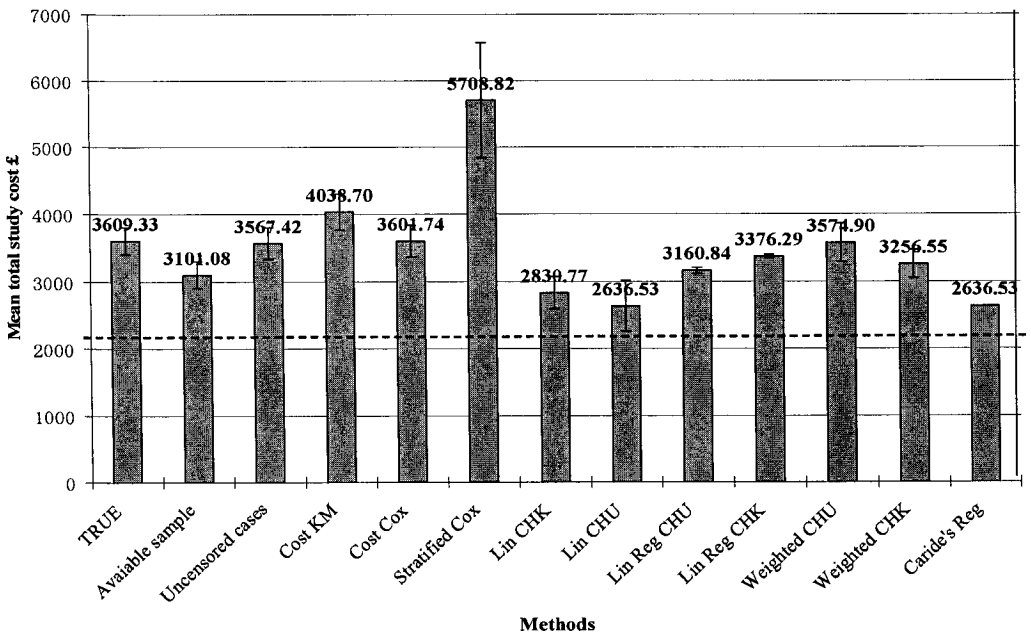


Figure 2: The estimated mean total study costs in the presence of censoring for the twelve different methods for estimating the mean study costs for 25% end-of-study censoring

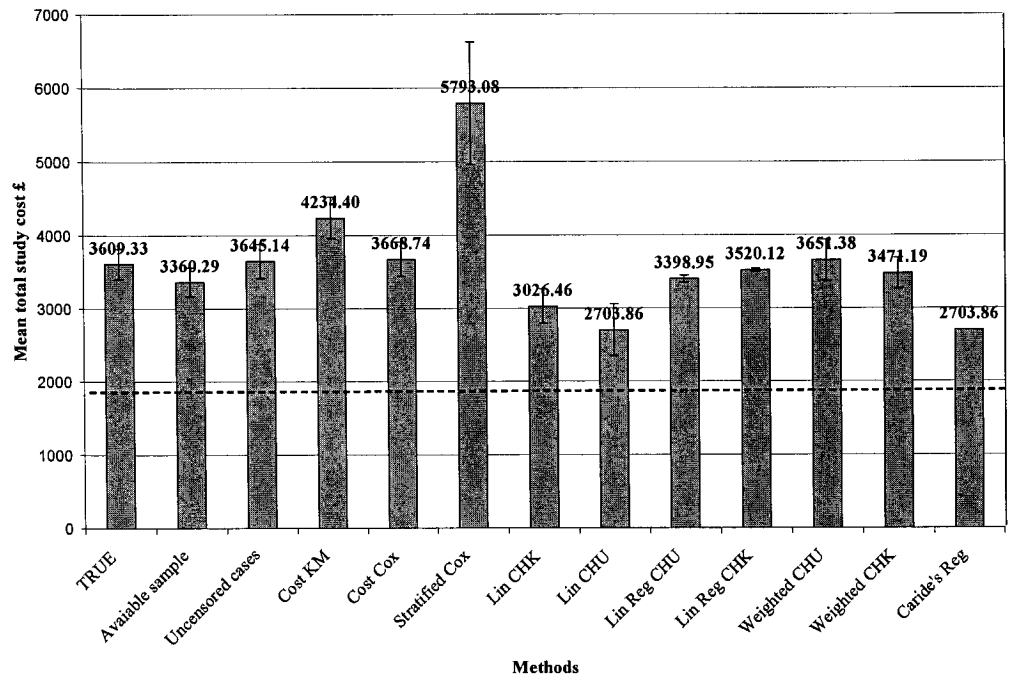
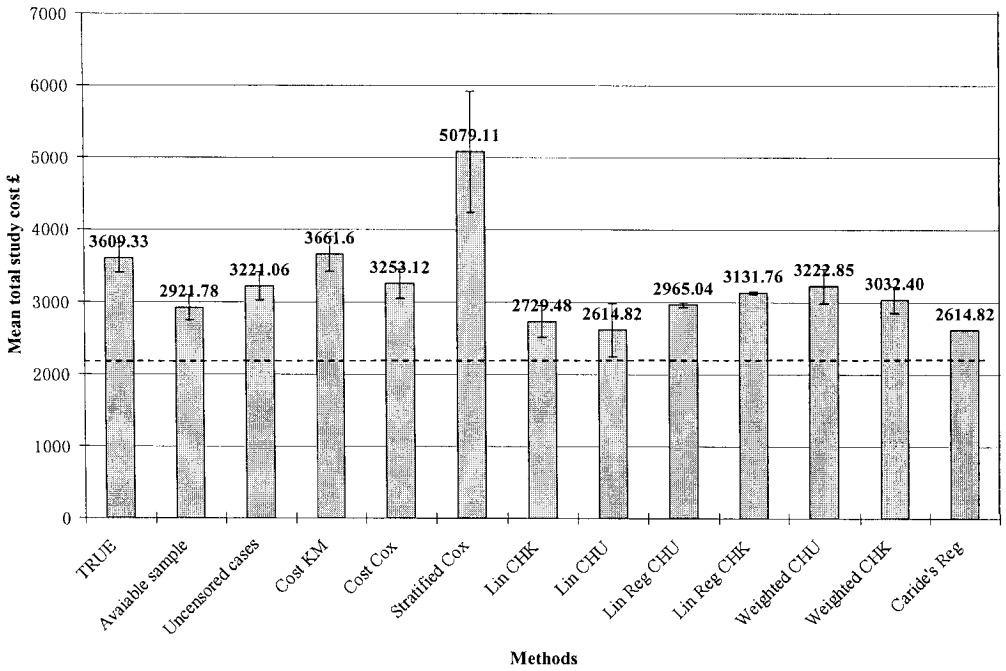


Figure 3: The estimated mean total study costs in the presence of censoring for the twelve different methods for estimating the mean study costs for 25% informative censoring



Of the twelve different methods, the following five methods consistently out performed the other methods to give the accurate estimates of the "true" cost of £3609.33 across all three censoring mechanisms: the Uncensored cases method, Cox's proportional hazard regression model, the Weighted method where cost histories were unknown, the Weighted method where cost histories were known and Lin's regression estimate where cost histories were known. The method that gave the least accurate estimate of mean total costs was the Stratified Cox model where the estimate of mean total costs was at least over 1.4 times greater than the "true" cost, and the prediction of standard error was over 4 times higher than the "true" standard error regardless of the underlying censoring mechanism. A statistical comparison of the twelve different methods, after adjusting for censoring mechanisms, showed that there was a

significant difference in mean cost estimates across methods (Friedman's χ^2_{11}=30.27, p=0.0014).

For all three censoring mechanisms (random, end-of-study and informative censoring), all methods except the Stratified Cox model predicted the mean total costs within £1000 (27.71%). A statistical comparison of the three censoring mechanisms, after adjusting for twelve methods, showed that there was no significant difference in mean total cost estimates across the three censoring mechanisms (Friedman's χ^2_2=3.65, p=0.1615).

As expected, the Kaplan-Meier cost method and the Stratified Cox model tended to overestimate the "true" cost for random censoring and end-of-study censoring due to the assumption of independence between the censoring mechanism and costs has been violated. However, the Kaplan-Meier cost method gave the most accurate estimate of the "true" cost for informative censoring where patients with ill health were censored. This is probably because in our study the censored patients were likely to be the sicker patients in the cohort and thus incurred the highest costs. When these patients were censored before the high costs were incurred, the Kaplan-Meier cost method would not overestimate the "true" cost. Thus, the Kaplan-Meier cost method gave the more accurate estimate of mean total costs than any other methods in this case.

As noted from the Tables 2 to 4, the Stratified Cox model, Lin's two regression methods and Carides' regression method gave poor predictions of the "true" standard error (101.89) where the Stratified Cox model tended to overestimate it and Lin's two regression methods and Carides' regression method tended to underestimate it. All

other methods gave the accurate measure of the "true" standard error term.

4.2.2 Analysis of secondary objective

In order to do statistical analysis for the secondary objective, we compared the twelve different methods given censoring under different levels of censoring to assess whether any particular method consistently performed better than the other methods. For random censoring and end-of-study censoring, the estimated mean total costs and standard errors in the presence of censoring for the twelve different methods under different levels of censoring (10%, 25%, and 50%) were calculated and tabulated respectively. We referred to 10%, 25% and 50% as light, medium and heavy censoring respectively. For informative censoring, the estimated mean total costs and standard errors in the presence of censoring for the twelve different methods under different levels of censoring (8%, 15%, and 21%) were calculated and tabulated respectively using the simulation method described in section 3.4.2.1 for generating the simulated data sets with different levels of informative censoring. The accuracy of methods was again evaluated by comparing the difference between estimates and the "true" cost for the cohort which means bias and using sampling standard errors.

Random censoring

Table 5 presents the estimated mean total costs and standard errors from the twelve methods under different levels of random censoring. For random censoring, the estimates of mean total costs from all twelve methods deteriorated as the proportion

of censored cases increased, where the estimates of the Kaplan-Meier cost method tending to be biased upward and the estimates of the other methods tending to be biased downward as the level of censoring increased. However, the poorer estimates are still within £1150 (31.86%) of the "true" cost except for the Stratified Cox model. In addition, the estimates of standard errors became less precise as the level of censoring increased.

Table 5: The estimated mean total costs and standard errors for the twelve methods under different levels of random censoring

	Mean Total Cost* (Standard Error**)		
	10% Censoring	**25% Censoring**	**50% Censoring**
True cost	3609.33 (101.89)	3609.33 (101.89)	3609.33 (101.89)
Available sample	3432.65 (101.93)	3101.08 (99.70)	2719.57 (93.65)
Uncensored cases	3610.64 (108.29)	3567.42 (119.56)	3512.78 (133.23)
Kaplan-Meier cost	3748.60 (112.30)	4038.70 (137.80)	4745.50 (194.60)
Cox's PH model	3638.85 (108.66)	3601.74 (120.07)	3532.03 (133.26)
Stratified Cox model	5801.47 (411.25)	5708.82 (442.40)	5298.46 (444.84)
Lin's method,CHK	3062.83 (118.41)	2830.77 (122.47)	2585.03 (122.22)
Lin's method,CHU	2651.08 (166.38)	2636.53 (191.71)	2570.06 (206.33)
Lin's regression CHU	3442.64 (24.21)	3160.84 (22.38)	2927.67 (24.27)
Lin's regression CHK	3566.30 (12.98)	3376.29 (12.75)	3201.13 (14.94)
Weighted CHU	3614.19 (117.75)	3574.90 (145.33)	3515.53 (188.29)
Weighted CHK	3489.28 (104.00)	3256.55 (107.09)	3005.33 (107.27)
Carides' regression	2651.08 (·)	2636.53 (·)	2570.06 (·)

*estimates of true mean total costs **estimates of true standard error

End-of-study censoring

Table 6 presents the estimated mean total costs and standard errors for the twelve methods under different levels of end-of-study censoring. For end-of-study censoring,

the estimates of mean total costs from all twelve methods deteriorated as the proportion of censored cases increased, where the estimate of the Kaplan-Meier cost method tending to be biased upward and the estimates of the Available sample method, the Stratified Cox model, Lin's two nonparametric methods, Lin's two regression methods, the Weighted method where cost histories were known, and Carides' regression method tending to be biased downward as the level of censoring increased. The estimates fluctuated slightly less using the Uncensored cases method, Cox's PH regression model and the Weighted method where cost histories were unknown. However, the poorer estimates are still within £1150 (31.86%) of the "true" cost except for the Stratified Cox model. Furthermore, the estimates of standard errors became less precise as the level of censoring increased.

Table 6: The estimated mean total costs and standard errors for the twelve methods under different levels of end-of-study censoring

	Mean Total Cost* (Standard Error)**		
	10% Censoring	**25% Censoring**	**50% Censoring**
True cost	3609.33 (101.89)	3609.33 (101.89)	3609.33 (101.89)
Available sample	3514.96(102.95)	3360.29(100.89)	2628.00(93.19)
Uncensored cases	3630.99(111.41)	3645.14(119.08)	3545.95(129.92)
Kaplan-Meier cost	3878.30(119.90)	4234.40(144.10)	4683.90(179.10)
Cox's PH model	3642.83(111.73)	3668.74(119.38)	3540.80(130.19)
Stratified Cox model	5910.78(422.56)	5793.08(423.89)	5234.64(377.52)
Lin's method,CHK	3118.04(120.20)	3026.46(121.15)	2527.62(122.95)
Lin's method,CHU	2721.01(174.60)	2703.86(181.39)	2478.61(170.03)
Lin's regression CHU	3518.60(23.64)	3398.95(21.95)	2866.66(19.36)
Lin's regression CHK	3595.00(12.48)	3520.12(11.86)	3212.22(12.18)
Weighted CHU	3629.66(121.71)	3651.38(143.71)	3554.91(190.23)
Weighted CHK	3564.56(105.46)	3471.19(106.61)	2942.88(105.93)
Carides' regression	2721.01(·)	2703.86(·)	2478.61(·)

*estimates of true mean costs **estimates of true standard error

Informative censoring

Table 7 shows the estimated mean total costs and standard errors for the twelve methods under different levels of informative censoring. For informative censoring, the estimates of mean total costs from the Kaplan-Meier cost method and Lin's method where cost histories were known were biased upward as the proportion of censored cases increased, whereas the estimates of mean total costs from the Uncensored cases method, Cox's PH regression model, Lin's method where cost histories were unknown, the Weighted method where cost histories were unknown and Carides' regression method were biased downward as the proportion of censored cases increased. The estimates fluctuated slightly less using the Available sample method, the Stratified Cox model, Lin's two regression methods and the Weighted method where cost histories were known. Generally, the estimates of mean total costs from the twelve methods deteriorated as the proportion of censored cases increased for informative censoring, though the poorer estimates are all within £1150 (31.86%) of the "true" cost except for the Stratified Cox model. However, the estimates of standard errors did not substantially change as the level of censoring increased.

Table 7: The estimated mean total costs and standard errors for the twelve methods under different levels of informative censoring

	Mean Total Cost* (Standard Error**)		
	8% Censoring	**15% Censoring**	**21% Censoring**
True cost	3609.33 (101.89)	3609.33 (101.89)	3609.33 (101.89)
Available sample	3477.53(97.90)	3449.03(94.01)	3507.71(95.93)
Uncensored cases	3495.10(103.40)	3402.39(101.95)	3274.06(102.41)
Kaplan-Meier cost	3720.50(113.20)	3934.30(126.80)	4298.10(159.80)
Cox's PH model	3521.47(103.73)	3433.15(102.15)	3305.56(102.64)
Stratified Cox model	5757.06(426.50)	5605.46(401.36)	5679.39(416.02)
Lin's method,CHK	3101.5(116.57)	3109.75(113.63)	3143.69(115.93)
Lin's method,CHU	2679.26(182.68)	2671.20(186.92)	2538.9(165.90)
Lin's regression CHU	3480.04(21.55)	3454.93(23.24)	3490.40(23.19)
Lin's regression CHK	3508.91(10.97)	3452.85(11.95)	3514.91(11.86)
Weighted CHU	3498.14(110.51)	3412.68(116.58)	3260.33(119.40)
Weighted CHK	3505.51(99.62)	3486.67(96.18)	3546.00(98.10)
Carides' regression	2679.26(·)	2671.20(·)	2538.90(·)

*estimates of true mean total costs **estimates of true standard error

In summary, the results from the analyses corresponding to the first two objectives have shown that the following five methods for estimating the mean total study costs given censoring, consistently, gave the mean total costs more accurately than the other methods regardless of the underlying censoring mechanism and censoring level: the Uncensored cases method, Cox's proportional hazard regression model, the Weighted method where cost histories were unknown, the Weighted method where cost histories were known and Lin's regression estimate where cost histories were known. As noted, the accuracy of mean total cost estimates depended not only on the censoring mechanism but also on the method chosen to estimate mean total costs. Furthermore, the accuracy of mean total cost estimates was also influenced by the proportion of

censored cases. Although the estimates of mean total study costs and standard errors did deteriorate as the level of censoring increased for all twelve methods, the difference from the "true" mean estimate was not substantially large except for the Stratified Cox model. However, the biases of the twelve methods from the "true" mean estimate can be of either direction depending on the different censoring mechanisms.

4.2.3 Analysis of third objective

The analysis for the third objective was to determine the consistency of the first two objectives through the Monte Carlo (MC) simulation analysis. We conducted our MC study for the purpose of answering our questions about the sampling distribution of the estimated mean total costs from the different methods to adjust for censored cost data, in order to determine what factors may affect such distributional characteristics. Corresponding to the analysis results from the first two objectives, the following six methods were selected to determine the consistency of the first two objectives under each combination of censoring mechanism and censoring level conditions through the MC simulation analysis: the Uncensored cases method, Cox's PH regression model, the Weighted method where cost histories were unknown, the Weighted method where cost histories were known, Lin's regression estimate where cost histories were known and the Available sample method. In order to evaluate the uncertainty of the censoring of the simulated data sets, 5000 alternate simulated data sets for each combination of censoring mechanism (random /end-of-study censoring) and censoring

level (10%, 25% or 50%) conditions were generated using the MC simulation method. The estimated mean total costs to adjust for censored cost data for the six different methods across the 5000 samples under each combination of censoring mechanism and censoring level conditions were calculated and shown in tables and frequency figures respectively, in order to assess which method was the most accurate through the MC simulation analysis.

4.2.3.1 Descriptive statistics for mean total cost estimates sample distribution

Random censoring

Table 8 presents the descriptive statistics for the mean total cost estimates sample distributions from the six different methods under the three levels (10%, 25% and 50%) of random censoring. Two observations were noted from Table 8. First, the means of 5000 mean total cost estimates from the six different methods under each of the three censoring levels conditions were very close to the corresponding mean total cost estimates calculated in the first two objectives for random censoring. This makes perfect sense, because this confirms that the complex statistical calculations can be reproduced on 5000 alternate simulated data sets through the MC simulation analysis, so the methods were shown to be robust. Thus the consistency of the first two objectives was determined. As a result, although mean total cost estimates may vary within certain ranges, the means of the sample mean total cost estimates from the six different methods should converge on the "true" mean cost for the cohort. Second, the standard deviation of the sampling distribution of the mean total cost estimates under

random censoring was smaller when the censoring level was lower, and it increased with an increase in the censoring level. This indicated that when the censoring level was high, there was more variability in the sample mean total cost estimates than there was when the censoring level was low. The same phenomenon was reflected by the width of range defined as the difference between maximum and minimum. As shown in Table 8, the width of range for the sampling distribution of the mean total cost estimates was smaller when the censoring level was lower, and it increased with an increase in the censoring level. This also gave the evidence that the variability of the estimates of mean total costs from the six different methods increased as the proportion of censored cases increased for random censoring. Furthermore, the accuracy of the six different methods was also evaluated by comparing the difference between the means of 5000 sample mean total cost estimates and the "true" cost for the cohort. Of the six different methods, the Uncensored cases method, Cox's PH regression model, and the Weighted method where cost histories were unknown, consistently, out performed the other three methods to give the accurate estimates of the "true" cost of £3609.33 across all three censoring levels.

Table 8: Descriptive statistics for the mean total cost estimates sample distributions under three levels of random censoring

Censoring level	10%					
	N	Mean*	Standard** Deviation	Minimum	Maximum	Width of Range
Available sample method	5000	3426.80	34.76	3284.52	3542.42	257.90
Uncensored cases method	5000	3609.62	33.54	3475.46	3713.56	238.10
Cox's PH regression model	5000	3635.35	33.69	3502.07	3741.90	239.83
Weighted method, CHU	5000	3611.64	33.59	3476.64	3715.90	239.26
Weighted method, CHK	5000	3485.32	32.00	3357.30	3588.38	231.08
Lin's regression estimate CHK	5000	3505.77	32.85	3368.18	3615.13	246.95
Censoring level	25%					
	N	Mean*	Standard** Deviation	Minimum	Maximum	Width of Range
Available sample method	5000	3152.21	51.37	2910.01	3312.29	402.28
Uncensored cases method	5000	3610.99	59.25	3398.60	3822.33	423.73
Cox's PH regression model	5000	3636.05	59.44	3419.86	3844.17	424.31
Weighted method, CHU	5000	3615.98	59.42	3401.09	3826.66	425.57
Weighted method, CHK	5000	3308.85	49.54	3083.64	3465.11	381.47
Lin's regression estimate CHK	5000	3330.10	51.00	3100.06	3489.46	389.40
Censoring level	50%					
	N	Mean*	Standard** Deviation	Minimum	Maximum	Width of Range
Available sample method	5000	2695.74	63.69	2466.39	2957.22	490.83
Uncensored cases method	5000	3609.36	104.10	3254.92	3986.12	731.20
Cox's PH regression model	5000	3628.02	103.70	3271.87	4019.25	747.38
Weighted method, CHU	5000	3618.69	105.24	3270.52	4000.51	729.99
Weighted method, CHK	5000	2991.43	66.40	2758.42	3241.04	482.62
Lin's regression estimate CHK	5000	3013.91	68.67	2776.11	3271.79	495.68

*the average of 5000 sample mean cost estimates

**the standard deviation of sample distribution

End-of-study censoring

Table 9 shows the descriptive statistics for the mean total cost estimates sample distributions from the six different methods under the three levels (10%, 25% and 50%) of end-of-study censoring. Two observations were noted from Table 9. First, the means of 5000 mean total cost estimates from the six different methods under each of the three censoring levels conditions were very close to the corresponding mean total cost estimates calculated in the first two objectives for end-of-study censoring. Thus, this confirms that the complex statistical calculations can be reproduced through the MC simulation analysis, so the methods were shown to be robust, and the consistency of the first two objectives was determined. As a result, although mean total cost estimates may vary within certain ranges, the means of the sample mean total cost estimates from the six different methods should converge on the "true" mean cost estimate for the cohort. Second, the standard deviation of the sampling distribution of the mean total cost estimates under end-of-study censoring was smaller when the censoring level was lower, and it increased with an increase in the censoring level. This indicated that when the censoring level was high, there was more variability in the sample mean total cost estimates than there was when the censoring level was low. The same result was also reflected by the width of range. As noted in Table 9, the width of range for the sampling distribution of the mean total cost estimates was smaller when the censoring level was lower, and it increased with an increase in the censoring level. This confirmed that the variability of the estimates of mean total costs from the six different methods increased as the proportion of censored cases increased

under end-of-study censoring. Furthermore, the accuracy of the six different methods was again evaluated by comparing the difference between the means of 5000 sample mean total cost estimates and the "true" cost for the cohort. Of the six different methods, the Uncensored cases method, Cox's PH regression model, and the Weighted method where cost histories were unknown, consistently, performed better than the other three methods to give the accurate estimates of the "true" cost of £3609.33 regardless of the underlying censoring level. These findings are consistent with the findings from the random censoring.

Table 9: Descriptive statistics for the mean total cost estimates sample distributions under three levels of end-of-study censoring

Censoring level	10%					
	N	Mean*	Standard** Deviation	Minimum	Maximum	Width of Range
Available sample method	5000	3495.11	19.72	3403.15	3546.64	143.49
Uncensored cases method	5000	3609.85	36.40	3478.46	3720.03	241.57
Cox's PH regression model	5000	3635.64	36.57	3505.59	3741.73	236.14
Weighted method, CHU	5000	3615.71	36.60	3480.62	3726.73	246.11
Weighted method, CHK	5000	3543.83	20.62	3449.11	3597.86	148.75
Lin's regression estimate CHK	5000	3563.24	21.66	3464.17	3625.36	161.19
Censoring level	25%					
	N	Mean*	Standard** Deviation	Minimum	Maximum	Width of Range
Available sample method	5000	3385.49	26.53	3263.77	3461.09	197.32
Uncensored cases method	5000	3608.96	54.94	3345.13	3794.72	449.59
Cox's PH regression model	5000	3634.02	54.82	3368.73	3823.41	454.68
Weighted method, CHU	5000	3637.10	56.04	3363.83	3831.16	467.33
Weighted method, CHK	5000	3489.94	29.36	3354.49	3572.79	218.30
Lin's regression estimate CHK	5000	3508.50	30.72	3367.46	3598.39	230.93
Censoring level	50%					
	N	Mean*	Standard** Deviation	Minimum	Maximum	Width of Range
Available sample method	5000	2631.04	44.42	2478.02	2785.64	307.62
Uncensored cases method	5000	3608.48	104.33	3266.09	3998.24	732.14
Cox's PH regression model	5000	3626.50	104.23	3275.24	4035.46	760.22
Weighted method, CHU	5000	3702.45	109.48	3347.43	4106.73	759.30
Weighted method, CHK	5000	2958.41	54.19	2777.68	3147.29	369.61
Lin's regression estimate CHK	5000	2973.00	56.93	2784.06	3170.93	386.87

* the average of 5000 sample mean cost estimates

**the standard deviation of sample distribution

4.2.3.2 Frequency of the rank of difference under different levels of censoring

Random censoring

In order to illustrate the accuracy of the six different methods measured by comparing the difference between each of 5000 sample mean total cost estimates and the "true" cost for the cohort which means biase under different levels of random censoring, Table 10 shows the frequency of the rank of difference for the six different methods across 5000 samples under three levels of random censoring (10%, 25% and 50%). The frequencies of the rank shown in Table 10 were only used for illustration of the accuracy of the different methods, not for decision-making regarding the best estimator. For random censoring, the Uncensored cases method consistently gave the most accurate estimate of the "true" cost of £3609.33 across 5000 samples regardless of the underlying censoring level in this MC analysis. Furthermore, the Weighted method where cost histories were unknown and Cox's PH regression model also gave consistently better mean total cost estimates under different levels of random censoring. The method that gave the least accurate estimate of mean total costs was the Available sample method across 5000 samples.

Table 10: Frequency (%) of the rank of difference for six methods under three levels of random censoring

Censoring level	10%					
Rank of difference	1	2	3	4	5	6
Available sample method	—	—	—	—	0.18	99.82
Uncensored cases method	48.46	16.62	34.92	—	—	—
Cox's PH regression model	33.14	1.02	56.52	5.56	3.58	0.18
Weighted method, CHU	15.18	81.06	2.70	1.06	—	—
Weighted method, CHK	—	1.00	0.06	2.70	96.24	—
Lin's regression estimate CHK	3.22	0.30	5.80	90.68	—	—
Censoring level	25%					
Rank of difference	1	2	3	4	5	6
Available sample method	—	—	—	—	—	100.00
Uncensored cases method	51.96	10.60	37.44	—	—	—
Cox's PH regression model	37.70	1.88	60.20	0.12	0.10	—
Weighted method, CHU	10.24	87.46	2.26	0.04	—	—
Weighted method, CHK	—	0.04	—	0.06	99.90	—
Lin's regression estimate CHK	0.12	—	0.10	99.78	—	—
Censoring level	50%					
Rank of difference	1	2	3	4	5	6
Available sample method	—	—	—	—	—	100.00
Uncensored cases method	42.26	22.02	35.72	—	—	—
Cox's PH regression model	38.24	22.22	39.52	0.02	—	—
Weighted method, CHU	20.42	55.60	23.98	—	—	—
Weighted method, CHK	—	—	—	0.28	99.72	—
Lin's regression estimate CHK	—	—	0.02	99.70	0.28	—

In order to illustrate graphically the accuracy of the six different methods under each combination of random censoring and censoring level conditions, Figures 4~6 present the frequency distribution of the rank of difference for the six different methods under three levels of random censoring (10%, 25% and 50%) respectively. As noted in Figures 4~6, the Uncensored cases method, the Weighted method where cost histories were unknown and Cox's PH regression model, consistently, out-performed the other three methods (the Available sample method, the Weighted method where cost histories were known and Lin's regression estimate where cost histories were known) across 5000 samples under the three levels of random censoring through the MC analysis. The method that gave the least accurate estimate of mean total costs regardless of the underlying censoring level was the Available sample method. The findings from the MC analysis correspond to the findings from the first two objectives. Thus, this confirms that the complex statistical calculations for random censoring can be reproduced through the MC analysis, so the methods were shown to be robust, and the consistency of the first two objectives for random censoring was determined through the MC analysis.

Figure 4: Frequency distribution of the rank of difference for six methods under 10% random censoring

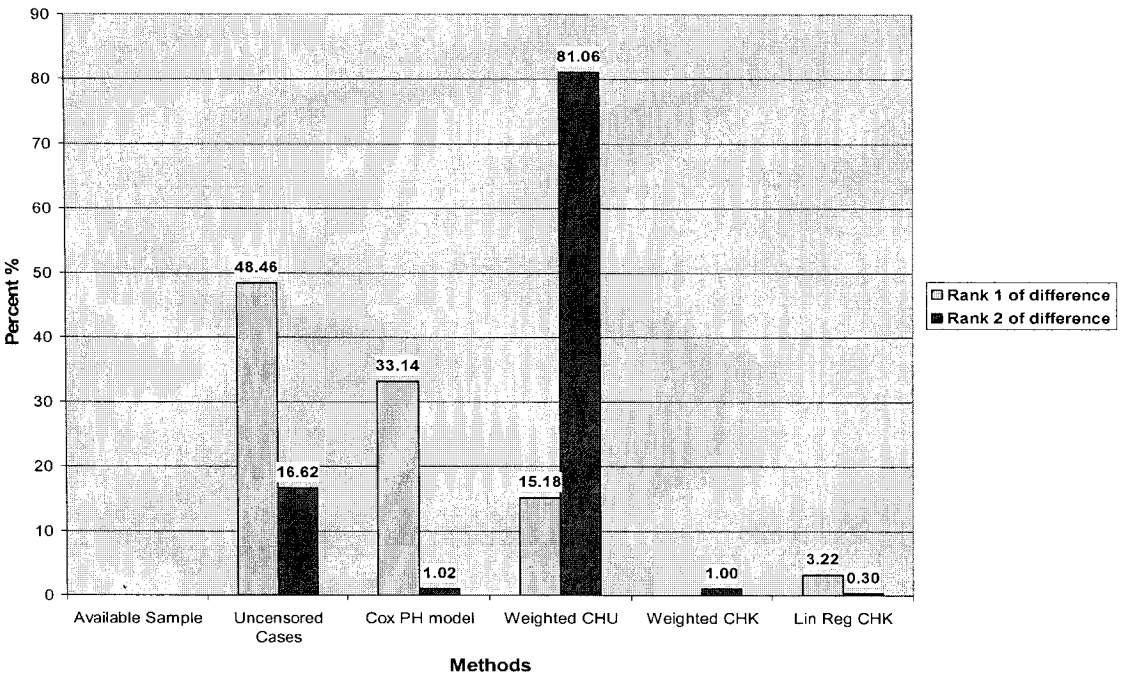


Figure 5: Frequency distribution of the rank of difference for six methods under 25% random censoring

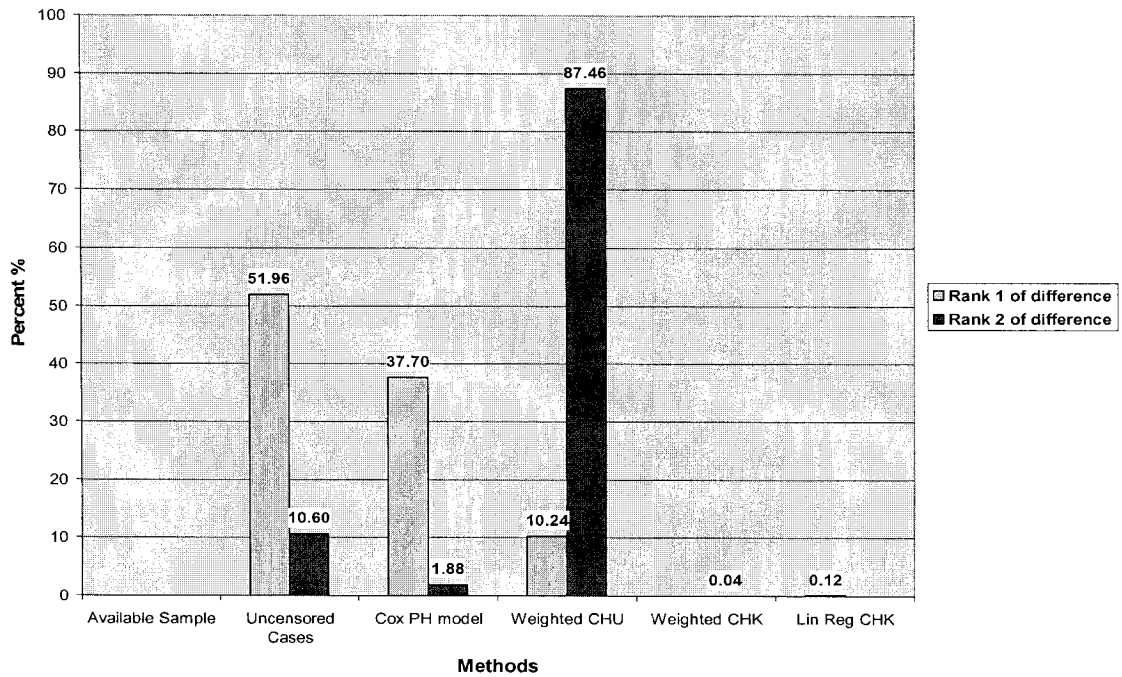
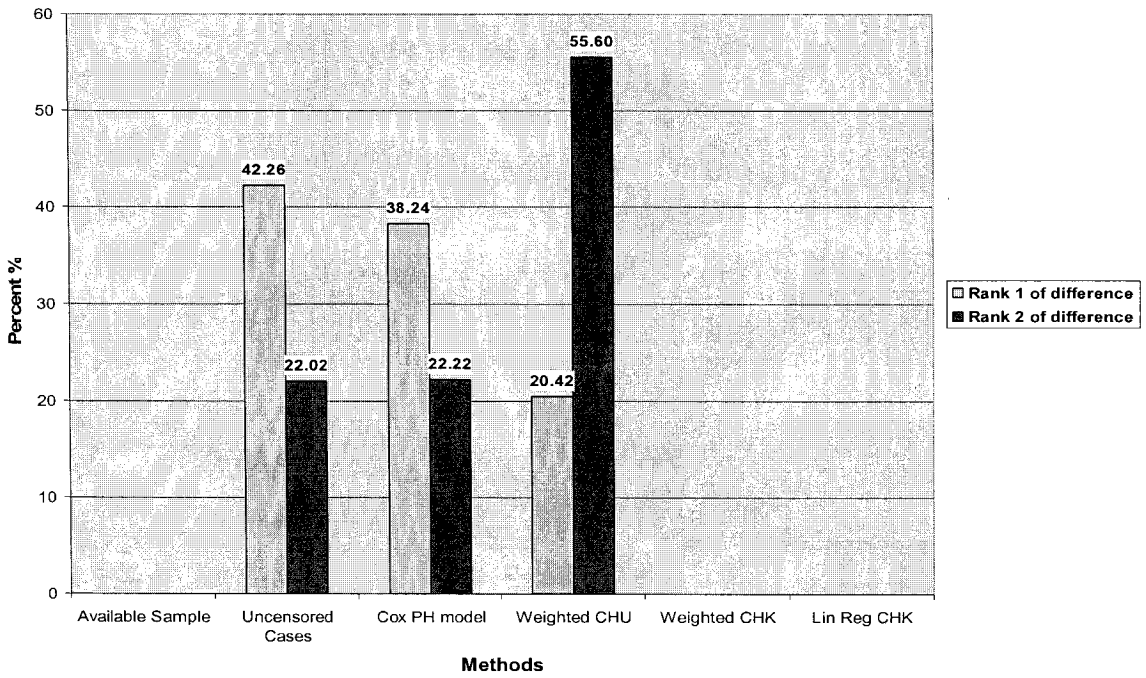


Figure 6: Frequency distribution of the rank of difference for six methods under 50% random censoring



End-of-study censoring

Similarly, in order to illustrate the accuracy of the six different methods measured by comparing the difference between each of 5000 sample mean total cost estimates and the "true" cost for the cohort which means bias under different levels of end-of-study censoring, the analysis was repeated for end-of-study censoring using three levels of censoring (10%, 25%, and 50%). Table 11 presents the frequency of the rank of difference for the six different methods across 5000 samples under the three levels of end-of-study censoring (10%, 25% and 50%). The frequencies of the rank shown in Table 11 were only used for illustration of the accuracy of the different methods, not for decision-making regarding the best estimator._For end-of-study censoring, the Uncensored cases method, Cox's PH regression model and the Weighted method

where cost histories were unknown, consistently, gave better accurate estimates of the "true" cost of £3609.33 than the other three methods across 5000 samples regardless of the underlying censoring level in this MC analysis. The method that gave the least accurate estimate of mean total costs was the Available sample method across 5000 samples.

Table 11: Frequency (%) of the rank of difference for six methods under three levels of end-of-study censoring

Censoring level	10%					
Rank of difference	**1**	**2**	**3**	**4**	**5**	**6**
Available sample method	—	—	0.82	0.24	2.18	96.76
Uncensored cases method	34.36	20.12	39.38	4.84	1.30	—
Cox's PH regression model	31.26	3.82	28.20	13.26	20.22	3.24
Weighted method, CHU	13.20	59.80	13.02	12.90	1.08	—
Weighted method, CHK	—	10.38	3.50	10.88	75.22	—
Lin's regression estimate CHK	21.16	5.88	15.08	57.88	—	—
Censoring level	**25%**					
Rank of difference	**1**	**2**	**3**	**4**	**5**	**6**
Available sample method	—	—	0.06	0.16	1.00	99.68
Uncensored cases method	48.60	7.16	42.92	0.96	0.36	—
Cox's PH regression model	20.12	46.84	20.58	8.82	3.56	—
Weighted method, CHU	21.44	34.44	28.32	7.24	8.32	0.24
Weighted method, CHK	—	5.70	4.56	2.08	87.66	—
Lin's regression estimate CHK	9.84	5.90	3.52	80.74	—	—
Censoring level	**50%**					
Rank of difference	**1**	**2**	**3**	**4**	**5**	**6**
Available sample method	—	—	—	—	—	100.00
Uncensored cases method	49.72	19.58	30.70	—	—	—
Cox's PH regression model	21.46	76.42	2.12	—	—	—
Weighted method, CHU	29.90	3.28	66.80	—	0.02	—
Weighted method, CHK	—	—	—	3.94	96.06	—
Lin's regression estimate CHK	—	—	0.02	96.06	3.92	—

In order to illustrate graphically the accuracy of the six different methods under each combination of end-of-study censoring and censoring level conditions, Figures 7~9 show the frequency distribution of the rank of difference for the six different methods under three levels of end-of-study censoring (10%, 25% and 50%) respectively. As shown in Figures 7~9, the Uncensored cases method, Cox's PH regression model and the Weighted method where cost histories were unknown, consistently, out-performed the other three methods (the Available sample method, the Weighted method where cost histories were known and Lin's regression estimate where cost histories were known) across 5000 samples under the three levels of end-of-study censoring through the MC analysis. The Available sample method, consistently, gave the least accurate estimate of mean total costs regardless of the underlying censoring level through the MC analysis. The findings from this MC analysis are consistent with the findings from the first two objectives. Thus, this confirms that the complex statistical calculations for end-of-study censoring can be reproduced through the MC analysis, so the methods were shown to be robust, and the consistency of the first two objectives for end-of-study censoring was determined through the MC analysis.

Figure 7: Frequency distribution of the rank of difference for six methods under 10% end-of-study censoring

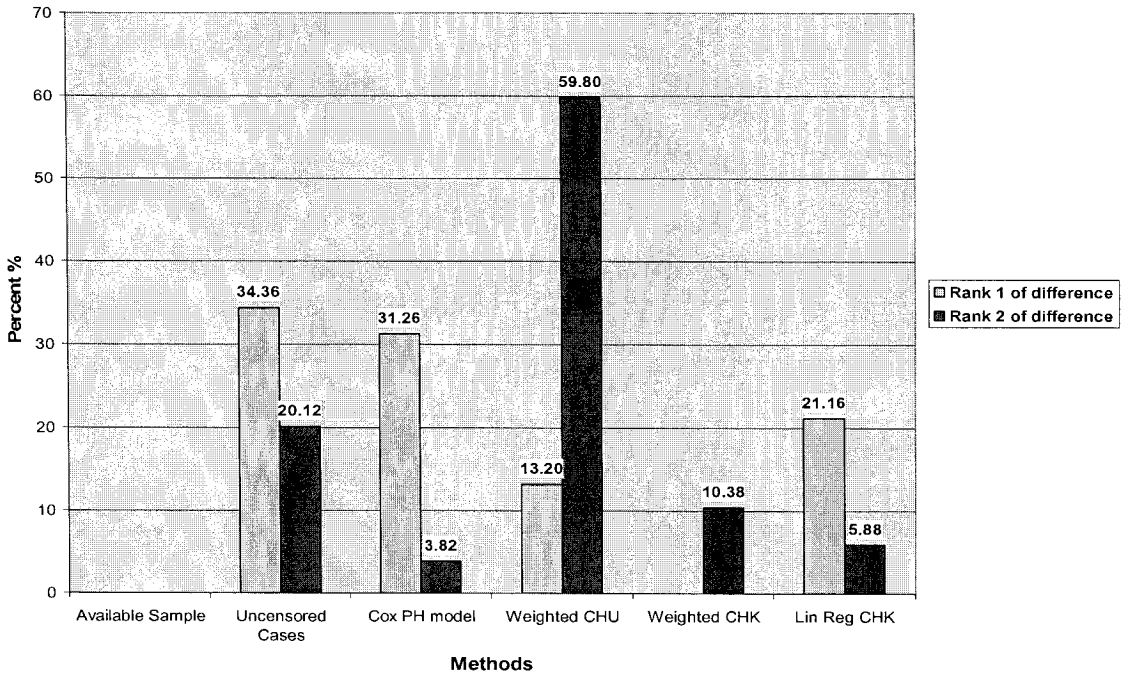


Figure 8: Frequency distribution of the rank of difference for six methods under 25% end-of-study censoring

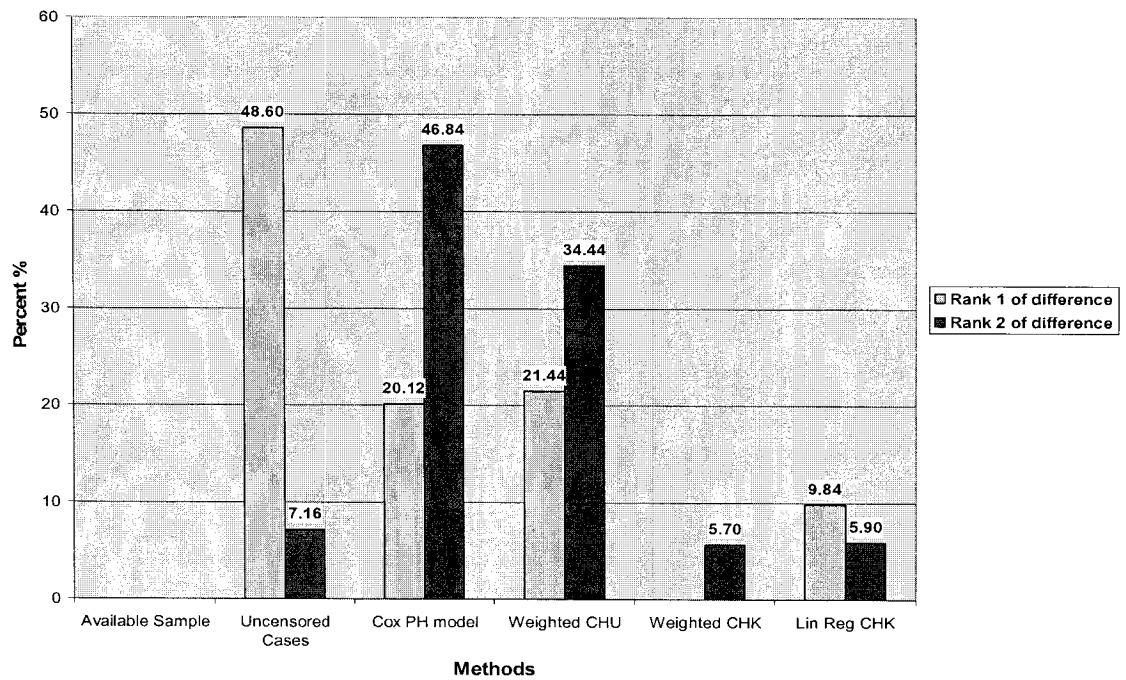
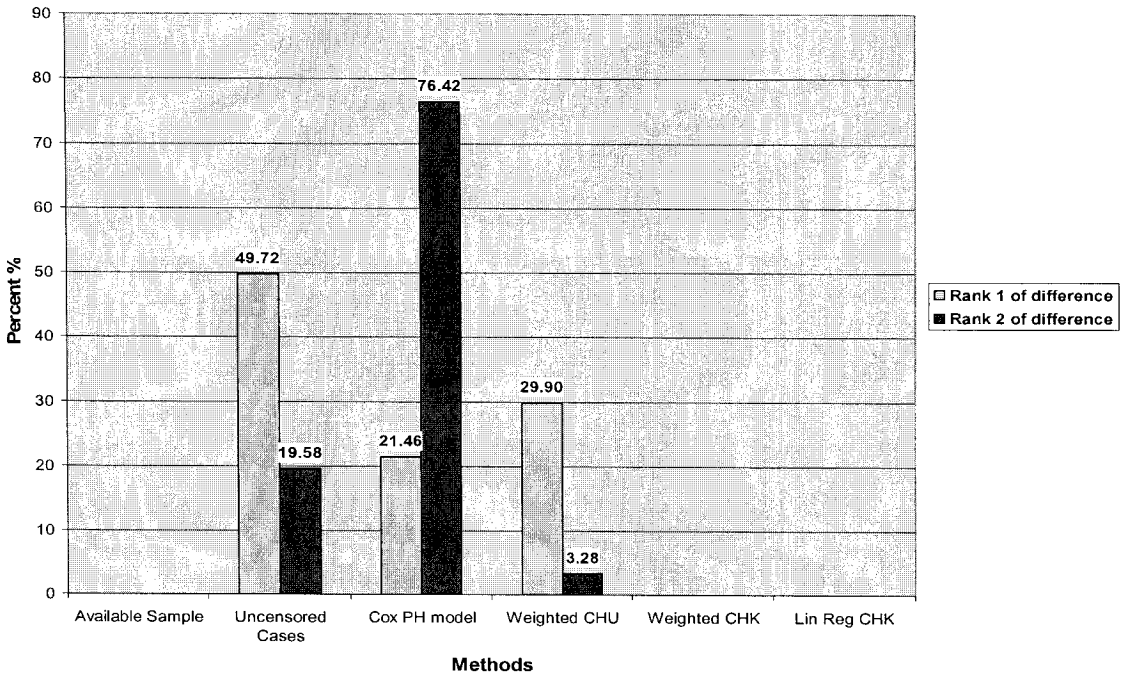


Figure 9: Frequency distribution of the rank of difference for six methods under 50% end-of-study censoring



4.2.3.3 Direction of the bias

Random censoring

In order to determine the direction of the biases of the six different methods from the "true" mean cost given random censoring across 5000 samples in this MC analysis, Figures 10~12 present graphical illustrations of the direction of the biases from the six different methods across 5000 samples under three levels (10%, 25% and 50%) of random censoring respectively. As noted from the Figures 10~12, Cox's PH regression model and the Weighted method where cost histories were unknown tended to overestimate the "true" cost, whereas the Available sample method, the Weighted method where cost histories were known and Lin's regression estimate where cost histories were known tended to underestimate the "true" cost across 5000 samples

under three levels of random censoring. However, the bias of the Uncensored cases method from the "true" mean cost can be of either direction across 5000 samples.

Figure 10: Direction of the biases from six methods under 10% random censoring

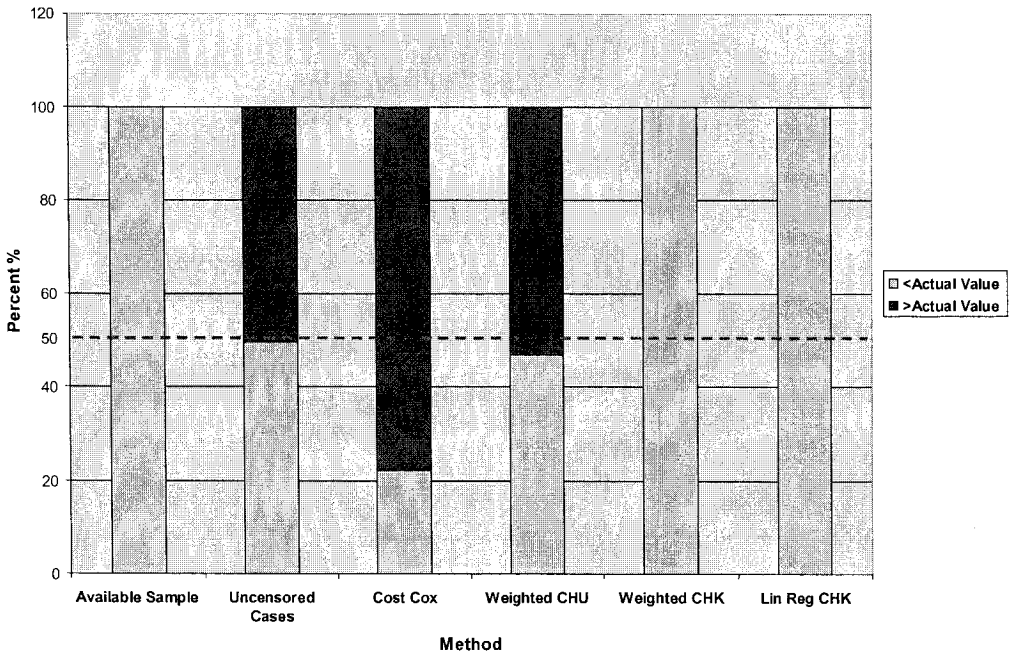


Figure 11: Direction of the biases from six methods under 25% random censoring

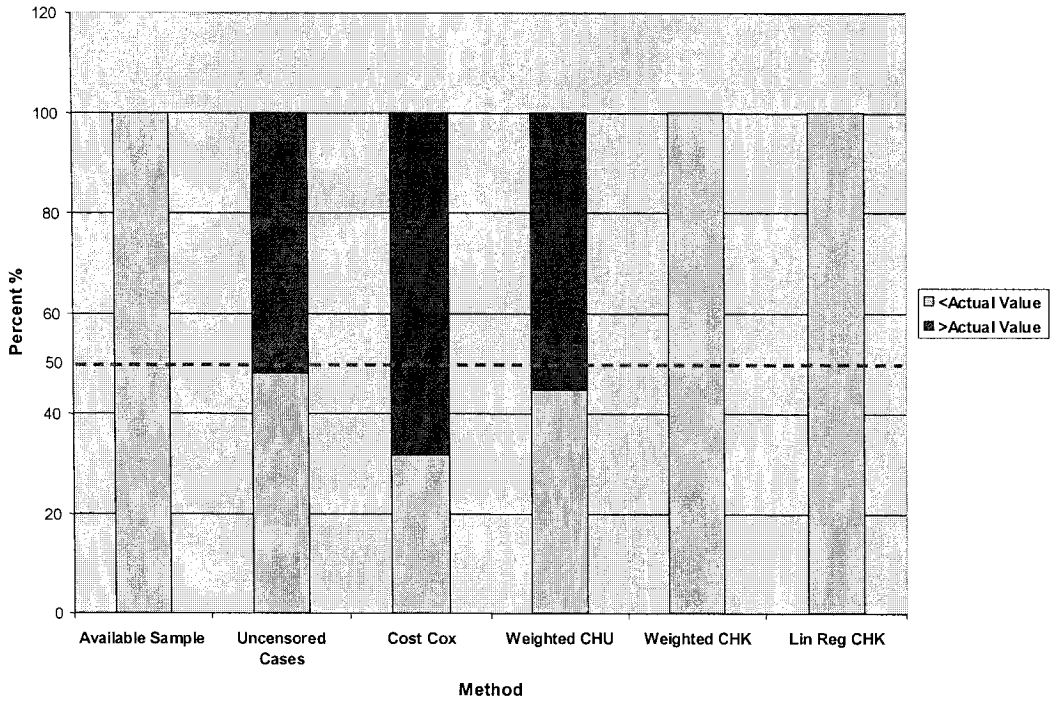
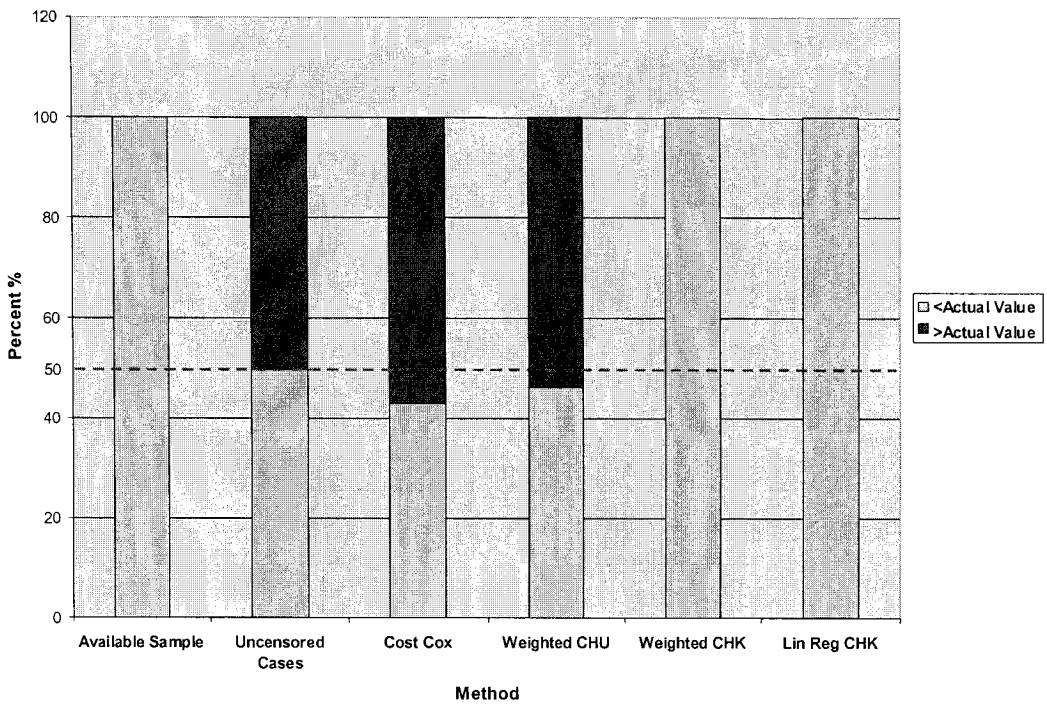


Figure 12: Direction of the biases from six methods under 50% random censoring



End-of-study censoring

Similarly, in order to determine the direction of the biases of the six different methods from the "true" mean cost under end-of-study censoring across 5000 samples in this MC analysis, Figures 13~15 present graphical illustrations of the direction of the biases from the six different methods across 5000 samples under three levels (10%, 25% and 50%) of end-of-study censoring respectively. As noted from the Figures 13~15, the estimates of Cox's PH regression model and the Weighed method where cost histories were unknown tended to be biased upward, whereas the estimates of the Available sample method, the Weighted method where cost histories were known and Lin's regression estimate where cost histories were known tended to be biased downward across 5000 samples under three levels of end-of-study censoring. However, the bias of the estimates fluctuated slightly less using the Uncensored cases method across 5000 samples. The findings are in agreement with the findings from random censoring.

Figure 13: Direction of the biases from six methods under 10% end-of-study censoring

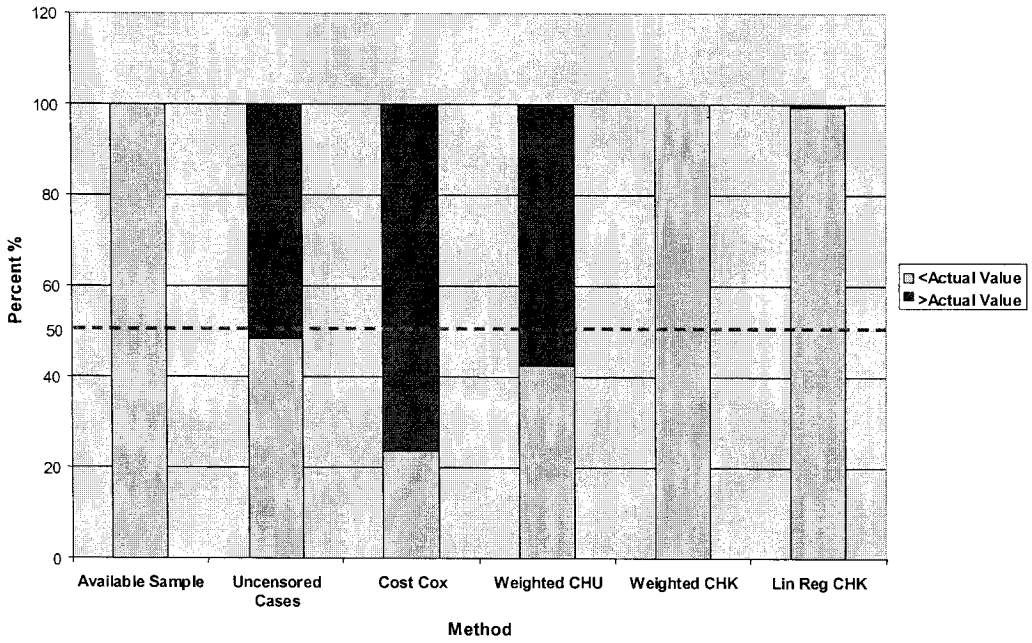


Figure 14: Direction of the biases from six methods under 25% end-of-study censoring

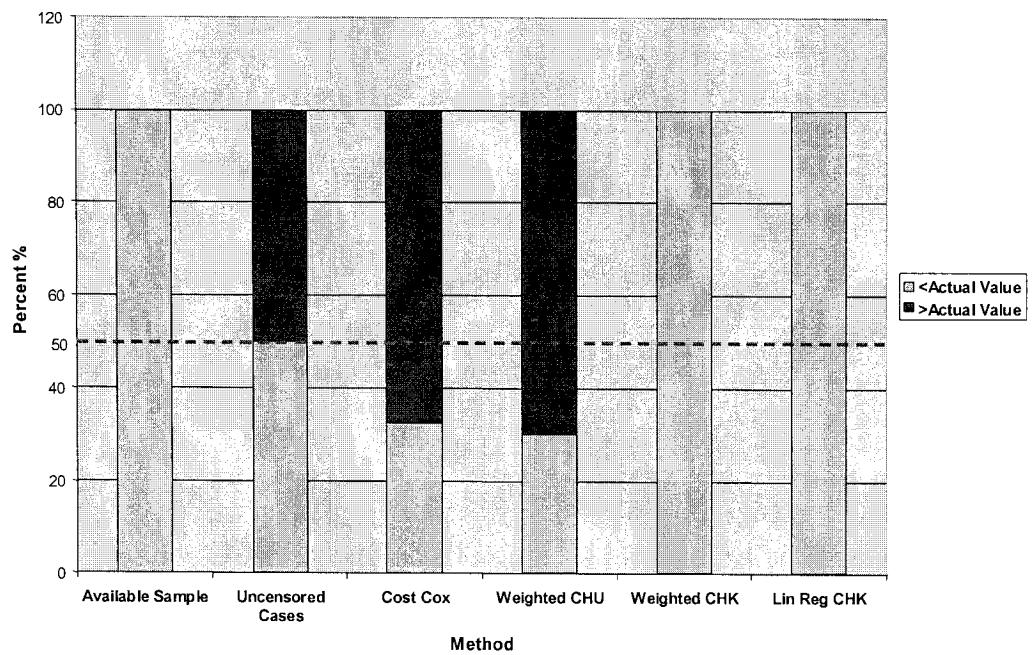
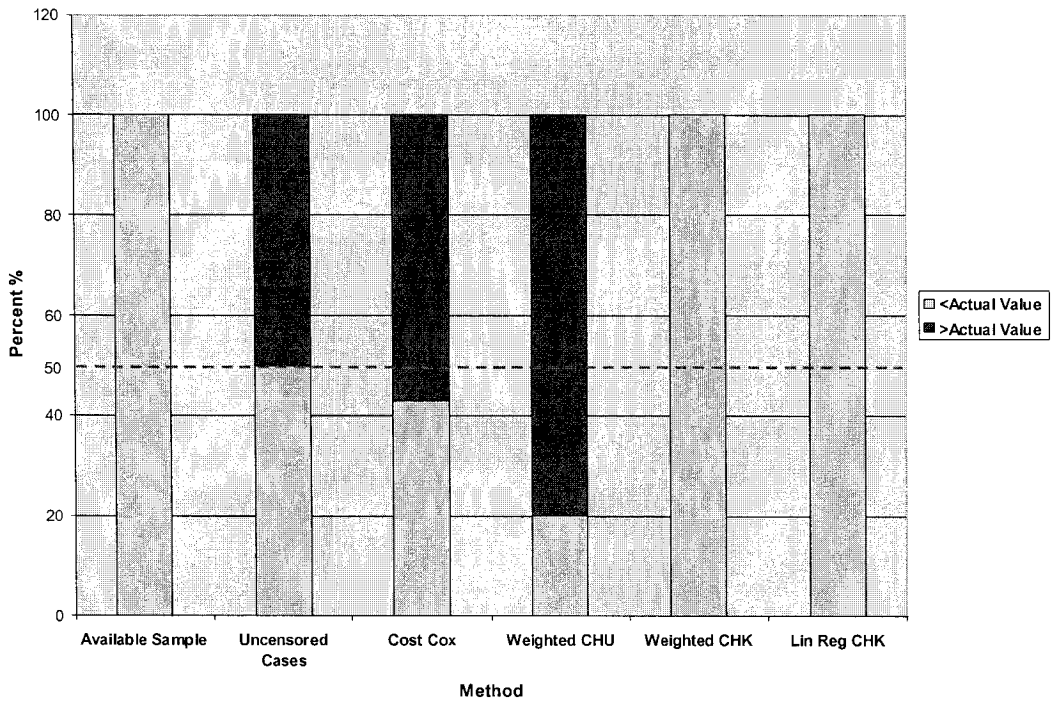


Figure 15: Direction of the biases from six methods under 50% end-of-study censoring



4.2.3.4 Analysis of MSE, variance and bias

Random censoring

The MSE is the expected squared deviation between the sample mean and the parameter it is desired to estimate. The MSE summarizes all the information about bias (accuracy) and variance of the estimator (precision) under study, so this is a criterion that incorporates both unbiasedness and efficiency. The purpose of the analysis of MSE, variance and bias was to evaluate the precision and the accuracy of the MC simulated estimates of mean total costs from the six different methods given different levels of random censoring. Table 12 shows the MSEs, variances and biases from the six different methods across 5000 samples under different levels (10%, 25% and 50%) of random censoring. As shown in Table 12, the MSEs were all lower for

the Uncensored cases method, Cox's PH regression model and the Weighted method where cost histories were unknown. The differences from the other methods were large enough to exclude sampling errors. These results suggested that we can improve the mean cost estimation by adjust for censoring using these three methods. The Available sample method which had the largest MSE, consistently, gave the least accurate estimate of mean total costs. For random censoring, the MSEs were smaller when the censoring level was lower, and they increased with an increase in the censoring level. This indicated that the variability of the estimates of mean total costs from the six different methods increased as the proportion of censored cases increased for random censoring, whereas the accuracy of the estimates of mean total costs from the six different methods decreased as the level of censoring increased.

Table 12: MSEs, variances and biases for six methods under different levels of random censoring

Censoring level	10%			
	N	MSE	Variance	$Bias^2$
Available sample method	5000	34524.58	1207.99	33316.59
Uncensored cases method	5000	1125.12	1125.03	0.08
Cox's PH regression model	5000	1811.99	1134.71	677.28
Weighted method, CHU	5000	1133.35	1128.01	5.34
Weighted method, CHK	5000	16402.59	1024.30	15378.30
Lin's regression estimate CHK	5000	11803.64	1079.44	10724.21
Censoring level	25%			
	N	MSE	Variance	$Bias^2$
Available sample method	5000	211600.40	2639.07	208961.40
Uncensored cases method	5000	3512.73	3509.99	2.74
Cox's PH regression model	5000	4246.45	3532.67	713.78
Weighted method, CHU	5000	3575.45	3531.28	44.17
Weighted method, CHK	5000	92739.66	2454.24	90285.43
Lin's regression estimate CHK	5000	80572.16	2600.90	77971.26
Censoring level	50%			
	N	MSE	Variance	$Bias^2$
Available sample method	5000	838699.90	4056.14	834643.70
Uncensored cases method	5000	10836.00	10836.00	0.00
Cox's PH regression model	5000	11102.76	10753.52	349.24
Weighted method, CHU	5000	11163.78	11076.20	87.58
Weighted method, CHK	5000	386210.18	4409.13	381801.04
Lin's regression estimate CHK	5000	359240.58	4714.99	354525.58

End-of-study censoring

The analysis of MSEs, variances and biases was repeated for end-of-study censoring. The purpose of the analysis was to determine the precision and the accuracy of the MC simulated estimates of mean total costs from the six different methods under different levels of end-of-study censoring. Table 13 shows the MSEs, variances and biases from the six different methods across 5000 samples under different levels (10%, 25% and 50%) of end-of-study censoring. As noted in Table 13, the Uncensored cases method, the Weighted method where cost histories were

unknown and Cox's PH regression model where the MSEs were lower, predicted better estimates of mean total costs than the other methods across 5000 samples under three levels of end-of-study censoring. The differences from other methods were large enough to exclude sampling errors. These results suggested that we can improve the mean cost estimation in the presence of censoring by using these three methods. The Available sample method with the largest MSE consistently predicted the least accurate estimate of mean total costs. For end-of-study censoring, the MSEs were smaller when the censoring level was lower, and they increased with an increase in the censoring level. This also gives the evidence that when the censoring level was high, there was more variability and less accuracy in the sample mean total cost estimates than there was when the censoring level was low. The findings are consistent with the findings from random censoring.

Table 13: MSEs, variances and biases for six methods under different levels of end-of-study censoring

Censoring level	10%			
	N	MSE	Variance	$Bias^2$
Available sample method	5000	13435.17	388.97	13046.19
Uncensored cases method	5000	1325.41	1325.14	0.27
Cox's PH regression model	5000	2029.65	1337.51	692.14
Weighted method, CHU	5000	1380.31	1339.62	40.69
Weighted method, CHK	5000	4714.84	425.05	4289.79
Lin's regression estimate CHK	5000	2593.66	469.30	2124.37
Censoring level	25%			
	N	MSE	Variance	$Bias^2$
Available sample method	5000	50806.56	703.99	50102.57
Uncensored cases method	5000	3018.81	3018.68	0.13
Cox's PH regression model	5000	3615.19	3005.46	609.73
Weighted method, CHU	5000	3911.93	3140.97	770.97
Weighted method, CHK	5000	15114.98	861.98	14253.00
Lin's regression estimate CHK	5000	11109.56	943.69	10165.86
Censoring level	50%			
	N	MSE	Variance	$Bias^2$
Available sample method	5000	959019.60	1972.91	957046.70
Uncensored cases method	5000	10885.23	10884.50	0.73
Cox's PH regression model	5000	11157.93	10862.99	294.94
Weighted method, CHU	5000	20657.00	11986.00	8671.00
Weighted method, CHK	5000	426634.94	2936.11	423698.83
Lin's regression estimate CHK	5000	408160.62	3241.29	404919.33

5. DISCUSSION

5.1 Overview

With the rapid escalation of costs of medical therapies, estimation of mean costs for these therapies is of increasing importance in order to allow economic evaluation of health care interventions. Despite the tremendous interest in the analysis of medical costs, there has been little progress in the development of formal statistical methods for such evaluation. A main difficulty lies in the incompleteness of the available data. Censored economic data are common in practice, because the issues related to the collection of economic data still are substantial. Thus more efficient methods should be adopted to estimate the mean total costs to adjust for censored cost data. In this present study, existing strategies for estimating the mean total costs in the presence of censoring were reviewed and assessed using simulated data sets generated from a real clinical data set. The goal was not to be exhaustive in reviewing this area, but to compare the performance of the different methods for estimating the mean total costs to adjust for censored cost data across different censoring mechanisms and levels of censoring, and to encourage the consideration of more efficient methods.

In this present study, the findings suggest that the following three methods for estimating the mean total study costs given censoring, consistently, gave the mean total costs more accurately than the other methods regardless of the underlying censoring mechanism and censoring level: the Uncensored cases method, Cox's PH regression model and the Weighted method where cost histories were unknown. If it was assumed that cost histories were collected, Lin's regression estimate where cost histories were known and the Weighted method where cost histories were known tended to give better

estimates of mean total costs. The method that gave the least accurate estimate of mean total costs was the Stratified Cox model where the assumption of independence between costs and the censoring mechanism was violated.

As expected, the accuracy of mean total cost estimates was found to depend not only on the method chosen to estimate the mean total costs but also on the proportion of censored cases. The estimates of mean total costs from all methods illustrated here deteriorated as the proportion of censored cases increased and the estimates of standard error became less precise as the level of censoring increased. This indicates that when the censoring level was high, there was more variability in the mean total cost estimates than there was when the censoring level was low. However, the biases of the different methods from the "true" mean estimate can be of either direction.

The findings from this present study were further studied through the Monte Carlo (MC) simulation analysis. The findings through the MC analysis indicate that the complex statistical calculations of mean total costs can be reproduced on 5000 alternate simulated data sets, so the methods illustrated here were shown to be robust. Thus the consistency of the findings from this study was determined.

5.2 Methodological approach

5.2.1 Mean total costs

In economic analysis, the most useful measure of central tendency is the mean, as it is necessary to allow for those patients who consume a disproportionate amount of resources (49). Medical cost distributions often exhibit a mass at zero representing nonusers of medical resources, and relatively small numbers of extremely high users

which tend to highly skew nonzero costs. The mean is intended to provide not only the overall distribution of costs in a population of patients, but also the extremely high users. These extremely high users are expected to be influential in decision-making about the adoption of alternative medical interventions. Therefore the mean is more efficient than the median which is much less sensitive to extremely high values. Thus, costs often are presented as some average, such as the average cost per case or the average cost per treatment. And comparisons of the average costs associated with the alternative therapies may lead to substantial cost reduction. For health policy analysis, information on mean total costs is also crucial. Due to the prominence of using mean total costs as measures in cost evaluation, most of the existing methods are developed to estimate the mean total study costs. Thus, the mean total costs were used as the measure of central tendency in this present study rather than the median costs.

5.2.2 Using the EXCEL system for conducting Monte Carlo studies

The EXCEL system has the combination of a powerful variety of built-in statistical procedures, mathematical functions, and the versatile programming capabilities associated with Visual Basic. This combination makes the EXCEL system ideal for conducting the Monte Carlo simulation analysis for this present study, especially the study related to complex statistical techniques. Such a combination of built-in statistical procedures and versatile programming capabilities makes it much more convenient for MCS researchers. In addition, the EXCEL system offers great flexibility in data generation, data transformation, obtaining and saving simulation results, etc. As noted by Wittwer et al., the completeness and the flexibility of the EXCEL system have convinced

us that currently no other system makes Monte Carlo research, especially research involving statistical techniques, easier and more efficient than the EXCEL system does (48).

5.2.3 Simple methods

Generally, the mean total costs are frequently estimated using the average total costs from all study patients, even though some of these costs are incomplete. This method, referred to as the Available sample method, was previously demonstrated to be always biased downward because the costs incurred after censoring times were not accounted for (10, 19). Results obtained in this study are consistent with previously published findings. The Available sample method was found to consistently underestimate the "true" cost regardless of the underlying censoring mechanism and level of censoring through the MC simulation analysis. For the Available sample method estimator in this study, the costs from the censoring times to the terminal time point (12 months) were ignored entirely, which resulted in substantial underestimation of the mean total costs unless all the censoring times were close to the terminal time point. However, the error associated with the Available sample method would not be too large, if number of observations with censoring was small and was balanced between the two treatment groups.

Alternatively, the mean total costs can be calculated from only the uncensored patients. The approach is called the Uncensored cases method. The findings in our present study suggest that the Uncensored cases method for estimating the mean total study costs given censoring produced the most accurate estimate compared to the other methods regardless of the underlying censoring mechanism and censoring level. These

results are not in agreement with the findings documented in previous studies (10, 19, 21). The previous studies reported that the Uncensored cases method was likely to be biased towards the costs of the patients with shorter survival times, because longer survival times were more likely to be censored. However, it is important to note that the findings in the previous reports were all obtained using the data sets from long-term studies (at least more than 12 months) which investigated long-term survivors. In contrast, the data analysis in this study only concentrated on the period up to 12 months from the start of treatment and focused on the patients' survival within 12 months from recruitment. As we know, the resulting mean cost estimator using the Uncensored cases method will be unbiased if all the patients who are under observation at the start of treatment have the same probability of being censored over the entire time period. This condition guarantees that the costs from uncensored cases are representative of all the costs from the trial patients over the entire time period. If censoring occurs during the entire time period, then the mean cost estimate tends to be driven by the costs of the patients who die early on in the study because, given the same censoring distribution, larger survival times are more likely to be censored. However, if the entire time period is narrow, the costs incurred over a small time interval (such as 12 months in this case) are stochastically similar between the censored and uncensored cases. Thus, the sample average of the costs from the patients who are under observation at the start of treatment and who are not censored over the entire time period provides a reasonable estimator of mean costs. So the corresponding estimator of mean costs obtained in this study was always less biased than the estimators documented in previous reports. Until recently, the Uncensored cases method is the default method in most statistical software packages. The

advantages of using this method are that it is easy to use and that it yields a same rectangular data file which would have resulted from a complete experiment. In practice, the Uncensored cases method may be an acceptable method for estimating the mean total costs in short-term studies (e.g. 12 months).

5.2.4 Survival analysis methods

In an attempt to adjust for the effects of censoring, the standard survival analysis technique (Kaplan-Meier estimator) has been applied to the problem of cost evaluation by treating costs as potentially right-censored survival times. This strategy, however, has been previously reported to be invalid unless all patients accumulate costs with a common rate function over time— which is certainly not the case for our study and unlikely to be true in most clinical trials (9, 10, 19). Results obtained in this analysis are generally in agreement with previously published findings. In practice, the cost functions vary among patients. Thus, a patient who accrues costs at higher rates tends to generate larger total costs at both the survival time and censoring time, which implies that the total cost at the survival time is positively correlated with the total cost at the censoring time. This correlation implies that "censored" total costs cannot be analyzed by the Kaplan-Meier analysis method which requires independence between the variable of interest and its censoring variable. Patients with lower costs will always be censored sooner than the patients with higher costs on the survival cost scale. Because the patients with lower costs are censored sooner, the Kaplan-Meier cost method estimator will overestimate the true mean total costs. However, we found that the Kaplan-Meier cost method gave an accurate estimate of the "true" cost under informative censoring where patients with ill health

were censored. This is probably because in our study the censored patients under informative censoring were the sicker patients in the cohort and there was a systematic tendency for the intervention to increase costs for the censored patients more than for the uncensored, thus incurring the highest costs. When these patients were censored before the high costs were incurred, the Kaplan-Meier cost method would not overestimate the "true" cost and gave the more accurate estimate of mean total costs within the constraints established by our data. Therefore, we need to extend the analysis to the other clinical studies to see if this observation is generalizable.

A natural extension to the Kaplan-Meier cost method is to apply Cox's proportional hazard regression model on a cost scale, in order to adjust for factors that may influence patient costs. In one published assessment of the predictive validity of Cox's proportional hazard regression model when applied to costs, Dudley et al. (20) found that Cox's PH regression model led to fairly accurate predictions of mean costs, median costs, and the proportion of patients with high costs within their sample of 155 patients undergoing coronary artery bypass graft surgery. Findings of the accuracy of Cox's PH regression model in this study are consistent with the literature, so we can lend credence to our results. Cox's PH regression model was detected to be applicable for the analysis of censored cost data in this study. The full distribution of costs can be analyzed without influence by extreme values. In essence, the extreme values pull the entire distribution towards the high cost patients; this phenomenon leads to overestimation of the high cost portion of the distribution. Since cost data often contain high cost extreme values, Cox's PH model appears better suited to analyses of cost data. Besides, Cox's PH model makes no assumptions concerning the distribution of costs, so the model can be used to fit

skewed cost distributions. An additional advantage of Cox's PH model is that it can handle censored observations. Longitudinal cost studies always face incomplete follow-up problem as patients are followed in time. Ability to address censored data may be useful in analysis of such longitudinal data. Thus, Cox's PH regression technique appears to be particularly helpful in analysis of skewed, censored data often encountered in the evaluation of hospital costs as our CHART study. Therefore, it appears that Cox's PH regression approach may be reasonable in some practical settings, though as noted previously, the assumption of independence between the censoring mechanism and costs strictly necessary for its validity may not be satisfied in the cost estimation. It remains to be seen whether general conditions can be derived under which the approach will provably yield a reasonable result in the costs setting.

Similarly, one previous study proposed the Stratified Cox model and found that a Stratified Cox regression model performed well in estimating the 2-year costs of Medicare patients admitted to the hospital for ischemic stroke (24). An attractive feature of the Stratified Cox model is that there are no assumptions relating to the independence of censoring. A proportional-hazard model is fitted within each stratum, so that the random variable of interest is the total costs incurred within a pre-specified time interval. In this way, the "one-to-one" link between the flow of costs and the flow of time is avoided. However, the Stratified Cox model was shown to consistently produce the least accurate estimate of mean total costs regardless of the underlying censoring mechanism and censoring level in our analysis. As noted, the censoring mechanism differs for this method as patient costs are censored within strata if they are incomplete for the full strata, and reasons for censoring include death or incomplete cost collection due to patient

censoring. Censoring the costs of patients who die is an option with survival methods, but it is still open to debate whether such censoring should be performed. One key concern is that a patient death in the Cox model is "informative censoring" which may introduce an element of bias in the cost estimation. Recently, one other study also reported that the Stratified Cox model produced the poorest estimate of mean total costs in a comparison with other methods for estimating the mean total costs in the presence of censoring (19).

5.2.5 Non-parametric methods

Lin et al. (10) acknowledge the difficulties of survival analysis methods and propose two methods which attempt to resolve these issues, called Lin's method CHK and Lin's method CHU. They introduce two estimators of mean costs under conditions of censoring which rely on the study period being partitioned into a number of subintervals such that censored observations occur at the boundaries of these intervals. Later studies have demonstrated that under such circumstances, these two approaches were found to give consistent estimators of mean costs and the associated variances were analytically derived (19, 44). However, results obtained in our analysis are in disagreement with previously published reports. We note that the validity of these two approaches depend on the pattern of the censoring distribution being of such a form to allow censoring times to correspond to the boundaries of the intervals of the partition — which is certainly not the case for our CHART study and unlikely to be true in most applications. If the censoring distribution is discrete, the interval boundaries can in theory be chosen to correspond to the possible censoring times and therefore the estimators are still going to be consistent. If the censoring distribution is continuous, the shorter the interval length,

that is the finer the partition of the study period, then the more unbiased the estimators. There is however a constraint associated with this point with reference to Lin's method CHU, which requires that the length of the intervals of the partition is such that it allows a reasonable number of deaths to be observed in each subinterval. In our present study, the hospital cost data were collected in 3, 6, 9 and 12 months after treatment started. Thus, a patient censoring in month 5 would have a recorded month 3 cost data but a censored cost at month 6. So it may not be possible however to ensure that the censoring times were confined to the boundaries of the intervals of the partition as required for consistency. In practice, there is no a priori reason however to expect censoring to conform to any such pattern and therefore in most applications consistency will be violated to some degree. As noted, the accuracy of Lin's method CHU estimate relies on the number of uncensored individuals in each subinterval and on the number who have complete costs at the largest observed time. A recent published study has evaluated the impact of varying the duration of analysis (45). It was found that the estimator of Lin's method CHK which uses intermediate individual cost histories appeared to be more accurate under a wide variety of conditions as opposed to the estimator of Lin's method CHU which is sensitive to the number of individuals contributing cost information. In general, the absolute mean cost estimates of Lin's two nonparametric methods decrease as the duration decreases since they are estimating the average cost over a shorter time period. This gives a strong indication that the issue of short-term (12 months in this case) is primarily responsible for Lin's two nonparametric estimators' poor performances in our results.

The limitations of Lin's two nonparametric methods have led to a further set of estimators proposed by Bang and Tsiatis (13), called the Weighted method CHK and the Weighted method CHU. Their estimators were shown to be consistent regardless of the censoring pattern and their variances were analytically derived in a recently published report (45). In our data analysis, the findings confirm that the Weighted method CHU for estimating the mean total study costs given censoring, consistently, gave the mean total costs more accurately than Lin's two nonparametric methods regardless of the underlying censoring mechanism and censoring level. If it was assumed that cost histories were collected, the Weighted method CHK tended to give better estimates of the mean total costs. Results are in agreement with recently published findings. In contrast, the set of estimators proposed by Bang and Tsiatis do not impose any restrictions on the distribution of censoring times. For instance, the idea underlying the partitioned weighted estimator is similar to that proposed by Lin et al., but the advantage of this method is that the consistency and asymptotic normality of the proposed estimator, unlike Lin's, does not depend on the choice of the partition or the discreteness of the censoring times. That is the asymptotic properties of this estimator are independent of the censoring pattern. Furthermore, it was detected that there was a similarity between estimators of Lin's method CHU and the Weighted method CHU in that they both use only the complete cost observations in estimating the mean costs. Lin's method CHU explicitly states that it relies on a "reasonable" number of deaths in each sub-interval of the partition and suggests a minimum number of five deaths in each subinterval. By contrast, the Weighted method CHU estimator does not rely on the pattern of the censoring distribution and therefore the small number of complete cost observations does not affect the estimates in

the same manner as in Lin's method CHU. However, caution should be exercised when applying all two weighted estimators in circumstances where there was heavy censoring in the tails of the distribution with small sample sizes which was common in real clinical trials. With regard to the two partitioned estimators of Lin's method CHK and the Weighted method CHK, they are similar in that they both divide the study period into subintervals and make use of individual intermediate cost history within each subinterval and in that they both use a weight to adjust interval costs for censoring. They are different both in the choice of this weight and in the interval costs that are adjusted by it. In Lin's method CHK, the weight is the Kaplan-Meier probability of survival to the start of the interval that adjusts estimates of mean costs in the interval, whereas the Weighted method CHK estimator uses the inverse of the probability of an individual not being censored evaluated at a given point in time to adjust individual observed costs in the interval.

Basically, Lin's two nonparametric approaches and the two weighted approaches all require the same amount of cost information, but the two weighted approaches are not restricted by the pattern of the censoring distribution and are therefore more general, they might be preferred to estimate the mean costs in most applications.

5.2.6 Regression methods

The cost of treating disease depends on patient characteristics, but standard tools for analyzing the clinical predictors of costs have not been well developed. Until recently, regression-based methods proposed by Lin are the only available methods that allow the analyst to control for confounding while addressing the problem of censoring (11). Lin's models provide expressions for coefficient estimates and their variances and are simple

extensions of the ordinary least-squares regression model. Results in our analysis are general consistent with Lin's findings. It was detected that Lin's regression estimate CHK tended to give the better estimate of mean total costs, if it was assumed that cost histories were recorded. The advantage of this method over Lin's regression estimate CHU is that there is an increase in the cost information being used by this estimator, as individuals who were treated as censored in Lin's regression estimate CHU, will be uncensored in some of the intervals of the partition in which their costs will contribute to the estimates. In contrast, individuals whose survival times were censored do not contribute any cost information to the estimator of Lin's regression estimate CHU. Thus, the estimator of Lin's regression estimate CHK may be more efficient when there is heavy censoring. Furthermore, Lin's regression approaches allow arbitrary censoring patterns which is more likely to be true in most applications and certainly the case for our CHART trial data, whereas that of Lin's nonparametric approaches require censoring to occur only at the boundaries of intervals.

However, there are two major problems with the application of ordinary least-squares to cost data. First, the assumptions underlying the ordinary least-squares may not hold for cost data. In particular, since there is a tendency for cost data to be highly skewed, neither the data nor the residuals from the regression may be normally distributed. The second major problem is that a few extremely high cost observations may dominate the results of ordinary least-squares regression. Thus, there are several reasons to consider alternative methods based on plausible parametric assumptions. Carides et al. (12) fit either parametric or nonparametric regression relationships between cumulative treatment cost and survival time, for uncensored patients. A particular advantage of this approach is the

gain in efficiency over purely nonparametric methods which results from the estimation and use of this relationship. Our present study demonstrates the same findings previously found by Carides. The two-stage estimator of Carides' regression method was shown to be close to Lin's method CHU estimator in our data analysis. And the standard error estimates were somewhat lower for the two-stage estimator, indicating an improvement in efficiency. In theory, Lin's method CHU estimator can be thought of as a special case of the two-stage estimator, where the total treatment cost is estimated by a step function rather than a smooth curve. This is the simplest form of nonparametric estimator for a mean function and was originally discussed by Tukey (46) in the context of exploratory data analysis. Furthermore, the two-stage estimator was found not to assume any knowledge about the stream of costs over time for individual patient. These cost histories may not be known in some applications, such as in a retrospective utilization study. However, a caveat with respect to the use of the parametric form of the two-stage estimator is the potential for misspecification of the functional form of the relationship or error structure and consequent bias and loss of efficiency.

Generally, use of the regression-based estimators may lead to further insights about the mechanism generating the data. For example, if censoring depends on age, then it is possible to incorporate this into the modeling of the censored cost data. If censored data exists then it is important to try to establish the reason why the data is censored before any modeling is carried out. This may help when deciding if any censoring assumptions are reasonable. In addition, the regression-based methods could model the cost-accrual and survival process over time, and so allow extrapolation beyond the length of the trial.

Fenn et al. (21) point out the importance of this to meet the long-term perspective needed for practical economic evaluation.

5.2.7 Censoring mechanisms

Methods for addressing censored cost data depend upon the censoring mechanism, that is, the reason for the data being censored. It is important to determine whether the censoring is related to the study variables or is predictable by study variables. Previous studies have pointed out that the appropriate strategy for analyzing censored cost data depends on mechanisms that give rise to censoring data (18, 19). Within the context of the analysis in our present study, it was found that the accuracy of mean total cost estimates depended not only on the method chosen to estimate the mean costs but also on the censoring mechanism. Results are generally consistent with previously published findings. It was shown that the accuracy of mean cost estimates under informative censoring was more complicated compared with random censoring and end-of-study censoring. Informative censoring occurs when patients drop out of a study for reasons that are related to the event of interest. However, it is difficult to distinguish between the informative censoring and the other censoring mechanisms. Because no censored cost data are observed for non-respondents, without external information there will be no way to judge whether the non-respondents censoring data are systematically different from the respondents observed data. Thus, the informative censoring may not be detected, and so the analysis with informative censoring requires some care. Some methods illustrated here rely on the assumption of independent censoring, such as Lin's two nonparametric methods and the two weighted methods. This assumption is clearly not satisfied if

patients are withdrawn from the study for health or cost related reasons. It is very difficult, if not impossible, to deal with such informative censoring even for the survival time distribution itself. Therefore, we must carefully examine the independent censoring assumption before applying the proposed methodologies to the other data sets. Further work is needed to investigate whether the censoring mechanism affects the size of the bias.

5.2.8 Levels of censoring

From theory it is expected that the degree of censoring will have a direct impact on the estimators' performance with this deteriorating as censoring level increases, although this impact will vary among the approaches. Several previous studies have pointed out that the accuracy of mean total cost estimates may be influenced by the proportion of censored cases (10, 13, 19, 44). However, all previously published reports were all based on one or two methods. The present study is the first to compare the twelve different methods to adjust for censored cost data under different levels of censoring to assess whether any particular method consistently performed well. It was determined that the estimates of mean total costs from all twelve methods deteriorated as the proportion of censored cases increased and the estimates of standard error became less precise as the level of censoring increased in our study. A direct consequence of heavy censoring is that when the censoring level is high, there is more variability in the underlying data set than there is when the censoring level is low. It may render the estimators increasingly unstable, especially with the small sample size. Thus, the caution should be exercised when apply all the methods illustrated here in circumstances where there is heavy

censoring with small sample size which is common in many real clinical trials, and also precisely the case in our simulated data sets.

5.3 Limitations

This study has a number of limitations:

First, it was conducted using simulated data sets generated from a single Randomized Controlled Trial (RCT) data. Although this would give more "real life" significance to the research, this approach may have the potential limitation due to the unique features of the single RCT presented in all simulations. Thus, the results may not be generalized to other applications. In general, the RCT data generation under a pre-determined set of parameters (e.g. variability in cost, co-morbidity) is theoretically a long and time-consuming process. This is so because you need to generate data from a distribution with all known characteristics, transform the data so that the data have desired shapes, and transform the data so that the simulated variables can be considered as samples randomly drawn from a population with all known inter-variable relationship pattern. It is obvious that this approach typically requires complex programming. Thus, our present study adopted only one single RCT in all simulations. Nevertheless, we believe that the proposed methods here for estimating the mean total costs in the presence of censoring should be generalized to other short-term trials similar to the CHART study.

Second, we compared various methods for predicting the mean total costs under each of the three censoring mechanisms (random, end-of-study and informative censoring) separately in this study. All methods except the Stratified Cox model produced favorable estimates of mean total study costs in the presence of each censoring mechanism. In

practice, censoring will not arise solely due to one type of censoring mechanism, since data may be censored due to random, end-of-study and informative causes all within one data set and it will not always be possible to distinguish what the underlying censoring mechanisms might be.

Third, several methods illustrated here for estimating the mean total costs in the presence of censoring (e.g. the Stratified Cox model, Lin's nonparametric methods, Lin's regression estimate CHK, the Weighted method CHK and Carides' regression method) all involved partitioning the study period into a number of subintervals. The choice of intervals for these methods will lead to the different degrees of accuracy, and the optimum number and lengths of intervals will rely on the choice of method. The previously published work has demonstrated that choosing the smaller interval lengths tended to give more accurate estimates for most of these methods (38).

Finally, for informative censoring, the proportion of censored cases (8%, 15%, 21%) depended on the choices of level of MET scores at which the data was censored. Although the estimates of mean total costs were very reasonable under informative censoring, varying these choices may result in the different degrees of accuracy among the methods.

5.4 Conclusions and future work

Despite the limitations associated with this study, the present analysis has identified the methods whose performance is deemed satisfactory for estimating the mean total costs regardless of the underlying censoring mechanism and level of censoring. Consequently, their application to the analysis of censored cost data is appropriate when

the estimates of mean total costs over the study period are sought. Furthermore, we offer the following recommendations on appropriate estimation of mean costs in the presence of censoring.

1. In practice, the Uncensored cases method, Cox's proportional hazard regression model and the Weighted method where cost histories were unknown may be appropriate methods for estimating the mean total costs in the presence of censoring in short-term studies (not more than 12 months). Under similar situations as our CHART study, the three proposed methods performed more than adequately. However, the proposed methods may not be suitable for estimating the mean total costs to adjust for censored cost data in the long-term studies. Research is needed on other longer-term data sets to assess the performance of these proposed methods and whether the results shown here using the CHART study can be generalized.
2. When only total costs are available on each patient, the Weighted method CHU is recommended. However, more efficiency can be obtained from having more information on cost accrual, and we recommend that costs per patient in each of a number of intervals should always be collected in trials with censoring. If enough cost histories are collected, Lin's regression estimate CHK and the Weighted method CHK are then recommended.
3. Other approaches were not identified to give the accurate estimates of mean total costs in our present study, but may be more efficient in long-term studies. However, future work is needed to assess whether these approaches can perform well in longer-term data.

4. When interest extends however beyond the length of the trial or when questions regarding the effect of covariates on cost occur, parametric models become a necessary alternative. It is clearly important that such parametric models make adjustment for censoring. However, we are not aware of any general work of this kind in the literature, but suggest that this is an emerging area for research.

With respect to the limitations discussed above, some future work will be needed. First, in order to avoid the unique features of the single RCT in all simulations, we need to extend the analysis here to the simulations under various scenarios to assess whether the results shown here using the CHART study can be generalized. Second, the future research needs to extend to other data sets to explore what might happen when the underlying censoring mechanism is unknown. Third, for the methods which involve dividing the study period into a number of subintervals, we are in process of future work to determine the optimal choice of interval lengths for each method, and to determine how this might vary across different data sets. Finally, we plan to explore how varying the choices of level of MET scores varies the accuracy of mean total costs for informative censoring in a further work.

REFERENCE

1. Drummond MF, O'Brien B, Stoddart GL, et al. Methods for the Economic Evaluation of Health Care Programs. 2nd ed. Oxford University Press: Oxford; 1997.
2. Hlatky MA, Rogers WJ, Johnstone I, et al. Medical care costs and quality of life after randomization to coronary angioplasty or coronary bypass surgery. N Engl J Med. 1997; 336: 92-9.
3. Mark DB, Hlatky MA, Califf RM, et al. Cost effectiveness of thrombolytic therapy with tissue plasminogen activator as compared with streptokinase for acute myocardial infarction. N Engl J Med.1995; 332:1418-24.
4. Gage BF, Cardinallli AB, Albers GW, et al. Cost-effectiveness of warfarin and aspirin for prophylaxis of stroke in patients with nonvalvular atrial fibrillation. JAMA. 1995; 274:1839-45.
5. Warrant JL, Feuer E, Potosky AL, et al. Use of Medicare hospital and physician data to assess breast cancer incidence. Med Care 1999; 37:445-456.
6. Newton KM, Wagner EH, Ramsey SD, et al. The use of automated data to identify complications and comorbidites of diabetes: a validation study. J Clin Epidemiol 1999; 52:199-207.
7. Fisher ES, Whaley FS, Krushat WM, et al. The accuracy of Medicare's hospital claim data: progress has been made, but problems remain. Am J Public Health 1992; 82: 243-248.
8. Duan N. Smearing estimate: A nonparametric retransformation method. J Am Stat Assoc 1983; 78:605-610.

9. Hallstrom AP, Sullivan SD. On estimating costs for economic evaluation in failure time studies. Medical Care 1998; 36(3):433-434.

10. Lin DY, Feuer EJ, Etzioni R, et al. Estimating medical costs from incomplete follow-up data. Biometrics 1997; 53:419-434.

11. Lin DY. Linear regression analysis of censored medical costs. Biostatistics 2000; 1:35-47.

12. Carides GW, Heyse JF, Lglewicz B. A regression-based method for estimating mean treatment cost in the presence of right-censoring. Biostatistics 2000; 1(3):299-313.

13. Bang H, Tsiatis AA. Estimating medical costs with censored data. Biometrika 2000; 87(2):329-343.

14. Coyle D., Drummond M.F. Costs of conventional radical radiotherapy versus continuous hyperfractionated accelerated radiotherapy (CHART) in the treatment of patients with head and neck cancer or carcinoma of the bronchus. Clinical Oncology 1997; 9:313-321.

15. Adams M, McCall N, Gray D, et al. Economic analysis in randomized control trials. Med Care 1992; 30: 231.

16. Hull RD, et al. Cost-effectiveness of primary and secondary prevention of fatal pulmonary embolism in high-risk surgical patients. Canadian Medical Association Journal 1982; 127: 990.

17. Lawless JF. Statistical Models and Methods for lifetime Data. 2nd ed. New York: Wiley; 2003.

18. Rubin DB. Inference and missing data. Biometrika 1976; 63: 581-592.

19. Young T, Buxton M. A comparison of methods to adjust for censored cost data under differing censoring mechanisms. Presented at the health economics study group meeting, Canterbury; July 9~11, 2003.

20. Dudley RA, Harrell FE Jr., Smith LR, et al. Comparison of analytic models for estimating the effect of clinical factors on the cost of coronary artery bypass graft surgery. J Clin Epidemiol. 1993; 46:261-71.

21. Fenn P, McGuire A, Phillips V, et al. The analysis of censored treatment cost data in economic evaluation. Med Care 1995; 33: 851.

22. Hiatt RA, Quesenberry CP, Selby JV, et al. The cost of acquired immunodeficiency syndrome in Northern California: the experience of a large prepaid health plan. Arch Intern Med 1990; 150:833-38.

23. Quesenberry C.P., Fireman B., Hiatt R.A., et al. A survival analysis of hospitalization among patients with acquired immunodeficiency syndrome. Am J. of Public Health 1989; 79:1643-1647.

24. Lipscomb J, Ancukiewicz M, Parmigiani G, et al. Predicting the cost of illness: A comparison of alternative models applied to stroke. Medical Decision Making 1998; 18 (Supplement):S39-S56.

25. Zhao H, Tsiatis A. A consistent estimator for the distribution of quality adjusted survival time. Biometrika 1997; 84(2):339-348.

26. Califf RM, Harrel FE, Lee KL, et al. The evaluation of medical and surgical therapy for coronary artery disease: A 15-year perspective. JAMA 1989; 261: 2077-2086.

27. Fisher LD, Kennedy JW, Davis KB, et al. Association of sex, physical size, and operative mortality after coronary artery bypass surgery in the coronary artery surgery study (CASS). J Thorac Cardiovasc Surg 1982; 84:334-341.

28. Merriam-Wester, Inc. Merriam-Webster's Collegiate Dictionary.10th ed. Springfield, MA: Merriam-Webster, Inc; 1994.

29. Reuven Y. Rubinstein. Simulation and the Monte Carlo method. New York: A John Wiley & Sons, Inc; 1981.

30. Gill RD. Censoring and Stochastic Integrals, Mathematical Centre Tracts 124. Amsterdam: Mathematisch Centrum; 1980.

31. Fleming TR, Harrington DP. Counting Process and Survival Analysis. New York: Wiley; 1991 Ch5.

32. Andersen P.K., Borgan Ø., Gill R.D. et al. Statistical Models based on Counting Processes. New York: Springer; 1993.

33. Horvitz DG, Thompson DJ. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 1952; 47:663-685.

34. Koul H, Susarla V, Van Ryzin J. Regression analysis with randomly right-censored data. The Annals of Statistics 1981; 9:1276-1288.

35. Robins JM, Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. In AIDS Epidemiology: Methodological Issues, Eds N.P. Jewell, K. Dietz and V.T. Farewell 1992, pp.297-331. Boston, MA:Birkhauser.

36. Lin DY, Ying Z. A simple nonparametric estimator of the bivariate survival function under univariate censoring. Biometrika 1993; 80:573-581.

37. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986; 73:13-22.

38. Eubank RL. Nonparametric Regression and Spline Smoothing. New York: Marcel Dekker; 1999.

39. Jennrich RI. Asymptotic properties of non-linear least squares estimators. The Annals of Mathematical Statistics 1969; 40:633-643.

40. Saunders MI, Dische S, Barrett A, et al. Randomised multicentre trials of CHART vs. conventional radiotherapy in head and neck and non-small-cell lung cancer-an interim report. Br J Cancer 1996; 73:1455-62.

41. Zigmond A., Snaith R. The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavia 1983; 67:361-370.

42. De Haes H., Van Knippenberg F, Neijt J. Measuring the psychological and physical distress of cancer patients: Structure and application of the Rotterdam Symptom Checklist. British Journal of Cancer 1990; 62:1034-1038.

43. Briggs AH, Goeree R, Blackhouse B, et al. Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for Gastro-Esophageal Reflux Disease. Medical Decision Making 2002; 22: 290-308.

44. O'Hagan A., Stevens J. On estimators of medical costs with censored data. J. of Health Economics 2004; 23:615-625.

45. Raikou M., McGuire A. Estimating medical care costs under conditions of censoring. J. of Health Economics 2004; 23:443-470.

46. Tukey J. Curves as parameters and touch estimation. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics 1961; 1:681-694.

47. Drummond MF, Menzin J, Oster G. Methodological issues in economic assessments of new therapies. The case of colony-stimulating factors. Pharmacoeconomics 1994;6(Suppl 2):18-26.

48. Wittwer, J.W., Monte Carlo Simulation in Excel: A Practical Guide" From Vertx42.com, June 1,2004.