WORKPLACE PROJECT PROTFOLIO

Submitted in accordance with the requirements for the Masters of Biostatistics

(Biostatistics Collaboration of Australia)

Project A

Evaluation of Pain Management Education Strategies in Cancer Patients

Project B

Comparison Hazard Ratios between

Cox Regression Model (the Noninterval-Censored Proportional Hazards Model)

and Interval-Censored Proportional Hazards Model

Huijun He

(SID 200338175)

The University of Sydney

5 September 2007

PREFACE

Student's Role

Overall contribution and value added to projects

My involvement was to extend a sound understanding of the theory and methods used in the projects. Both projects provided me a balance between theory and practical issues. I prepared the data for analyses, conducted the statistical analyses, presented and interpreted the results, and wrote up the reports. Under the supervision at each stage of the projects, my questions and feedback were adequately addressed.

Reflections on Learning

Both projects provided good experience for me to build on what I had learnt through the BCA coursework units. Project A (the effect of intervention trial) provided me a good opportunity to obtain practical experience in the management and analysis of data using the statistical software STATA. Since this was my first encounter with clinical trail data and my first experience with a factorial design, I was encouraged to learn about both of these concepts and the most common approach for analysis (for a 2x2 factorial). Project B (a simulation exercise) provided a project to develop theoretical knowledge and required a sound understanding of survival data. This project compared two models in the analysis of time-to-event outcomes. The principle and method used in Project B extended the material covered in Survival Analysis (SVA) unit, and I found that writing programmes to generate the simulated data and the censored data was a challenge for me. This project afforded me with experience in the use of simulations to develop theory. I had a better understanding the characteristics of an exponential distribution of survival times with censored observations and the censoring mechanism with its influence on the model selection and interpretation. Overall, my knowledge and skills learnt in both projects were extended and the experience was invaluable.

Communication skills

During these projects I would discuss any problem and obtain advice from my supervisors. This meant that I had regular face-to-face meetings with both supervisors,

as well as communication via email.

Work patterns/planning

These projects were undertaken outside my normal working hours. Time management was a challenge to me. Most of time, I had weekly meetings to discuss the Project A progress with my supervisor. This project was broken into smaller tasks and regularly reviewed to provide clear planning objectives and continuously review the project progress. Project B started later than Project A, a whole idea was discussed at the beginning and developed during the time. However, my supervisor provided me extra time for discussion and necessary advice when it was required.

Statistical principles and statistical methods

I was keen to take on projects which gave me opportunity to further my understanding of statistical principles and methods. These two projects combined both practical issues and theoretical topics and consolidated almost of the theory learnt through the BCA units which I had taken.

Statistical computing

The statistical software package STATA was used in both projects. Although I have been using STATA in some BCA units, I was still a reasonably naive STATA user prior to the commencement of these projects. In particular, Project B was a challenge for me, as writing a programme to simulate data was not easy at the beginning and I felt frustrated. I had to read user book and tried to get information from the relevant topics. However, since undertaking these projects, I have become more familiar with STATA and feel comfortable now with its programming.

Consolidation of BCA knowledge

Both projects required the knowledge and skills gained through the BCA units. Without any background and previous experience in mathematics and clinical trials, I improved the learning process through these projects, which helped to ties some of the BCA units together. When I encountered problems, I tried to read the course notes to seek the

background knowledge and possible solutions (or at least directions for solutions). Project A involved regression modeling in a clinical trial setting, which used the knowledge and skill gained from several BCA units. Project B consolidated my understanding of survival data, which I had gained from SVA.

Teamwork

I was undertaking both projects directly supervised by the project supervisors only. There was no other opportunity to involve teamwork.

Ethical Considerations

There were no major ethical considerations for both projects. For Project A, the data was de-identified with the use of a unique subject ID for each patient with no personally identifying information. Since Project B involved simulation data from a chosen distribution, there was no actual patient data information was used.

PROJECT A

Project Title

Evaluation of Pain Management Education Strategies in Cancer Patients

Location and dates

NHMRC Clinical Trials Centre, University of Sydney

August, 2006-January, 2007

Context

A randomised trial was conducted to investigate if education materials are effective in reducing barrier behaviours against effective pain management for patients with cancer. Two educational interventions were assessed in a 2x2 factorial design - a video/DVD and a booklet. Patients were assessed on pain knowledge and attitudes, quality of life and anxiety. A questionnaire, known as the Barriers Questionnaire (BQ), was used to evaluate barrier attitudes to pain management.

Contribution of Student

The student was required to understand how to conduct the primary analysis of the Barrier Questionnaire data. The student carried out the related data management tasks that were defined by the project supervisor. The student also wrote a report describing the work and interpretation of results. All the calculations in this project were using STATA and the results were verified by SAS software.

Statistical Issues

- Selection of suitable outcome measurement
  (distribution, treat as continuous/categorical or convert to dichotomous)
- Factorial design issues
- Regression model including or excluding interaction

Declaration by Student

I declare that this project is evidence of my own work, with direction and assistance provided by my project supervisor, Peta Forder. This work has not been previously submitted for academic credit.

………………………………………………

Huijun He

………………………………………………

Date

Comments and Declaration by Project Supervisor

The declared contributions by this student are true and correct. To my knowledge, the involvement and effort of this student for this project is satisfactory for the requirements of a BCA Workplace Project.

………………………………………………

Peta Forder

………………………………………………

Date

Table of Contents

Acknowledgements

1.0 BACKGROUND

1.1 Background and Rationale

<u>*Background*</u>

Ferrell (1999) states that pain is an important problem for the patients with cancer. Pain arising from cancer affects many patients in a variety of ways and depends on severity and type of pain. Since pain is subjective, the knowledge and attitudes regarding effective pain management may have an important impact on the patient's quality of life and as such, patients should be involved in their own pain management strategy.

In this study, the Barriers Questionnaire (BQ) is used to assess obstacles in patients' attitude that may hinder effective pain management. The BQ is a 27-item self-report question designed to measure the extent to which patients perceive particular issues (namely addiction, tolerance, side effect, progression of disease, fatalism, injections, good patient behaviors and distraction of doctor) as potential barriers to pain management. Each issue is assessed through a number of questions, with a mean score obtained for each subscale issue and a total mean score evaluated over all questions on the Barrier Questionnaire. The Barriers Questionnaire was used by Ward (1993) to assess the extent to which patients have concerns about reporting pain and using pain medication. Ward found that there was an inverse relationship between education level and the total score measured from the Barriers Questionnaire, suggesting that pain education can be effective in improving the pain management in the patients with cancer. Additionally, higher BQ scores were related to higher pain levels, indicating changes in pain levels reflect in the BQ scores due to the impact of the pain education.

<u>*Rationale*</u>

The investigators for this study wished to determine if specific educational interventions would improve pain knowledge and attitudes, quality of life and anxiety in patients with pain due to advanced cancer compared with standard care. If there is significant

improvement on patients' perceptions of barriers towards effective pain management when educational strategies are employed, the educational interventions could be implemented for all cancer patients to optimize effect pain management.

1.2 Objectives and Aims

The aims for this project were to determine if an educational intervention improves perceptions regarding pain management compared with standard care, specifically with respect to:

a. an educational booklet and/or

b. an educational video/DVD

The primary outcome was measured through the Barrier Questionnaire (BQ), a validated instrument designed to measure knowledge and attitudes to pain. From this questionnaire, it is possible to obtain outcome measures for:

- Total BQ score (primary outcome)
- BQ subscales, including patient perceptions concerning:
  - addiction,
  - tolerance,
  - side effects,
  - progression of disease,
  - fatalism,
  - injections,
  - behavior of a 'good patient' and
  - distraction of doctors' attention from clinical management of the disease.

The Barrier Questionnaire was administered at baseline and then again two weeks and four weeks after baseline. The primary outcome was focused on the short-term responses (i.e., 2 weeks after baseline).

1.3 Trial Design

This study was designed as a randomized 2x2 factorial design with the following structure:

- Factor A : Booklet versus standard care
- Factor B : Video/DVD versus standard care

Standard care is currently defined as use of no educational interventions (e.g., no provision of booklets, pamphlets, video/DVDs, etc).

1.4 Patients

Patients were recruited through oncology and radiation oncology outpatient clinics within New South Wales. Patients were considered eligible for the trial if they satisfied the following criteria:

- metastatic or locally advanced cancer,
- cancer pain score greater than or equal to 2/10 in the last week,
- treated through an outpatient clinic,
- possessed proficient English skills,
- a survival prognosis of greater than one month (to allow for measurement of outcomes),
- receiving cancer treatment at participating hospitals,
- age greater than or equal to 18.

Once eligible patients were identified, they were invited to participate and received an information sheet. If agreeable, patients signed the consent form and completed the demographic and baseline questionnaires. The first set of questionnaires was administered in an outpatient clinic setting so that patients were provided with standard instructions and guidance for completing the forms. Subsequent questionnaires were able to be completed by the patient in their own home.

In this randomized study, there were 217 patients recruited and randomized into four treatment groups: (a) standard care; (b) video/DVD alone; (c) booklet alone; or

(d) video/DVD+ booklet.

There were 140 patients who completed questionnaires at both week 0 (baseline), week 2 and week 4, with an additional 16 patients completing the questionnaires at only week 0 and week 2. Therefore, the data of those 156 patients were assessed with respect to the short-term changes in total BQ score (primary outcome) and for the eight BQ subscale scores. There were 61 patients who were randomized but not included in the analysis due to missing follow-up data. These patients were evenly balanced between the four treatment groups and had reasonably similar baseline characteristics, suggesting no systematic bias due to treatment allocation for missing follow-up data.

2.0 METHODS

2.1 A randomized 2x2 factorial design

Theory

Steven Piantadosi (1997) states that "Factorial clinical trials are experiments that test the effect of more than one treatment using a design that permits an assessment of interactions among the treatments." When conducting a trial with a factorial design, there are two classes of effects that need to be considered, which include interactions between the factors versus the main effects of the factors tested. In a standard 2x2 factorial design, we consider the effects of factor A and factor B in the following context:

- Whether there is evidence of an interaction between the two factors, implying that the effect of factor A depend on the levels (presence/absence) of factor B (i.e., the combination of effects is multiplicative, which is more/less than the expected sum of the separate effects).

- In the absence of an interaction, whether there is evidence of an effect for factor A and/or factor B. If these two factors are independent, the combined effects of factor A and factor B will equal the simple sum of the separate effects of factor A and factor B (the combination of effects is additive).

We can apply linear modeling techniques to estimate the main effects of the factors and their interactions of factor A and factor B. The form of the linear model is presented as below:

$$E\{Y\} = \beta_0 + \beta_A X_A + \beta_B X_B + \beta_{AB} X_A X_B$$

Where X_A and X_B are representative for the individual factors A and B, $X_A X_B$ is the interaction term of factor A and factor B. The coefficients of β_A, β_B and β_{AB} represent the effects of factor A, factor B and the interaction effect of factor A and factor B on the outcome.

Trial Application

For our trial, we used regression models to examine the effects of the two factors

(booklet, video/DVD). To begin with, we examined whether there was evidence of an interaction effect between the two treatment factors. If there was no evidence of a significant interaction (at the 5% level), the main effects were addressed in another model omitting the inclusion of an interaction term.

As there were four treatment groups in this study (standard care; video/DVD alone; booklet alone; or video/DVD + booklet), the following binary indicators were used to code the two factors:

- When the treatment groups were video/DVD alone or video/DVD plus booklet, the binary indicator for Video/DVD was set to 1; when the treatment groups were booklet alone or standard care, it was coded as 0.
- When the treatment groups were booklet alone or video/DVD plus booklet, the binary indicator for Booklet was set to 1; when the treatment groups were video/DVD alone or standard care, it was coded as 0.

As a result, we fitted the following two models, including or excluding the interaction between booklet and video/DVD, and the results of these two models were assessed.

- Model 1: including interaction

$$E\{Y\} = \beta_0 + \beta_{(Video/DVD)} X_{(Video/DVD)} + \beta_{(Booklet)} X_{(Booklet)} + \beta_{(Video/DVD*Booklet)} X_{(Video/DVD)} X_{(Booklet)}$$

- Model 2: excluding interaction

$$E\{Y\} = \beta_0 + \beta_{(Video/DVD)} X_{(Video/DVD)} + \beta_{(Booklet)} X_{(Booklet)}$$

2.2 Selection of appropriate outcome for analysis

When applying standard linear regression models, it is assumed that the deviations of individual values around the mean value will be modeled as arising from a Normal probability distribution. In this study, the total BQ score and the corresponding BQ subscales at week 0 and week 2 were skewed, and as such, the error distribution was not normally distributed. This is a violation of one of the fundamental assumptions of the regression model and as such, any inference made from such model may be incorrect. However, it was observed that the change in BQ score between week 2 and week 0 were normally distributed (less skewed), with residuals following a Normal distribution after a linear model was imposed. As a result, it was decided to use the change in BQ scores (week 2 - week 0) as the outcome for the following analyses. An example of the distribution of total BQ scores at week 0, week 2 and its difference at (week 2 - week 0) is shown in Figure 2.2.1

Figure 2.2.1 The distribution of total BQ score at week 0, week 2 and its difference for (week 2 - week 0)
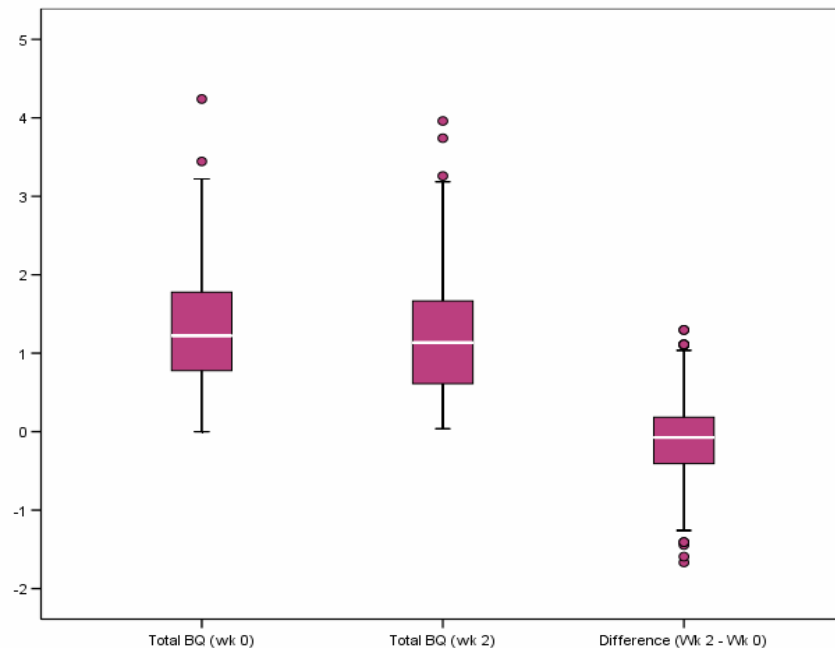
3.0 RESULTS

3.1 The summary of baseline characteristics of the treatment groups

The baseline data were collected once the eligible patients were identified and before the patients were randomly allocated to the treatment groups. If the randomization has been conducted correctly, the study groups should be similar in baseline characteristics (except for the play of chance). The baseline data included age, sex, type of cancer, the year first diagnosed, highest education level, marital status, country of birth and the first language are presented in Table 3.1.1. The patients had an average age of 62 years. There was higher percentage of female patients (57%) than male patients (42%) in the study. The major types of cancer were breast cancer (26%) and lung cancer (19%). The majority of patients (73%) were first diagnosed with cancer after 2000 year and 63% patients were married. There were 30% patients who had completed university/tertiary study, with around 76% patients born in Australia and 92% patients using English as their first language. The four treatment groups appeared reasonably similar in the distribution of these baseline characteristics, particularly when considering the small sample size of each group. Hence, in our study we did not include the baseline covariates in the model.

Table 3.1.1 The baseline characteristics of the treatment groups

Baseline Characteristics	Standard Care (n=38)	Booklet (n=36)	Video/DVD (n=37)	Video/DVD & Booklet (n=45)	Total (N=156)
Age: Mean (Std. Dev)	60.5 (12.0)	61.9 (12.3)	62.3 (15.3)	63.1 (11.6)	62.0 (12.7)
Median [Q1, Q3]	61 [52, 69]	63 [53, 72]	66 [51, 74]	63 [56, 71]	63 [54, 71]
[Min, Max]	[29, 88]	[34, 89]	[26, 87]	[36, 86]	[26, 89]
Sex: Female	22 (56%)	20 (56%)	25 (69%)	22 (49%)	89 (57%)
Male	16 (41%)	16 (44%)	11 (31%)	23 (51%)	66 (42%)
Type of Cancer					
Bowel	4 (10%)	2 (6%)	2 (6%)	3 (7%)	11 (7%)
Breast	9 (23%)	8 (22%)	14 (39%)	9 (20%)	40 (26%)
Genito-Urinary	3 (8%)	2 (6%)	3 (8%)	6 (13%)	14 (9%)
Prostate	5 (13%)	3 (8%)	2 (6%)	7 (16%)	17 (11%)
Skin	6 (15%)	6 (17%)	6 (17%)	4 (9%)	22 (14%)
Lung/Mesothelioma	6 (15%)	9 (25%)	5 (14%)	9 (20%)	29 (19%)
Other	4 (10%)	4 (11%)	3 (9%)	6 (13%)	17 (11%)
Year first diagnosed					
<2000 year	12 (31%)	10 (28%)	12 (33%)	7 (16%)	41 (26%)
≥2000 year	26 (67%)	26 (72%)	24 (67%)	38 (84%)	114 (73%)
Highest Education					
No certificate	7 (18%)	10 (28%)	8 (22%)	8 (18%)	33 (21%)
Junior certificate	8 (21%)	9 (25%)	6 (17%)	13 (29%)	36 (23%)
Senior certificate	8 (21%)	4 (11%)	3 (8%)	5 (11%)	20 (13%)
TAFE/Trade	3 (8%)	3 (8%)	5 (14%)	5 (11%)	16 (10%)
University/Tertiary	11(28%)	9 (25%)	14 (39%)	13 (29%)	47 (30%)
Marital Status					
Single	2 (5%)	2 (6%)	3 (8%)	1(2%)	8 (5%)
Married	24 (62%)	27 (75%)	20 (56%)	28 (62%)	99 (63%)
De facto	3 (8%)	1 (3%)	6 (17%)	4 (9%)	14 (9%)
Separated or divorced	5 (13%)	3 (8%)	3 (8%)	6 (13%)	17 (11%)
Widowed	4 (4%)	3 (8%)	4 (11%)	5 (11%)	16 (10%)
Country of Birth					
Australia	30 (77%)	28 (78%)	28 (78%)	33 (73%)	119 (76%)
Other	8 (21%)	8 (22%)	8 (22%)	11(24%)	35 (22%)
First Language					
English	36 (92%)	35 (97%)	33 (92%)	40 (89%)	144 (92%)
Other	2 (5%)	1 (3%)	3 (8%)	4 (9%)	10 (6%)

3.2 Assessment of the effect of video/DVD/booklet on the total BQ score and the BQ subscales changes between week 2 and week 0 (baseline)

To determine the effect of a video/DVD and/or booklet intervention on pain barrier

issues for the changes of the total BQ score and the BQ subscales of addiction, tolerance, side effect, progression of disease, fatalism, injections, good patient behaviors and distraction of doctor, a linear regression model was fitted, including and excluding an interaction term for the two factors (Video/DVD*Booklet). If there was little evidence of an interaction between the two factors (tested at the conventional 5% level), the model fitted with main effects only was then considered. The results from these analyses are shown in Table 3.2.1 and Table 3.2.2

Table 3.2.1 Model results for change in the total BQ score and BQ subscales (week 2-week 0) when including an interaction term for the two factors

Change in Outcome (week 2-week 0)	Covariate	Coefficient (95% CI)	P-value
Total BQ	Video/DVD	-0.08 (-0.36, 0.20)	0.57
	Booklet	0.02 (-0.26, 0.29)	0.91
	Video/DVD * Booklet	-0.0005 (-0.38, 0.38)	0.99
Addiction	Video/DVD	-0.78 (-1.36, -0.20)	0.009
	Booklet	-0.49 (-1.08, 0.09)	0.10
	Video/DVD * Booklet	0.87 (0.068, 1.68)	0.034
Tolerance	Video/DVD	-0.39 (-0.88, 0.11)	0.12
	Booklet	0.07 (-0.43, 0.56)	0.79
	Video/DVD * Booklet	0.33 (-0.36, 1.01)	0.35
Side effects	Video/DVD	-0.06 (-0.47, 0.35)	0.77
	Booklet	-0.03 (-0.45, 0.38)	0.88
	Video/DVD * Booklet	0.13 (-0.44, 0.70)	0.66
Progressive disease	Video/DVD	0.21 (-0.31, 0.73)	0.42
	Booklet	-0.07 (-0.58, 0.45)	0.80
	Video/DVD * Booklet	-0.51 (-1.23, 0.21)	0.16
Fatalism	Video/DVD	-0.16 (-0.55, 0.24)	0.43
	Booklet	0.04 (-0.35, 0.44)	0.84
	Video/DVD * Booklet	-0.10 (-0.65, 0.45)	0.72
Injections	Video/DVD	0.16 (-0.29, 0.61)	0.47
	Booklet	-0.22 (-0.66, 0.23)	0.34
	Video/DVD * Booklet	-0.04 (-0.65, 0.58)	0.91
Good patients	Video/DVD	0.14 (-0.32, 0.60)	0.55
	Booklet	0.02 (-0.44, 0.48)	0.93
	Video/DVD * Booklet	0.05 (-0.59, 0.69)	0.88
Distract Dr.	Video/DVD	0.08 (-0.38, 0.54)	0.73
	Booklet	0.08 (-0.38, 0.54)	0.73
	Video/DVD * Booklet	-0.12 (-0.76, 0.52)	0.72

Table 3.2.2 Model results for change in the total BQ score and BQ subscales (week 2-week 0) when excluding an interaction term for the two factors

Change in Outcome (week 2-week 0)	Covariate	Coefficient (95% CI)	P-value
Total BQ	Video/DVD	-0.08 (-0.27, 0.11)	0.41
	Booklet	0.01 (-0.17, 0.20)	0.88
Addiction	Video/DVD	-0.33 (-0.74, 0.08)	0.11
	Booklet	-0.03 (-0.44, 0.37)	0.87
Tolerance	Video/DVD	-0.22 (-0.56, 0.12)	0.21
	Booklet	0.24 (-0.11, 0.58)	0.17
Side effects	Video/DVD	0.005 (-0.28, 0.29)	0.97
	Booklet	0.04 (-0.25, 0.32)	0.81
Progressive disease	Video/DVD	-0.05 (-0.41, 0.31)	0.77
	Booklet	-0.33 (-0.69, 0.03)	0.07
Fatalism	Video/DVD	-0.21 (-0.48, 0.07)	0.14
	Booklet	-0.01 (-0.28, 0.26)	0.93
Injections	Video/DVD	0.14 (-0.16, 0.45)	0.35
	Booklet	-0.23 (-0.54, 0.07)	0.13
Good patients	Video/DVD	0.16 (-0.15, 0.48)	0.31
	Booklet	0.04 (-0.27, 0.36)	0.79
Distract Dr.	Video/DVD	0.02 (-0.30, 0.34)	0.91
	Booklet	0.02 (-0.30, 0.34)	0.91

For the short-term changes in total BQ score (week 2 - week 0) and the short-term changes in most of the BQ subscales (e.g., tolerance, side effect, progression of disease, fatalism, injections, good patient behaviors and distraction of doctor), there was little evidence of an interaction between video/DVD and booklet treatments (all p> 0.05 for the interaction terms, Table 3.2.1). Furthermore, even after excluding the possible interaction, the results indicated that there was little evidence to suggest a benefit of either video/DVD or booklet for patients' total BQ score or for the other BQ subscales

of tolerance, side effect, progression of disease, fatalism, injections, good patient behaviors and distraction of doctor (all $p>0.05$, Table 3.2.2).

However, for the change in the addiction BQ subscale (week 2 - week 0), there was evidence of an interaction between video/DVD and booklet treatment ($p= 0.034$), which indicates that these two factors (video/DVD and booklet) are not independent. Hence we need to compare the effect of the treatments against standard care (no intervention). It should be noted that some caution is required to interpret this conclusion in view of the multiple comparisons/tests that have been conducted. Keeping this in mind, the result is investigated further. Table 3.2.3 presents the results for each of the four treatment groups. The corresponding graph (Figure 3.2.1) indicates that the change in the addiction BQ subscale depends on the specific combination of treatment interventions (booklet alone, video/DVD alone, or the combination of booklet and video/DVD). From observation, if the subject received video/DVD treatment only (compared to standard care treatment), it appeared that the subject could obtain the most benefit in reducing the barrier perception of addiction. If the subject received booklet treatment only, there appeared to be a smaller reduction in the addiction subscale. It was good to have either video/DVD or booklet treatment, however, the combination of video/DVD and booklet appeared to have less of reduction in the addiction subscale when compared with video/DVD or booklet treatment alone.

Table 3.2.3 Descriptive statistics for the BQ Addiction subscale at week 0 and week 2

Treatment Group	Week 0 (95% CI)	Week 2 (95% CI)	Mean change for (week 2-week 0)(95% CI)	
Standard Care	1.577 (1.130, 2.033)	1.757 (1.316, 2.198)	0.180	(-0.222, 0.582)
Booklet	1.781 (1.334, 2.229)	1.467 (1.043, 1.891)	-0.314	(-0.727, 0.099)
Video/DVD	1.843 (1.324, 2.362)	1.241 (0.827, 1.677)	-0.602	(-1.007, -0.197)
Video/DVD & Booklet	1.355 (0.983, 1.728)	1.133 (0.773, 1.494)	-0.222	(-0.635, 0.191)

Figure 3.2.1 Changes of BQ subscale for Addition for (week 2 - week 0)
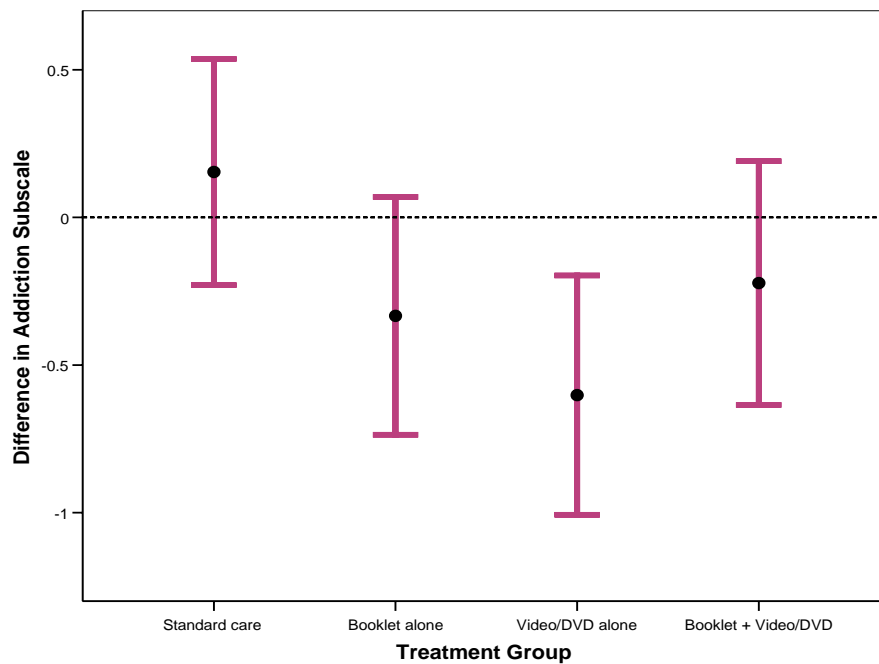
Three comparisons were constructed to formally assess this: a) standard care versus booklet alone; b) standard care versus video/DVD alone; c) standard care versus booklet and video/DVD together.

Compared to standard care treatment (no intervention), there is strong change in the BQ subscale of addiction for patients allocated to video/DVD treatment alone (p=0.009), indicating a considerable mean reduction in addiction anxiety of 0.78 when patients received an educational video/DVD intervention. Along similar lines, when compared to standard treatment of no educational intervention, there appeared to be a lessening in the anxiety score (mean change -0.49) when patients received an educational booklet alone, although this reduction was not statistically significant (p=0.10). Similarly, there was little evidence of a noteworthy reduction in the change of anxiety scores for

patients who received both the booklet and the video/DVD against patient who received no education inventions (p=0.15).

From these results (Table 3.2.3 and Figure 3.2.1), it would appear that the video/DVD is likely to an effective educational instrument with respect of attitudes toward drug addiction. This analysis suggests that the provision of an educational video/DVD will help to reduce patients' concerns about drug addiction for pain management, although there is insufficient evidence to support the same position with respect to the provision of an education booklet.

3.3 Visual assessment of trends in BQ score and the BQ subscales from baseline (week 0) to 4 weeks post-randomisation

Trajectory plots were made on a random selection of patients from each of the four treatment groups (using a quarter of the patients, selected at random). Figure 3.3.1 and Figure 3.3.2 display the data graphically for the trend in BQ score and the BQ subscale of addiction from week 0 to week 4, respectively, the individual lines connecting the repeated observations over time for each patient, as displayed for each of the four treatment groups. As can been seen from the plots, there is substantial variation between patients in each treatment group and there appears little trend of BQ score changes over time in each treatment group. The changes monitored over the 4-week period were not investigated with statistical models as the clinical investigators were principally interested in the short-term changes only.

Figure 3.3.1 Trajectory plot of total BQ score changes score over weeks (at 0,2 & 4) for each treatment group
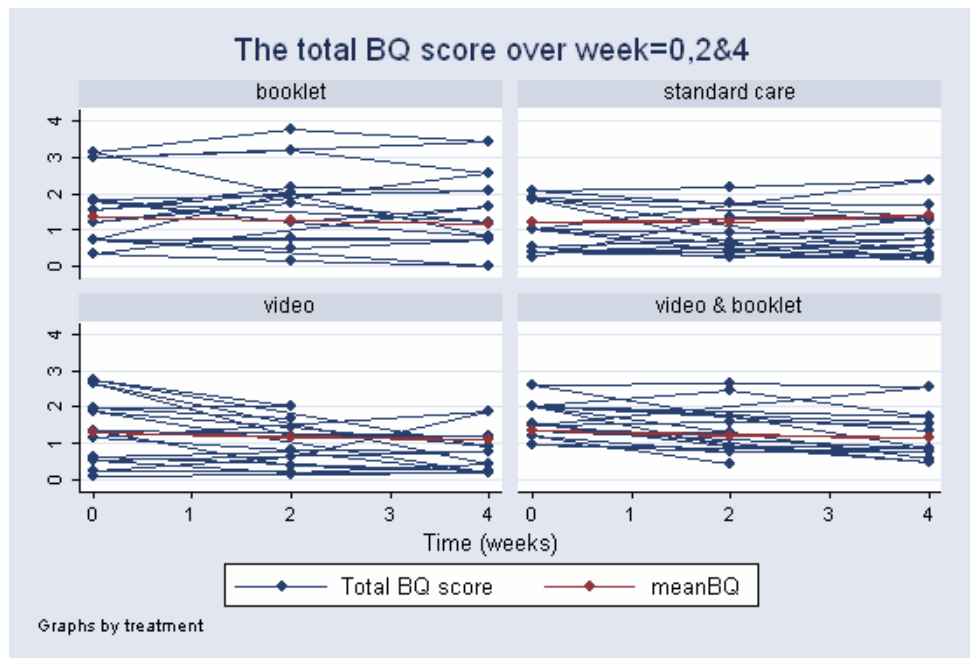
Figure 3.3.2 Trajectory plot of BQ subscale changes for addiction over weeks (at 0,2 & 4) for each treatment group
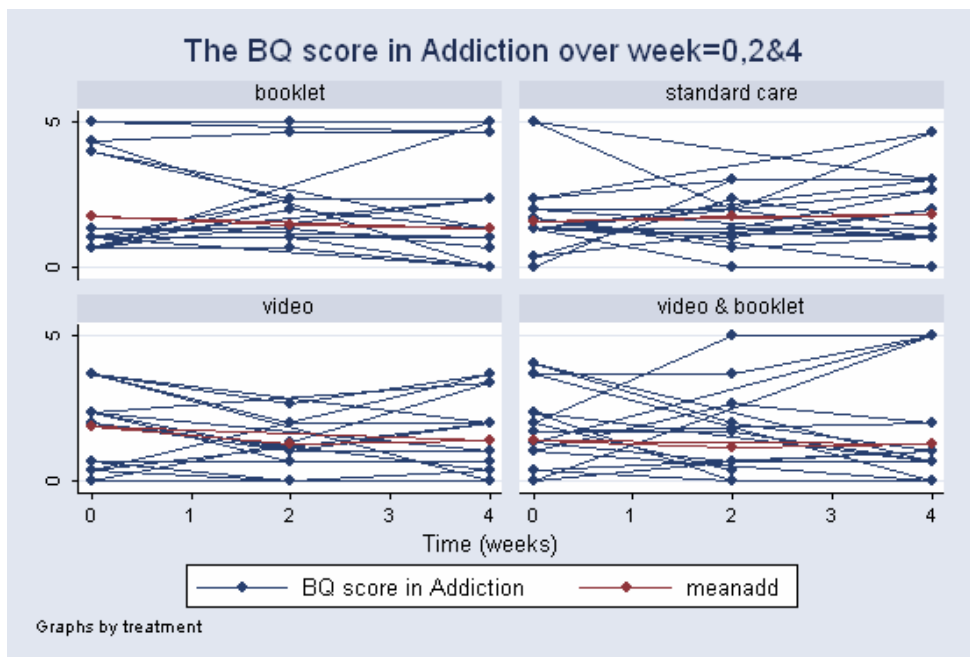
3.4 Worst reported pain at baseline and two weeks post-randomisation

Patients were asked to provide a daily rating of their pain on a scale of 0 to 10, where 0 indicated no pain and 10 indicated unbearable pain. These daily pain ratings were summarized in 14 days intervals, where the baseline pain inventory was measured 14 days prior to randomization and the two weeks following randomization represented the

week 2 measurements. For each interval, the worst reported daily pain score was obtained from the 14 days of observations. From a patient management perspective, clinicians were interested if patients were scoring moderate to severe pain (i.e., pain scored greater than 5 out of 10) as this indicates pain high enough to require medical treatment (e.g., use of opioids or similar).

Worst reported pain scores were re-coded into a binary variable (pain=0 if worst pain score ≤ 5 and pain=1 if worst pain score > 5). The baseline status regarding proportion of patients reporting worst pain greater than five is presented below in Table 3.4.1 and is similar for the treatment groups under consideration.

Table 3.4.1 Baseline status (week 0) for worst reported pain score > 5

Treatment group	N	Patients with a worst reported pain score > 5	
		n	%
Standard care	38	22	(58%)
Booklet only	36	18	(50%)
Video/DVD only	36	25	(69%)
Booklet & Video/DVD	46	32	(70%)
No Video/DVD	74	40	(54%)
Video/DVD	82	57	(70%)
No Booklet	74	47	(64%)
Booklet	82	50	(61%)

A similar assessment was conducted using the worst pain scores reported two weeks after randomisation. There was no substantial change compared to baseline. This data was analysed using a chi-squared test to examine the association between the worst pain score and the treatment groups. Table 3.4.2 displays the results among various treatment groups. At 2 weeks after randomisation, there was no evidence of an association between worst reported pain score greater than five and the four treatment groups (p=0.45). Similarly, there was no evidence of an association between worst reported pain score greater than five and video/DVD (p=0.24) or booklet (p=0.45).

Table 3.4.2 Post-randomisation status (week 2) for worst reported pain score > 5

Treatment group	N	Patients with a worst reported pain score > 5		χ2(df), p
		n	%	
Standare care	38	16	42	
Booklet only	36	15	42	χ2(3)=2.67, p=0.45
Video/DVD only	36	21	58	
Booklet & Video/DVD	46	21	46	
No Video/DVD	74	31	42	χ2(1)=1.36, p=0.24
Video/DVD	82	42	51	
No Booklet	74	37	50	χ2(1)=0.58, p=0.45
Booklet	82	36	44	

3.5 Assessment of the change in the worst pain scores between week 0 (baseline) and week 2

Summary pain scores (such as worst reported pain) exhibited skew distributions and were not normally distributed. However, similar to the BQ scores, the difference in pain scores between week 2 and week 0 were observed to be approximately normally distributed and as a result, the changes were used as the outcome to assess the effect of booklet and/ or video/DVD on worst reported pain scores.

The results of the mean changes between week 2 and week 0 for the treatment with video/DVD or without video/DVD and the treatment with booklet or without booklet are shown in Table 3.5.1. There was no evidence of an interaction between video/DVD and booklet interventions on changes in worst report pain score (p=0.91, see Table 3.5.2). As such, video/DVD and booklet intervention were considered as possible independent influences on the change in worse pain score after 2 weeks. However, there was no evidence of that the change of worse pain scores between baseline (week 0) and a fortnight later for patients assigned to video/DVD was different to the change in pain scores for patients assigned to no video/DVD treatment (p=0.11). Similarly, there was no evidence of differences for patients allocated to a booklet intervention versus no booklet (p=0.40).

Table 3.5.1 Descriptive statistics for the worst reported pain score at week 0 and week 2

Treatment group	Mean at Week 0 (95% CI)	Mean at Week 2 (95% CI)	Mean change for (week 2 - week 0) (95% CI)
No Video/DVD	5.373 (4.707, 6.039)	4.819 (4.242, 5.396)	-0.554 (-1.269, 0.161)
Video/DVD	6.199 (5.659, 6.738)	4.884 (4.308, 5.461)	-1.315 (-1.882, -0.747)
No Booklet	5.727 (5.098, 6.356)	5.004 (4.445, 5.563)	-0.723 (-1.282, -0.164)
Booklet	5.879 (5.292, 6.467)	4.717 (4.128, 5.306)	-1.162 (-1.864, -0.460)

Table 3.5.2 The coefficients of the worst pain scores changed for (week 2 - week 0) with p-value when fitting regression model including or excluding interaction between video/DVD and booklet

Change in Outcome (week 2 - week 0)	Model	Covariate	Coefficient (95% CI)	P-value
Worst pain scores	including interaction	Video/DVD	-0.67 (-1.98, 0.63)	0.31
		Booklet	-0.33 (-1.63, 0.98)	0.62
		Video/DVD * Booklet	-0.11 (-1.92, 1.70)	0.91
	excluding interaction	Video/DVD	-0.73 (-1.63, 0.17)	0.11
		Booklet	-0.38 (-1.29, 0.52)	0.40

DISCUSSION

This 2x2 factorial design study examined whether the educational intervention strategies (either by a booklet and/or a video/DVD) improves barriers, behaviors and perceptions towards effective pain management in patients with cancer. Short-term changes in patient perceptions were assessed in 156 patients who had completed follow-up data two weeks after baseline. Patients were assessed in terms of their barriers/perceptions towards overall/total pain management, as well as specific attitudes towards perceived addiction of pain medications, pain tolerance, side effects, progression of disease, fatalism, injections, good patient behaviors and doctor distractions. The analysis indicated that there was little evidence of benefit with the use of the educational interventions for nearly all the pain management barriers. However, there was one exception observed indicating that there was some evidence of an improvement in attitudes concerning addiction to pain medications after the use of an educational intervention of video/DVD. The potential effects of the interventions were also investigated in terms of reported pain, although it was found that there was no evidence that a pain education campaign improves the worst-reported pain scores.

This study was my first exposure to 'real' data and also my first exposure to clinical trial data. While the analysis approach was (relatively) simple, this project allowed me the opportunity to consider the statistical implications for sub-optimal trial methodology (such as an insufficient sample size) and analysis issues (such as multiple testing of multiple outcomes).

In this project, analyses were conducted over 9 outcomes for pain barriers (the total BQ score and 8 subscales) with only one significant result observed (the addiction subscale), with no adjustments for multiply comparisons. This single observed significant result may be an artifact of an inflated Type I error rate due to the multiplicity issue (using multiple outcomes and multiple tests). As a result, this single significant result should be treated with caution. No other statistically significant results were observed which may support the use of a video/DVD education strategy to alleviate concerns/barriers for effective pain management, suggesting that this single significant result may be a

chance occurrence.

To accurately estimate the treatment effect by comparing the treatment groups, we need to obtain the optimal sample size in a study. Consequently, it is important to define the required clinical effect of the treatments or the significant treatment difference at the beginning of the trial design. For this study, there are two aspects we need to consider for the implications concerning statistical power.

In order to detect whether or not a difference is statistically significant (at the conventional 5% level), a sufficient sample size calculation with the desired power is ideally required to be considered during the trial design stage. When the sample size was calculated in this study, there was little published information on what was considered a clinically relevant difference/reduction in BQ score and as such, a reduction of at least 20% on the total score and the subscales was (arbitrarily) chosen to be sufficient to consider the education intervention clinically useful. Therefore, the sample size calculation was based on detecting a 20% change in the total BQ score. A similar study was conducted in another country earlier, using a different patient setting (female patients with breast cancer). From this previous work, the estimated outcomes measures and variability in the change in the total BQ score was used for designing the sample size. If a trial design assumes a greater difference would be detected with a smaller variance, the study would be underpowered to detect a smaller true difference between treatments. Unfortunately in this study, the observed difference between treatment groups was smaller than what had been assumed and the variance was much greater, effectively reducing the statistical power of the study and making it unlikely that statistically significant results (at the 5% level) would be achieved if the interventions were truly effective at reducing concerns about affective pain management. The ability to detect significant interactions was dramatically affected, making it very unlikely that interactions would be detected. With hindsight, the investigators should have been counseled to run a small pilot study to compare their planned population against the previous study.

One should also consider whether it was reasonable to assume that the two interventions

could be assumed to act independently. In the trial design stage, it was assumed that both booklet and video/DVD treatment were independent factors with little interaction between these two treatments. However, after the trial was conducted and the clinicians were consulted, it was discovered that it was thought likely that the combination of both booklet and video would help patients more than simply the provision of one intervention alone and it was hoped that the combination of treatments would produce vastly greater reductions in barrier perceptions (due to the consolidation of knowledge provided to these patients). This sounded like the investigators were actually hoping/anticipating for an interaction between the two interventions. However, the trial was not specifically designed to investigate this aspect, limiting the usefulness of the trial data to provide evidence of this hypothesis.

In summary, the aspects discussed above from this study demonstrate that attention to trial design issues is critical for useful trial results and the ability of a trial to answer a hypothesis. Unfortunately, this particularly trial did not approach a statistician for advice prior to commencement. If a well-designed study is properly designed and conducted, the trial is able to answer the initial question posed by the clinicians. In order to have a well-designed study, it is important that the clinicians communicate effectively and correctly with the statisticians during the stage of the study design and that this communication translates into a suitable design which will achieve the trial objectives.

REFERENCES

- Associate Professor Andrew Forbes/Professor John Carlin, Biostatistics Collaboration of Australia, Longitudinal and Correlated Data (LCD) study notes (2005)

- Melanie Lovell, Staff Specialist, Palliative Care St Vincent's Hospital Clinical Research Fellow, Prince of Wales Hospital 2005. Presentation "Patient Barriers to Cancer Pain Management".

- Professor Annette Dobson, Biostatistics Collaboration of Australia, Categorical Data Analysis and Generalized Linear Models (CDA) study notes (2005)

- Sandra E.Ward, Nancy Goldberg, Vickie Miller-McCauley, Carol Mueller, Ann Nolan, Darlene Pawlik-Plank, Alice Robbins, Doris Stormoen and David E. Weissman. Pain, 52(1993) 319-324.

- Steven Piantadosi, Clinical Trials- Chapter 15 Factorial Designs. John Wiley & Sons, Inc. (1997)

APPENDIX

Major STATA commands used in this study

. sort sid week

. by sid: generate diffadd=add[2]-add[1]

. by sid: generate diffBQ=BQ[2]-BQ[1]

/*generate the difference between week2 and week0 for the BQ subscale of addiction, repeat the command to generate the total BQ score and other subscales*/

gen Video/DVD=0 if randomisation==5 | randomisation==6

/*generate variale Video/DVD for those subjects received video/DVD treatment*/

replace Video/DVD=1 if randomisation==3 | randomisation==4

gen Booklet=0 if randomisation==4|randomisation==6

/*generate variable Booklet for those subjects received booklet treatment*/

replace Booklet=1 if randomisation==3|randomisation==5

gen treatment=6-randomisation   /*recode the treatments to 0, 1,2,3*/

bysort Video/DVD: summarize diffadd difftol diffside diffpro difffat diffinj diffgood diffdis diffBQ, detail

/*obtain the statistical description for the difference of total BQ score and BQ subscales in video/DVD treatment*/

bysort Booklet: summarize diffadd difftol diffside diffpro difffat diffinj diffgood diffdis diffBQ, detail

/*obtain the statistical description for the difference of total BQ score and BQ subscales in booklet treatment*/

xi:regress diffadd i.Video/DVD i.Booklet i.Video/DVD*i.Booklet

/*regression including interaction for the difference of addiction between week2 and week0, repeated this command in the total BQ score and other BQ subscales */

xi:regress diffadd i.Video/DVD i.Booklet

/*regression excluding interaction for the difference of addiction between week2 and week0, repeated this command in the toal BQ score and other BQ subscales*/

scatter add week,by(treatment, title("The BQ score in Addiction over week=0,2&4")) c(l)

table trt week, c(mean add)

/*graph the trajectory plots for the addiction subscale, repeated this command to obtain the plots for the total BQ score*/

```
gen code_add=0 if add<=0.5

    /*recode the addiction subscale and check the general association at baseline*/

replace code_add=1 if add>0.5 & add<=1.5

replace code_add=2 if add>1.5&add<=2.5

replace code_add=3 if add>2.5&add<=3.5

replace code_add=4 if add>3.5&add<=4.5

replace code_add=5 if add>4.5 & add!=.

tabulate code_add treatment if week==0,cchi2 chi2 exact expected

use "G:\Pain Education Study\WPP\Dataset\Pain worst.dta", clear

    /*create the dataset for pain scores analysis*/

merge sid using "G:\Pain Education Study\WPP\Dataset\Randomisation.dta"

sort sid

sort sid week

by sid: generate diff_painworst=pain_worst[2]-pain_worst[1]
```

PROJECT B

Project Title

Comparison Hazard Ratios between Cox Regression Model (the Noninterval-Censored Proportional Hazards Model) and Interval-Censored Proportional Hazards Model

Location

NHMRC Clinical Trials Centre, University of Sydney

The department of Radiology and Oncology, Westmead Hospital, Sydney

August, 2006-January, 2007

Context

This project has been developed by A/Professor Val Gebski to compare two regression approaches in survival analysis. The challenging issues in this project were to generate simulated data, obtain the interval-censored data and compare an interval-censored approach to classical method.

Contribution of Student

The student was involved in the survival data handling. The student was required to understand the censoring mechanism with its influence on the model selection and interpretation.

Statistical Issues

- How to generate the original time to event that follows an exponential distribution with a constant hazard rate
- How to obtain the interval-censored data
- Use of the statistical software STATA for analysing time to event data
- Use of the simulation in survival data analysis

Declaration by Student

I declare that this project is evidence of my own work, with direction and assistance provided by my project supervisor, A/Professor Val Gebski. This work has not been previously submitted for academic credit.

…………………………………………………

Huijun He

…………………………………………………

Date

Comments and Declaration by Project Supervisor

This work was performed by Huijun under my supervision. She performed all the simulations herself and was responsible for compiling this report. I believe that this report constitutes a satisfactory contribution to the workplace project component of the Master of Biostatistics course at the University of Sydney.

…………………………………………………

A/Professor Val Gebski

…………………………………………………

Date

Table of Contents

Acknowledgements

It is not possible to write this report without the advice and supervision from A/Professor

Val Gebski. I am most grateful to him for his extra time throughout this project.

1.0 BACKGROUND

1.1 Overview the time to event problem

In survival analysis of clinical studies an outcome of interest is the time to an event, such as the time to disease progression, subjects may not have experienced the event of interest during the observation period of the study. With this type of time to event data, for the subjects who will not have experienced the event under investigation at the time of analysis, the time to event is classified as being censored. Thus we do not know when they will experience the event, only that at the time of analysis of the trial they are still event free. Censoring may also occur due to the fact that the subject has left the study before they experience the event, or the subject is "lost to follow up" during the study, or the subject experienced a competing event in advance of the event of interest in the study. For the subjects who experienced the event in the study, a censoring indicator is set to the value of one, while for the censored observations this censoring indicator is given a value of zero. Hence, the time to event data in a study contains the times to event of the subjects who experienced the event of interest as well as the times to event of the subjects who are censored. We are interested what kind of data can estimate the effect of treatment on the time to event of interest.

For example, in a survival study subjects are recruited over a period of 2 years (24 months) with a further 18 months follow-up period, the overall study period being 42 months. Figure 1.1 shows some subjects (e.g., #3, #4, #5 and #6) were observed to have the event of interest (denoted by X's) during the study period. Subject #1 was not observed to experience the event at the end of the study. Subject #2 was lost to follow up before the end of study. Subjects #1 and #2 are censored (denoted by circles) as their event status is unknown at the time of the study period.

Figure 1.1 recruited, follow-up, and censoring on an event time study
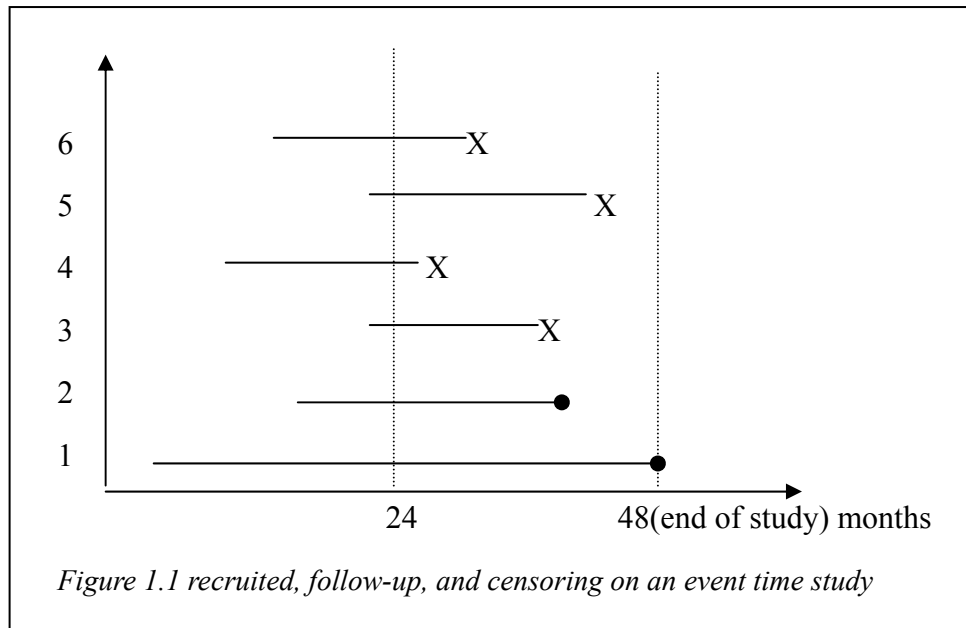
1.2 Interval Censoring

In some situations, the time to event of interest may only be known to have occurred within some time interval. Subjects are followed up at scheduled visits but the event of interest may have occurred between visits, and hence, we do not know the exact time to the event. Consider the example in an oncology study to evaluate the time to disease progression in patients with advanced colorectal cancer. The tumor progression is the term for the gradual cellular and molecular changes that occur during the development of tumors (i.e., cancer cell growth, invasion and metastasis, etc). Computerised tomography (CT) scans are used to detect the tumor progression during the study period. The time to tumor progression is defined as the time from the date of the original CT scan to the date of tumor progression. The subject will be diagnosed as having tumor progression if the CT scan shows the evidence in a new tumor or the increasing size of the original tumor, which is indicating how deep the tumor has penetrated, whether it has invaded the adjacent organs, and whether it has spread to distant organs, etc. For those subjects who are not observed as having a new tumor, or remain the same size/ reduce the size of the original tumor comparing with the original CT scan, we regard these subjects do not have tumor progression at the end of the observation period.

Since the tumor progression of the subjects is regularly checked using CT scans at fixed time point (e.g., every 6 weeks until the end of study), if the tumor progression of a subject has not observed at one CT scan but has been deemed to have progressed at the next scan, the time to tumor progression will only know to be within this interval (6 weeks) but we do not know the actual time to tumor progression for this subject. Data of this type is often referred to interval-censored data. For example, if the CT scan at week 12 indicated the subject without tumor progression but the next scan at week 18 was showing progression indicating tumor progression having occurred between week 12 and week 18. Standard methods to analyse such data assume that the event occurred at the end of interval. In this situation, such analysis would set the time to tumor progression at week 18. This raises the question as to the setting of the event time to the time of assessment (i.e., follow-up visits) will have impact on the estimates of the hazard ratio.

1.3 Objectives

To investigate the question mentioned above, we performed a simulation study to compare the estimates of the hazard ratio using:

- The original times to disease progression (which were known from the simulated data);
- A time to event analysis (Cox proportional hazards model) where the subjects' time to disease progression were recorded as at the time of visit when the CT scan was performed;
- An interval-censored approach (interval-censored proportional hazards model) where the subjects who were identified as having disease progression at a scheduled scan were assumed to have failed in the interval between their previous scan and the current scan.

2.0 METHODS

The exponential distribution is often used to model survival data (i.e., time to disease progression with censored observations). We will assume that the times to disease progression follow an exponential distribution for our investigation. The exponential distribution has density function:

$$f(x,\lambda) = \lambda e^{-\lambda x} \quad (x \geq 0) \qquad (1)$$

Where λ is a rate of failure per unit time (i.e., in our case, the number of failures per month), the exponential variable x is the time to disease progression (i.e., in our case, it is measured in months) of a random subject from the population. The exponential distribution is unique in that it has a constant hazard rate λ.

In order to generate the times to event (disease progression) of subjects from an exponential distribution with a hazard rate λ, we need to find a relationship between uniform (0,1) random number $(U(0,1))$ and the exponential distribution. Hence, this is required to transfer the $U(0,1)$ random number into the form of $f(x,\lambda)$. If we find the cumulative distribution function of the exponential distribution F(x,λ), we know that this lies between 0 and 1 as F(0,λ) =0, F(∞,λ) = 1. Now, if given a random variable U drawn from the uniform distribution on the unit interval (0,1), by using inverse transformation of the cumulative distribution function we can obtain the corresponding exponential variable x. The exponential distribution with different hazard rate λ will give different cumulative distribution function. The cumulative distribution function is determined by integrating formula (1)

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x} \quad (x \geq 0) \qquad (2)$$

In order to get the exponential variable x from the cumulative distribution function, we simply use formula (2) to obtain

$$U = 1 - e^{-\lambda x} \qquad (3)$$

Then the exponential variable is

$$x = \frac{-\ln(1-U)}{\lambda} \qquad (4)$$

If we generate U_i variable from $U\,(0,\,1)$, we can then create the exponential variable x_i for any fixed λ using the relationship

$$x_i = \frac{-\ln(1-U_i)}{\lambda} \qquad (5)$$

We use a two-arm randomised trial (colorectal cancer example) to evaluate the time to disease progression in patient tumor. We know that group 1 (control group) has a median time to tumor progression of 12 months. Now if a new treatment used in group 2 (treatment group) can change this median time to tumor progression to 18 months with a recruitment period of 24 months and an 18 months follow-up period (42 months for the overall study period). We assume the time to tumor progression follows an exponential distribution with a fixed hazard rate λ; hence, the original times to tumor progression can be generated by using formula (5) with a fixed hazard rate λ. Before we generate the original time to tumor progression we need to know the hazard rate λ for each group. As the median time to tumor progression $t_{(median)}$ is given in this study, this is the time to tumor progression at which the cumulative distribution function is equal to 0.5, then the relationship between the median time to tumor progression $t_{(median)}$ and λ can be obtained from formula (2)

$$F(x) = \int_0^x \lambda e^{-\lambda t}\, dt = 0.5$$

When x is equal to the median time to tumor progression ($x = t_{(median)}$)

$$\int_0^{t_{(median)}} \lambda e^{-\lambda t}\, dt = -e^{-\lambda t}\Big|_0^{t_{(median)}} = 1 - e^{-\lambda . t_{(median)}} = 0.5$$

Then the relationship between the median time to tumor progression $t_{(median)}$ and the hazard rate λ is

$$e^{-\lambda \cdot t_{(median)}} = 0.5 \qquad (6)$$

To get a simple relationship we rearrange formula (6) to get

$$\lambda = \frac{\ln 2}{t_{(median)}} \qquad (7)$$

By using the corresponding median time $t_{(median)}$ to tumor progression in formula (7), we may obtain the hazard rate λ for each group, which is $\lambda_{control} = \ln 2/12 = 0.058$ for control group and $\lambda_{treatment} = \ln 2/18 = 0.039$ for treatment group. Hence, in our study the true hazard ratio is $\frac{\lambda_{treatment}}{\lambda_{control}} = \frac{0.039}{0.058} = 0.672$. Using the formula (5) we may generate the original times to tumor progression for each group through (8) and (9). (Appendix 6.3 shows the results of the times to tumor progression which were known from the simulated data)

$$\text{Control group} \quad x_{(control)_i} = \frac{-\ln(1-U_i)}{0.058} \qquad (8)$$

$$\text{Treatment group} \quad x_{(treatment)_i} = \frac{-\ln(1-U_i)}{0.039} \qquad (9)$$

In order to obtain 80% power and 5% for the significance level, the sample size of this trial requires 150 subjects per treatment group (the total sample size is 300). The sample size calculation is based on using the exponential distribution with a set accrual time (patients being uniformly recruited during this period) and a fixed follow-up time (in order to allow the last subject recruited to have sufficient time to experience an event). If θ is the hazard ratio then the number of events required (assuming equal sized group) is given by formula (S Piantadosi, 1997)

$$n_{event} = 4(Z_{1-\alpha/2} + Z_{1-\beta})^2 / (\ln\theta)^2 \qquad (10)$$

In our case $\theta = \frac{12}{18} = 0.667, Z_{1-\alpha/2} = 1.96, Z_{1-\beta} = 0.84$, then the number of events required is

$$n_{event} = \frac{31.36}{(\ln 0.667)^2} = 191.2$$

Figure 2.1 Example of accrual and follow-up time

Accrual Period

Start   Stop   Follow-up

Enter $(t_{(acc)} + t_{(fu)} - x)$

0   x $t_{(acc)}$ $t_{(acc)} + t_{(fu)}$
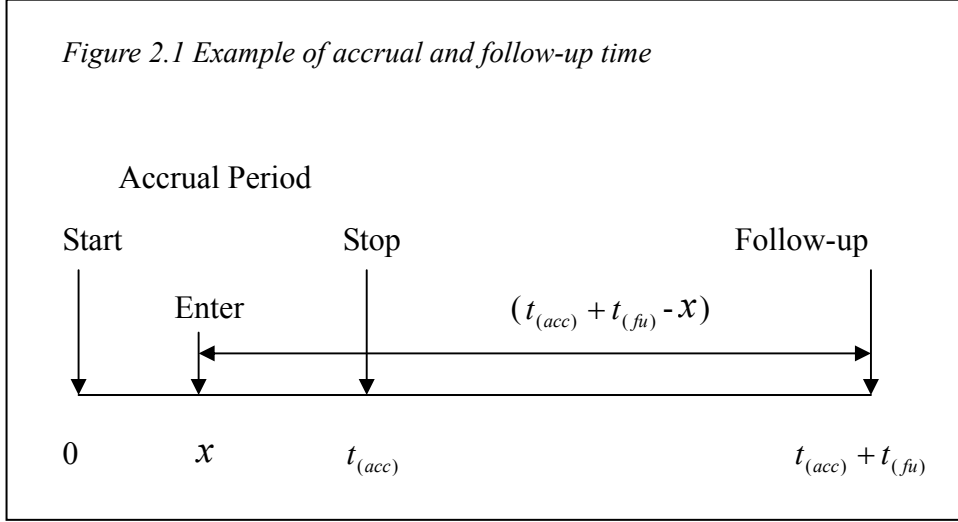
When the accrual time is $t_{(acc)}$ and subjects are uniformly accrued during this period (Figure 2.1), we know a subject enters the study at the time point $x(x \sim U[0, t_{(acc)}])$, the probability of the subject enters the study at the time point x is $1/t_{(acc)}$. When the follow-up time is $t_{(fu)}$ and the total period of study is $t_{(acc+fu)} = t_{(acc)} + t_{(fu)}$, the subjects entering the study at its commencement could have a maximum follow-up time of $t_{(acc+fu)}$ and those subjects entering at $t_{(acc)}$ would have a follow-up time of $t_{(fu)}$. The probability of no event occurring at the time x is:

$$\Pr(no \cdot event) = \int_0^{t_{(acc)}} P(0 \cdot event \mid enters \cdot at \cdot time \cdot x) * \frac{1}{t_{(acc)}} dx$$

$$= \frac{1}{t_{(acc)}} \int_0^{t_{(acc)}} P\{survive \cdot for \geq (t_{(acc)} + t_{(fu)} - x)\} dx$$

$$= \frac{1}{t_{(acc)}} \int_0^{t_{(acc)}} S(t_{(acc)} + t_{(fu)} - x) dx \qquad (11)$$

When $u = t_{(acc)} + t_{(fu)} - x$, the survival function is set as $S(t_{(acc)} + t_{(fu)} - x) = S(u)$.

Since $dx = -du$, when $x = 0, u = t_{(acc)} + t_{(fu)}; x = t_{(acc)}, u = t_{(fu)}$ then :

$$\int_0^{t_{(acc)}} S(u)dx = \int_{(t_{(acc)}+t_{(fu)})}^{t_{(fu)}} -S(u)du = \int_{t_{(fu)}}^{(t_{(acc)}+t_{(fu)})} S(u)du$$

Hence, the formula (11) is transferred into formula (12):

$$\Pr(no \cdot event) = \frac{1}{t_{(acc)}} \int_{t_{(fu)}}^{t_{(acc+fu)}} S(u)du = \frac{1}{t_{(acc)}} \int_{t_{(fu)}}^{(t_{(acc)}+t_{(fu)})} S(u)du \qquad (12)$$

The survival function $S(u)$ at any time point u is the reverse cumulative distribution function of (2), which is $S(u) = e^{-\lambda \cdot u}$.

Then,

$$\Pr(no \cdot event) = \frac{1}{t_{(acc)}} \int_{t_{(fu)}}^{(t_{(acc)}+t_{(fu)})} \lambda e^{-\lambda \cdot x} dx = \frac{1}{\lambda \cdot t_{(acc)}} \left[-e^{-\lambda \cdot x} \right]_{t_{(fu)}}^{(t_{(acc)}+t_{(fu)})}$$

$$= \frac{1}{\lambda \cdot t_{(acc)}} \left[e^{-\lambda \cdot t_{(fu)}} - e^{-\lambda \cdot (t_{(acc)}+t_{(fu)})} \right] \qquad (13)$$

and

$$\Pr(at.least.one.event) = 1 - \frac{1}{\lambda \cdot t_{(acc)}} \left[e^{-\lambda \cdot t_{(fu)}} - e^{-\lambda \cdot (t_{(acc)}+t_{(fu)})} \right] \qquad (14)$$

Hence, the probability of at least one event occurring in the control group is given by:

$$\Pr_{control} = 1 - \frac{\left[e^{-\lambda_{(conrol)} t_{(fu)}} - e^{-\lambda_{(control)} (t_{(acc)}+t_{(fu)})} \right]}{\lambda_{(control)} \times t_{(acc)}} \qquad (15)$$

Similarly the probability of at least one event occurring in the treatment group is given by:

$$\Pr_{treatment} = 1 - \frac{\left[e^{-\lambda_{(treatment)} t_{(fu)}} - e^{-\lambda_{(treatment)} (t_{(acc)}+t_{(fu)})} \right]}{\lambda_{(treatment)} \times t_{(acc)}} \qquad (16)$$

As we know

$$\lambda_{(control)} = 0.058$$

$$\lambda_{(treatment)} = 0.039$$

$$t_{(fu)} = \text{Follow-up time (month)} = 18$$

$$t_{(acc)} = \text{Accrual time (month)} = 24$$

$$t_{(acc+fu)} = \text{Total study period (month)} = t_{(acc)} + t_{(fu)} = (24+18) = 42$$

By using formula (15) and (16) we obtain the probability of an event occurring in the control group and treatment group

$$Pr_{control} = 0.810$$

$$Pr_{treatment} = 0.678$$

The sample size is required

$$N_{SampleSize} = \frac{n_{event}}{\frac{1}{2}(Pr_{control} + Pr_{treatment})} = \frac{191.2}{\frac{1}{2}(0.810+0.678)} = 257$$

So the sample size of 300 for the trial is sufficient.

To capture the information of time to tumor progression, whether censored or not, it is necessary to record data for each individual subject, which includes the event time and its corresponding censoring indicator (whether the event time records the interval to an event or to a censored point). The definition and calculation of the relevant variables used in this study are shown as below:

- Entry time

  The overall study period for this trial is 42 months (equal to 182 weeks). Within the recruitment period of 24 months, the recruitment rate of eligible subjects into this study for each group is 150/24 = 6.25/month, which is round at 6 subjects per month. We assume subject is recruited in a random order (e.g., the 1st 6 subjects enter study in month 1; the 2nd 6 subjects enter study in month 2, etc.) Hence, in each group, subjects are recruited until month 25 to obtain the sample size of 150 for each group.

- Follow-up time

  After month 26, there is a period of additional follow-up lasting to month 42. Attached to each subject the follow-up time is (42- entry time) months, e.g., the 1st 6 subjects entered at the time of month 1 will receive 41 months follow-up, the 2nd

6 subjects entered in month 2 will receive 40 months follow-up, the last 6 subjects recruited in month 25 will have 17 months follow-up, etc.

- The original time to disease progression and its corresponding censoring indicator

  The original times to tumor progression for both control group and treatment group can be generated by using the formula (7) and (8). As the study period is 42 months, the time to tumor progression for the subject needs to be replaced by the corresponding follow-up time if the time to tumor progression is greater than its follow-up time, which means this subject will not be observed to have experienced tumor progression at the end of study. This subject is censored (a censoring indicator is set to the value of zero). If the time to tumor progression is less than or equal to the follow-up time, the subject will have experienced tumor progression, hence this subject is failed (the censoring indicator is set to the value of one).

- Scanning visit interval

  In this study we assessed the difference of hazard ratios by changing the time interval when a CT scan is conducted to every 3,6,9,12 and 15 weeks. For example, when the scans occur every 6 weeks until the end of study period (42 months), the 1st scanning visit is at week 6 after the patient is randomised into the study, the 2nd scanning visit is at week 12, and the 30th scanning visit is at week 180 which is the end of the study period (182 weeks). Note that there is a gap of two weeks between this last scan time and the end point of study period when the scanning visit interval is set at every 6 weeks.

- Interval-censored data

  The example of data with interval censoring in this study is presented in Figure 2.2 (e.g., the scanning visit interval at every 6 weeks). We record the time to tumor progression in the following three ways: (a) at time t_0 (beginning of the interval); (b) at time t_1 (end of the interval); (c) between time t_0 and t_1 (interval falling).

  - Subject #1: the subject experienced tumor progression at the exact scanning visit time, we record this subject's time to tumor progression at time t_0;

  - Subject #2: the time to tumor progression is falling in the time interval of (t_0, t_1),

this subject is censored at time t_0 at the previous scanning visit but was observed to have experienced tumor progression at the current scanning visit. We record this subject's time to tumor progression at time t_1;

- Subject #3: the subject is supposed to have experienced tumor progression at week 181, this subject is censored at week180 which is not observed to have tumor progression at that time, however, the next scanning visit is at week 186 which is after the study end point (week 182), in this case, this subject is censored before the end of study period but it is failed in the actual study.

- Subject #4 and #5: these subjects are censored in the actual study and at the last scanning visit at week 180 as well, which means that the tumor progression is not observed at the end of study.

Figure 2.2 Example of date with interval censoring in this study
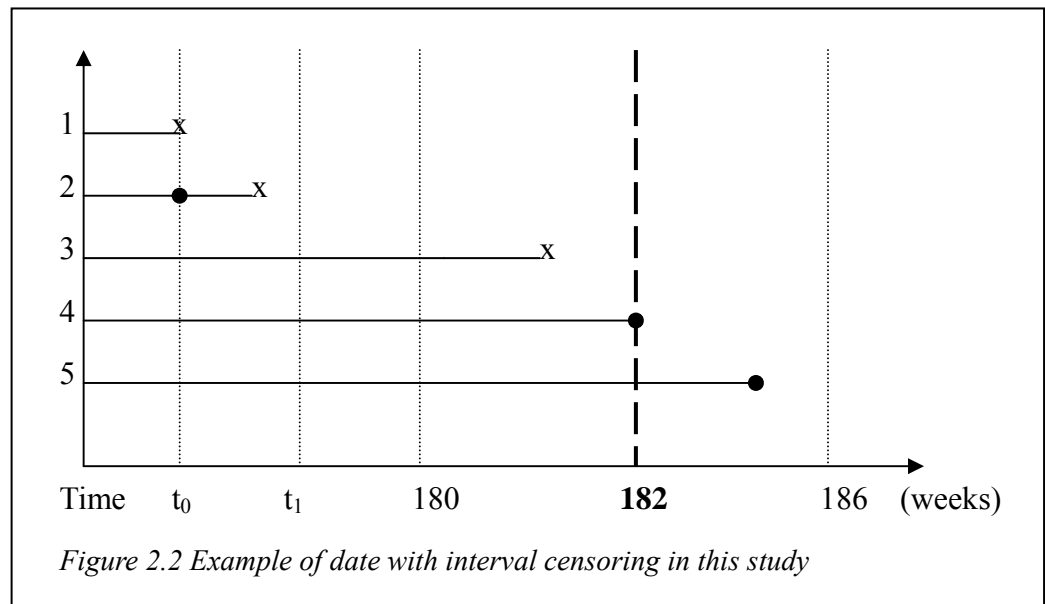
3.0 RESULTS

3.1 Comparison of the hazard ratios based on Cox regression model with different scanning visit intervals

The Cox regression model in the statistical software STATA is used to produce the estimated coefficients. In calculating these parameters, seed was initially set at 300 to generate the uniform random data. Running the simulation trials N=100 to obtain the corresponding estimated coefficient and its standard error of Cox regression model with different scanning visit intervals, to display the results of fitting the model, we obtain the estimated hazard ratios, their standard errors and their associated 95% confidence

intervals through the relationship between the estimated coefficient and the hazard ratio. Hence, the estimated hazard ratio is:

$$HR = \exp(\hat{\beta})$$

The associated 95% confidence intervals of the estimated hazard ratio are:

$$\exp[\hat{\beta} \pm 1.96SE(\hat{\beta})]$$

The standard error of the estimated hazard ratio is obtained using the delta method approximation (D W.Hosmer, JR & S Lemeshow, 1999):

$$SE(HR) = SE(\exp(\hat{\beta})) = \exp(\hat{\beta}).SE(\hat{\beta})$$

Table 3.1 contains the values of the estimated hazard ratio, its standard error and its associated 95% confidence intervals for the required models. The comparison graphs are presented in Figure 3.1.

Table 3.1 The estimated hazard ratios and 95% confidence intervals for the treatment effect based on Cox Regression with different scanning visit interval at every 3,6,9,12 and 15 weeks.

Model	Hazard Ratio (SE)	95% CI
Cox Regression when the original times to tumor progression were used	0.669 (0.0032)	(0.662, 0.675)
Cox Regression when scanning visit interval is at every 3 weeks	0.674 (0.0030)	(0.668, 0.680)
Cox Regression when scanning visit interval is at every 6 weeks	0.679 (0.0031)	(0.673, 0.685)
Cox Regression when scanning visit interval is at every 9 weeks	0.683 (0.0031)	(0.677, 0.689)
Cox Regression when scanning visit interval is at every 12 weeks	0.689 (0.0032)	(0.682, 0.695)
Cox Regression when scanning visit interval is at every 15 weeks	0.692 (0.0033)	(0.686, 0.698)

Figure 3.1 The graph for the hazard ratios based on Cox Regression with different scanning visit interval at every 3,6,9,12 and 15 weeks (N=100).
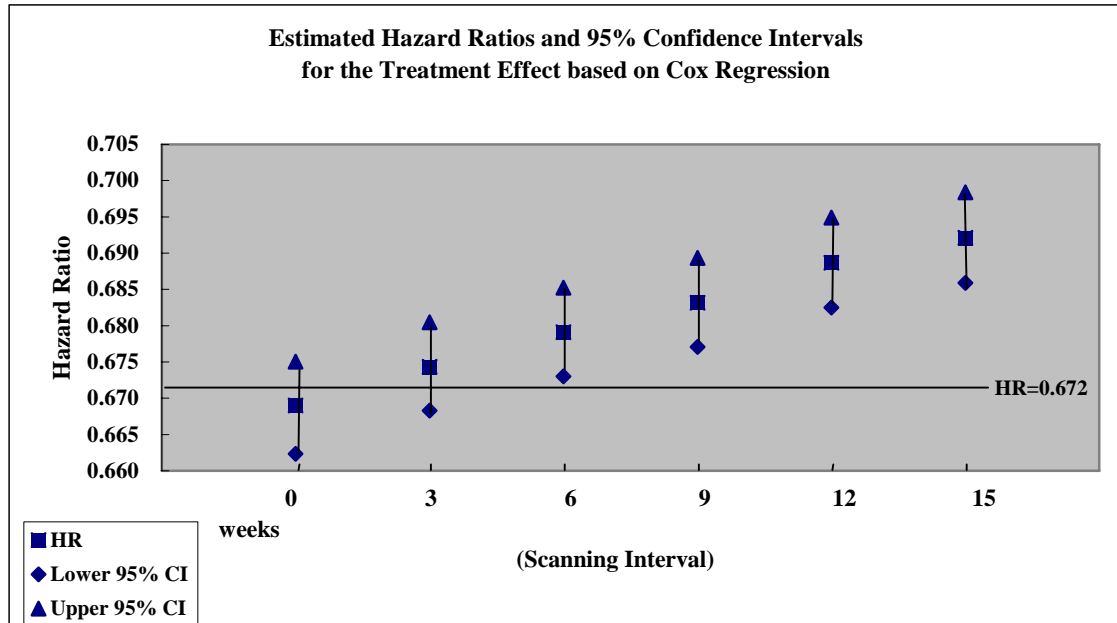
Note:

1. When the scanning visit interval is equal to 0 in the graph, it means the original times to tumor progression were fitted into Cox regression model.

2. The horizontal line at HR=0.672 is the true hazard ratio of $\dfrac{\lambda_{treatment}}{\lambda_{control}} = \dfrac{0.039}{0.058} = 0.672$

3.2 Comparison of the hazard ratios based on interval-censored (IC) proportional hazards model with different scanning visit intervals

In order to obtain the hazard ratio using the interval-censored proportional hazards model, we created the expanded data set. This data set contained a new variable, which was a binary outcome indicator of the censor status in the specified time interval. Table 3.2 is presented a few specific examples from the expanded data set when the scanning visit interval was at every 6 weeks. The first three lines of data in this table represented the subject #1 experienced tumor progression between week 12 and week 18. This was noted by week=18. This subject contributed three lines of data to the expanded data set. The first line corresponds to that this subject had not experienced tumor progression at week 6. For this line the interval=1 and binary outcome indicator=0. The second line is for the second 6-week period, denoted by interval=2, and had not experienced tumor

progression with binary outcome indicator=0. The third line is for the third 6-week period, the subject had tumor progression corresponding to the interval=3 and binary outcome indicator=1.

The second block of 14 lines represented subject #6, who had tumor progression between week 78 and week 84. During the first 13 intervals of follow-up, this subject had not experienced tumor progression hence the binary outcome indicator =0. The value was changed to 1 in the 14th interval, which was indicating the subject experienced tumor progression at this interval.

The third block of 20 lines represented subject #89, who did not experienced tumor progression at the end of study and kept the binary outcome indicator=0 at 20th interval.

In STATA the expanded data set was performed and fitted the interval-censored proportional hazards model with this expanded data set. The simulation trials were conducted 100 times when initiating to set seed at 300. We obtained the corresponding estimated coefficients at different scanning visit intervals. Table 3.3 is shown the hazard ratios and the associated 95% confidence intervals at different scanning visit interval. The comparison graph of the hazard ratios is shown in Figure 3.2.

Table 3.2 Examples of the expanded data set required to fit the interval-censored(IC) proportional hazards model with scanning visit interval at every 6 weeks.

sid	scanning time (weeks)	interval	censor	binary outcome indicator
1	18	1	1	0
1	18	2	1	0
1	18	3	1	1
6	84	1	1	0
6	84	2	1	0
6	84	3	1	0
6	84	4	1	0
6	84	5	1	0
6	84	6	1	0
6	84	7	1	0
6	84	8	1	0
6	84	9	1	0
6	84	10	1	0
6	84	11	1	0
6	84	12	1	0
6	84	13	1	0
6	84	14	1	1
89	120	1	0	0
89	120	2	0	0
89	120	3	0	0
89	120	4	0	0
89	120	5	0	0
89	120	6	0	0
89	120	7	0	0
89	120	8	0	0
89	120	9	0	0
89	120	10	0	0
89	120	11	0	0
89	120	12	0	0
89	120	13	0	0
89	120	14	0	0
89	120	15	0	0
89	120	16	0	0
89	120	17	0	0
89	120	18	0	0
89	120	19	0	0
89	120	20	0	0

Table 3.3 The estimated hazard ratios and 95% confidence intervals for the treatment effect based on interval-censored (IC) proportional hazards model with different scanning visit interval at every 3,6,9,12 and 15 weeks.

Model	Hazard Ratio (SE)	95% CI
Cox Regression when the original times to tumor progression were used	0.669 (0.0032)	(0.662, 0.675)
IC Regression when scanning visit interval is at every 3 weeks	0.670 (0.0030)	(0.664, 0.676)
IC Regression when scanning visit interval is at every 6 weeks	0.670 (0.0030)	(0.662, 0.676)
IC Regression when scanning visit interval is at every 9 weeks	0.669 (0.0035)	(0.662, 0.676)
IC Regression when scanning visit interval is at every 12 weeks	0.670 (0.0034)	(0.663, 0.676)
IC Regression when scanning visit interval is at every 15 weeks	0.670 (0.0034)	(0.663, 0.676)

Figure 3.2 The graph for the hazard ratios based on interval-censored(IC) proportional hazards model with different scanning visit interval at every 3,6,9,12 and 15 weeks (N=100)
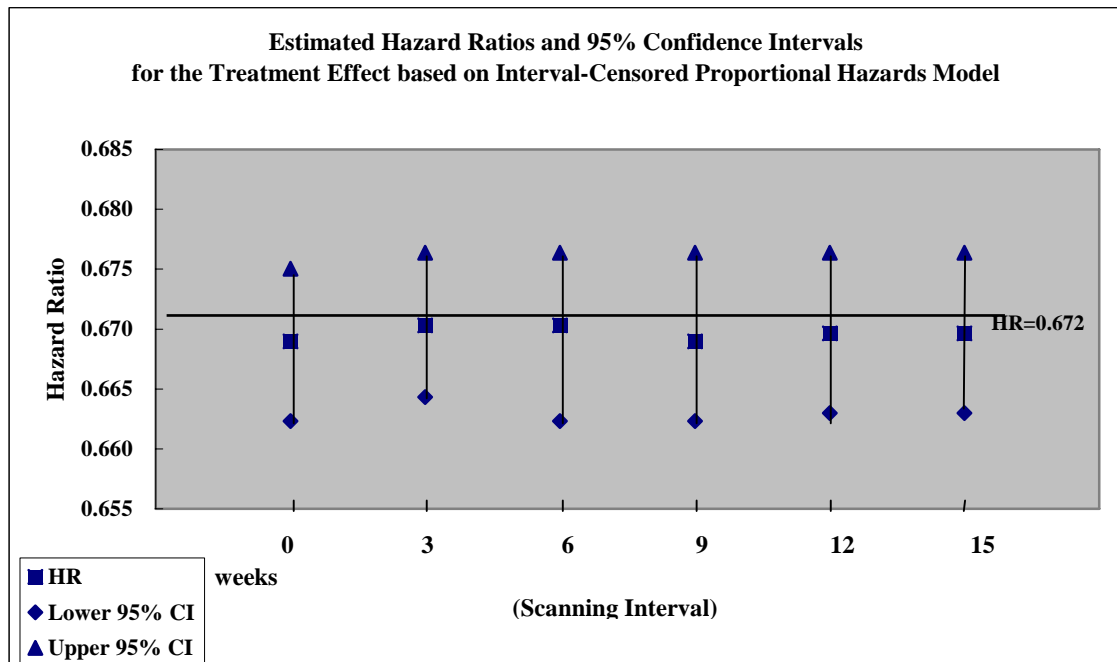
Note:

1. when the scanning visit interval is equal to 0 week, it means the original times to tumor progression were fitted into interval-censored proportional hazards model.

2. The horizontal line at HR=0.672 is the true hazard ratio of $\dfrac{\lambda_{treatment}}{\lambda_{control}} = \dfrac{0.039}{0.058} = 0.672$

4.0 DISCUSSIONS

In this study we firstly fitted the Cox regression model with data, where the subjects' times to tumor progression were recorded as at the time of visit when the CT scan was

performed with different scanning visit interval at every 3, 6, 9, 12 and 15 weeks. Using the scanning visit interval to fit Cox regression model will change the actual time to tumor progression forward to the right end of the specific scanning visit interval.

Figure 4.1 shows these changes (circle denoted as censoring and X denoted as failure). For example, the subject had actual time to tumor progression at week 25. When the scanning visit interval was set at every 3 or 9 weeks, this subject was set to have experienced tumor progression at week 27; when the scanning visit interval was set at every 6 weeks or 15 weeks, this subject experienced tumor progression at week 30; when the scanning visit interval was set at every 12 weeks, they were set to have tumor progression at week 36 during the study period. These changes would lead to conservative results for the estimation of hazard ratio. Hence, changing the scanning visit interval may have impact on recording the time to tumor progression fitted into the Cox regression model, and the hazard ratio is affected accordingly. When compared with the result of the actual study (Cox regression model when the original times to tumor progression were used), the scanning visit interval increases, and the estimated hazard ratio becomes conservative (Table 3.1 and Figure 3.1).

Figure 4.1 Example of the changes of recording time to tumor progression in the Cox regression model with different scanning visit interval at every 3,6,9,12 and 15 weeks.
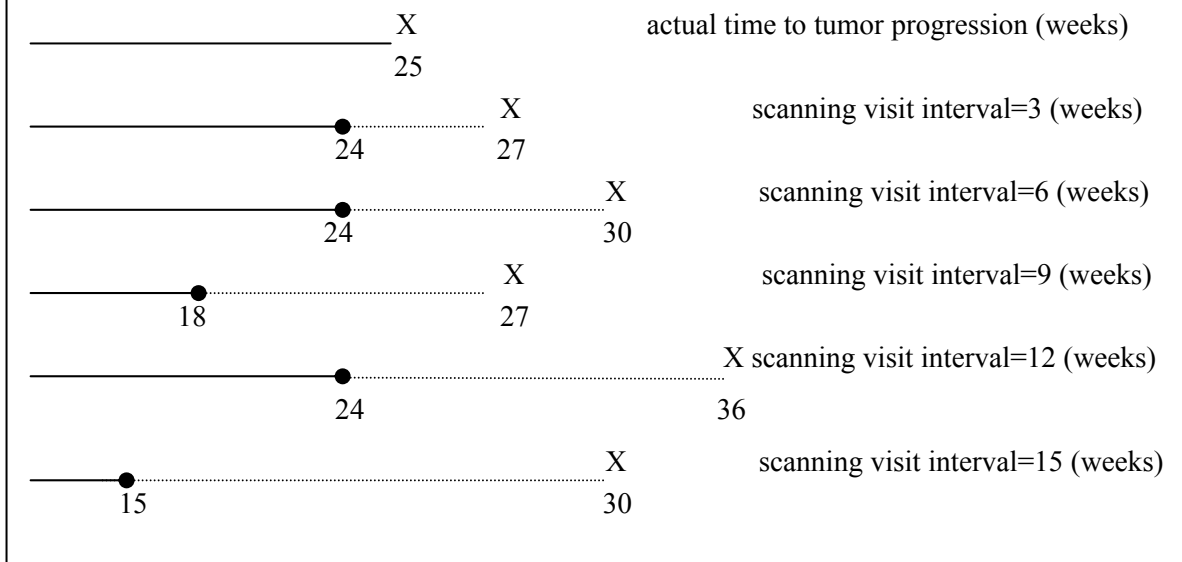
When changing the scanning visit intervals, the number of subjects' failures corresponding to the specific scanning visit interval would change, and the number of censored observations changes accordingly. For example, one simulation trial was conducted for each scanning visit interval. The number of failures in this data set for the corresponding scanning visit interval is presented in Table 4.1. From the results we obtained that the numbers of failures and the corresponding censored observations changed with the scanning visit intervals. The wider the scanning visit intervals, the less number of failures, consequently, the greater number of censored observations. Since the last scanning visit interval for the specific subject is outside of the study endpoint, in the actual study this subject was failed, however, it was censored in the scanning visit interval. The total number of these subjects may change the number of failures in the associated model. Hence, there is a bias when we fit the Cox regression with different scanning visit intervals. In Figure 3.1 the 95% confidence intervals do not include the true hazard ratio of 0.672 for the scanning visit intervals at every 6,9,12 and 15 weeks. While compared to the horizontal line at HR=0.672, the bias of the hazard ratio increases with increasing scanning visit interval. With the scanning visit interval at every 15 weeks the power to detect the hazard ratio of 0.692 is 77%, it decreases when compared to the actual study with actual times to tumor progression used, which has

80% power to detect the hazard ratio of 0.667.

Table 4.1 Example of the number of failures and censored observations in the corresponding Cox regression model

Model	Number of	
	Failures	Censored Observations
Cox Regression when the original times to tumor progression are used	207	93
Cox Regression when scanning visit interval is at every 3 weeks	203	97
Cox Regression when scanning visit interval is at every 6 weeks	203	97
Cox Regression when scanning visit interval is at every 9 weeks	202	98
Cox Regression when scanning visit interval is at every 12 weeks	201	99
Cox Regression when scanning visit interval is at every 15 weeks	200	100

Secondly we fitted the interval-censored proportional hazards model with different scanning visit interval at every 3,6,9,12 and 15 weeks. The hazard ratios were almost constant and close to the estimate that was obtained from the actual study data using a Cox regression model (Table 3.3) While compared to the horizontal line at HR=0.672, the 95% confidence intervals include the true hazard ratio of 0.672 for all scanning visit intervals (Figure 3.2). This shows that the use of interval-censored proportional hazards model when analysing time to event data at different intervals we obtain more accurate estimates of the hazard ratios.

5.0 REFERENCES

- A/Prof Gillian Heller, Prof. Malcolm Hudson. Department of Statistics, Macquarie University. Unit of Study Notes-Survival Analysis (STAT827/SVA 2006)

- David W.Hosmer, Jr. Stanley Lemeshow. Applied Survival Analysis. Regression Modeling of Time to Event Data.Chapter 7 "Extensions of the Proportional Hazards Model" and Chapter 8 "Parametric Regression Models". John Wiley & Sons, Inc. (1999).

- Emmanuel Lesaffre, Arnost Komarek. An Overview of Methods for Interval-Censored Data with an Emphasis on Applications in Dentistry, Statistical Methods in Medical Research 2005: 14: 539-552.

- Jane C. Lindsey and Louise M. Ryan. Tutorial in Biostatistics Methods for Interval-Censored Data, Statistics in Medicine, Vol. 17, 219-238 (1998).

- Steven Piantadosi. Clinical Trials- A Methodologic Perspective. Chapter 6 "Objectives and Endpoints". John Wiley & Sons, Inc. (1997).

6.0 APPENDIX

Major STATA Programmes used in the project

Appendix 6.1 Simulation based on Cox Regression with scanning visit interval at every 3,6,9,12 and 15 weeks

```
set memory 700m
set seed 300                    /*set seed so get same random data in future runs of same program*/
        program define sim_Cox, rclass          /*define the program named"sim_Cox"*/
        drop _all                               /*drop previous data*/
        set obs 150                             /*change sample size by change set obs*/
        count                                   /*count the total observations*/
        generate record=_n                      /*assign ID for the observations*/
        gen u=uniform()                 /*generate uniform random numbers*/
        generate svw=(-ln(1-(u))/0.058)*52/12
                                        /*generate time to tumor progression in weeks for the control group*/
        expand 2                        /*use the same uniform random numbers*/
        generate sid=_n                 /*assign ID for all observations*/
        egen treatment=cut(sid),group(2)    /*generate treatment variable*/
        replace svw=svw*0.058/0.039 if treatment==1
                                /*generate time to tumor progression in weeks for the treatment group*/
        egen obs=cut(record),group(25)
        replace obs=obs+1                                       /*accrual period is 25 months*/
        generate followupW=(42-obs)*52/12               /*follow up time in weeks*/
        generate diff_svw=svw-followupW         /*generate Cox censor data*/
        generate cox_censor=1 if diff_svw<0
        replace cox_censor=0 if diff_svw>0
        replace svw=followupW if cox_censor==0
                /*recode all subjects who have times>study period 182 weeks*/

        generate scan_3=svw if mod(svw,3)==0       /*generate scan interval=3 weeks*/
        replace scan_3=svw-mod(svw,3) +3 if mod(svw,3)>0
        generate censor_3=1 if scan_3<followupW   /*generate Interval censor data*/
        replace censor_3=0 if scan_3>=followupW
                /*recode all subjects whose scan interval times>180 weeks*/

        generate scan_6=svw if mod(svw,6)==0       /*generate scan interval= 6 weeks*/
        replace scan_6=svw-mod(svw,6)+6 if mod(svw,6)>0
        generate censor_6=1 if scan_6<followupW
        replace censor_6=0 if scan_6>=followupW

        generate scan_9=svw if mod(svw,9)==0        /*generate scan interval=9 weeks*/
        replace scan_9=svw-mod(svw,9)+9 if mod(svw,9)>0
        generate censor_9=1 if scan_9<followupW
```

```
        replace censor_9=0 if scan_9>=followupW

        generate scan_12=svw if mod(svw,12)==0    /*generate scan interval=12weeks*/
        replace scan_12=svw-mod(svw,12)+12 if mod(svw,12)>0
        generate censor_12=1 if scan_12<followupW
        replace censor_12=0 if scan_12>=followupW

        generate scan_15=svw if mod(svw,15)==0          /*generate scan interval=15weeks*/
        replace scan_15=svw-mod(svw,15)+15 if mod(svw,15)>0
        generate censor_15=1 if scan_15<followupW
        replace censor_15=0 if scan_15>=followupW

        stset svw,failure(cox_censor)                              /*Cox Regression*/
        stcox treatment
        return scalar b_Cox=_b[treatment]                /*save results for Cox*/

        stset scan_3, failure(censor_3)       /*Cox Regression when scan interval=3 weeks*/
        stcox treatment
        return scalar b_IC_3=_b[treatment]          /*save results for IC_scan_3*/

        stset scan_6, failure(censor_6)       /*Cox Regression when scan interval=6 weeks*/
        stcox treatment
        return scalar b_IC_6=_b[treatment]               /*save results for IC_scan_6*/

        stset scan_9, failure(censor_9)       /*Cox Regression when scan interval=9 weeks*/
        stcox treatment
        return scalar b_IC_9=_b[treatment]          /*save results for IC_scan_9*/

        stset scan_12, failure(censor_12)     /*Cox Regression when scan interval=12 weeks*/
        stcox treatment
        return scalar b_IC_12=_b[treatment]               /*save results for IC_scan_12*/

        stset scan_15, failure(censor_15)     /*Cox Regression when scan interval=15 weeks*/
        stcox treatment
        return scalar b_IC_15=_b[treatment]               /*save results for IC_scan_15*/
end
simulate "sim_Cox" coef_Cox=r(b_Cox) coef_IC_3=r(b_IC_3) coef_IC_6=r(b_IC_6)
coef_IC_9=r(b_IC_9) coef_IC_12=r(b_IC_12) coef_IC_15=r(b_IC_15), dots reps(100)
/*do the simulations, change number of simulations by changing the number of reps*/

describe
summarize
```

Appendix 6.2 Simulation based on Interval-Censored Proportional Hazards Model with scanning visit interval at every 6 weeks (this programme is applied to change the scanning visit interval at every 3, 9, 12 and 15 weeks by changing the corresponding variables)

```
clear
set memory 700m
set matsize 800
set seed 300                    /*set seed so get same random data in future runs of same program*/
        program define sim_IC, rclass                    /*define the program named"sim_IC"*/
        drop _all                               /*drop previous data*/
        set obs 150
        count
        generate record=_n
        gen u=uniform()                         /*generate uniform random numbers*/
        generate svw=(-ln(1-(u))/0.058)*52/12
        expand 2
        generate sid=_n
        egen treatment=cut(sid),group(2)
        replace svw=svw*0.058/0.039 if treatment==1
        egen obs=cut(record),group(25)
        replace obs=obs+1
        generate followupW=(42-obs)*52/12
        generate diff_svw=svw-followupW
        generate cox_censor=1 if diff_svw<0                    /*generate Cox censor data*/
        replace cox_censor=0 if diff_svw>0
        replace svw=followupW if cox_censor==0
        generate scan_6=svw if mod(svw,6)==0     /*generate scanning visit interval at every 6 weeks*/
        replace scan_6=svw-mod(svw,6) +6 if mod(svw,6)>0
        gen count=scan_6/6                          /*generate the expnded data set*/
        sort sid
        expand count
        sort sid
        quietly by sid: gen intnew = _n                     /*generate the binary outcome indicator*/
        quietly by sid: gen censor1 = cond( _n == _N, cox_censor, 0)
        tab intnew,gen(int)

        quietly glm censor1 treatment    int1 int2 int3 int4 int5 int6 int7 int8 int9 int10 int11 int12 int13
        int14 int15 int16 int17 int18 int19 int20 int21 int22 int23 int24 int25 int26 int27 int28 int29
        int30, nocons nolog family(bin) link(cl)                  /*fitting the binary regression model*/
        return scalar b_IC_6=_b[treatment]                  /*save results for IC_scan_6*/
end
simulate "sim_IC" coef_IC_6=r(b_IC_6), dots reps(100)
/*do the simulations, change number of simulations by changing the number of reps*/
summarize
```

Appendix 6.3 The summary results of the failure times simulated using formula (8) and (9) for

control group and treatment group

Note: $Mean_{control} = \dfrac{1}{\lambda_{control}} = \dfrac{1}{0.058} \times (52/12) = 74.7(weeks)$

$Mean_{treatment} = \dfrac{1}{\lambda_{treatment}} = \dfrac{1}{0.039} \times (52/12) = 111.1(weeks)$

Simulation	Group	Mean	SD	Min	Max
	Control group	79.196	78.862	0.250	388.476
1	Treatment group	117.778	117.282	0.371	577.733
	Control group	70.263	77.684	0.966	563.496
2	Treatment group	104.493	115.530	1.437	838.020
	Control group	79.928	76.249	0.017	344.940
3	Treatment group	118.867	113.396	0.025	512.988
	Control group	67.284	59.715	0.239	252.934
4	Treatment group	100.064	88.807	0.355	376.159
	Control group	70.322	77.055	0.024	498.110
5	Treatment group	104.581	114.594	0.036	740.779
	Control group	73.573	74.216	0.252	416.855
6	Treatment group	109.417	110.373	0.374	619.939
	Control group	76.011	67.274	0.073	309.804
7	Treatment group	113.042	100.048	0.109	460.734
	Control group	75.427	84.330	0.069	468.221
8	Treatment group	112.174	125.415	0.102	696.329
	Control group	74.874	71.898	0.506	397.778
9	Treatment group	111.351	106.925	0.752	591.568
	Control group	86.398	80.876	1.115	374.290
10	Treatment group	128.490	120.277	1.659	556.637
	Control group	69.773	68.631	0.017	553.732
11	Treatment group	103.764	102.067	0.025	823.499
	Control group	68.836	71.663	0.097	374.506
12	Treatment group	102.372	106.576	0.144	556.957
	Control group	77.470	69.183	1.445	319.165
13	Treatment group	115.211	102.887	2.149	474.655
	Control group	73.005	71.506	0.827	474.386
14	Treatment group	108.571	106.343	1.230	705.497
	Control group	86.627	89.793	0.019	674.852
15	Treatment group	128.830	133.539	0.028	1003.626
	Control group	84.201	87.315	0.010	586.140
16	Treatment group	125.223	129.854	0.015	871.696
Simulation	Group	Mean	SD	Min	Max
	Control group	80.841	78.339	0.618	479.174
17	Treatment group	120.225	116.505	0.919	712.617
18	Control group	76.182	78.430	1.872	508.018

Simulation	Group	Mean	SD	Min	Max
	Treatment group	113.296	116.640	2.783	755.514
19	Control group	71.778	69.546	1.118	356.336
	Treatment group	106.747	103.428	1.663	529.936
20	Control group	83.181	83.302	0.567	483.499
	Treatment group	123.705	123.884	0.844	719.050
21	Control group	71.259	79.991	0.306	639.630
	Treatment group	105.974	118.961	0.454	951.244
22	Control group	83.993	76.125	1.081	338.657
	Treatment group	124.912	113.211	1.607	503.644
23	Control group	71.417	63.821	0.599	437.507
	Treatment group	106.210	94.913	0.891	650.652
24	Control group	85.362	82.536	0.145	444.303
	Treatment group	126.949	122.746	0.215	660.758
25	Control group	74.875	69.847	0.314	386.859
	Treatment group	111.352	103.875	0.467	575.329
26	Control group	78.113	84.541	0.316	534.308
	Treatment group	116.168	125.728	0.470	794.612
27	Control group	74.208	69.336	0.054	342.025
	Treatment group	110.360	103.115	0.080	508.653
28	Control group	78.562	83.984	0.095	643.718
	Treatment group	116.836	124.899	0.141	957.324
29	Control group	82.924	98.018	0.046	683.498
	Treatment group	123.322	145.770	0.069	1016.484
30	Control group	77.228	77.756	0.521	501.544
	Treatment group	114.852	115.637	0.775	745.886
31	Control group	66.840	83.068	0.017	508.584
	Treatment group	99.403	123.537	0.025	756.356
32	Control group	65.944	67.863	0.495	296.828
	Treatment group	98.071	100.924	0.737	441.436
33	Control group	80.004	79.333	0.138	557.588
	Treatment group	118.980	117.982	0.205	829.234
34	Control group	69.518	66.809	0.533	304.572
	Treatment group	103.386	99.357	0.792	452.954
35	Control group	72.108	72.039	0.018	423.204
	Treatment group	107.238	107.135	0.027	629.381
36	Control group	62.300	53.560	0.426	221.991
	Treatment group	92.651	79.654	0.633	330.140

Simulation	Group	Mean	SD	Min	Max
37	Control group	76.427	79.742	0.235	630.445
	Treatment group	113.661	118.591	0.349	937.585
38	Control group	75.639	71.910	1.802	333.211
	Treatment group	112.489	106.943	2.679	495.544

Simulation	Group	Mean	SD	Min	Max
39	Control group	70.297	67.285	0.170	423.471
	Treatment group	104.544	100.065	0.253	629.777
40	Control group	67.332	69.443	0.297	367.117
	Treatment group	100.135	103.275	0.442	545.969
41	Control group	79.374	66.538	0.027	355.270
	Treatment group	118.044	98.953	0.040	528.351
42	Control group	73.322	78.681	0.578	404.618
	Treatment group	109.044	117.012	0.860	601.739
43	Control group	85.996	94.843	0.832	528.002
	Treatment group	127.891	141.048	1.238	785.233
44	Control group	69.948	66.073	0.120	366.515
	Treatment group	104.025	98.262	0.178	545.074
45	Control group	72.756	67.026	0.969	404.903
	Treatment group	108.201	99.679	1.441	602.163
46	Control group	79.088	84.946	0.803	508.102
	Treatment group	117.618	126.330	1.195	755.639
47	Control group	72.964	71.024	0.911	385.948
	Treatment group	108.510	105.626	1.355	573.974
48	Control group	68.363	74.839	0.200	506.401
	Treatment group	101.668	111.298	0.297	753.109
49	Control group	80.284	85.301	0.080	515.415
	Treatment group	119.396	126.857	0.120	766.515
50	Control group	86.868	110.809	0.141	840.253
	Treatment group	129.188	164.792	0.210	1249.607
51	Control group	74.134	70.917	1.151	324.308
	Treatment group	110.251	105.466	1.712	482.304
52	Control group	83.406	77.948	0.443	390.064
	Treatment group	124.040	115.923	0.659	580.096
53	Control group	63.958	69.680	0.310	440.826
	Treatment group	95.117	103.626	0.462	655.588
54	Control group	77.633	80.048	0.723	409.157
	Treatment group	115.454	119.046	1.075	608.491
55	Control group	72.773	67.736	0.148	329.409
	Treatment group	108.226	100.735	0.220	489.891
56	Control group	74.850	69.062	0.144	367.784
	Treatment group	111.315	102.707	0.214	546.961

Simulation	Group	Mean	SD	Min	Max
57	Control group	76.844	79.683	0.102	416.138
	Treatment group	114.280	118.503	0.152	618.872
58	Control group	82.858	89.985	0.456	488.928
	Treatment group	123.225	133.824	0.678	727.124
59	Control group	74.157	79.545	0.333	605.078

	Treatment group	110.284	118.298	0.495	899.859
60	Control group	68.327	66.743	0.217	479.250
	Treatment group	101.615	99.259	0.322	712.730
61	Control group	85.271	82.819	1.293	474.838
	Treatment group	126.813	123.167	1.922	706.170
62	Control group	75.237	73.147	0.202	349.057
	Treatment group	111.891	108.783	0.301	519.111
63	Control group	67.644	63.609	0.085	358.815
	Treatment group	100.598	94.598	0.126	533.622
64	Control group	71.008	75.003	0.167	366.527
	Treatment group	105.602	111.543	0.248	545.091
65	Control group	75.041	80.682	0.068	367.599
	Treatment group	111.600	119.989	0.101	546.685
66	Control group	73.488	71.673	0.146	352.680
	Treatment group	109.290	106.591	0.217	524.498
67	Control group	85.618	75.102	0.415	380.927
	Treatment group	127.329	111.691	0.618	566.507
68	Control group	80.438	85.819	1.501	470.673
	Treatment group	119.625	127.628	2.232	699.975
69	Control group	74.753	80.953	0.346	467.006
	Treatment group	111.172	120.392	0.514	694.522
70	Control group	74.716	80.600	0.470	498.299
	Treatment group	111.117	119.866	0.698	741.060
71	Control group	62.245	57.833	0.200	244.447
	Treatment group	92.569	86.008	0.297	363.537
72	Control group	66.688	65.514	0.116	392.603
	Treatment group	99.177	97.431	0.173	583.871
73	Control group	76.211	75.264	0.097	449.978
	Treatment group	113.340	111.931	0.144	669.198
74	Control group	74.658	65.242	1.426	332.546
	Treatment group	111.029	97.026	2.121	494.556
75	Control group	76.528	82.179	0.704	358.850
	Treatment group	113.811	122.215	1.047	533.674
76	Control group	70.929	67.926	0.233	352.743
	Treatment group	105.484	101.018	0.347	524.592

Simulation	Group	Mean	SD	Min	Max
77	Control group	69.340	63.837	0.107	425.971
	Treatment group	103.121	94.937	0.159	633.496
78	Control group	66.788	57.875	0.665	284.417
	Treatment group	99.326	86.070	0.989	422.980
79	Control group	73.599	71.591	0.753	331.062
	Treatment group	109.455	106.468	1.119	492.348

Simulation	Group	Mean	SD	Min	Max
	Control group	75.800	74.871	0.330	424.877
80	Treatment group	112.728	111.346	0.491	631.868
	Control group	71.802	80.462	0.651	582.252
81	Treatment group	106.782	119.662	0.968	865.913
	Control group	79.284	72.572	2.379	332.991
82	Treatment group	117.909	107.928	3.537	495.218
	Control group	82.525	86.377	0.283	533.147
83	Treatment group	122.729	128.458	0.421	792.885
	Control group	80.661	86.342	0.790	454.368
84	Treatment group	119.958	128.406	1.175	675.727
	Control group	74.260	63.361	2.403	277.752
85	Treatment group	110.438	94.229	3.574	413.067
	Control group	72.538	70.563	0.349	416.641
86	Treatment group	107.877	104.940	0.519	619.620
	Control group	71.643	67.247	1.852	291.207
87	Treatment group	106.546	100.009	2.754	433.076
	Control group	78.215	70.138	0.595	349.010
88	Treatment group	116.319	104.309	0.886	519.040
	Control group	82.109	74.843	0.476	387.498
89	Treatment group	122.111	111.305	0.708	576.280
	Control group	66.363	68.033	0.672	410.685
90	Treatment group	98.694	101.177	0.999	610.763
	Control group	71.660	65.195	1.890	358.646
91	Treatment group	106.571	96.956	2.811	533.371
	Control group	62.962	66.812	1.020	392.721
92	Treatment group	93.635	99.361	1.517	584.046
	Control group	82.903	79.088	0.137	474.513
93	Treatment group	123.292	117.618	0.204	705.687
	Control group	72.186	78.604	0.191	520.547
94	Treatment group	107.353	116.898	0.284	774.147
	Control group	63.984	65.105	1.550	367.117
95	Treatment group	95.156	96.823	2.305	545.969
	Control group	68.950	67.368	1.164	451.675
96	Treatment group	102.541	100.189	1.732	671.722

Simulation	Group	Mean	SD	Min	Max
	Control group	67.768	66.162	0.126	393.975
97	Treatment group	100.783	98.395	0.187	585.912
	Control group	89.914	86.512	0.539	461.968
98	Treatment group	133.719	128.659	0.802	687.029
	Control group	72.783	74.115	0.003	450.324
99	Treatment group	108.242	110.223	0.005	669.713
100	Control group	72.127	80.091	0.248	412.516

Treatment group	107.266	119.110	0.369	613.486