

國立中央大學

統計研究所

碩士論文

調整共變數兩組存活中位數差異之區間估計

指導教授：陳玉英博士

研究生：吳玉華

中華民國九十四年六月



# 國立中央大學圖書館 碩博士論文電子檔授權書

(93年5月最新修正版)

本授權書所授權之論文全文電子檔，為本人於國立中央大學，撰寫之碩/博士學位論文。(以下請擇一勾選)

- ()同意 (立即開放)  
()同意 (一年後開放)，原因是：\_\_\_\_\_。  
()同意 (二年後開放)，原因是：\_\_\_\_\_。  
()不同意，原因是：\_\_\_\_\_。

以非專屬、無償授權國立中央大學圖書館與國家圖書館，基於推動讀者間「資源共享、互惠合作」之理念，於回饋社會與學術研究之目的，得不限地域、時間與次數，以紙本、微縮、光碟及其它各種方法將上列論文收錄、重製、公開陳列、與發行，或再授權他人以各種方法重製與利用，並得將數位化之上列論文與論文電子檔以上載網路方式，提供讀者基於個人非營利性質之線上檢索、閱覽、下載或列印。

研究生簽名：\_\_\_\_\_ 吳 玉 華 \_\_\_\_\_

論文名稱：\_\_\_\_\_ 調整共變數兩組存活中位數差異之區間估計 \_\_\_\_\_

指導教授姓名：\_\_\_\_\_ 陳 玉 英 \_\_\_\_\_

系所：\_\_\_\_\_ 統 計 \_\_\_\_\_ 所 博士 碩士班

學號：\_\_\_\_\_ 92225015 \_\_\_\_\_

日期：民國\_\_\_\_\_ 94 \_\_\_\_\_ 年\_\_\_\_\_ 6 \_\_\_\_\_ 月\_\_\_\_\_ 27 \_\_\_\_\_ 日

備註：

1. 本授權書請填寫並親筆簽名後，裝訂於各紙本論文封面後之次頁（全文電子檔內之授權書簽名，可用電腦打字代替）。
2. 請加印一份單張之授權書，填寫並親筆簽名後，於辦理離校時交圖書館（以統一代轉寄給國家圖書館）。
3. 讀者基於個人非營利性質之線上檢索、閱覽、下載或列印上列論文，應依著作權法相關規定辦理。

## 摘要

本文針對接受兩種不同處理且具共變數的病人，探討如何在調整共變數後建立兩組存活函數中位數差異之區間估計。假設兩處理組內共變數對相對風險的影響一致且與時間無關，則可以分層 Cox 模式 (stratified Cox model) 描述病人的風險函數，藉此了解處理間的風險比例。共變數對兩個處理組內的相對風險影響不同，則分別就各處理組資料配適具共變數的 Cox 模式描述病人的風險狀況。然後，在設置模型之下估計各處理組之存活函數並求其變異數估計，代入 Su and Wei(1993) 的估計函數內求得調整共變數後，存活中位數差異之信賴區間。本文除以實例說明上述方法之應用，並且採用模擬方法研究比較本文所提方法與文獻既有方法之優劣。

## 誌謝辭

很感謝這兩年多來指導老師 陳玉英教授在專業知識上以及研究態度上的悉心指導，讓我受益良多，感謝口試委員嵇允嬋教授、謝國文教授及鄭秀麗教授給予本篇論文寶貴意見，使得本篇論文得以順利完成。

在研究所的兩年間，謝謝所上的老師們如此認真的在教學，讓我真的有學到一些知識，對統計有更深入的了解。感謝 陳玉英老師，在這將近三年當中，很努力的教導我，不僅僅是學業上的知識，還要有如何做研究的態度。雖然我常常都很搞不清楚狀況，對自己也很沒有信心，但是老師您還是不斷的想要把最好的東西給我，願意教我，真的很感謝您。謝謝玉媚學姊當我在研究上碰到很大難關的時候，不倦的指導我，幫助我學習。感謝啟軒學長，在我論文碰到很大難關的時候，給予我非常多的指導與幫助，非常的感謝你。感謝助理 劉碧華小姐，您就像母親一樣，每次有好的東西就會請我們，還會教我很多做人處世的道理，真的很謝謝您的照顧。謝謝楊文熙學長和許根寧學長，這幾個月，和你們一起討論很多問題與想法，給予我很多的啟發與幫助。以及我的同學們，非常感謝這兩年有你們和我一同學習。

謝謝福智團體的常師父和法師們，這兩年是我生命啟發很重要的兩年，謝謝你們在這之中，不斷的教導我何謂人生的真相，要如何的

去面對生命，如何在境界當中學習與突破，把握生命的宗旨。謝謝長華老師、依芸老師、婷尹、婉禎、鐘鐸社、福青社的大家，還有新竹大專班所有的老師同學們，沒有你們，這兩年不會如此精采，謝謝大家無窮的關懷與鼓勵，畢生難忘。

謹將本文獻給我在天上的父親，女兒從來沒有忘記您對我的教誨：要做一個社會上有用的人，好好學習。獻給我最深愛的母親、啟軒及弟弟，你們是我生命當中很重要的支柱，謝謝你們的包容與關愛。獻給師父，謝謝您教我學習廣論，佛法讓我的生命每一天都有不同的價值，每一個人事物開始有很特殊的意義，感謝您。

吳玉華 謹識  
于國立中央大學統計研究所  
中華民國九十四年六月

# 目 錄

第一章 緒論.....	1
第二章 文獻回顧.....	4
2.1 單樣本存活中位數之信賴區間.....	4
2.2 雙樣本存活中位數差異之信賴區間.....	8
第三章 統計方法.....	14
3.1 分層Cox模式.....	14
3.2 分組Cox模式.....	16
第四章 實例分析.....	19
第五章 模擬研究.....	32
3.1 模擬方法 .....	32
3.2 模擬結果 .....	37
第六章 結論與討論.....	39
參考文獻.....	41

## 圖表目錄

圖一：單一療法組與混合療法組之存活函數·····	20
圖二：單一療法組與混合療法組之對數累積風險圖·····	21
圖三：根據共變數為一般條件(C)的分層 Cox 模式建立存活中位數差異之信賴區間圖 ·····	27
圖四：根據共變數為一般條件 C=1 時的分層 Cox 模式建立存活中位數差異之信賴區間圖 ·····	30
圖五：根據共變數為一般條件 C=2 時的分層 Cox 模式建立存活中位數差異之信賴區間圖 ·····	30
圖六：根據共變數為一般條件 C=3 時的分層 Cox 模式建立存活中位數差異之信賴區間圖 ·····	31
圖七：各結構之存活函數與對數累積風險圖·····	34
表一：接受單一療法病人風險模式配適之 $-2\log \hat{L}$ 值·····	22
表二：接受混合療法病人風險模式配適之 $-2\log \hat{L}$ 值·····	23
表三：分層 Cox 模式配適之 $-2\log \hat{L}$ 值·····	24
表四：就兩組資料配適具處理、一般條件及一般條件與處理組的交互作用後之參數，及其參數之變異數，與對應之 p 值·····	26
表五：根據共變數為一般條件(C)的分層 Cox 模式建立存活中位數差異之信賴區間 ·····	26
表六：根據共變數為一般條件(C)與最初量測之腫瘤直徑大小(T)的分層 Cox 模式建立存活中位數差異之信賴區間 ·····	28
表七：根據共變數為一般條件(C)與最初量測之腫瘤直徑大小(T)的分層 Cox 模式建立存活中位數比例之信賴區間 ·····	29
表八：模擬之結構表·····	33
表九：固定 $x=0.5$ ， $n_1=n_2=50$ 時，各種結構之 95% 信賴區間的涵蓋機率及其平均區間長度。 ·····	37

表十：固定  $x = 0.5$ ， $n_1 = n_2 = 100$  時，各種結構之 95% 信賴區間的涵蓋  
    機率及其平均區間長度。 ..... 38

表十一：固定  $x = 0.5$ ， $n_1 = n_2 = 200$  時，各種結構之 95% 信賴區間的涵  
    蓋機率及其平均區間長度。 ..... 38

# 第一章 緒論

就對稱分布而言，平均數即可反應分布位置的訊息，因此一般的雙樣本問題經常就平均數的差異加以推論。但是就偏歪分布而言，平均數已不具母體位置代表性，此時，中位數才能反應母體的位置特性。因此，母體分布偏歪的雙樣本問題多數就中位數差異加以推估。一般而言，存活時間的分布為偏歪的，應該考慮的是兩個存活中位數之差異的推論。

過去已經有很多文獻探討單樣本的存活中位數之區間估計，可參考 Brookmeyer and Crowley(1982), Efron(1981), Emerson(1982), Reid(1981), Simon and Lee(1982) 及 Slud, Byar, and Green(1984) 等。針對具存活資料的雙樣本問題，Wang and Hettmansperger (1990)，利用兩個單樣本存活中位數之信賴區間，建立一個雙樣本存活中位數之信賴區間，但是應用此一方法必須要估計存活時間的密度函數。因此，Su and Wei (1993) 提出了一個無母數方法建立兩組存活中位數之差異或比例的信賴區間，此一方法不必估計存活時間的密度函數，因此計算較為簡易。

在臨床試驗中，除存活時間外，經常得知病人與病情相關的共變數。若這些共變數影響病人的存活時間，則稱為診斷因子(prognostic factor)。為研究共變數對於存活時間的影響，一般會先考慮採用 Cox

(1972) 比例風險模式配適存活資料。在 Cox 模式之下，具某種共變數病人相對於基準組病人的風險比例 (hazard ratio) 與時間無關。因為在分層設計中，各層 Cox 模式的相對風險比例一致且與時間無關，Zucker (1998) 提出運用分層 Cox 比例風險模式描述各層的風險比例。Kim (2001) 則運用分層 Cox 比例風險模式，分別就基準組存活中位數差異及對應平均存活函數中位數差異建立其信賴區間。值得一提的是，就分層 Cox 模式而言，具共變數的兩處理組病人，相對風險與兩基準組病人相對風險一致，但是兩處理組病人的存活函數及存活中位數與其共變數有關，所以 Kim (2001) 所提出的對應平均存活函數，即將共變數的影響平均後，所建立的存活中位數差異信賴區間並不恰當。鑑此，本文研究具相同共變數病人接受兩種不同處理的存活中位數差異之區間估計。

假如兩個處理組具共變數的病人與基準組病人之相對風險一致且與時間無關，本文以分層 Cox 模式描述病人的風險函數，藉此了解處理間的風險比例。若兩處理組具共變數的病人與基準組病人之相對風險不一致，但與時間無關，則本文建議分別就各處理組資料配適 Cox 模式描述病人的風險狀況，藉此了解處理間的風險比例。

本文第二章，回顧有關單樣本存活中位數之信賴區間及雙樣本存活中位數差異的區間估計，第三章提出兩種調整共變數兩組存活中位

數差異之信賴區間估計方法。第四章根據口腔及咽喉癌的研究資料 (Kalbfleisch and Prentice, 1980)，進行實例的分析。最後在第五章為結論及討論未來可能之研究。

## 第 2 章 文獻回顧

### 2.1 單樣本存活中位數之信賴區間

定義  $t(p)$  為存活函數的第  $p$  百分位數，例如  $t(50)$  即是中位數。令  $S(t)$  為在時間為  $t$  的存活函數 (survivor function)。就單樣本的存活資料而言，若觀察的存活時間為  $t_{(1)} < \dots < t_{(r)}$ ，則存活中位數  $t(50)$  之估計量為

$$\hat{t}(50) = \min\{t_{(j)} : \hat{S}(t_{(j)}) < 0.5, j = 1, 2, \dots, r\},$$

其中  $\hat{S}(t)$  為存活函數  $S(t)$  的 Kaplan-Meier 估計量。此一存活中位數之變異數估計為

$$\text{Var}\{\hat{t}(50)\} = \left( \frac{1}{\hat{f}\{\hat{t}(50)\}} \right)^2 \hat{\text{Var}}[\hat{S}\{\hat{t}(50)\}],$$

其中，若  $t_{(k)} \leq \hat{t}(50) < t_{(k+1)}$ ，則  $\text{Var}[\hat{S}\{\hat{t}(50)\}]$  的 Greenwood 估計式為

$$\hat{\text{Var}}\{\hat{S}\{\hat{t}(50)\}\} = \{\hat{S}\{\hat{t}(50)\}\}^2 \sum_{j=1}^k \frac{d_j}{n_j(n_j - d_j)},$$

而

$$\hat{f}\{\hat{t}(p)\} = \frac{\hat{S}\{\hat{u}(p)\} - \hat{S}\{\hat{l}(p)\}}{\hat{l}(p) - \hat{u}(p)}$$

為在時間  $\hat{t}(p)$  之機率密度估計，其中

$$\hat{u}(p) = \max\{t_{(j)} \mid \hat{S}(t_{(j)}) \geq 1 - \frac{p}{100} + \varepsilon, j = 1, 2, \dots, r\},$$

$$\hat{l}(p) = \min\{t_{(j)} \mid \hat{S}(t_{(j)}) \leq 1 - \frac{p}{100} - \varepsilon, j = 1, 2, \dots, r\}.$$

所以當  $p=50$  時，中位數  $t(50)$  的  $100(1-\alpha)\%$  信賴區間為

$$\hat{t}(50) \pm z_{\alpha/2} se\{\hat{t}(50)\}, \quad (2.1)$$

其中， $z_{\alpha/2}$  為標準常態分布的上  $\alpha/2$  百分位數， $se\{\hat{t}(50)\} = \sqrt{\hat{Var}(\hat{t}(50))}$ 。

若資料中另有共變數，則引用 Cox 比例風險模式配適存活資料。

假設  $T_1, \dots, T_n$  為互相獨立的存活時間， $C_1, \dots, C_n$  為對應的設限時間，此

外，對應  $T_i$  的共變數(診斷因子)為  $X_i = (X_{i1}, \dots, X_{ip})$ ，此時資料為

$(Y_i, \delta_i, X_i)$ ， $i=1, \dots, n$ ，其中觀測到的時間為  $Y_i = \min(T_i, C_i)$ ，右設限指

標為  $\delta_i = I(T_i \leq C_i)$ 。令  $\lambda(t|\mathbf{x})$  為給定  $X=\mathbf{x}$  時病人的風險函數，在此

$\lambda(t|\mathbf{X}) = \lambda_0(t) \exp(\boldsymbol{\beta}'\mathbf{X})$ ，其中  $\lambda_0(t)$  為基準( $x=0$ )風險函數，使得累積基準

風險函數  $\Lambda_0(t)$  存在。假設在給定  $\mathbf{X}$  之下， $T$  與  $C$  獨立，並且針對每一

個病人， $i=1, 2, \dots, n$ ，定義計數過程(counting process)

$$N_i(t) = I(T_i \leq t, \delta_i = 1)$$

$$\text{和 } J_i(t) = I(T_i \geq t)。$$

根據 Anderson et al. (1993) 結果，令

$$S^{(0)}(\boldsymbol{\beta}, t) = \sum_{j=1}^n J_j(t) \exp(\boldsymbol{\beta}'\mathbf{X}_j)$$

$$S^{(1)}(\boldsymbol{\beta}, t) = \sum_{j=1}^n \mathbf{X}_j J_j(t) \exp(\boldsymbol{\beta}'\mathbf{X}_j)$$

$$S^{(2)}(\boldsymbol{\beta}, t) = \sum_{j=1}^n \mathbf{X}_j^{\otimes 2} J_j(t) \exp(\boldsymbol{\beta}'\mathbf{X}_j)$$

$$E(\boldsymbol{\beta}, t) = \frac{S^{(1)}(\boldsymbol{\beta}, t)}{S^{(0)}(\boldsymbol{\beta}, t)}$$

$$V(\boldsymbol{\beta}, t) = \frac{S^{(2)}(\boldsymbol{\beta}, t)}{S^{(0)}(\boldsymbol{\beta}, t)} - E(\boldsymbol{\beta}, t)^{\otimes 2}$$

此處對任何一個向量  $\mathbf{a} = (a_1, \dots, a_r)$  而言， $\mathbf{a}^{\otimes 2}$  為一個  $r \times r$  的矩陣，其中每一個  $(i, j)$  項為  $a_i a_j$ 。

若 Anderson et al. (1993) 的 VII.2.1(a) 所列條件成立，且  $\boldsymbol{\beta}$  在  $\boldsymbol{\beta}_0$  的鄰域 (neighborhood)，則在觀察存活時間內 ( $0 \leq t \leq \tau$ )， $S^{(k)}(\boldsymbol{\beta}, t) \xrightarrow{u.p.} s^{(k)}(\boldsymbol{\beta}, t)$ ，其中  $\xrightarrow{u.p.}$  代表均勻機率收斂， $s^{(k)}(\boldsymbol{\beta}, t)$  為  $S^{(k)}(\boldsymbol{\beta}, t)$  的期望值。所以，

$$e(\boldsymbol{\beta}, t) = \frac{s^{(1)}(\boldsymbol{\beta}, t)}{s^{(0)}(\boldsymbol{\beta}, t)}$$

$$v(\boldsymbol{\beta}, t) = \frac{s^{(2)}(\boldsymbol{\beta}, t)}{s^{(0)}(\boldsymbol{\beta}, t)} - e(\boldsymbol{\beta}, t)^{\otimes 2}$$

進一步由 Anderson et al. (1993) 定理 VII.2.2 得知，

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{D} N(0, \boldsymbol{\Sigma}^{-1})，$$

其中  $\xrightarrow{D}$  代表分布收斂，

$$\boldsymbol{\Sigma} = \int_0^{\tau} v(\boldsymbol{\beta}_0, u) s^{(0)}(\boldsymbol{\beta}_0, u) \lambda_0(u) du$$

$\boldsymbol{\Sigma}$  為一  $k \times k$  正定矩陣。此一風險模式尚需估計基準風險函數  $\lambda_0(t)$ ，

Breslow (1972) 提出一個累積基準風險函數估計式，為

$$\hat{\Lambda}_0(t) = -\log \hat{S}_0(t) = \int_0^t (S^{(0)}(\hat{\boldsymbol{\beta}}, s))^{-1} dN(s)$$

其中存活函數為  $S(t)$  估計式為  $\hat{S}(t) = \hat{S}_0(t) \exp(\hat{\boldsymbol{\beta}}' \mathbf{x})$ 。

定義在給定共變數  $\mathbf{x}$  之下，第  $i$  組的存活時間  $T$  之存活中位數為

$$\xi(50|\mathbf{x}) = \inf\{t: S(t|\mathbf{x}) \leq \frac{1}{2}\} \quad ,$$

則其估計式為

$$\hat{\xi}(50|\mathbf{x}) = \inf\{t: \hat{S}(t|\mathbf{x}) \leq \frac{1}{2}\} \quad . \quad (2.2)$$

由 Doss and Gill (1992) 結果得知

$$\sqrt{n}\{\hat{\xi}(50) - \xi(50)\} \approx \frac{\sqrt{n}\{S(\hat{\xi}(50)) - S(\xi(50))\}}{0.5\lambda(\xi(50))} \quad . \quad (2.3)$$

且 Anderson and Gill (1982) 證明  $\sqrt{n}\{\hat{S}_0(t) - S_0(t)\}$  將會弱收斂 (weakly converge) 至一個高斯過程 (Gaussian process)，其平均數為 0，變異數為

$$\sigma^2(t) = S^2(t)\{(a(t)) + b(t)'\Sigma^{-1}b(t)\} \quad .$$

其中

$$\begin{aligned} a(t) &= \int_0^t \frac{\lambda_0(u)}{S^{(0)}(\boldsymbol{\beta}, u)} du \quad , \quad b(t) = \int_0^t e^{(0)}(\boldsymbol{\beta}, u)\lambda_0(u) du \quad , \\ \Sigma &= \int_0^t v(\boldsymbol{\beta}, u) s^{(0)}(\boldsymbol{\beta}, u)\lambda_0(u) du \quad . \end{aligned} \quad (2.4)$$

由 (2.3) 與 (2.4)，可推知  $\sqrt{n}\{\hat{\xi}(p) - \xi(p)\}$  亦會弱收斂至一個高斯過程，其平均數為 0，變異數為

$$V(\xi(p)) = \frac{\sigma^2(\xi(p))}{(1-p)^2 \lambda^2(\xi(p))} = \frac{[(a(t)) + b(t)'\Sigma^{-1}b(t)]}{\lambda^2(\xi(p))} \quad .$$

所以，得到  $\hat{\xi}(p)$  的變異數的估計式為

$$\hat{V}(\hat{\xi}(p)) = \frac{\hat{\sigma}^2(\hat{\xi}(p))}{(1-p)^2 \hat{\lambda}^2(\hat{\xi}(p))} = \frac{[(\hat{a}(t)) + \hat{b}(t)'\hat{\Sigma}^{-1}\hat{b}(t)]}{\hat{\lambda}^2(\hat{\xi}(p))}$$

其中  $a(t)$  ,  $b(t)$  ,  $\Sigma$  ,  $\xi(p)$  的估計式為

$$\begin{aligned}\hat{a}(t) &= n \int_0^t \frac{d\hat{\Lambda}_0(u)}{S^{(0)}(\hat{\beta}, u)} , \\ \hat{b}(t) &= \int_0^t \frac{S^{(1)}(\hat{\beta}, u) d\hat{\Lambda}_0(u)}{S^{(0)}(\hat{\beta}, u)} , \\ \hat{\Sigma} &= \frac{1}{n} \int_0^t v(\hat{\beta}, u) S^{(0)}(\hat{\beta}, u) \hat{\lambda}_0(u) du\end{aligned}\quad (2.5)$$

而  $\hat{\lambda}_0(t) = h^{-1} \int_0^\infty K\{h^{-1}(t-s)\} d\hat{\Lambda}_0(s)$  為 Wells (1994) 提出之  $\lambda_0(t)$  的估計式，且  $K(t) = 0.75(1-|t|^2)I(|t| \leq 1)$  為 Epanechnikov 核函數(kernel function)。

綜合以上所述，得知共變數為  $\mathbf{x}$  時，存活中位數  $\hat{\xi}(50|\mathbf{x})$  (2.2) 的信賴區間為

$$\hat{\xi}(50|\mathbf{x}) \pm z_{\alpha/2} \sqrt{V(\hat{\xi}(50|\mathbf{x}))} \quad (2.6)$$

其中  $z_{\alpha/2}$  為標準常態分配的上  $\alpha/2$  百分位數。

## 2.2 雙樣本存活中位數差異之信賴區間

在不考慮共變數資訊之下，可由前面(2.1)之單樣本的結果，求得以下雙樣本中位數差異之信賴區間：

$$\hat{t}_2(50) - \hat{t}_1(50) \pm z_{\alpha/2} \sqrt{\hat{V}ar[\hat{t}_1(50)] + \hat{V}ar[\hat{t}_2(50)]} ,$$

其中  $\hat{t}_i(50)$  為第  $i$  組存活中位數之估計量， $\hat{V}ar\{\hat{t}_i(p)\}$  則為  $\hat{t}_i(p)$  的變異數估計量。Su and Wei(1993)則提出一個較為簡便的無母數方法。令  $\theta_1$  為第一組的存活中位數，即  $S_1(\theta_1) = 1/2$ 。考慮函數  $g(\tau, \theta_1) = \tau + \theta_1$  或

$g(\tau, \theta_1) = \tau\theta_1$ 。若  $\theta_2 = g(\tau, \theta_1)$  為第二組的存活中位數，則  $g(\tau, \theta_1) = \tau + \theta_1$  代表  $\tau = \theta_2 - \theta_1$ ，為兩組存活中位數差異；若  $g(\tau, \theta_1) = \tau\theta_1$ ，則代表  $\tau = \theta_2/\theta_1$ ，為兩組存活中位數的比例。Su and Wei (1993)建議如下估計式：

$$W(\tau, \theta_1) = \frac{(\hat{S}_1(\theta_1) - \frac{1}{2})^2}{\sigma_1^2(\theta_1)} + \frac{(\hat{S}_2(g(\tau, \theta_1)) - \frac{1}{2})^2}{\sigma_2^2(g(\tau, \theta_1))} \quad (2.7)$$

其中  $\sigma_i^2$  為  $\hat{S}_i(t)$  的變異數之 Greenwood 估計式， $i=1,2$ 。得知  $G(\tau) = \min_{\theta_1} W(\tau, \theta_1)$  為 c(Basawa and Koul, 1988)。所以， $\tau$  的  $100(1-\alpha)\%$  信賴區間為

$$I = \{\tau : G(\tau) < \chi_1^2(\alpha)\} ,$$

其中  $\chi_1^2(\alpha)$  為卡方分布的上  $\alpha$  百分位。

Kim(2001)考慮針對具共變數  $\mathbf{x}$  的兩個處理組資料，配適分層 Cox 比例風險模式：

$$\lambda_i(t; \mathbf{x}) = \lambda_{i0}(t) \exp(\boldsymbol{\beta}'\mathbf{x}) , \quad i=1,2 \quad (2.8)$$

其中  $\lambda_{i0}(t)$  為第  $i$  組之基準風險函數， $\boldsymbol{\beta}$  則為對應  $\mathbf{x}$  的未知參數。並且藉由基準存活函數與對應平均存活函數求出調整共變數後兩組存活中位數差異之信賴區間。由(2.8)得知

$$\frac{\lambda_{21}(t; \mathbf{x})}{\lambda_{20}(t)} = \frac{\lambda_{11}(t; \mathbf{x})}{\lambda_{10}(t)} , \quad \frac{\lambda_{21}(t; \mathbf{x})}{\lambda_{11}(t; \mathbf{x})} = \frac{\lambda_{20}(t)}{\lambda_{10}(t)}$$

亦即在各處理組內，具共變數  $\mathbf{x}$  病人與其基準病人的相對風險一致；此外，兩個處理組具相同共變數病人的相對風險與兩個處理組基

準病人的相對風險一樣。因此，Kim (2001) 首先針對兩個基準組病人存活中位數差異建立信賴區間。令  $S_{i_0}(t)$  為第  $i$  組的基準存活函數 (baseline survival function)，則第  $i$  組的存活函數為  $S_i(t) = S_{i_0}(t)\exp(\boldsymbol{\beta}'\mathbf{x})$ ，其估計式為  $\hat{S}_i(t) = \hat{S}_{i_0}(t)\exp(\hat{\boldsymbol{\beta}}'\mathbf{x})$ 。由 Andersen et al. (1993) 的推論 VII.2.4 可知， $\sqrt{n}\{\hat{S}_{i_0}(t) - S_{i_0}(t)\}$ ， $i=1,2$ ，弱收斂到一個二維的高斯過程，平均數為 0，變異數為

$$\sigma^2(s,t) = S_{i_0}(s)S_{i_0}(t)[\delta_{i'i'}(a_i(s \wedge t)) + b_i(s)'\boldsymbol{\Sigma}^{-1}b_i(t)] \quad , \quad i, i' = 1, 2$$

其中

$$a_i(t) = \int_0^t \frac{\lambda_{0i}(s)}{S_{i_0}^{(0)}(\boldsymbol{\beta}_0, s)} ds \quad , \quad b_i(t) = \int_0^t e_i(\boldsymbol{\beta}_0, s)\lambda_{i_0}(s) ds$$

$$\boldsymbol{\Sigma} = \sum_{i=1}^2 \int_0^t v_i(\boldsymbol{\beta}_0, s) S_{i_0}^{(0)}(\boldsymbol{\beta}_0, s) \lambda_{i_0}(s) ds$$

因為存活時間  $T$  之基準存活中位數  $\xi_i(50)$  (baseline survivor median)

為

$$\xi_i(50) = \inf\{t : S_{i_0}(t) \leq 0.5\} \quad i=1,2 \quad (2.9)$$

其估計式為

$$\hat{\xi}_i(50) = \inf\{t : \hat{S}_{i_0}(t) \leq 0.5\} \quad i=1,2 \quad (2.10)$$

假設兩組之存活中位數差異為： $\Delta = \xi_1(50) - \xi_2(50)$  且其估計式為

$$\hat{\Delta} = \hat{\xi}_1(50) - \hat{\xi}_2(50) \quad .$$

設

$$\Theta(t_1, t_2) = \frac{b_1(t_1)}{\lambda_{10}(t_1)} - \frac{b_2(t_2)}{\lambda_{20}(t_2)}$$

則由 Kim(2001)之定理一可得

$$v(\xi_1(p), \xi_2(p)) = \sum_{i=1}^2 \frac{a_i(\xi_i(p))}{\lambda_{i0}(\xi_i(p))} + \Theta(\xi_1(p), \xi_2(p))' \Sigma^{-1} \Theta(\xi_1(p), \xi_2(p)) \quad (2.11)$$

其中  $a_i(t)$  ,  $b_i(t)$  ,  $\Sigma$  ,  $\xi_i(p)$  的估計式分別為

$$\begin{aligned} \hat{a}_i(t) &= n \int_0^t \frac{d\hat{\Lambda}_{i0}(s)}{S_i^{(0)}(\hat{\beta}, s)} , \\ \hat{b}_i(t) &= \int_0^t \frac{S_i^{(1)}(\hat{\beta}, s) d\hat{\Lambda}_{i0}(s)}{S_i^{(0)}(\hat{\beta}, s)} , \\ \hat{\Sigma} &= \frac{1}{n} \sum_{i=1}^2 \int_0^t V_i(\hat{\beta}, s) S_i^{(0)}(\hat{\beta}, s) \hat{\lambda}_{i0}(s) ds \end{aligned} \quad (2.12)$$

及事實上， $\hat{\lambda}_{i0}(t)$  為 Wells(1994)提出之估計式

$$\hat{\lambda}_{i0}(t) = \frac{1}{h_i} \int_0^\infty K_i\left(\frac{t-s}{h_i}\right) d\hat{\Lambda}_{i0}(s) , \quad (2.13)$$

其中， $K(t) = 0.75(1-|t|^2)I(|t| \leq 1)$  為 Epanechnikov 核函數。由上述推導可

知兩組存活中位數差異  $\hat{\Delta} = \hat{\xi}_1(50) - \hat{\xi}_2(50)$  之  $100(1-\alpha)\%$  的信賴區間為：

$$\hat{\Delta} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{v}}{n}} \quad (2.14)$$

其中  $\hat{v}$  為 (2.11)  $v$  之估計式， $z_{\frac{\alpha}{2}}$  為標準常態分布之上  $\alpha/2$  百分位數。

Kim(2001) 也藉由對應平均存活函數求得中位數，並且建立存活中位數差異的信賴區間。令  $S_i(t|\mathbf{x})$  為考慮共變數  $\mathbf{x}$  之下，第  $i$  組的存活

函數， $i=1,2$ 。則在分層 COX 模式下(2.8)， $S_i(t|\mathbf{x}) = \exp\{-\exp(\boldsymbol{\beta}'\mathbf{x})\Lambda_{i0}(t)\}$ ，其估計式為  $\hat{S}_i(t|\mathbf{x}) = \exp\{-\exp(\hat{\boldsymbol{\beta}}'\mathbf{x})\hat{\Lambda}_{i0}(t)\}$ 。定義第  $i$  組的平均存活函數為

$$S_{\cdot i}(t|\mathbf{x}) = \frac{1}{n} \sum_{g=1}^2 \sum_{j=1}^{n_g} S_i(t|\mathbf{x}_{gj}) , \quad i=1,2 ,$$

則第  $i$  組的平均存活函數  $S_{\cdot i}(t|\mathbf{x})$  之中位數為

$$\xi_{\cdot i}(50) = \inf\{t : S_{\cdot i}(t) \leq 0.5\} \quad i=1,2$$

其估計式為

$$\hat{\xi}_{\cdot i}(50) = \inf\{t : \hat{S}_{\cdot i}(t) \leq 0.5\} \quad i=1,2 .$$

假設兩組之存活中位數差異為  $\Delta_{\cdot} = \xi_{\cdot 1}(50) - \xi_{\cdot 2}(50)$ ，且其估計式為

$$\hat{\Delta}_{\cdot} = \hat{\xi}_{\cdot 1}(50) - \hat{\xi}_{\cdot 2}(50) .$$

在組別為  $i=1,2$  時，設

$$\Gamma_i(t) = \frac{1}{n} \sum_{g=1}^2 \sum_{j=1}^{n_g} S_i(t|\mathbf{x}_{gj}) \exp(\boldsymbol{\beta}'\mathbf{x}_{gj}) ,$$

$$\Omega_i(t) = \frac{1}{n} \sum_{g=1}^2 \sum_{j=1}^{n_g} S_i(t|\mathbf{x}_{gj}) \exp(\boldsymbol{\beta}'\mathbf{x}_{gj}) \mathbf{x}_{gj} ,$$

$$\Phi_i(t) = -\Lambda_{i0}(t)\Omega_i(t) + \Gamma_i(t)b_i(t) ,$$

$$\Psi(t_1, t_2) = \frac{\Lambda_{10}(t_1)\Omega_1(t_1)}{\lambda_{10}(t_1)\Gamma_1(t_1)} - \frac{\Lambda_{20}(t_2)\Omega_2(t_2)}{\lambda_{20}(t_2)\Gamma_2(t_2)}$$

由 Andersen et al.(1993) 的推論 VII.2.2 與 VII.2.3 可知，

$\sqrt{n}\{\hat{S}_{\cdot i}(t) - S_{\cdot i}(t)\}$ ， $i=1,2$ ，弱收斂到一個二維的高斯過程，其平均數為 0，變異數為

$$\tau(s, t) = S_{i' i'} \Gamma_i(s) \Gamma_{i'}(t) a_i(s \wedge t) + \Phi_i(s)' \boldsymbol{\Sigma}^{-1} \Phi_{i'}(t) , \quad i, i' = 1, 2$$

由 Kim(2001)之定理二則可知

$$\begin{aligned}
 v(\xi_1(p), \xi_2(p)) &= \sum_{i=1}^2 \frac{a_i(\xi_1(p))}{\lambda_{i0}^2(\xi_1(p))} + \{\Theta(\xi_1(p), \xi_1(p)) - \Psi(\xi_1(p), \xi_1(p))\}' \\
 &\quad \times \Sigma^{-1} \Theta(\xi_1(p), \xi_1(p)) - \Psi(\xi_1(p), \xi_1(p))\}
 \end{aligned} \tag{2.15}$$

其中  $a_i(t)$  ,  $b_i(t)$  ,  $\Sigma$  ,  $\xi_i(p)$  ,  $\lambda_{i0}(t)$  的估計式同(2.13)及(2.12)。由上

述推導可知兩組存活中位數差異  $\hat{\Delta}$  之  $100(1-\alpha)\%$  信賴區間為

$$\hat{\Delta} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{v}}{n}} \tag{2.16}$$

其中  $\hat{v}$  為  $v(2.15)$  之估計式 ,  $z_{\frac{\alpha}{2}}$  為標準常態分布之上  $\alpha/2$  百分位數。

## 第三章 統計方法

### 3.1 分層 Cox 模式

假設不同處理組內具相同共變數  $\mathbf{x}$  的病人與其基準病人之相對風險一致且與時間無關，則可以分層 Cox 模式(2.8)描述病人的風險函數，藉此了解處理間的風險比例。本文首先推廣 Su and Wei 雙樣本存活中位數差異信賴區間 (2.7)，求出調整共變數後，存活中位數差異之信賴區間。

在分層 Cox 模式之下，第  $i$  組的存活函數為， $S_i(t) = \{S_{i0}(t)\}^{\exp(\beta' \mathbf{x})}$ ，其估計式  $\hat{S}_i(t) = \{\hat{S}_{i0}(t)\}^{\exp(\hat{\beta}' \mathbf{x})}$ ，其中  $S_{i0}(t)$  與  $\hat{S}_{i0}(t)$  為第  $i$  組的基準存活函數及其估計數。在共變數  $\mathbf{X} = \mathbf{x}$  之下，令估計方程式為

$$W(\tau, \xi_1 | \mathbf{x}) = \frac{(\hat{S}_1(\xi_1 | \mathbf{x}) - 0.5)^2}{\sigma_1^2(\hat{\xi}_1 | \mathbf{x})} + \frac{(\hat{S}_2(g(\tau, \xi_1) | \mathbf{x}) - 0.5)^2}{\sigma_2^2(g(\tau, \xi_1) | \mathbf{x})} \quad (3.1)$$

由 Klein and Zhang(2001)結果可得知，

$$U(t | \mathbf{x}) = \sqrt{n} \{ \hat{S}_i(t | \mathbf{x}) - S_i(t | \mathbf{x}) \} = \sqrt{n} \{ \exp(-\hat{\Lambda}_i(t | \mathbf{x})) - \exp(-\Lambda_i(t | \mathbf{x})) \}$$

弱收斂於

$$U_1(t | \mathbf{x}) = \sqrt{n} \{ S_i(t | \mathbf{x}) [\hat{\Lambda}_i(t | \mathbf{x}) - \Lambda_i(t | \mathbf{x})] \}$$

令  $M_{ij}$  為平均數為 0 的鞅 (martingale)， $i=1,2$ ， $j=1,2,\dots,n_i$  且

$\bar{M}_i(u) = \sum_{j=1}^{n_i} M_{ij}$ 。進一步由 Andersen and Gill(1982)結果得知  $U_1(t | \mathbf{x})$

弱收斂於  $\tilde{U}(t | \mathbf{x})$ ，其中

$$\begin{aligned}\tilde{U}(t|\mathbf{x}) &= \sqrt{n}S_i(t|\mathbf{x}) \int_0^t \frac{\exp(\beta'x)d\bar{M}_i(u)}{n_i S_i^{(0)}(\hat{\boldsymbol{\beta}}, u)} \\ &\quad + \{S_i(t\mathbf{x})r_i(t|\mathbf{x})i\}'\boldsymbol{\Sigma}^{-1} \left\{ \frac{1}{\sqrt{n_i}} \sum_{j=1}^{n_i} \int_0^{\infty} [\mathbf{x}_{ij} - E_i(\hat{\boldsymbol{\beta}}, u)] dM_{ij}(u) \right\}\end{aligned}\tag{3.2}$$

其中

$$\begin{aligned}S_i^{(0)}(\boldsymbol{\beta}, t) &= \frac{1}{n_i} \sum_{j=1}^{n_i} J_j(t) \exp(\boldsymbol{\beta}'\mathbf{X}_j) \\ r_i(t|\mathbf{x}) &= \int_0^t \exp(\boldsymbol{\beta}'_0\mathbf{x})[\mathbf{x} - e_i(\boldsymbol{\beta}, u)]\lambda_{i0}(u)du ,\end{aligned}$$

其估計式為

$$\hat{r}_i(t|\mathbf{x}) = \int_0^t \exp(\hat{\boldsymbol{\beta}}'_0\mathbf{x})[\mathbf{x} - E_i(\hat{\boldsymbol{\beta}}, u)] \frac{dN_i(u)}{n_i S_i^{(0)}(\hat{\boldsymbol{\beta}}, u)} .$$

由 Anderson et al. (1993) II.5.1 的鞅中央極限定理 (Rebolledo's theorem) 得知，就  $0 \leq t \leq \tau$ ， $\tau < \inf_{t>0} \{t : ET_{ij} = 0\}$  時， $\tilde{U}(t|\mathbf{x})$  弱收斂於一個平均數為 0 的高斯鞅 (Gaussian vector martingale)，其變異數可估計為

$$\hat{\sigma}_i^2(t|\mathbf{x}) = \hat{S}_i^2(t|\mathbf{x}) \left\{ \frac{n}{n_i^2} \int_0^t \frac{\exp(2\hat{\boldsymbol{\beta}}'\mathbf{x}) dN_i(u)}{[S^{(0)}(\hat{\boldsymbol{\beta}}, u)]^2} + [\hat{r}_i(t|\mathbf{x})]' \hat{\boldsymbol{\Sigma}}^{-1} [\hat{r}_i(t|\mathbf{x})] \right\} \quad i=1,2$$

其中  $N_i(s) = \sum_j N_{ij}(s)$ 。

所以， $\hat{S}_i(t|\mathbf{x})$  的變異數估計式為

$$\hat{\sigma}_i^2(t|\mathbf{x}) = \hat{S}_i^2(t|\mathbf{x}) \left\{ \frac{n}{n_i^2} \int_0^t \frac{\exp(2\hat{\boldsymbol{\beta}}'\mathbf{x}) dN_i(u)}{[S^{(0)}(\hat{\boldsymbol{\beta}}, u)]^2} + [\hat{r}_i(t|\mathbf{x})]' \hat{\boldsymbol{\Sigma}}^{-1} [\hat{r}_i(t|\mathbf{x})] \right\} , \quad i=1,2\tag{3.3}$$

同理可知，在  $\mathbf{X}=\mathbf{x}$  之下， $G(\tau|\mathbf{x})=\min_{\xi_1} W(\tau, \xi_1|\mathbf{x})$  服從自由度為 1 的卡方分布 (Basawa and Koul, 1988)。因此，在  $\mathbf{X}=\mathbf{x}$  之下，存活中位數差異或比例  $\tau$  的  $100(1-\alpha)\%$  信賴區間為

$$I = \{\tau : G(\tau) < \chi_1^2(\alpha)\},$$

其中  $\chi_1^2(\alpha)$  為卡方分布的上  $\alpha$  百分位。

### 3.2 分組 Cox 模式

當碰到非等比例風險的資料，假設不同處理組內具相同共變數  $\mathbf{x}$  的病人與其基準病人之相對風險不一致，但與時間無關時，則可以分別就各處理組資料配適 Cox 模式描述病人的風險函數，藉此了解處理間的風險比例。本文推廣 Su and Wei 雙樣本存活中位數差異信賴區間 (2.7)，求出調整共變數後，存活中位數差異之信賴區間。首先，在共變數為  $\mathbf{x}$  之下，於各組中考慮以下 Cox 比例風險模式

$$\lambda_i(t; \mathbf{x}) = \lambda_{i0} \exp(\boldsymbol{\beta}'_i \mathbf{x}), \quad i=1,2 \quad (3.4)$$

其中  $\lambda_{i0}(t)$  為第  $i$  組的基準風險函數， $\boldsymbol{\beta}_i = (\beta_{i1}, \dots, \beta_{ip})'$  為第  $i$  組的未知參數， $i=1,2$ 。在此模式 (3.4) 之下，第  $i$  組的存活函數為  $S_i(t) = \{S_{i0}(t)\}^{\exp(\boldsymbol{\beta}'_i \mathbf{x})}$ ，其估計式  $\hat{S}_i(t) = \{\hat{S}_{i0}(t)\}^{\exp(\hat{\boldsymbol{\beta}}'_i \mathbf{x})}$ ，其中  $S_{i0}(t)$  與  $\hat{S}_{i0}(t)$  為第  $i$  組的基準存活函數及其估計量。在共變數  $\mathbf{X}=\mathbf{x}$  之下，令估計方程式為

$$W(\tau, \xi_1 | \mathbf{x}) = \frac{(\hat{S}_1(\xi_1 | \mathbf{x}) - 0.5)^2}{\sigma_1^2(\xi_1 | \mathbf{x})} + \frac{(\hat{S}_2(g(\tau, \xi_1) | \mathbf{x}) - 0.5)^2}{\sigma_2^2(g(\tau, \xi_1) | \mathbf{x})} \quad (3.5)$$

由 Tsiatis (1981) 結果可得知

$$\sqrt{n} \{ \hat{\Lambda}_0(t) - \Lambda_0(t) \} \xrightarrow{D} W(a(t)) - b(t)^T \Sigma^{-1/2} \mathbf{Z} \quad (3.6)$$

其中， $\xrightarrow{D}$  代表分布收斂。就  $0 \leq t < \infty$ ， $W(\cdot)$  為一標準的布朗運動 (Brownian motion)， $\mathbf{Z}$  為一個  $q$  維的常態分布隨機向量，其平均值為 0，變異數為單位矩陣 (identity covariance matrix)， $\mathbf{Z}$  與  $W(\cdot)$  獨立。且由以上可知

$$\hat{V}ar[\sqrt{n} \{ \hat{\Lambda}_0(t) - \Lambda_0(t) \}] = \hat{a}(t) + \hat{b}(t)^T \hat{\Sigma}^{-1} \hat{b}(t)$$

此處，

$$\hat{a}(t) = \int_0^t \frac{d\hat{\Lambda}_0(u)}{S^{(0)}(\hat{\boldsymbol{\beta}}, u)}, \quad \hat{b}(t) = \int_0^t \frac{S^{(1)}(\hat{\boldsymbol{\beta}}, u) d\hat{\Lambda}_0(u)}{S^{(0)}(\hat{\boldsymbol{\beta}}, u)},$$

$$\hat{\Sigma} = \int_0^t v(\hat{\boldsymbol{\beta}}, u) S^{(0)}(\hat{\boldsymbol{\beta}}, u) \lambda_0(u) du$$

由 Anderson and Gill (1982) 結果得知  $\sqrt{n} \{ \hat{\Lambda}_0(t | \mathbf{x}) - \Lambda_0(t | \mathbf{x}) \}$  之變異數估計式為

$$\begin{aligned} \hat{V}ar[\sqrt{n} \{ \hat{\Lambda}_0(t | \mathbf{x}) - \Lambda_0(t | \mathbf{x}) \}] &= \hat{V}ar[\sqrt{n} \{ \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}) \hat{\Lambda}_0(t) - \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}) \Lambda_0(t) \}] \\ &= \exp(2\hat{\boldsymbol{\beta}}' \mathbf{x}) \{ \hat{a}(t) + \hat{H}'(t) \Sigma^{-1} \hat{H}(t) \}, \end{aligned}$$

其中，

$$\hat{H}(t) = \int_0^t \left[ \frac{S^{(1)}(\hat{\boldsymbol{\beta}}, u)}{S^{(0)}(\hat{\boldsymbol{\beta}}, u)} - \mathbf{x}' \right] d \hat{\Lambda}_0(u)$$

則由(3.6)式可得知

$$\sqrt{n}\{\hat{S}_0(t) - S_0(t)\} \xrightarrow{D} S_0(t) \left[ W(a(t)) - b(t)^T \Sigma^{-1/2} \mathbf{Z} \right]$$

且  $\sqrt{n}\{\hat{S}_0(t) - S_0(t)\}$  之變異數估計式為

$$\hat{V}ar[\sqrt{n}\{\hat{S}_0(t) - S_0(t)\}] = \hat{S}_0^2(t) [(\hat{a}(t)) + \hat{b}(t)^T \Sigma^{-1} \hat{b}(t)]$$

並由 Anderson and Gill (1982) 可得知(3.5)式中  $\hat{S}_i(t|\mathbf{x})$  的變異數估計式為

$$\begin{aligned} \hat{\sigma}_i^2(t|\mathbf{x}) &= \hat{V}ar[\sqrt{n}\{\hat{S}_i(t) - S_i(t)\}] \\ &= \hat{V}ar[\sqrt{n}\{\exp(-\hat{\Lambda}(t|\mathbf{x})) - \exp(-\Lambda(t|\mathbf{x}))\}] \\ &= [\hat{S}_i(t|\mathbf{x})]^2 \exp(2\hat{\boldsymbol{\beta}}'_i \mathbf{x}) \{\hat{a}_i(t) + \hat{Q}'_i(t) \Sigma_i^{-1} \hat{Q}_i(t)\}, \end{aligned}$$

其中，

$$\hat{Q}_i(t) = \int_0^t \left[ \mathbf{x} - \frac{\mathbf{S}^{(1)}(\hat{\boldsymbol{\beta}}, u)}{\mathbf{S}^{(0)}(\hat{\boldsymbol{\beta}}, u)} \right]^T d\hat{\Lambda}_{i0}(u),$$

$$\hat{a}_i(t) = \int_0^t \frac{d\hat{\Lambda}_{i0}(u)}{S_i^{(0)}(\hat{\boldsymbol{\beta}}, u)},$$

$$S_i^{(0)}(\boldsymbol{\beta}, t) = \frac{1}{n_i} \sum_{j=1}^n J_j(t) \exp(\boldsymbol{\beta}' \mathbf{X}_j).$$

同理可知，在  $\mathbf{X} = \mathbf{x}$  之下， $G(\tau|\mathbf{x}) = \min_{\xi_1} W(\tau, \xi_1|\mathbf{x})$  服從自由度為 1 的卡方

分布(Basawa and Koul, 1988)。因此，在  $\mathbf{X} = \mathbf{x}$  之下，存活中位數差

異或比例  $\tau$  的  $100(1-\alpha)\%$  信賴區間為

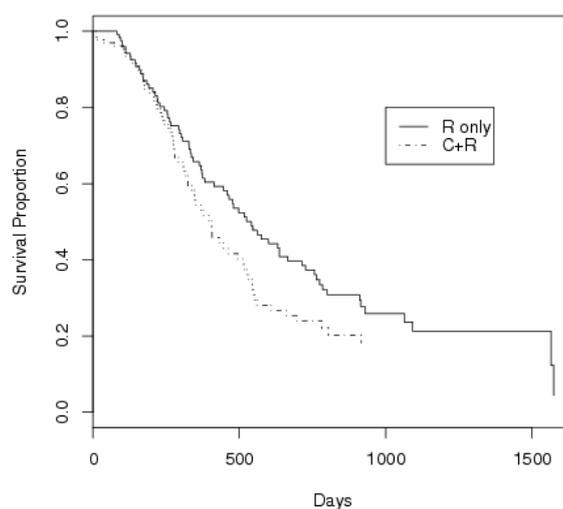
$$I = \{\tau : G(\tau) < \chi_1^2(\alpha)\},$$

其中  $\chi_1^2(\alpha)$  為卡方分布的上  $\alpha$  百分位。

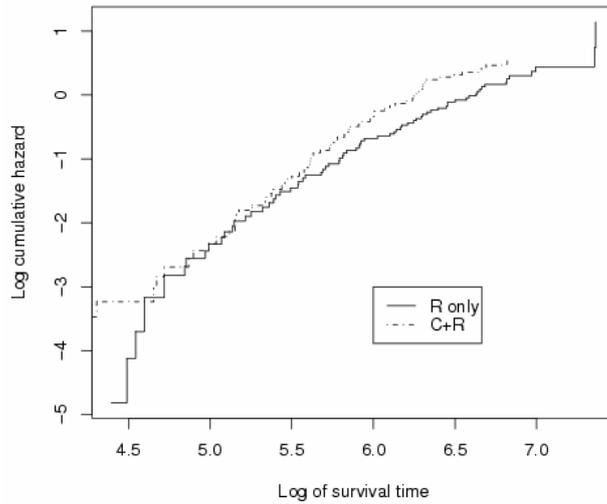
## 第四章 實例分析

腫瘤學放射性治療組織 (Radiation Therapy Oncology Group) 所做的隨機臨床實驗 (Kalbfleisch and Prentice, 1980) 主要是比較放射線(單一)療法(radiation therapy)，與化學療法配合放射線(混合)療法(chemotherapy followed by radiation therapy)，在治療口腔與咽喉癌上的差異。此一實驗中有 100 位病人接受單一療法，其他 95 位病人接受混合療法。研究中有 6 個診斷因子，分別為性別 (Sex)，腫瘤等級 (tumor grade)，病人一般狀況 (general condition)，腫瘤位置 (tumor site)，最初量測之腫瘤直徑大小 (T stage) 和淋巴結轉移狀況 (N stage)。此處腫瘤等級(簡記為 G)是腫瘤細胞(tumor cell)與原始細胞(host cell)的相似程度，分為可良好區別 (well differentiation)、適度的可區別 (moderately differentiation) 與不易區別 (poorly differentiation) 三類。病人的一般狀況(簡記為 C)，為病人身體機能(functional capacity)的測量值，分為非無行為能力或非殘疾者(no disability)、只能做限定的行為動作(restricted work)、與在生活上需要特別的照顧 (requires assistance with self care) 三類。腫瘤位置(簡記為 S) 分為咽喉(faucal arch)、扁桃腺小窩(tonsillar fossa)與咽部的

舌頭(pharyngeal tongue)三類。量測之主要腫瘤直徑大小(簡記為T)分為較小的主要腫瘤(primary tumor)-腫瘤直徑小於等於兩公分、二種中等程度，分別為腫瘤直徑在二到四公分之間與腫瘤直徑大於四公分, 和大且具侵略性並連接組織的腫瘤(massive invasive tumor with extension to adjoining tissue)四類。最後淋巴節轉移狀況(簡記為N)分為無臨床上之淋巴節轉移(no clinical evidence of lymph node metastases)、單一且直徑小於三公分的淋巴節、單一且直徑大於三公分的淋巴節與複合(multiple)淋巴節點四類。實驗的詳細介紹可參考 Karrison (1987), Zucker (1998) 和 Kim (2001) 文章。



圖一：單一療法組與混合療法組之存活函數。



圖二：單一療法組與混合療法組之對數累積風險圖。

由圖一、二得知接受單一療法與混合療法兩組存活函數在前期有相交的狀況發生，顯示處理風險並非等比例。因此，我們考慮配適分層 Cox 模式，或是分別在兩處理組之下分別配適 Cox 模式。

首先根據概似比檢定(likelihood ratio test)選擇共變數。由於腫瘤等級(G)、病人一般狀況(G)、腫瘤位置(S)、最初量測之腫瘤直徑大小(T) 和淋巴結轉移狀況(N)五個變數皆為離散資料，所以在配適模型時使用虛擬變數(dummy variables)。腫瘤等級(G)的虛擬變數為 G1、G2；病人一般狀況(C)的虛擬變數為 C1、C2；腫瘤位置(S)的虛擬變數為 S1、S2；最初量測之腫瘤直徑大小(T)有三個虛擬變數:T1、T2、T3；淋巴結轉移狀況(N)亦有三個虛擬變數:N1、N2、N3。

各變數之虛擬變數代表意義如下所述：

$$\begin{cases} G_1 = 1, \text{若 } G = 1 \\ G_2 = 1, \text{若 } G = 2 \end{cases},$$

$$\begin{cases} C_1 = 1, \text{若 } C = 1 \\ C_2 = 1, \text{若 } C = 2 \end{cases},$$

$$\begin{cases} S_1 = 1, \text{若 } S = 1 \\ S_2 = 1, \text{若 } S = 2 \end{cases},$$

$$\begin{cases} T_1 = 1, \text{若 } T = 1 \\ T_2 = 1, \text{若 } T = 2 \\ T_3 = 1, \text{若 } T = 3 \end{cases},$$

$$\begin{cases} N_1 = 1, \text{若 } N = 1 \\ N_2 = 1, \text{若 } N = 2 \\ N_3 = 1, \text{若 } N = 3 \end{cases}。$$

表一：接受單一療法病人風險模式配適之 $-2\log\hat{L}$ 值

模型中的共變數	$-2\log\hat{L}$ 值	P value
null	579.2272	
Sex	576.1290	0.078
G1, G2	576.3767	0.24
C1, C2	565.5773	0.0011
S1, S2	577.9075	0.5169
T1, T2, T3	571.3340	0.0483
N1, N2, N3	573.7651	0.1409
C1, C2, T1, T2, T3	557.7882	

由表一得知，對於單一療法而言，加入一般條件(C1, C2) 的模型顯著 (p 值=0.00108)，加入最初量測之腫瘤直徑大小(T1, T2, T3)的模型亦顯著(p 值=0.0483)，並含一般條件(C1, C2)的模型加入最初量測之腫瘤直徑大小(T1, T2, T3)的邊際貢獻亦顯著(概似比為 7.7891，對應之 p 值=0.051)。所以考慮模式加入共變數 C1, C2 及 T1, T2, T3。

表二：接受混合療法病人風險模式配適之  $-2\log \hat{L}$  值

模型中的共變數	$-2\log \hat{L}$ 值	P value
null	550.1659	
Sex	549.9891	0.6741
G1, G2	547.0781	0.2135
C1, C2	530.2141	$4.65 \times 10^{-5}$
S1, S2	546.6871	0.1756
T1, T2, T3	545.6781	0.2134
N1, N2, N3	544.5929	0.1343

由表二得知，對於混合療法而言，模型加入一般條件(C1, C2)後顯著 (p 值=0.0000465)，加入其他變數皆不顯著。所以，對於混合療法組我們只用一個共變數，即一般條件(C1, C2)來配適此一資料。

表三：分層 Cox 模式配適之  $-2\log\hat{L}$  值

模型中的共變數	$-2\log\hat{L}$ 值	P value
null	1129.393	
Sex	1128.408	0.321
G1, G2	1123.506	0.053
C1, C2	1096.453	$7.03 \times 10^{-8}$
S1, S2	1128.530	0.65
T1, T2, T3	1117.854	0.0091
N1, N2, N3	1119.207	0.017
C1, C2, T1, T2, T3	1086.346	
C1, C2, N1, N2, N3	1089.685	
T1, T2, T3, N1, N2, N3	1111.267	
C1, C2, T1, T2, T3, N1, N2, N3	1082.094	

由表三得知，對於分層 Cox 模式而言，模型加入一般條件(C1, C2)顯著(p 值= $7.03 \times 10^{-8}$ )，模型加入最初量測之腫瘤直徑大小(T1, T2, T3)顯著(p 值=0.00914)，模型加入淋巴結轉移狀況(N1, N2, N3)也顯著(p 值=0.017)，並且含一般條件(C1, C2)的模型加入最初量測之腫瘤直徑大小(T1, T2, T3)的邊際貢獻亦顯著(概似比為 10.1071，p 值=0.018)。但是模式中含一般條件(C1, C2)後，淋巴結轉移狀況(N1, N2, N3)的邊際貢獻並不顯著(概似比為 6.768，p 值=0.08)。綜上所述，在配適分層 Cox 模式我們考慮一般條件(C1, C2)與最初量測之腫瘤直徑大小(T1, T2, T3)兩個共變數。

根據表一、表二、表三之分析與結果，一般條件(C1, C2)及最初量測之腫瘤直徑大小(T1, T2, T3)在單一療法組與分層 Cox 模型皆顯著，但是在混合療法時，最初量測之腫瘤直徑大小(T1, T2, T3)並不顯著。為了能比較在具共變數之下，分組配適之方法與分層配適之方法，我們先考慮只用一個共變數，即一般條件(C1, C2)配適此一資料。

根據分組配適之模式， $\lambda_{i1}(t|\mathbf{C}) = \lambda_{i0}(t) \exp(\boldsymbol{\beta}_i \mathbf{C})$ ， $i=1,2$ ，檢驗此一資料是否能合理的以分層 Cox 模式加以配適，亦即檢定  $H_0: \beta_1 = \beta_2$ 。因為  $\hat{\beta}_1 = 1.07$ ， $\hat{Var}(\hat{\beta}_1) = 0.069$ ， $\hat{\beta}_2 = 0.93$ ， $\hat{Var}(\hat{\beta}_2) = 0.0368$ 。得知

$$\frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\hat{Var}(\hat{\beta}_1) + \hat{Var}(\hat{\beta}_2)}} = 0.4299$$

其 p 值近似 0.672。因此，保留  $H_0: \beta_1 = \beta_2$ ，亦即可考慮配適分層 Cox 模式。

在配適分層 Cox 模式情形下，此處先檢驗是否需要加入共變數 C 與處理組 Trt 的交互作用，目的是要進一步判斷此一資料是適合配適分層 Cox 模式，或是應該分組配適 Cox 模式。若交互作用項顯著，則不適合配適分層 Cox 模式。表四顯示由於在加入 C×Trt 此項時，其 p 值為 0.78，故此一交互作用項不顯著，可知模型沒有加入一般條件(C)與處理組(Trt)交互作用項的必要。

表四：就兩組資料配適具處理、一般條件與一般條件與處理組的交互作用後之參數，及其參數之變異數，與對應之 p 值。

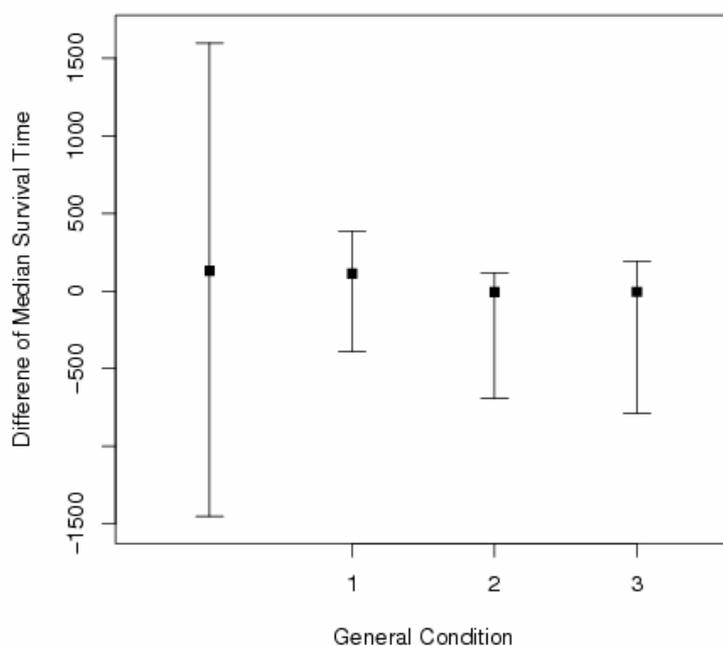
共變數	係數	標準差	P 值
C	1.1379	0.547	0.038
Trt	0.1647	0.451	0.720
C×Trt	-0.0874	0.316	0.780

根據以上的結果，我們發現此模型適合配適分層 Cox 模式，因為處理組(Trt)與共變數(C)沒有明顯的交互作用。所以再一次說明分層 Cox 模式的適用性。

根據共變數為一般條件(C)下，資料經由配適分層 Cox 模式可得知調整共變數兩組存活中位數差異之信賴區間，結果如表五及圖三。

表五：根據共變數為一般條件(C)的分層 Cox 模式建立存活中位數差異之信賴區間。

一般條件 (C)	存活中位數		存活中位數之差異 (單一療法-混合療法)	95%之信賴區間
	單一療法	混合療法		
1	637	526	111	(-391, 763)
2	266	273	-7	(-688, 324)
3	172	177	-5	(-788, 316)



圖三：根據共變數為一般條件(C)的分層 Cox 模式建立存活中位數差異之信賴區間圖。

由表五或圖三皆可得知，採用單一療法或是混合療法的病患，其存活中位數差異並無顯著差異。

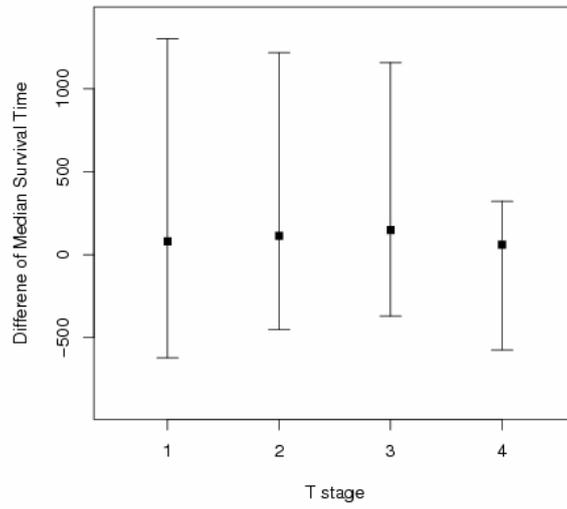
根據共變數為一般條件(C)與最初量測之腫瘤直徑大小(T)下，資料經配適分層 Cox 模式得知調整共變數兩組存活中位數差異之信賴區間，結果如表六及圖四-圖六，以及表七(調整共變數兩組存活中位數比例之信賴區間)。

表六：根據共變數為一般條件(C)與最初量測之腫瘤直徑大小(T)的分層 Cox 模式建立存活中位數差異之信賴區間。

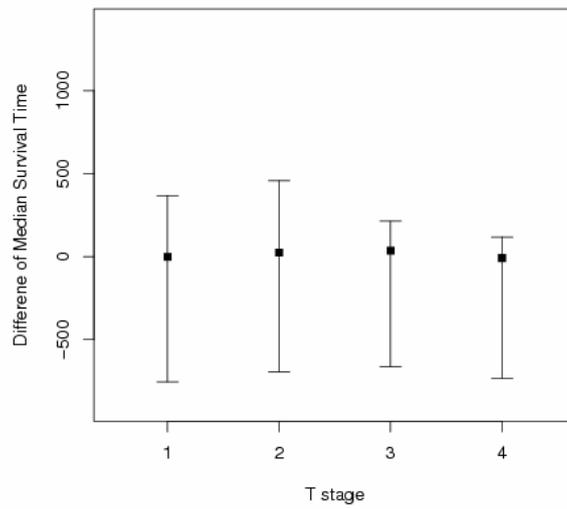
一般條件(C)	腫瘤直徑大小(T)	存活中位數		存活中位數之差異 (單一療法-混合療法)	95%信賴區間
		單一療法	混合療法		
1	1	541	461	80	(-620, 1562)
1	2	775	661	114	(-451, 1562)
1	3	757	608	149	(-370, 1553)
1	4	465	404	61	(-573, 625)
2	1	255	256	-1	(-754, 563)
2	2	334	310	24	(-694, 714)
2	3	327	291	36	(-662, 453)
2	4	222	230	-8	(-732, 284)
3	1	159	173	-14	(-804, 295)
3	2	191	208	-17	(-788, 362)
3	3	191	192	-1	(-772, 324)
3	4	144	154	-10	(-804, 216)

表七：根據共變數為一般條件(C)與最初量測之腫瘤直徑大小(T)的分層 Cox 模式建立存活中位數比例之信賴區間。

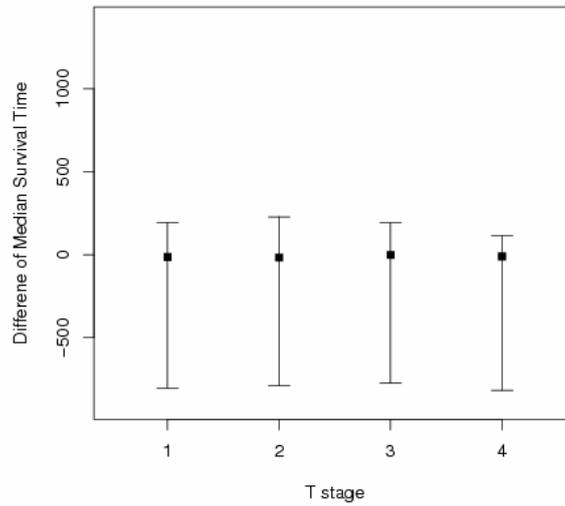
一般條件(C)	腫瘤直徑大小(T)	存活中位數		存活中位數之比例(單一療法/混合療法)	95%信賴區間
		單一療法	混合療法		
1	1	541	461	1.173536	(0.388, 143.09)
1	2	775	661	1.172466	(0.488, 143.09)
1	3	757	608	1.245066	(0.588, 143.09)
1	4	465	404	1.15099	(0.388, 4.488)
2	1	255	256	0.996094	(0.288, 2.988)
2	2	334	310	1.077419	(0.288, 2.888)
2	3	327	291	1.123711	(0.288, 1.888)
2	4	222	230	0.9652174	(0.688, 1.788)
3	1	159	173	0.9190751	(0.488, 2.688)
3	2	191	208	0.9182692	(0.488, 2.488)
3	3	191	192	0.9947917	(0.488, 2.288)
3	4	144	154	0.935065	(0.488, 2.288)



圖四：根據共變數為一般條件  $C=1$  時的分層 Cox 模式建立存活中位數差異之信賴區間圖。



圖五：根據共變數為一般條件  $C=2$  時的分層 Cox 模式建立存活中位數差異之信賴區間圖。



**圖六：**根據共變數為一般條件 C=3 時的分層 Cox 模式建立存活中位數差異之信賴區間圖。

由表六、表七或圖四-圖六皆可得知，採用單一療法或是混合療法的病患，其存活中位數差異並無顯著差異，但是，此一信賴區間之長度隨一般條件(C)的等級增加而縮短，這也許是病人一般狀況嚴重者，具較佳資訊可供區間估計單一療法與混合療法的存活中位數之差異。

## 第五章 模擬研究

### 5.1 模擬方法

本文針對存活資料服從不同尺度參數(scale parameter)與形狀參數(shape parameter)的韋伯分布(Weibull distribution)，其存活函數為

$$S_{0i}(t) = \exp\{-\alpha_i(t)^{\gamma_i}\} \quad i = 1, 2$$

其中， $\alpha_i > 0, \gamma_i > 0$ ，分別為韋伯分布的尺度參數與形狀參數。研究分別應用分層 Cox 模式與分組 Cox 模式，建構調整共變數兩組存活中位數差異信賴區間之涵蓋機率 (coverage probability) 以及平均區間長度(average confidence interval length)。

使用 International Mathematics and Statistical Library (簡稱 IMSL) 的附屬程式生成隨機變數。考慮四種結構的韋伯分布，分別為  $(\alpha_1, \alpha_2, \gamma_1, \gamma_2) = (0.276, 0.064, 0.75, 1.5)$ ， $(0.117, 0.117, 1.25, 1.25)$ ， $(0.086, 0.15, 1.25, 1.25)$ ， $(0.101, 0.095, 1.25, 1.75)$  四種狀況。

在不同處理組之下，考慮樣本數  $n_1 = n_2 = 50, 100, 200$ ，設限分布(censoring distribution)由  $U(6, 8)$  分布生成，考慮單一共變數  $X_1$  與  $X_2$ ，分別由  $U(0, 1)$  分布生成，且固定各處理組之共變數為  $X_1$  與  $X_2$ 。考慮各組的真實迴歸參數(true regression parameter)為

$(\beta_1, \beta_2) = (0.4, 0.4)$ ， $(0.8, 0.4)$  兩種，其中  $\beta_1 = \beta_2$  為適合配適分層 COX 模式之結構，其中  $\beta_1 \neq \beta_2$  為適合配適分組 COX 模式之結構。此時，真實存活中位數差異為

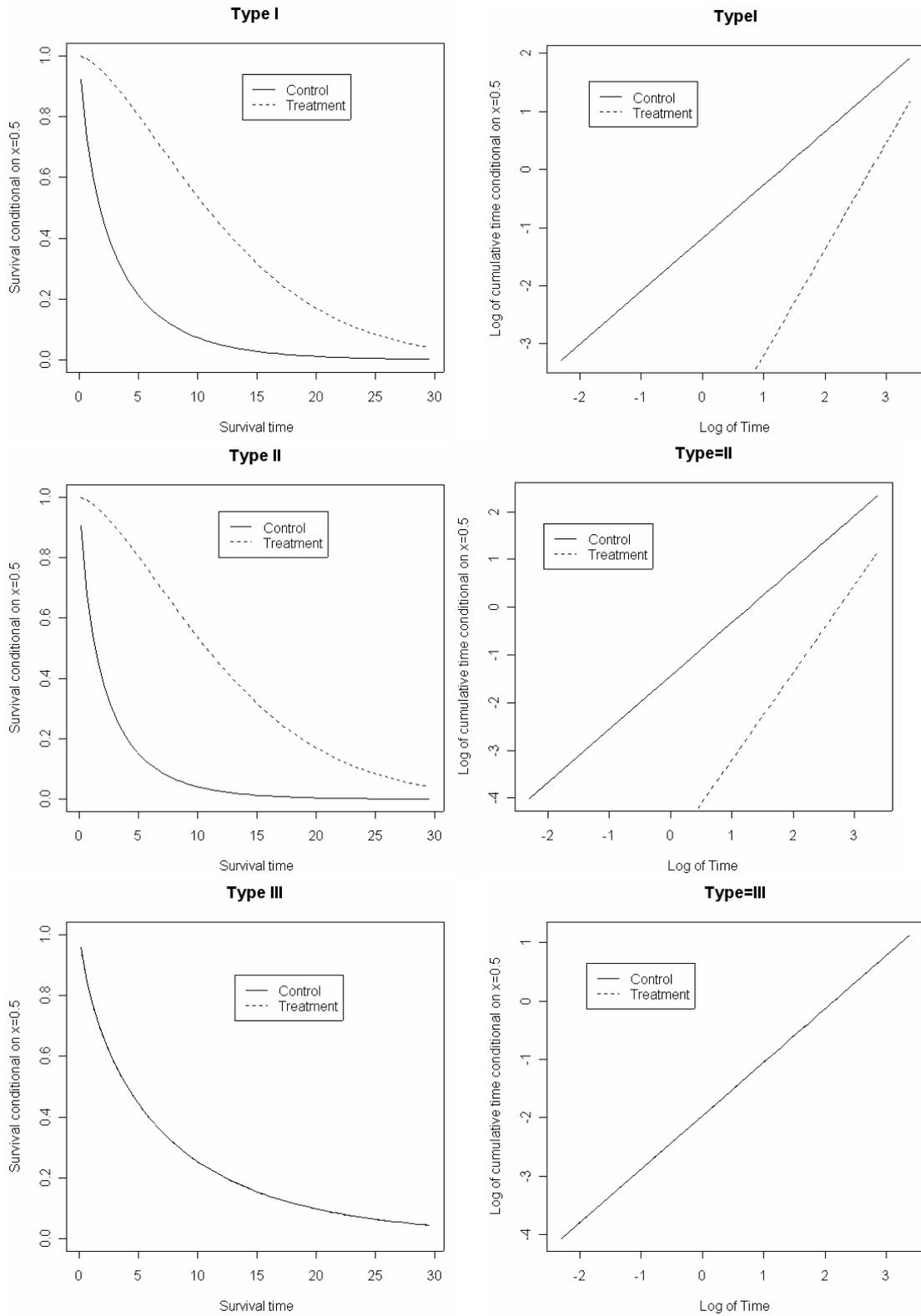
$$\left\{ \frac{-\ln(0.5)}{\alpha_2 \psi_2} \right\}^{\frac{1}{\gamma_2}} - \left\{ \frac{-\ln(0.5)}{\alpha_1 \psi_1} \right\}^{\frac{1}{\gamma_1}}, \quad \psi_i = \exp(\beta_i' x_i)。$$

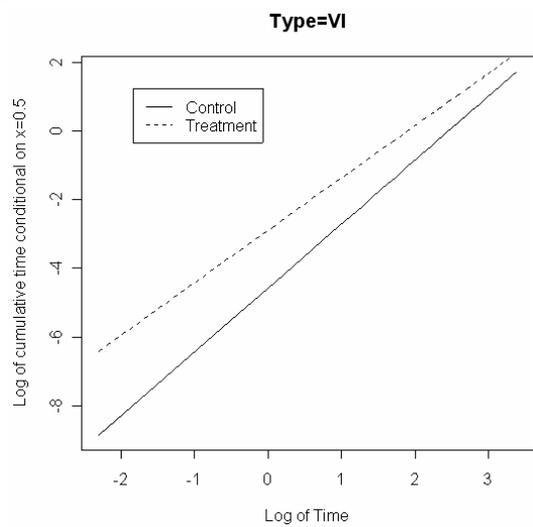
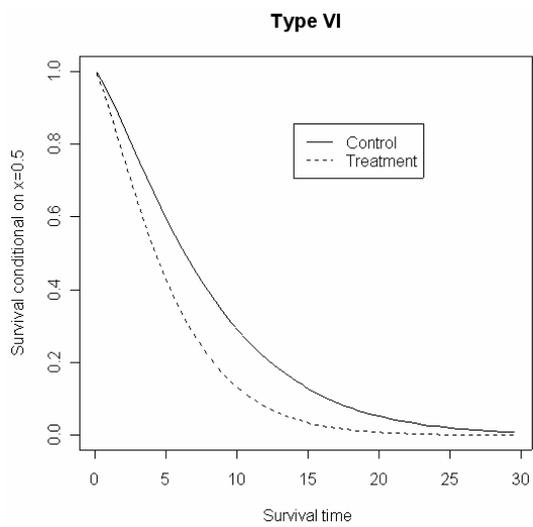
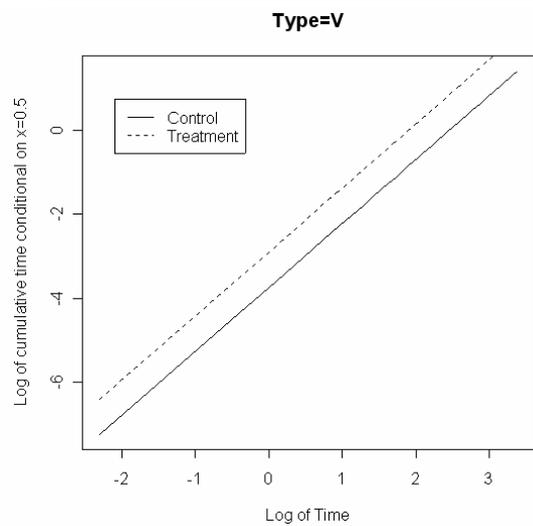
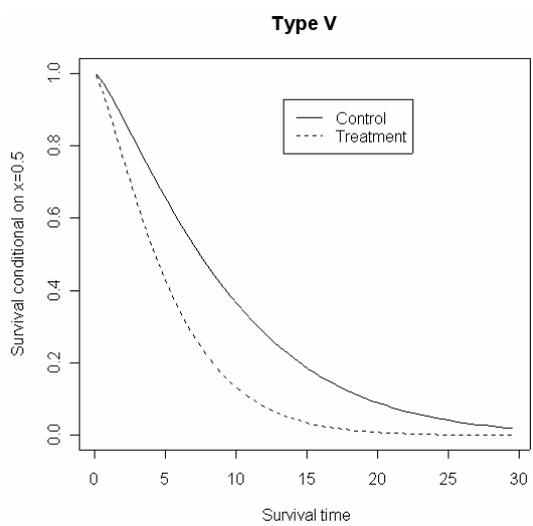
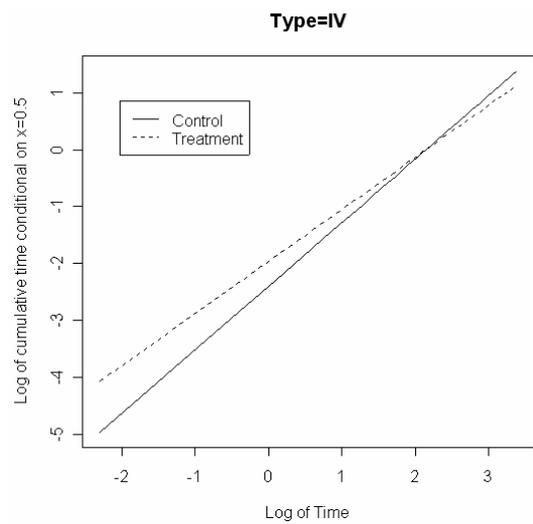
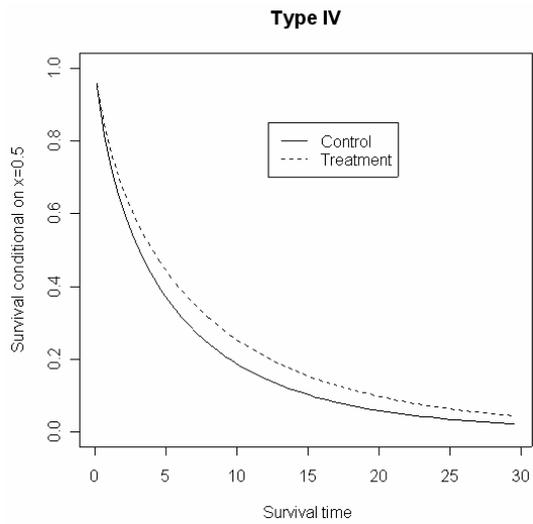
每一次的模擬次數為 10,00 次，根據模擬的雙樣本存活資料分別求出調整共變數兩組存活中位數差異的 95% 信賴區間，然後計數 10,00 個區間涵蓋真正存活中位數差異的比例，藉以估計 95% 信賴區間的涵蓋機率及其平均區間長度。

表八：模擬之結構表

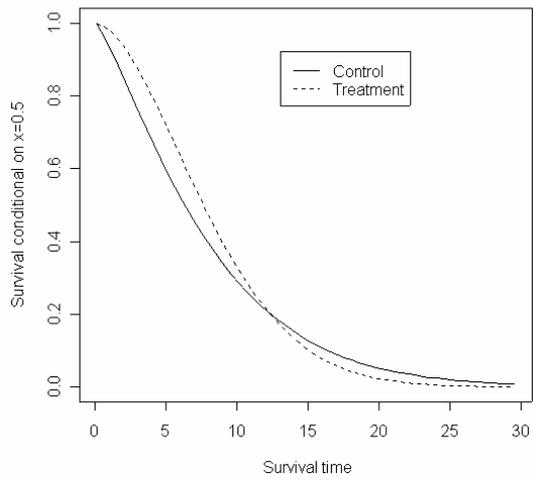
結構	$\alpha_1$	$\alpha_2$	$\gamma_1$	$\gamma_2$	$\beta_1$	$\beta_2$
I	0.276	0.064	0.75	1.5	0.4	0.4
II	0.276	0.064	0.75	1.5	0.8	0.4
III	0.117	0.117	1.25	1.25	0.4	0.4
IV	0.117	0.117	1.25	1.25	0.8	0.4
V	0.086	0.15	1.25	1.25	0.4	0.4
VI	0.086	0.15	1.25	1.25	0.8	0.4
VII	0.101	0.095	1.25	1.75	0.4	0.4
VIII	0.101	0.095	1.25	1.75	0.8	0.4

圖七:各結構之存活函數與對數累積風險圖

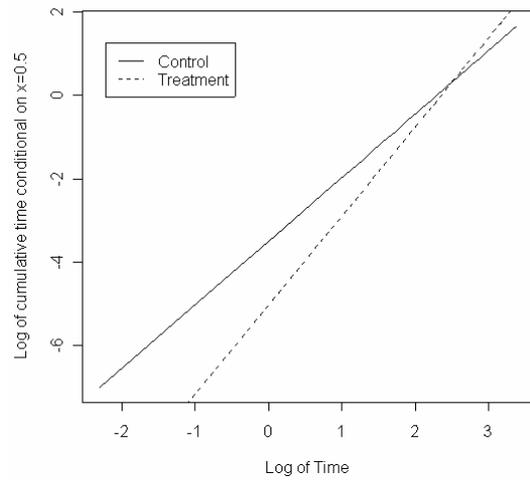




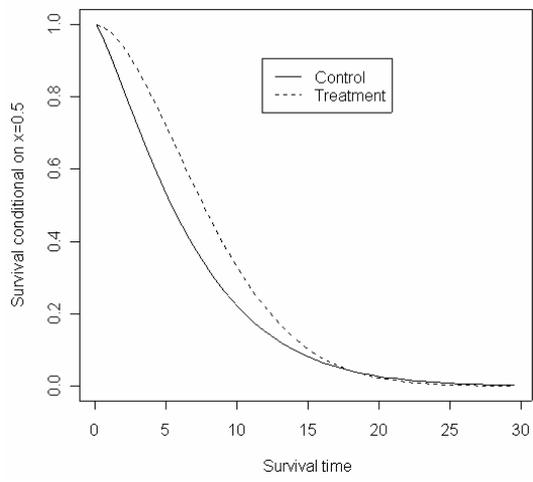
**Type VII**



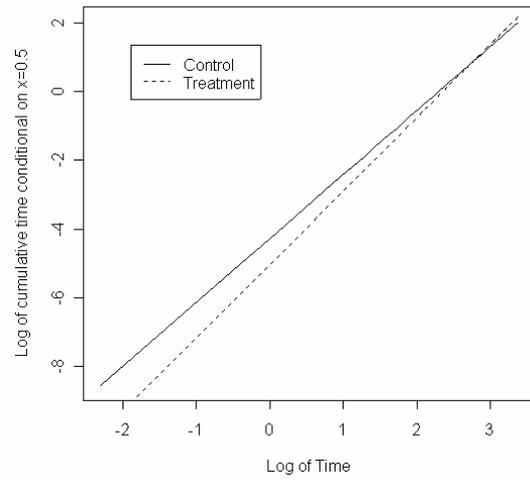
**Type=VII**



**Type VIII**



**Type=VIII**



## 5.2 模擬結果

根據表九至表十一可以清楚發現由推廣 Su and Wei(1993)方法，將存活資料配適分層 Cox 模式與配適分組 Cox 模式建立的信賴區間的涵蓋機率皆呈現高於信心水準保守的情形。

在  $\beta_1 = \beta_2$  或  $\beta_1 \neq \beta_2$  之下，兩種方法所建立的信賴區間的涵蓋機率沒有明顯差異，因為兩種方法皆相當保守。並且在不同樣本數之下，兩種方法所建立的信賴區間的涵蓋機率有略為下降的趨勢，但皆保持在高於信心水準的狀態。

表九: 固定  $x = 0.5$ ， $n_1 = n_2 = 50$  時，各種結構之 95% 信賴區間的涵蓋機率及其平均區間長度。

N	$\alpha_1$	$\alpha_2$	$\gamma_1$	$\gamma_2$	$\beta_1$	$\beta_2$	Coverage rate		Average length	
							separate	stratified	separate	stratified
100	0.276	0.064	0.75	1.5	0.4	0.4	0.986	0.987	6.416	6.435
100	0.276	0.064	0.75	1.5	0.8	0.4	0.981	0.983	4.584	4.663
100	0.117	0.117	0.75	1.25	0.4	0.4	0.986	0.989	4.708	4.730
100	0.117	0.117	1.25	1.25	0.8	0.4	0.983	0.984	3.917	3.979
100	0.086	0.150	1.25	1.25	0.4	0.4	0.989	0.991	5.813	5.822
100	0.086	0.150	1.25	1.25	0.8	0.4	0.987	0.989	4.807	4.906
100	0.101	0.095	1.25	1.75	0.4	0.4	0.986	0.988	4.458	4.496
100	0.101	0.095	1.25	1.75	0.8	0.4	0.988	0.987	3.429	3.494

表十:固定  $x = 0.5$  ,  $n_1 = n_2 = 100$  時, 各種結構之95%信賴區間的涵蓋機率及其平均區間長度。

N	$\alpha_1$	$\alpha_2$	$\gamma_1$	$\gamma_2$	$\beta_1$	$\beta_2$	Coverage rate		Average length	
							separate	stratified	separate	stratified
200	0.276	0.064	0.75	1.5	0.4	0.4	0.990	0.990	3.884	3.899
200	0.276	0.064	0.75	1.5	0.8	0.4	0.986	0.986	3.006	3.028
200	0.117	0.117	1.25	1.25	0.4	0.4	0.985	0.986	2.923	2.929
200	0.117	0.117	1.25	1.25	0.8	0.4	0.983	0.982	2.602	2.619
200	0.086	0.150	1.25	1.25	0.4	0.4	0.988	0.988	4.579	4.613
200	0.086	0.150	1.25	1.25	0.8	0.4	0.987	0.985	2.998	3.025
200	0.101	0.095	1.25	1.75	0.4	0.4	0.980	0.980	2.843	2.848
200	0.101	0.095	1.25	1.75	0.8	0.4	0.982	0.982	2.140	2.158

表十一:固定  $x = 0.5$  ,  $n_1 = n_2 = 200$  時, 各種結構之95%信賴區間的涵蓋機率及其平均區間長度。

N	$\alpha_1$	$\alpha_2$	$\gamma_1$	$\gamma_2$	$\beta_1$	$\beta_2$	Coverage rate		Average length	
							separate	stratified	separate	stratified
400	0.276	0.064	0.75	1.5	0.4	0.4	0.985	0.986	2.403	2.406
400	0.276	0.064	0.75	1.5	0.8	0.4	0.987	0.982	1.974	1.970
400	0.117	0.117	0.75	1.25	0.4	0.4	0.984	0.986	1.9863	1.9891
400	0.117	0.117	1.25	1.25	0.8	0.4	0.976	0.977	1.776	1.773
400	0.086	0.150	1.25	1.25	0.4	0.4	0.989	0.991	2.628	2.635
400	0.086	0.150	1.25	1.25	0.8	0.4	0.979	0.983	1.907	1.901
400	0.101	0.095	1.25	1.75	0.4	0.4	0.981	0.983	1.752	1.753
400	0.101	0.095	1.25	1.75	0.8	0.4	0.972	0.980	1.425	1.425

## 第六章 結論與討論

本文藉由推廣 Su and Wei(1993)的雙樣本存活中位數差異之信賴區間，探討如何針對具共變數的病人與接受兩種不同處理後，進行存活中位數差異之區間估計。就處理與共變數有無交互作用，分別以分別 Cox 模式或分層 Cox 模式描述病人的風險函數，藉此了解處理間的風險比例。相較於 Kim(2001)根據兩個處理基準組存活中位數差異及兩個處理組平均存活函數中位數差異分別建立的信賴區間而言，本文所提方法，反應了共變數對於兩組存活函數中位數差異之區間估計的影響。值得一提的，本文方法不僅可以建立存活函數中位數差異之信賴區間，亦可用於建立存活函數中位數比例之信賴區間。

實務上有許多問題皆可推廣雙樣本方法獲得解決，例如：多樣本的成對比較(pairwise)。本文方法經由適當的多重調整(multiplicity adjustment)，亦可獲得成對多重比較方法，探求成對組存活中位數的差異或比例之聯合信賴域(simultaneous confidence region)。此外，存活資料中若右設限情形嚴重時，本文所提方法亦可自然的延展為存活百分位差異的區間估計。最後，針對具共變數的病人與其基準組病人相對風險與時間相關時，Cox 模式不再適用，因此，如何選擇或應用更合理的模式，例如：加速失敗時間

模式 (accelerated failure time model) 或比例勝算模式 (proportional odds model) 等，推廣本文所提方法，也是未來可能的研究方向。

## 參考文獻

1. Anderson,P.K.,Borgan,O.,Gill,R.D.,and Keiding,N.,(1993). *Statistical models based on counting processes*. Springer-Verlag,New York.
2. Anderson,P.K. and Gill,R.D.,(1982). Cox's regression model for counting processes: A large sample study. *The Annals of statistics* **10**,1100-1120.
3. Basawa,I.V. and Koul,H.L.(1988). Large-sample statistics based on quadeatic dispersion. *International Statistical Review* **56**,199-219.
4. Breslow, N. (1975).Analysis of survival data under the proportional hazards model. *International Statistical Review* **43**,45-58.
5. Brookmeyer, R. and Crowley, J. (1982). A confidence interval for the median survival time. *Biometrics* **38**,29-41.
6. Collett, D.(2003) *Modelling survival data in medical research*. Chapman and Hall, New York.
7. Doss,H. and Gill,R.D.,(1992). An Elementary Approach to Weak Convergence for Quantile Processes,With Applications to censored Survival Data. *Journal of the American Statistical Association* **87**, 869-877.
8. Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association* **76**, 312-319.
9. Emerson, J.D. (1982). Nonparametric confidence intervals for the median in the presence of right censoring. *Biometrics* **38**,17-27.
10. Hardel,W.,(1991). *Smoothing Techniques with Implementation in*. Springer-Verlag, New York.
11. Kalbfleisch, J.D. and Prentice, R.L.,(1980) *The Statistical analysis of failure time data*. Wiley, New York.

12. Kaplan, E.L and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53**, 457-481.
13. Kim, J. (2001). Confidence Intervals for the Difference of Median Survival Times Using the Stratified Cox Proportional Hazards Model. *Biometrical Journal* **43**,781-790.
14. Klein, J.P. and Moeschberger, M.L. (1997). *Survival analysis: Techniques for censored and truncated data*. Springer-Verlag, New York.
15. Rebolledo, R. (1980a). Central limit theorems for local martingales. *Z.Wahrsh.verw.Geb.* **51**,269-286.
16. Reid, N. (1981). Estimating the median survival time. *Biometrika* **68**,601-608.
17. Simon, R. and Lee, Y.J. (1982). Nonparametric confidence limits for survival probabilities and median survival time. *Cancer Treatment Reports* **66**,37-42.
18. Slud, E. V., Byar, D.P., and Green, S.B. (1984). A comparison of reflected verses test-based confidence interval for the median survival time, based on censored data. *Biometrics* **40**,587-600.
19. Su, J.Q. and Wei, L.J. (1993). Nonparametric Estimation for the Difference or Ratio of Median Failure Times. *Biometrics* **49**,603-607.
20. Tsiatis, A.A. (1981).A large sample study of Cox's regression model. *The Annals of Statistics* **9**,93-108.
21. Wang, J.L. and Hettmansperger, T.P.(1990). Two-sample inference for median survival times based on one-sample procedures for censored survival data. *Journal of the American Statistical Association* **85**,529-536.
22. Wells, M.T. (1994) Nonparametric kernel estimation in counting processes with explanatory variables. *Biometrika* **81**,795-801.

23. Zhang, M.J. and Klein, J.P. (2001). Confidence Bands for the Difference of Two Survival Curves Under Proportional Hazards model. *Lifetime Data Analysis* 7, 243-254.
24. Zucker, D.M., (1998) Restricted Mean Life With Covariates: Modification and Extension of a Useful Survival Analysis Method. *Journal of the American Statistical Association* 93, 702-709.