

東海大學統計研究所

碩士論文

生活品質調整後存活時間差異之

機動檢定

研究生：陳柏尚

指導教授：張玉媚博士

中華民國一百年七月

東海大學碩士班研究生

論文口試委員審定書

統計學系碩士班陳柏尚君所提之論文

生活品質調整後存活時間差異之機動檢定

經本委員會審議，認為符合碩士資格標準。

論文口試委員召集人 陳春樹 (簽章)

委員 張玉媚

俞一凡

中華民國 100 年 7 月 14 日

致謝

首先誠摯的感謝指導教授張玉媚博士，由論文题目的發想，到之後的實地執行與分析撰寫，每次的 meeting 老師都不厭其煩的教導與提點，也讓我有持續的動力能夠完成這份論文，更對於統計學領域有進一步了解。而在與老師的互動過程中，儘管我的想法有時表達的不夠完整，老師仍清楚地指出我該走的方向，這份論文的完成要感謝張老師的付出與包容，讓學生能夠在研究階段如此順利。

此外，也要感謝兩位口試委員陳春樹博士與俞一唐博士對這份論文的建言。感謝兩位老師在百忙之中審核，同時以豐富的學術涵養與嚴謹的研究態度，給予這份論文精闢的見解與寶貴的建議，使得這份論文能夠更為完整。

在研究所修課期間，感謝統計系上曾經教導過的老師，也感謝系辦裡的學姐們在研究生事務上的幫助，使我能專心於課業之上，而免於雜務的紛擾。

兩年的研究所生涯，感謝共同努力的研究所同學們，從原文文獻的翻譯導讀到期末讓人焦頭爛額的報告，以及生死關頭的考試，因為有你們在一旁共同努力，讓艱辛的研究所學習過程變得較為有趣，因為有你們的笑聲，研究室才不會顯得寂寥。

最後感謝我的家人，讓我可以無後顧之憂的在外地求學，謝謝你們。

陳柏尚 謹誌

東海大學統計研究所

中華民國一百年七月

摘要

在臨床實驗中，評估藥效往往都是以延長壽命作為評量的指標，近年來隨著物質生活的提升，醫師與病患開始注重治療的生活品質，並以生活品質做為衡量藥效的指標之一。本文將針對生活品質資料，藉由讓資料選取最適合的生活品質權重，建構生活品質調整後存活時間差異之機動檢定，並利用模擬方式探討不同樣本數、存活函數與相關性下，觀察型 I 誤差與檢定力的表現，以及探討本研究所提之機動檢定中，所選取到最具顯著治療差異的生活品質權重分佈和其代表的治療狀態，並以實際資料加以說明其應用。

關鍵字：排序法、生活品質資料、生活品質調整後存活時間、限制平均壽命、機

動檢定

Abstract

In clinical trials, the evaluation of treatment effects is based on the length of survival time. In recent years, physicians and patients pay much attention to the quality of life during treatments with improved material life, and quality of life becomes one criterion for assessing the effect of drugs. In this thesis, we proposed a versatile test for the difference between two quality adjusted survival times, and select the most suitable weight parameter of quality by data by maximizing the test statistic. Some numerical studies are performed under various sample sizes and survival functions with different correlations, for evaluating the performance of the type I error and power. We also discuss the distribution of the weights selected in the versatile test. Finally, the proposed testing procedure is illustrated with two real data sets.

Key words : permutation, quality of life data, quality adjusted survival time, restricted mean life, versatile test

目錄

第一章 緒論	6
第二章 文獻回顧	8
第一節 生活品質調整後存活時間	8
第二節 生活品質調整後存活時間之共變異數矩陣	10
第三章 統計方法	13
第四章 模擬研究	15
第一節 模擬方法	15
第二節 模擬結果	16
第五章 資料分析	19
資料一:賀爾蒙治療對乳癌的影響	19
資料二:長短期化學治療對乳癌的影響	20
第六章 結論與未來研究方向	22
參考文獻	23
附錄	25

圖目錄

圖 1. 生活品質各階段存活曲線圖	8
圖 2. $(\mu_{12}, \mu_{22}) = (0.5, 0.5)$ 之下，各相關係數設定之生活品質資料存活函數圖 形	25
圖 3. $(\mu_{12}, \mu_{22}) = (0.5, 0.2)$ 之下，各相關係數設定之生活品質資料存活函數圖 形	26
圖 4. 德國乳癌研究資料之 Kaplan-Meier 存活函數估計圖	27
圖 5. 德國乳癌研究資料在生活品質調整後存活時間差異的信賴域圖形	27
圖 6. 國際乳癌研究組織第五期資料之 Kaplan-Meier 存活函數估計圖	28
圖 7. 國際乳癌研究組織第五期資料在生活品質調整後存活時間差異的信賴域圖 形	29
圖 8. 國際乳癌研究組織第五期資料在生活品質調整後存活時間差異的信賴域 上、下界等高線圖	30

表目錄

表 1. $n_{ig}=50, i=1,2,3, g=1,2$ 及 $\alpha = 0.05$ 之下，各個設置下生活品質資料差異的型 I 誤差率估計	31
表 2. $n_{ig}=100, i=1,2,3, g=1,2$ 與固定生活品質權重 $n_{ig}=100$ 及 $n_{ig}=200, \alpha = 0.05$ 之下，各個設置下生活品質資料差異的型 I 誤差率估計	32
表 3. $n_{ig}=100, i=1,2,3, g=1,2, \alpha = 0.05$ 之下，各個設置下生活品質資料差異的檢定力估計	33
表 4. $n_{ig}=50, i=1,2,3, g=1,2$ ，機動檢定法的型 I 誤差率估計值下，各個設置下 $(\hat{w}_{TOX}, \hat{w}_{REL})$ 個數分佈	34
表 5. $n_{ig}=100, i=1,2,3, g=1,2$ ，機動檢定法的型 I 誤差率估計值下，各個設置下 $(\hat{w}_{TOX}, \hat{w}_{REL})$ 個數分佈	35
表 6. $n_{ig}=100, i=1,2,3, g=1,2$ ，機動檢定法的檢定力估計值下，各個設置下 $(\hat{w}_{TOX}, \hat{w}_{REL})$ 個數分佈	36
表 7. 德國乳癌資料在各限制平均壽命的機動檢定與固定生活品質權重檢定統計量，其對應的單尾 p 值列於括號中	37
表 8. 國際乳癌研究組織第五期資料在各限制平均壽命的機動檢定與固定生活品質權重檢定統計量，其對應的單尾 p 值列於括號中	38
表 9. 國際乳癌研究組織第五期資料在生活品質調整後存活時間差異的信賴域中顯著差異 $(\hat{w}_{TOX}, \hat{w}_{REL})$ 範圍	39

第一章 緒論

在臨床實驗中，常常就以實驗組與對照組來比較不同治療方法的療效，且通常是以病人治療時間或治療後的存活時間來評估治療方法的好壞。近年來人們的平均壽命逐漸提高，也開始追求生活品質，隨著物質生活及醫療水準的進步，病患與醫生越來越重視治療的生活品質，而生活品質也成為評估治療方法的另一個考量指標，然而，同時將存活時間與生活品質列入考量來評估治療方法是有困難的，因為這兩種指標往往都是互相抵觸的，例如癌症病患，為了能夠讓存活時間延長而接受毒性很強的化學治療，因其副作用而感到不適，導致生活品質降低，病患須在延長存活時間與維持生活品質作個取捨或尋求一平衡點。

生活品質是一種整合性的概念，漸漸的受到重視，在近幾年的臨床研究，已經有相當多的生活品質文獻資料，Gelber et al. (1989)提出無疾病或毒性反應時間 (Time Without Symptoms of disease or toxicity of Treatment ,TWiST)的觀念，將接受實驗病患的存活時間分段並給予不同的權重討論其生活品質及存活時間。Glasziou et al. (1990)將生活品質與存活時間合併成為新的測量指標，生活品質調整後存活年 (Quality Adjusted Life Years, QALY)，並根據QALY提出了生活品質調整後存活分析 (Quality adjusted survival analysis)來找出不同治療方法之間無差異性的權重範圍。Gelber et al.(1996)將生活品質應用在直腸癌的化療與化療加上放射治療兩種治療方法的評估。Murray和Cole (2000)提出Q-TWiST的漸進變方的公式與樣本數的計算，並給予固定權重下，觀察其檢定力的表現。Sloan et al.

(2002)提出了生活品質平均存活圖 (Q-TWiST mean survival plot)，觀察給予不同的權重，如何影響其治療方法好壞的變化。

過去關於生活品質調整後存活時間(Quality Adjusted Survival Time)這方面的文獻，大部分都是直接將生活品質與存活時間相乘，結合得到生活品質調整存活時間分析的結果，而其結果是會小於未經過調整的整體存活平均時間。過去對於生活品質調整存活時間的檢定都是在固定權重下進行，無法觀察到所有權重的情形，本文將以Murray和Cole (2000)所提的漸進變方公式計算其Q-TWiST的統計量，並由資料選取其最顯著的統計量以及權重的範圍，再藉由排序法(Permutation)這個技巧計算其臨界值，然後提出生活品質調整後存活時間差異之機動檢定，觀察其處理組的治療效果優於對照組的權重。

第二章將簡單介紹生活品質的基本概念，並回顧近幾年Q-TWiST的相關文獻。第三章提出生活品質調整後存活時間差異之機動檢定，第四章為模擬研究，藉以評估本文所提統計方法之表現，並與生活品質權重固定時的檢定做比較。第五章我們將分析兩筆實際資料，說明本文所提之生活品質調整後存活時間差異之機動檢定在生活品質資料的實際應用。最後，在第六章根據研究結果加以總結及討論未來可再進一步的研究方向。

第二章 文獻回顧

第一節 生活品質調整後存活時間

Q-TWiST方法的第一步就是定義臨床試驗中病人於生活品質各階段時間所代表的治療狀況，令 T_1 為病人接受藥物治療所產生的毒性與副作用的時間(TOX)； T_2 為疾病復發的時間； T_3 為病人死亡的時間，因此，Gelber et al. (1989)定義 T_2-T_1 為TWiST，即病人治療後恢復到幾乎未罹病狀態的時間； T_3-T_2 為疾病復發的時間(REL)。綜合這三個互斥的狀態，TOX、TWiST和REL為描述病人於研究期間各階段的情況，並依照各階段的病人狀況分別給予一個適當的生活品質權重，再藉由觀察病人所得到的各階段存活時間，合併為生活品質調整後存活時間(Quality Adjusted Survival Time)，如圖1所示。

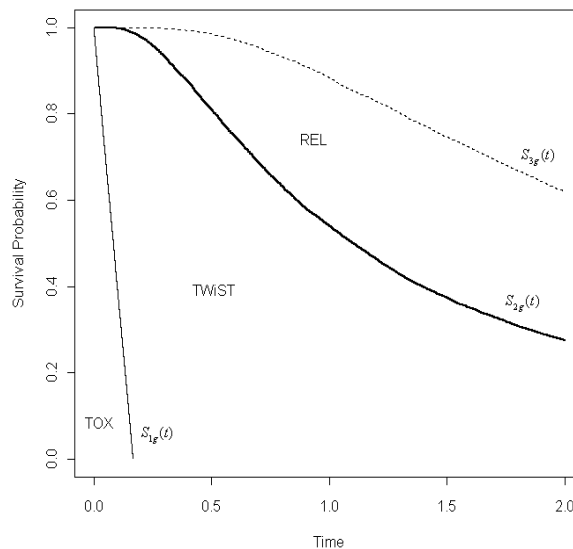


圖1. 生活品質各階段存活函數曲線圖

令 T_{ik} 與 T_{jk} 分別為生活品質資料中任意兩個階段內發生事件的時間點，其中 $i, j=1, 2, 3$ ， $i \neq j$ ， $k=1, \dots, n$ ，定義 U_{ik} 為設限時間， $\delta_{ik}=I(T_{ik} \leq U_{ik})$ 為右設限指標函數，

而我們實際觀測到的資料為 $X_{ik}=\min(T_{ik}, U_{ik})$ ，並且假設 T_{ik} 與 U_{ik} 獨立， T_{jk} 亦是同理，通常 $U_{1k}=U_{2k}=U_{3k}$ ，因為假設病人在TWiST階段已經設限時，在REL階段則沒有提供訊息。

為了檢定兩組的存活函數，我們令虛無假設 $H_0: Q_1(w_{TOX}, w_{REL})=Q_2(w_{TOX}, w_{REL})$ ，對立假設則設定為 $H_a: Q_1(w_{TOX}, w_{REL})\leq Q_2(w_{TOX}, w_{REL})$ 。假設 $g=1,2$ 為治療組別，我們令生活品質調整後存活時間為 $Q_g(w_{TOX}, w_{REL})$ ，其定義為

$$Q_g(w_{TOX}, w_{REL}) = w_{TOX} \int_0^\tau S_{1g}(t)dt + \int_0^\tau \{S_{2g}(t) - S_{1g}(t)\}dt + w_{REL} \int_0^\tau \{S_{3g}(t) - S_{2g}(t)\}dt,$$

則估計的生活品質調整後存活時間為

$$\hat{Q}_g(w_{TOX}, w_{REL}) = w_{TOX} \int_0^\tau \hat{S}_{1g}(t)dt + \int_0^\tau \{\hat{S}_{2g}(t) - \hat{S}_{1g}(t)\}dt + w_{REL} \int_0^\tau \{\hat{S}_{3g}(t) - \hat{S}_{2g}(t)\}dt, \quad (1)$$

其中 $\hat{S}_{ig}(t)$ 為 $S_{ig}(t) = P(T_{ig} > t)$ 的Kaplan-Meier (1958)存活函數估計， $i=1,2,3$ ， w_{TOX} 、 w_{REL} 為生活品質權重，且其值介於0和1之間，積分上限的 τ 為計算平均生活品質調整後存活時間的限制存活時間，Glasziou et al. (1990)認為受訪中位數 (median follow-up time)是一個合理的選擇，其他研究者可能會選擇存活曲線中最大的存活時間來做為限制存活時間，以觀察後期Q-TWiST的差異。

透過生活品質權重 w_{TOX} 與 w_{REL} 來調整 Q_g 中每個階段的存活狀態，如果兩個權重皆為1， Q_g 則回到無任何加權的整體存活曲線，且其分析是基於研究開始至病人死亡的時間。若權重皆為0， Q_g 只有提供TWiST這階段的時間，而沒有在TOX與REL提供訊息，此時對病人而言是注重在無病存活時間。在大多數的情況，生活品質權重的選擇反應在為了降低治療毒性或疾病復發而減少存活的利益。

令兩組生活品質調整後存活時間差異為

$$\Delta Q(w_{TOX}, w_{REL}) = Q_2(w_{TOX}, w_{REL}) - Q_1(w_{TOX}, w_{REL}).$$

則其估計式為

$$\Delta \hat{Q}(w_{TOX}, w_{REL}) = \hat{Q}_2(w_{TOX}, w_{REL}) - \hat{Q}_1(w_{TOX}, w_{REL}),$$

Glasziou et al. (1990)利用拔靴法(Bootstrap)求生活品質調整後存活時間的變異

數估計 $\widehat{Var}(\hat{Q}_g(w_{TOX}, w_{REL}))$, $g=1,2$ ，並令

$$T(w_{TOX}, w_{REL}) = \frac{\Delta \hat{Q}(w_{TOX}, w_{REL})}{\sqrt{\widehat{Var}(\Delta \hat{Q}(w_{TOX}, w_{REL}))}},$$

其中 $\widehat{Var}(\Delta \hat{Q}(w_{TOX}, w_{REL})) = \widehat{Var}(\hat{Q}_1(w_{TOX}, w_{REL})) + \widehat{Var}(\hat{Q}_2(w_{TOX}, w_{REL}))$ ，且證明

$T(w_{TOX}, w_{REL})$ 服從自由度為 $n_1 + n_2 - 2k$ 的 t 分配，其中 n_1 與 n_2 分別為 2 組的樣本

數， k 為治療狀態的數量。因此治療差異的 100 (1- α)% 信賴區間為

$$(\Delta_L, \Delta_U) = \Delta \hat{Q}(w_{TOX}, w_{REL}) \pm t_{\alpha/2, n_1+n_2-k} \sqrt{\widehat{Var}(\Delta \hat{Q}(w_{TOX}, w_{REL}))}.$$

將 Δ_L 與 Δ_U 令為 0 後，等式的兩邊平方可得

$$t_{\alpha/2}^2 [\widehat{Var}(\Delta \hat{Q}(w_{TOX}, w_{REL}))] - (\Delta \hat{Q}(w_{TOX}, w_{REL}))^2 = 0.$$

對其解等式便可得到無顯著治療差異的生活品質權重 (w_{TOX}, w_{REL}) 範圍。

第二節 生活品質調整後存活時間之共變異數矩陣

本節應用Murray和Cole(2000)的結果，得到生活品質調整後存活時間的變異

數共變異數為

$$V_{ijg} = \int_0^\tau \int_0^\tau \left[\int_v^\tau S_{jg}(t) dt \right] \left[\int_u^\tau S_{ig}(t) dt \right] G_{ijg}(u, v) dv du, i, j = 1, 2, 3, g = 1, 2, \quad (2)$$

其中

$$G_{ijg}(u, v) = \frac{P(X_i \geq u, X_j \geq v)}{P(X_i \geq u)P(X_j \geq v)} \times [\lambda_{ij}(u, v) - \lambda_{ij}(u|v)\lambda_j(v) - \lambda_{ji}(v|u)\lambda_i(u) + \lambda_i(u)\lambda_j(v)].$$

$\lambda_{ij}(u, v)$ 為生活品質調整後資料第*i*與*j*階段的病人在時間點*u*與*v*的風險函數；

$\lambda_{ij}(u|v)$ 為第*j*階段在時間點*v*還存活的病人，第*i*階段在時間點*u*的條件風險函

數； $\lambda_{ji}(v|u)$ 亦同樣道理。將 $S_{ig}(t)$ 、 $S_{jg}(t)$ 與 $G_{ijg}(u, v)$ 之一致估計式代入式(2)便

可以得到

$$\hat{V}_{ijg} = \sum_{\{u: X_{i(g)} \leq \tau\}} \sum_{\{v: X_{j(g)} \leq \tau\}} \left[\int_v^\tau \hat{S}_{jg}(t) dt \right] \left[\int_u^\tau \hat{S}_{ig}(t) dt \right] \\ \times n \frac{Y_{ij}(u, v)}{Y_i(u)Y_j(v)} \left[\frac{dN_{ij}(u, v)}{Y_{ij}(u, v)} - \frac{dN_{ij}(u|v)dN_j(v)}{Y_{ij}(u, v)Y_j(v)} - \frac{dN_{ji}(v|u)dN_i(u)}{Y_{ij}(u, v)Y_i(u)} + \frac{dN_i(u)dN_j(v)}{Y_i(u)Y_j(v)} \right].$$

其中 $Y_i(u) = \sum_{k=1}^n I(X_{ik} \geq u)$ 為第*i*階段在時間點*u*未設限也未死亡的人數； $Y_j(v)$ 則同

理；因此 $Y_{ij}(u, v) = \sum_{k=1}^n I(X_{ik} \geq u, X_{jk} \geq v)$ 為同時在第*i*、*j*階段分別在時間點*u*、*v*未

設限也未死亡的人數。 $dN_{ij}(u|v) = \sum_{k=1}^n \delta_{ik} I(X_{ik} = u) I(X_{jk} \geq v)$ 為第*j*階段在時間點*v*

未設限也未死亡的條件下，第*i*階段在時間點*u*死亡的條件人數。當*i=j*時，則

$\hat{G}_{ijg}(u, v) = \text{Var}(\hat{S}_{ig}(t))$ 。另外，我們可以將式子(1)改寫成

$$\hat{Q}_g(w_{TOX}, w_{REL}) = (w_{TOX} - 1) \int_0^\tau \hat{S}_{lg}(t) dt + (1 - w_{REL}) \int_0^\tau \hat{S}_{2g}(t) dt + w_{REL} \int_0^\tau \hat{S}_{3g}(t) dt.$$

整合上述式子可以得到生活品質調整後存活時間的變異數估計為

$$\hat{\sigma}_g^2(w_{TOX}, w_{REL}) = \sum_{i=1}^3 \sum_{j=1}^3 w_i w_j \hat{V}_{ijg} / n_g = \hat{V}_g / n_g. \quad (3)$$

其中 $w_1 = w_{TOX} - 1$ ， $w_2 = 1 - w_{REL}$ ， $w_3 = w_{REL}$ 。

因此，我們令

$$Z(w_{TOX}, w_{REL}) = \Delta \hat{Q}(w_{TOX}, w_{REL}) / \sqrt{\hat{\sigma}^2(w_{TOX}, w_{REL})}, \quad (4)$$

為生活品質調整後存活時間差異檢定統計式，其中

$$\hat{\sigma}^2(w_{TOX}, w_{REL}) = \hat{\sigma}_1^2(w_{TOX}, w_{REL}) + \hat{\sigma}_2^2(w_{TOX}, w_{REL}).$$

在給定生活品質權重下， $Z(w_{TOX}, w_{REL})$ 的極限分配會趨近於標準常態分配，因此

$p\text{-value} \approx P(Z(w_{TOX}, w_{REL}) > z(w_{TOX}, w_{REL}))$ ，其中 $z(w_{TOX}, w_{REL})$ 為觀察的檢定統計

值。當 $p\text{-value} \leq \alpha$ 時，我們宣稱第二組的處理比第一組的處理來的有效。

第三章 統計方法

Murray和Cole(2000)固定生活品質權重 w_{TOX} 、 w_{REL} 分別為 (0,0)、(0.5,0.5)、(1,1) 時，代表分別觀察只有TWiST階段、Tox與REL階段權重減半及整體存活時間，計算其檢定力與型I誤差，並非觀察到所有生活品質權重下的治療方法差異，此時，其他的生活品質權重所反應出的治療差異性將被忽略。因此，我們考慮由資料自行選取其適當生活品質權重，以反應兩組治療存活時間的差異，此時，我們定義

$$\Delta\hat{Q}(\hat{w}_{TOX}, \hat{w}_{REL}) = \hat{Q}_2(\hat{w}_{TOX}, \hat{w}_{REL}) - \hat{Q}_1(\hat{w}_{TOX}, \hat{w}_{REL}), \quad (5)$$

標準化後的檢定統計量則為

$$Z(\hat{w}_{TOX}, \hat{w}_{REL}) = \Delta\hat{Q}(\hat{w}_{TOX}, \hat{w}_{REL}) / \sqrt{\hat{\sigma}^2(\hat{w}_{TOX}, \hat{w}_{REL})}, \quad (6)$$

其中 $(\hat{w}_{TOX}, \hat{w}_{REL}) \equiv \arg \max_{(w_{TOX}, w_{REL}) \in [0,1]} \{Z(w_{TOX}, w_{REL})\}$ 。然而，欲觀察所有生活品質權重下治療方法的差別，其檢定統計量的極限分配不會趨近於標準常態分配，因此，本研究藉由讓資料選取最顯著的統計量，以觀察所選取到的生活品質權重所反應出的治療差異，並利用排序法(Permutation)造出其抽樣分配，根據其抽樣分配來判斷是否接受原始假設，並觀察其檢定力與型I誤差。

根據排序法建立 $Z(\hat{w}_{TOX}, \hat{w}_{REL})$ 之抽樣分配的步驟如下：

1. 由原始資料經過式子(6)的計算得到最顯著的生活品質權重 $(\hat{w}_{TOX}, \hat{w}_{REL})$
 $\equiv \arg \max_{(w_{TOX}, w_{REL}) \in [0,1]} \{Z(w_{TOX}, w_{REL})\}$ 及所對應的生活品質檢定統計量 $Z(\hat{w}_{TOX}, \hat{w}_{REL})$ 。
2. 將原始資料的兩組治療組別合併在一起，並利用抽出後不放回的方式，隨機給

予每個病患編號 1 至 $(n_1 + n_2)$ ，其中 n_1 與 n_2 分別為兩組的樣本數。

3. 將編號 1 到 n_1 的病患視為第一組(對照組)，編號 $(n_1 + 1)$ 至 n_2 視為第二組 (治療組)，將此資料視為新的資料，然後計算在治療差異最顯著的生活品質權重

$(\hat{w}_{TOX}, \hat{w}_{REL}) \equiv \arg \max_{(w_{TOX}, w_{REL}) \in [0,1]} \{Z(w_{TOX}, w_{REL})\}$ 時的檢定統計量(6)，稱它為

$Z^*(\hat{w}_{TOX}, \hat{w}_{REL})$ 。

4. 重複步驟 2 與步驟 3 執行 N 次，將得到的 N 個 $Z_i^*(\hat{w}_{TOX}, \hat{w}_{REL}), i = 1, \dots, N$ ，因此，

得到 $p\text{-value} \approx \sum_{i=1}^N I(Z_i^*(\hat{w}_{TOX}, \hat{w}_{REL}) > Z(\hat{w}_{TOX}, \hat{w}_{REL})) / N$ 。

在顯著水準 α 之下，當 $p\text{-value} \leq \alpha$ 時，我們宣稱當在最顯著的生活品質權重下，第二組的治療比第一組的來的有效。另外可以進一步計算生活品質調整後

存活時間的信賴域，可以求得在 $\Delta\hat{Q}(\hat{w}_{TOX}, \hat{w}_{REL})$ 的信賴域為

$$\Delta\hat{Q}(\hat{w}_{TOX}, \hat{w}_{REL}) \pm Z_c^* \sqrt{\hat{\sigma}^2(\hat{w}_{TOX}, \hat{w}_{REL})} \quad (7)$$

利用式子(7)可以找出生活品質權重影響治療方法顯著差異的範圍以及其代表的治療情形，其中 Z_c^* 為估計的臨界值。

第四章 模擬研究

第一節 模擬方法

本章節利用模擬方法，研究本文所提生活品質調整後存活時間差異之機動檢定中所選到最顯著的生活品質權重，觀察在不同情況下生活品質權重的分佈以及其代表的涵意，並與在固定生活品質權重的檢定做檢定力與型 I 誤差的比較。

我們參考 Murray 和 Cole (2000) 的模擬方式，考慮在一個對照組($g=1$)與一個處理組($g=2$)下，樣本數 n_{ig} 分別為 50 與 100， $i=1,2,3, g=1,2$ ，限制平均壽命時間點為 $\tau=2$ 之下進行生活品質資料模擬研究。考量生活品質資料中三種治療狀態：治療毒性反應、疾病復發以及死亡，分別令為 T_{1g}, T_{2g}, T_{3g} ，其中 $T_{1g} \leq T_{2g} \leq T_{3g}$ 。 T_{1g} 為治療毒性階段，考慮其存活時間服從均勻分布介於 0 到 $1/6$ ，記為 $U(0, 1/6)$ 。將 TWiST 與 REL 狀態分別令為 T_{2g}^* 與 T_{3g}^* ，其存活時間服從二元對數常態分配，其中心參數為 (μ_{1g}, μ_{2g}) 以及對應的離散參數為 $(\sigma_{1g}, \sigma_{2g})$ ， ρ_g 則為 T_{2g}^* 與 T_{3g}^* 的相關係數。因此疾病復發時間 $T_{2g} = T_{1g} + T_{2g}^*$ 和死亡時間 $T_{3g} = T_{2g} + T_{3g}^*$ 。將設限時間令為 $C = U \times I(B=1) + \tau \times I(B=0)$ ，其中 U 為均勻分配介於 $1/6$ 到 τ ， τ 為限制平均壽命時間點， B 為機率為 0.1 的 *Bernoulli* 分配。

此模擬研究中，考慮不同情形下的 TWiST 與 REL，兩組的二元對數常態分配中對數存活時間的中心參數分別為 $(\mu_{11}, \mu_{21}) = (\mu_{12}, \mu_{22}) = (0,0)$ 、 $(-0.3,-0.3)$ 、 $(0.3,0.3)$ ， $\sigma_{ig}=1, i=1,2,3, g=1,2$ ，樣本數 n_{ig} 分別為 50 與 100， $i=1,2,3, g=1,2$ ，相關係數分別考量 $\rho_1=\rho_2=0.2$ 、 0.5 、 0.9 ，與固定生活品質權重為 $(0,0)$ 、 $(0.5,0.5)$ 、 $(1,1)$

進行型 I 誤差率的比較，並觀察樣本數在 200 之下，其型 I 誤差的控制情形，結果列於表 1 和表 2。檢定力的研究，我們考慮在對照組對數存活時間的中心參數 $(\mu_{11}, \mu_{21}) = (0,0)$ 下，處理組的中心參數 $(\mu_{12}, \mu_{22}) = (0.5,0.5)$ 、 $(0.5,0.2)$ ，兩組的 $\sigma_{ig}=1$ ，樣本數 n_{ig} 為 100， $i=1,2,3, g=1,2$ ，相關係數 $\rho_1=\rho_2=0.2、0.5、0.9$ 和 $\rho_1=0.9, \rho_2=0.5$ 以及 $\rho_1=0.9, \rho_2=0.2$ ，對應的存活函數圖形繪於圖 2 與圖 3，與固定生活品質權重為 $(0,0)$ 、 $(0.5,0.5)$ 、 $(1,1)$ 進行檢定力的比較，模擬結果列於表 3。

模擬次數為 3,000 次，根據模擬的生活品質樣本資料，計算各種檢定統計量，並且記錄是否拒絕虛無假設。在虛無假設成立時，計算 3,000 次當中拒絕虛無假設的比例，此為型 I 誤差率估計。檢定力估計則為在對立假設成立時，3,000 次模擬當中拒絕虛無假設的比例。在 $\alpha=0.05$ 下，型 I 誤差率估計的標準誤為 $0.004(\approx \sqrt{0.05 \times 0.95 / 3000})$ 。此外，我們記錄每一組設定下生活品質權重的分配狀況，其生活品質權重分佈情形列於表 4 至表 6。

第二節 模擬結果

表 1 與表 2 分別在樣本數 50 與 100 之下，觀察生活品質資料各種設置型 I 誤差的表現。由表 1 可以得知，在樣本數為 50 時，各種設置的固定生活品質權重之檢定的型 I 誤差率皆無法控制在原本設定的 0.05 左右，而本研究所提之生活品質調整後差異之機動檢定所求出的型 I 誤差率在各設置下皆可控制得宜。將樣本數提高到 100 後，由表 2 發現部分設置的固定生活品質權重之檢定其型 I 誤差仍然無法控制在 0.05。而將固定生活品質權重之檢定的樣本數提高到 200 時可以

發現，在中心參數設定為 (0.3,0.3) 時，還是有部分生活品質權重的型 I 誤差無法控制在 0.05。表 4 與表 5 為虛無假設成立下，在樣本數分別為 50 與 100 時機動檢定所選取到生活品質權重 ($\hat{w}_{TOX}, \hat{w}_{REL}$) 的分佈狀況，可以發現，在樣本數為 50 時，生活品質權重 \hat{w}_{TOX} 不論在哪種設置下，其選取到的權重平均皆在 0.5 左右，而 \hat{w}_{REL} 的平均皆高過 0.5，且在 3 種 (μ_{1g}, μ_{2g})， $g=1,2$ 設置之下，隨著疾病復發與死亡時間相關性 ρ_g ， $g=1,2$ 越低，則 \hat{w}_{REL} 的平均越高。在樣本數為 100 時，所選取到生活品質權重 \hat{w}_{TOX} 的平均一樣皆在 0.5 左右， \hat{w}_{REL} 則在 (μ_{1g}, μ_{2g})=(0.3,0.3) 時其權重平均皆明顯的高過 0.5，且相同地在 3 種 (μ_{1g}, μ_{2g}) 的設置下，隨著疾病復發與死亡時間相關性 ρ_g 降低， \hat{w}_{REL} 的平均值而提高。

就檢定力而言，由表 3 可以發現，兩種的參數設定，在兩組治療中疾病復發與死亡時間的相關係數相等的情形下，我們所提的檢定方法其檢定力大過固定生活品質權重在(1,1)時的檢定力，而比生活品質權重在(0,0)、(0.5,0.5)的時候較小。而在對照組的疾病復發與死亡時間相關性遠大過治療組的相關性($\rho_1=0.9, \rho_2=0.2$)的時候，我們所提的機動檢定法卻比固定生活品質權重在(1,1)時較低，而比固定生活品質權重(0,0)時較高，其造成的原因由圖 2 與圖 3 中的(4)可以發現，在兩組的疾病復發與死亡時間相關性差距較大時，兩組死亡時間的存活函數曲線差距較大，生活品質權重在(1,1)時，其代表的涵意是觀察整體存活時間，因此在此時檢定力會較大，而生活品質權重在(0,0)代表只看 TWiST 階段，因此在此設置下檢定力較低。表 6 為對立假設成立下，在樣本數為 100 時各設置所選的生活品質權

重分佈情形，可以發現，在對照組的疾病復發與死亡時間相關性遠大過治療組的相關性的時候，生活品質權重 \hat{w}_{REL} 的平均值會比兩組的疾病復發與死亡時間相關性相等時來的高，表示在此種設置下，生活品質權重 \hat{w}_{REL} 將更被注重。在中心參數設定為(0.5, 0.2)時，由於給定治療組死亡時間的平均值較低，因此生活品質權重 \hat{w}_{REL} 應較不被注重，但在 $\rho_1=0.9, \rho_2=0.2$ 時，生活品質權重 \hat{w}_{REL} 的平均值又遠比其他相關性設置較高，更可以說明在對照組的疾病復發與死亡時間相關性遠大過治療組的相關性時，生活品質權重 \hat{w}_{REL} 將更被注重。本文所提之機動檢定法其檢定力在部分設置下低於固定生活品質權重下的檢定力，其造成的原因可能是我們所提的機動檢定法的型 I 誤差率偏低的影響，但整體而言，機動檢定法是將所有的生活品質權重都納入考量，而固定生活品質權重卻只考慮到部分的權重。

第五章 資料分析

資料一：賀爾蒙治療對乳癌的影響

德國乳癌研究組織(German Breast Cancer Study Group, GBCS) 於 1984 年 7 月至 1989 年 12 月招募淋巴結陽性反應的乳癌病患(Sauerbrei and Royston, 1999)，將病患以隨機跟非隨機的方式分成兩組，對照組是以 3 和 6 週期的 CMF 化學治療方法，處理組則以 3 和 6 週期的 CMF 化學治療方法加上性賀爾蒙調節劑 Tamoxifen 的治療方式(Schmoor et al., 1996)，並觀察其疾病復發與死亡的時間。若至實驗結束仍未發現疾病復發或死亡，則該筆資料視為設限資料。此實驗中總共有 686 位病患，440 位病人接受 CMF 化學治療(對照組)，246 位接受 CMF 化學治療跟賀爾蒙治療(處理組)，其 Kaplan-Meier 存活函數估計圖繪於圖 4。

此筆資料中並無觀測治療毒性反應的時間，因此我們只考慮生活品質權重 \hat{w}_{REL} 來調整存活時間。表 7 為不同的限制平均壽命 τ 之下，本文所提之機動檢定法與固定生活品質權重的檢定統計量與 p 值的比較。由 p 值可以發現機動檢定在所有的 τ 之下，其檢定都拒絕虛無假設，而固定生活品質權重之檢定在給定生活品質權重為 1 的時候，較小的 τ 其檢定都是不拒絕的。而在生活品質權重的選取中，大多數的 τ 所選取到的生活品質權重都為 0，其代表著只看 TWiST 這階段的差異，而限制平均壽命為 1000 與 1250 時，其選到的生活品質權重並不像其他的都為 0，表示隨著觀察的時間點不同，也會影響到生活品質權重的選取。另外，我們更進一步計算生活品質調整後存活時間差異的信賴域，得到生活品質權重影

響治療方法顯著差異的範圍，其結果列於表 7，並以 τ 為 750 與 1000 時為例將信賴域圖形展示於圖 5。然而，當生活品質權重 \hat{w}_{REL} 給定顯著差異範圍之外時，其信賴域將包含 0，表示治療方法無顯著差異。

資料二：長短期化學治療對乳癌的影響

國際乳癌研究組織(International Breast Cancer Study Group, IBCSG) 為了評估單一週期的輔助化學治療與長期性化學治療對於女性乳癌治療的效果，招募了 1229 位淋巴結陽性反應的乳癌病患，以隨機的方式將病人分成 2 組，對照組為單一週期的完整化學治療，處理組為 6 到 7 個週期的傳統定時性化學治療。並記錄治療毒性反應、疾病復發與死亡時間(Gelber et al., 1991)。

此筆資料由統計軟體 SAS 從 IBCSG Trial V 的原始資料中，隨機抽取出的子資料，因此樣本數為 1135，其中接受單一週期化療的病患為 378 位，接受長期性化療的病患為 757 位，其 Kaplan-Meier 存活函數估計圖繪於圖 6。表 8 為不同的限制平均壽命 τ 之下，本文所提之機動檢定法與固定生活品質權重的檢定統計量與 p 值的比較，可以發現，在所有考慮的限制時間點之下，機動檢定所得到的 p 值皆為 0，即代表著所建立的抽樣分配皆小於所得到的最顯著生活品質調整後存活時間差異統計量，且所選取到的生活品質權重皆為(1,0)，其代表著只注重 TOX 與 TWiST 階段，而忽略 REL 這階段所提供的訊息。在固定生活品質權重之檢定的 p 值中，可以看出生活品質權重給定(1,1)在限制平均壽命從 60 到 100 時，皆比權重(0,0)、(0.5,0.5)要大，表示權重給定(0,0)與(0.5,0.5)時，治療方法差

異顯著性比權重為(1,1)時較高，在限制平均壽命為 50 之下，儘管權重(1,1)比(0,0)與(0.5,0.5)較小，但 3 種生活品質權重下皆不顯著，而在機動檢定的方法中，卻能找出其顯著的生活品質權重。然而，我們進一步計算生活品質調整後存活時間差異的信賴域，得到生活品質權重影響治療方法顯著差異的範圍，並以限制平均壽命為 60、80、100 為例將結果列於表 9，且將信賴域圖形繪於圖 7，然而，生活品質權重影響治療方法顯著差異的範圍是一個平面範圍，因此，表 9 是將 \hat{w}_{TOX} 分成 11 個點後，再以每個點對應到 \hat{w}_{REL} 的顯著範圍而作呈現。圖 8 是 τ 為 60、80 與 100 時，信賴域上下界的等高線圖，在 τ 為 80 與 100 時，其信賴域上界皆在 0 之上，因此觀察信賴域下界若高於 0 則為生活品質權重影響顯著差異範圍，而在 τ 為 60 時，信賴域上界低於 0 的範圍為對照組優於處理處的生活品質權重範圍，信賴域包含 0 則為治療方法無顯著差異的生活品質權重範圍。

第六章 結論與未來研究方向

本文主要針對生活品質資料，藉由讓資料自行選取其適當生活品質權重，並根據排序法建構出生活品質調整後存活時間差異之機動檢定。由模擬結果中發現，機動檢定在每種設置下的型 I 誤差率皆控制在顯著水準，而固定生活品質權重之檢定卻隨著給定生活品質權重的不同，其型 I 誤差率也受到影響，且部分設置的型 I 誤差率無法控制在顯著水準。而隨著資料設置的不同，機動檢定的檢定力雖然比不同部分固定生活品質權重之檢定較低，但是在真實情況中，隨著治療方法差異性的不同，應綜觀所有生活品質權重的影響，而不能只考量部分生活品質權重，因此本文所提之機動檢定便具有此優勢。

雖然已有檢定方法可供進行存活函數差異的研究，但是在生活品質資料研究中，更重要的是生活品質權重的選擇。未來，可藉由配適 Cox 模型，引進共變數的訊息至生活品質調整後存活時間差異檢定，透過更多資訊來探討生活品質調整後存活時間的差異性。

參考文獻

- Cole, B. F., Gelber, R. D., and Goldhirsch, A. for the International Breast Cancer Study Group (IBCSG). (1993). Cox regression models for quality-adjusted survival analysis. *Statistics in Medicine* **12**, 975-987.
- Gelber, R. D., Gelman R. S., and Goldhirsch A. (1989). A quality-of- life-oriented endpoint for comparing therapies, *Biometrics* **45**, 781-795.
- Gelber, R. D., Goldhirsch, A., and Cavalli, F. for the International Breast Cancer Study Group (IBCSG). (1991). Quality-of-life-adjusted evaluation of a randomized trial comparing adjuvant therapies for operable breast cancer. *Annals of Internal Medicine* **114**, 621-628.
- Gelber, R. D., Cole, B. F., Gelber, S., and Goldhirsch, A. (1995). Comparing treatments using quality-adjusted survival: The Q-TWiST method. *American Statistician* **49**, 161-169.
- Gelber, R. D., Goldhirsch, A., Cole, B. F., Wieand, H. S., Schroeder, G., and Krook, J. E. (1996). A quality adjusted time without symptoms or toxicity (Q-TWiST) analysis of adjuvant radiation therapy and chemotherapy for resectable rectal cancer. *Journal of the National Cancer Institute* **88**, 1039-1045.
- Glasziou, P. P., Simes, R. J., and Gelber, R. D. (1990). Quality adjusted survival analysis. *Statistics in Medicine* **9**, 1259-1276.

- Kaplan, E. L. and Meier P. (1958). Nonparametric estimation from incomplete observations. *Annal of the Institute of Statistical Mathematics* **53**, 457-481.
- Murray, S. and Cole, B. (2000). Variance and sample size calculations in quality-of-life-adjusted survival analysis (Q-TWiST). *Biometrics* **56**, 173-182.
- Murray, S. (2001). Using weighted Kaplan-Meier statistics in nonparametric comparisons of paired censored survival outcomes. *Biometrics* **57**, 361-368.
- Sauerbrei, W. and Royston, P. (1999) Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. *Journal of the Royal Statistical Society Series A* **162** part 1, 71-94.
- Schmoor, C. Olschewski, M. and Schumacher, M. (1996). Randomized and non-randomized patients in clinical trials: experiences with comprehensive cohort studies. *Statistics in Medicine* **15**, 263-271.
- Sloan J. A., Sargent D. J., Lindman J., Allmer C., Vargas-Chanes D., Creagan E.T., Bonner J. A., O'Connell M. J., Dalton R.J., Rowland K. M., Brooks B. J. & Laurie J. A. (2002) A new graphic for quality adjusted life years (Q-TWiST) survival analysis: The Q-TWiST plot. *Quality of Life Research* **11**, 37-45.

附錄

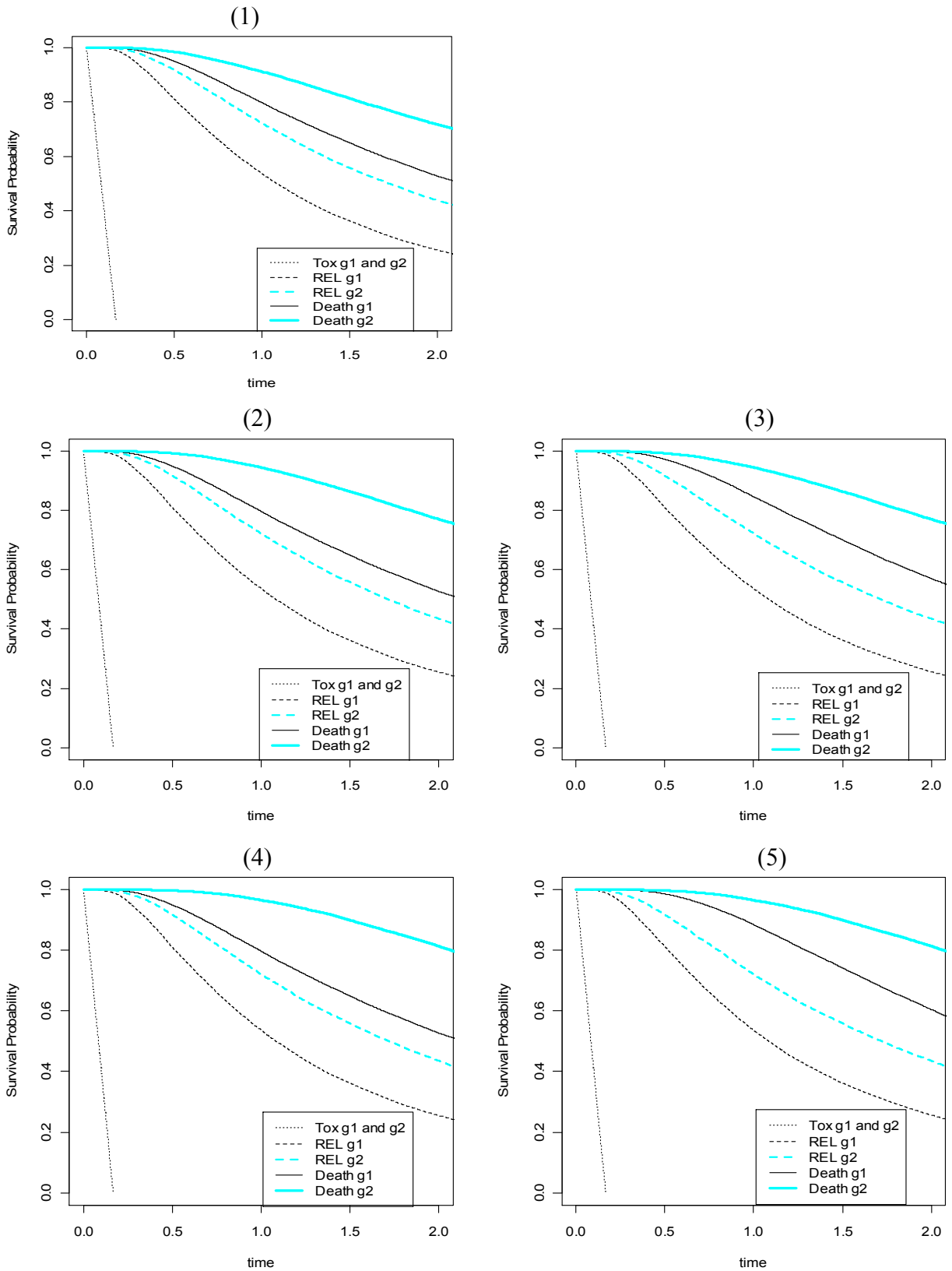


圖 2. $(\mu_{12}, \mu_{22}) = (0.5, 0.5)$ 之下，各相關係數設定之生活品質資料存活函數圖形：

(1) $\rho_1 = \rho_2 = 0.9$ (2) $\rho_1 = 0.9, \rho_2 = 0.5$ (3) $\rho_1 = \rho_2 = 0.5$ (4) $\rho_1 = 0.9, \rho_2 = 0.2$ (5) $\rho_1 = \rho_2 = 0.2$ 。

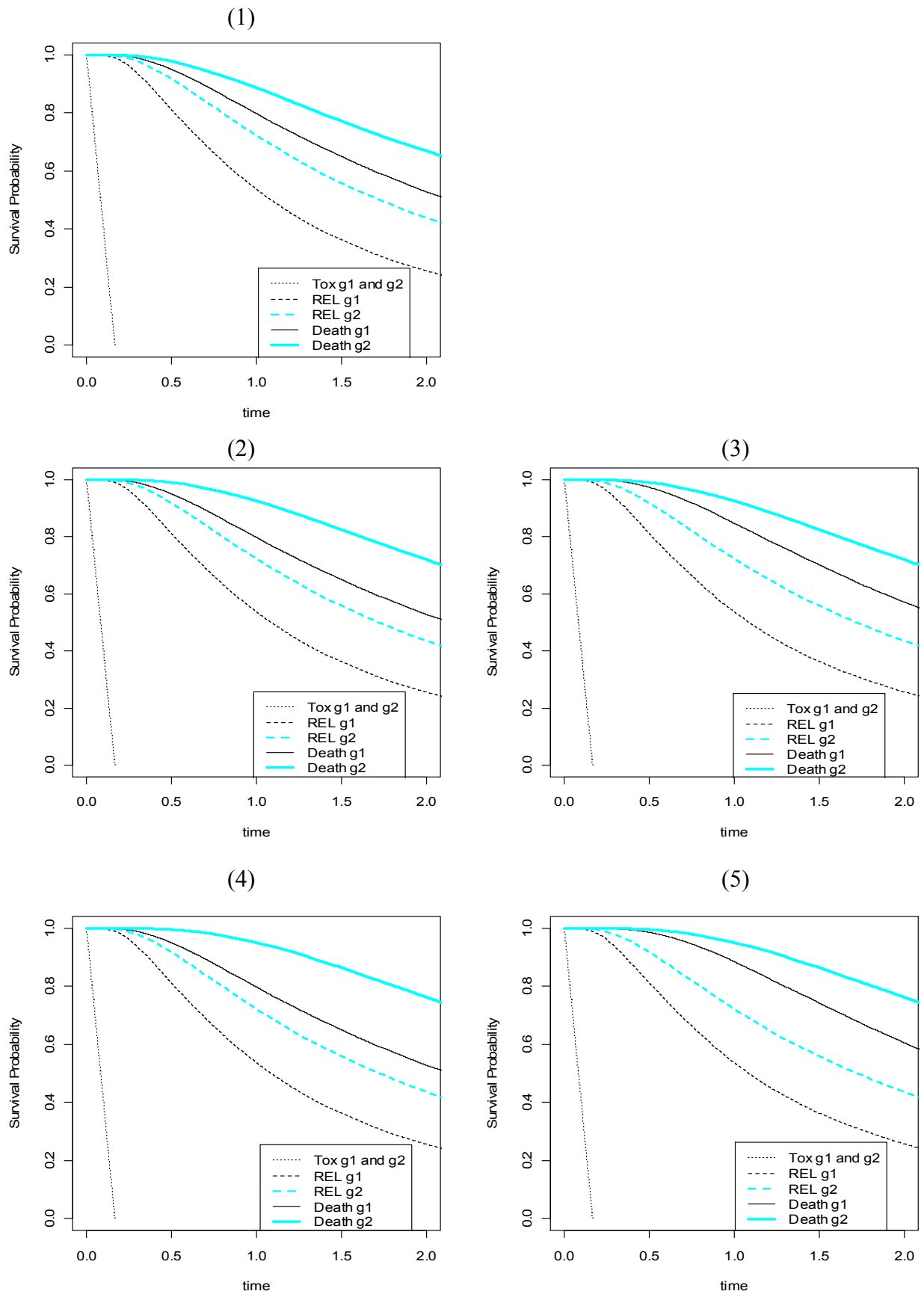


圖 3. $(\mu_{12}, \mu_{22}) = (0.5, 0.2)$ 之下，各相關係數設定之生活品質資料存活函數圖形：

(1) $\rho_1 = \rho_2 = 0.9$ (2) $\rho_1 = 0.9, \rho_2 = 0.5$ (3) $\rho_1 = \rho_2 = 0.5$ (4) $\rho_1 = 0.9, \rho_2 = 0.2$ (5) $\rho_1 = \rho_2 = 0.2$ 。

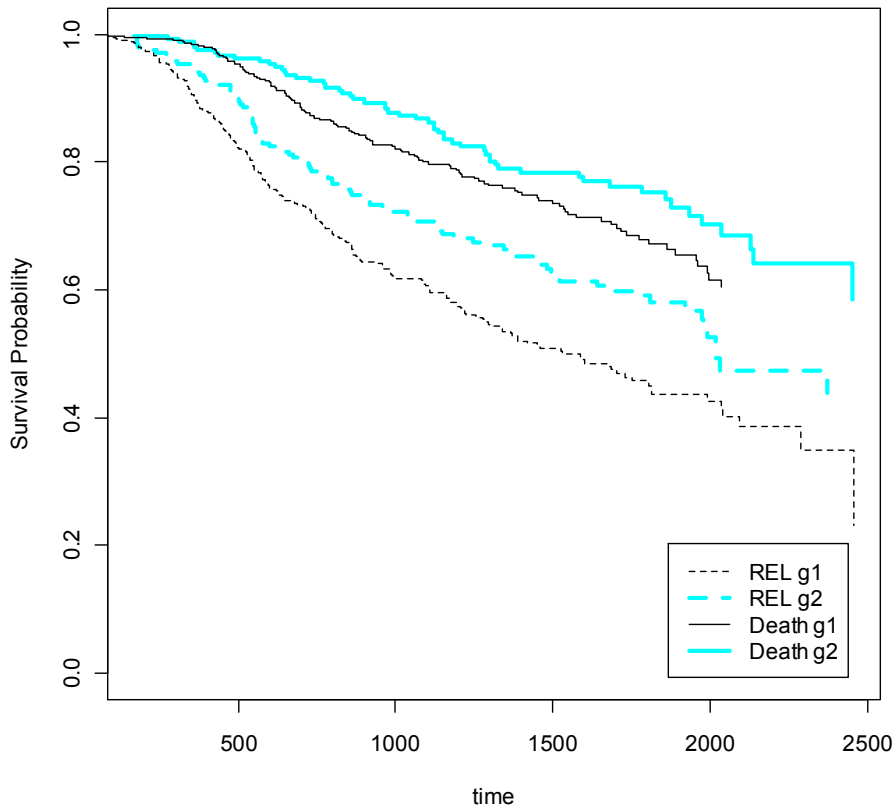


圖 4. 德國乳癌研究資料之 Kaplan-Meier 存活函數估計圖。

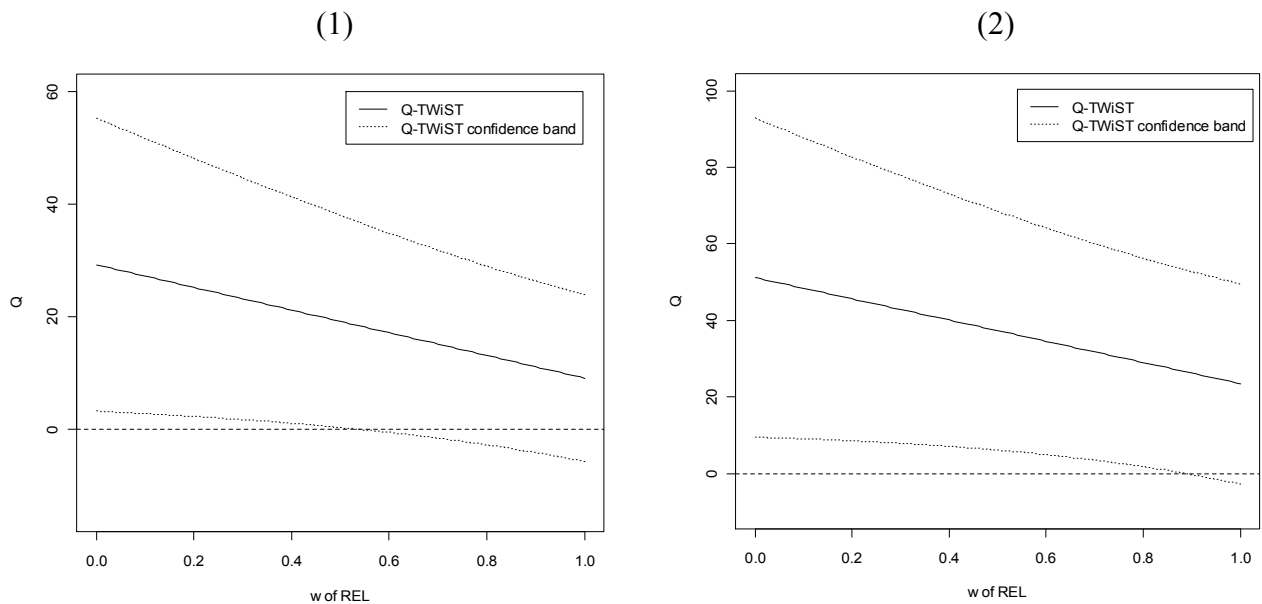


圖 5. 德國乳癌研究資料在生活品質調整後存活時間差異的信賴域圖形：

(1) $\tau=750$ (2) $\tau=1000$ 。

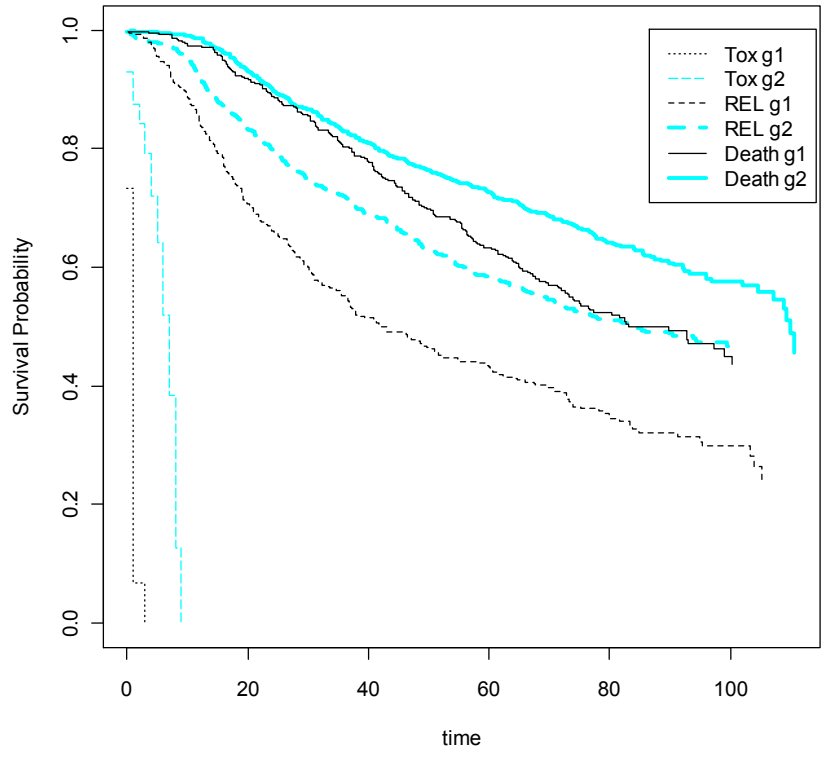


圖 6. 國際乳癌研究組織第五期資料之 Kaplan-Meier 存活函數估計圖。

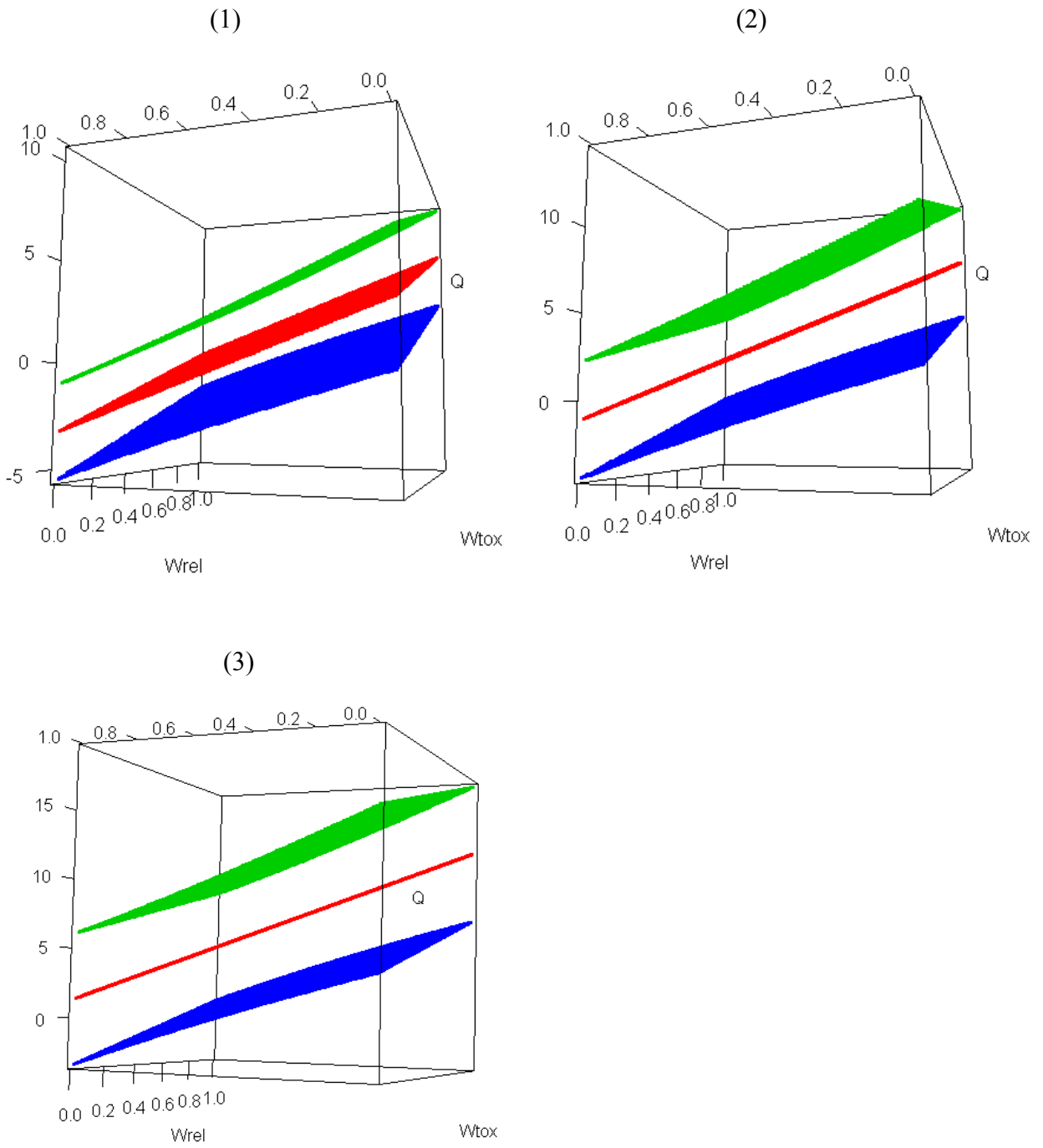


圖 7. 國際乳癌研究組織第五期資料在生活品質調整後存活時間差異的信賴域圖

形：(1) $\tau=60$ (2) $\tau=80$ (3) $\tau=100$ 。

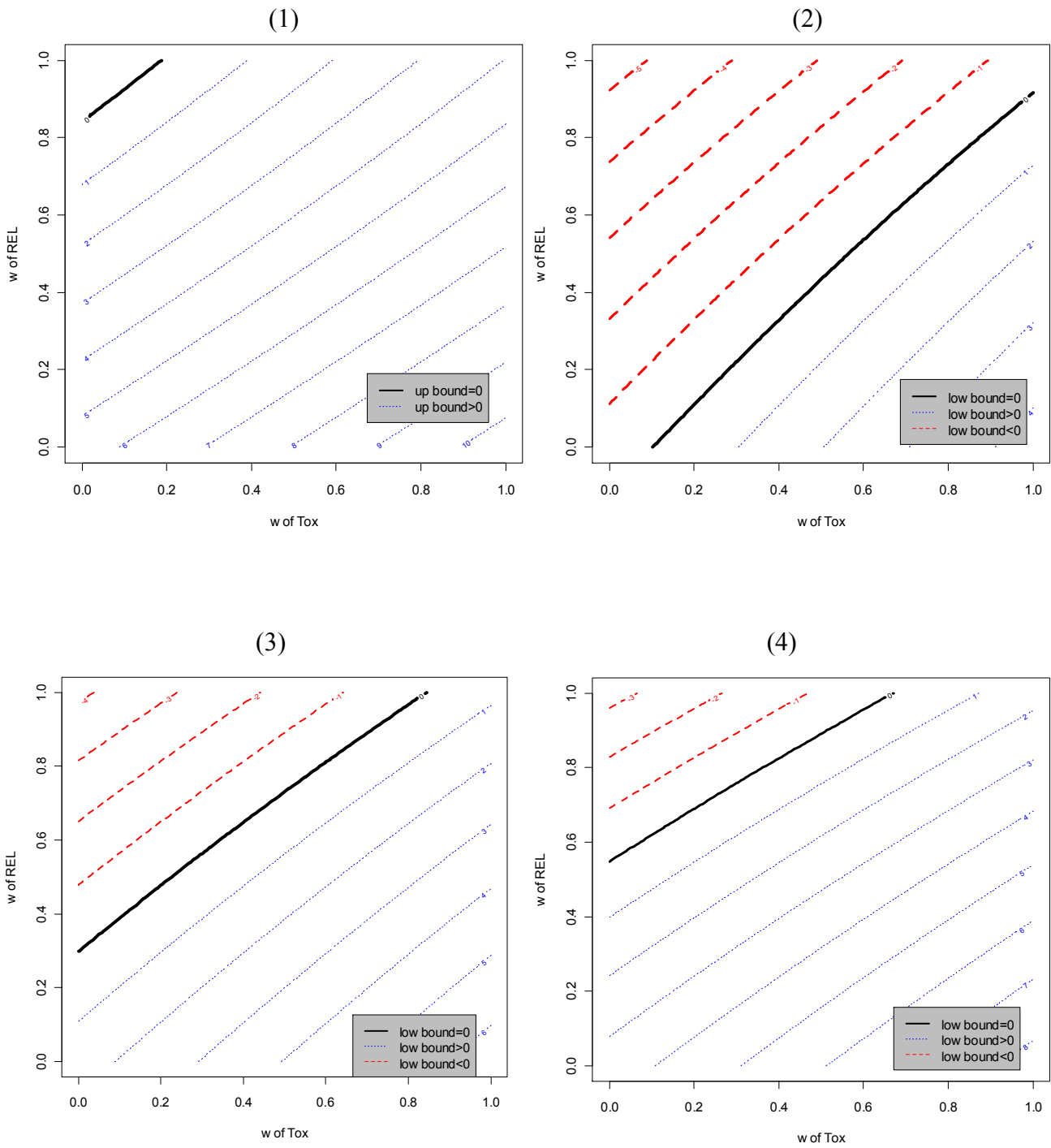


圖 8. 國際乳癌研究組織第五期資料在生活品質調整後存活時間差異的信賴域

上、下界等高線圖：(1) $\tau=60$ 上界 (2) $\tau=60$ 下界 (3) $\tau=80$ 下界 (4) $\tau=100$ 下界。

表 1. $n_{ig}=50, i=1,2,3, g=1,2$ 及 $\alpha = 0.05$ 之下，各個設置下生活品質資料差異的型 I

誤差率估計。

		$Z(\hat{w}_{TOX}, \hat{w}_{REL})$	$Z(0,0)$	$Z(0.5,0.5)$	$Z(1,1)$
(μ_{1g}, μ_{2g})	ρ_g				
(0,0)	0.9	0.058	0.0807	0.0737	0.106
(0,0)	0.5	0.0617	0.0773	0.081	0.116
(0,0)	0.2	0.061	0.0723	0.0813	0.121
(-0.3,-0.3)	0.9	0.0543	0.0777	0.0757	0.09
(-0.3,-0.3)	0.5	0.0587	0.0673	0.074	0.090
(-0.3,-0.3)	0.2	0.0543	0.0667	0.07	0.091
(0.3,0.3)	0.9	0.0617	0.0817	0.094	0.148
(0.3,0.3)	0.5	0.0513	0.0787	0.091	0.167
(0.3,0.3)	0.2	0.052	0.077	0.104	0.196

表 2. $n_{ig}=100, i=1,2,3, g=1,2$ 與固定生活品質權重 $n_{ig}=100$ 及 $n_{ig}=200, \alpha=0.05$ 之

下，各個設置下生活品質資料差異的型 I 誤差率估計。

(μ_{1g}, μ_{2g})	ρ_g	$Z(\hat{w}_{TOX}, \hat{w}_{REL})$	$n_{ig}=100$			$n_{ig}=200$		
			$Z(0,0)$	$Z(0.5,0.5)$	$Z(1,1)$	$Z(0,0)$	$Z(0.5,0.5)$	$Z(1,1)$
(0,0)	0.9	0.0417	0.062	0.0627	0.0773	0.05	0.0493	0.063
(0,0)	0.5	0.050	0.058	0.0587	0.076	0.057	0.0517	0.0563
(0,0)	0.2	0.0513	0.057	0.0583	0.082	0.055	0.0563	0.0687
(-0.3,-0.3)	0.9	0.0443	0.061	0.0593	0.069	0.0513	0.0523	0.06
(-0.3,-0.3)	0.5	0.051	0.0567	0.057	0.07	0.062	0.056	0.0597
(-0.3,-0.3)	0.2	0.0553	0.059	0.061	0.0667	0.0577	0.0587	0.058
(0.3,0.3)	0.9	0.044	0.0617	0.066	0.1023	0.0523	0.0607	0.0877
(0.3,0.3)	0.5	0.0473	0.058	0.0743	0.123	0.057	0.0637	0.0863
(0.3,0.3)	0.2	0.046	0.064	0.073	0.1323	0.0557	0.0673	0.1067

表 3. $n_{ig}=100, i=1,2,3, g=1,2, \alpha = 0.05$ 之下，各個設置下生活品質資料差異的檢
定力估計。

(μ_{12}, μ_{22})	(ρ_1, ρ_2)	$Z(\hat{w}_{TOX}, \hat{w}_{REL})$	$Z(0,0)$	$Z(0.5,0.5)$	$Z(1,1)$
(0.5,0.5)	(0.9,0.9)	0.8837	0.935	0.927	0.835
(0.5,0.5)	(0.5,0.5)	0.8743	0.9277	0.944	0.831
(0.5,0.5)	(0.9,0.5)	0.936	0.937	0.972	0.947
(0.5,0.5)	(0.2,0.2)	0.8573	0.9327	0.954	0.822
(0.5,0.5)	(0.9,0.2)	0.96	0.9367	0.986	0.980
(0.5,0.2)	(0.9,0.9)	0.8552	0.935	0.878	0.649
(0.5,0.2)	(0.5,0.5)	0.8497	0.9277	0.9	0.663
(0.5,0.2)	(0.9,0.5)	0.8993	0.937	0.947	0.881
(0.5,0.2)	(0.2,0.2)	0.8383	0.9327	0.925	0.690
(0.5,0.2)	(0.9,0.2)	0.9503	0.9367	0.977	0.961

表 4. $n_{ig}=50, i=1,2,3, g=1,2$ ，機動檢定法的型 I 誤差率估計值下，各個設置下

$(\hat{w}_{TOX}, \hat{w}_{REL})$ 個數分佈。

(μ_{1g}, μ_{2g})	ρ_g	$(\hat{w}_{TOX}, \hat{w}_{REL})$									平均	標準差
		(0,0)	(0,1)	(1,0)	(1,1)	(0,0~1)	(1,0~1)	(0~1,0)	(0~1,1)	(0~1,0~1)		
(0,0)	0.9	553	669	534	682	213	220	44	60	25	(0.4998,0.550)	(0.493,0.470)
(0,0)	0.5	485	659	504	646	287	267	42	72	38	(0.498, 0.569)	(0.492,0.461)
(0,0)	0.2	422	653	443	616	354	335	38	96	43	(0.495, 0.599)	(0.490,0.447)
(-0.3,-0.3)	0.9	543	656	596	606	245	231	44	47	32	(0.499, 0.521)	(0.493,0.470)
(-0.3,-0.3)	0.5	527	629	486	600	317	323	42	46	30	(0.489, 0.543)	(0.493,0.458)
(-0.3,-0.3)	0.2	487	625	473	570	350	352	36	78	29	(0.489, 0.553)	(0.492,0.454)
(0.3,0.3)	0.9	506	686	501	737	192	198	42	118	20	(0.510, 0.592)	(0.490, 0.468)
(0.3,0.3)	0.5	395	727	406	727	259	261	36	146	43	(0.503, 0.649)	(0.487, 0.444)
(0.3,0.3)	0.2	360	672	337	748	295	309	40	192	47	(0.513, 0.672)	(0.484, 0.431)

表 5. $n_{ig}=100, i=1,2,3, g=1,2$ ，機動檢定法的型 I 誤差率估計下，各個設置下

$(\hat{w}_{TOX}, \hat{w}_{REL})$ 個數分佈。

(μ_{1g}, μ_{2g})	ρ_g	$(\hat{w}_{TOX}, \hat{w}_{REL})$									平均	標準差
		(0,0)	(0,1)	(1,0)	(1,1)	(0,0~1)	(1,0~1)	(0~1,0)	(0~1,1)	(0~1,0~1)		
(0,0)	0.9	544	659	572	644	226	239	33	64	19	(0.504, 0.542)	(0.493, 0.470)
(0,0)	0.5	509	591	507	590	325	342	29	76	31	(0.503, 0.544)	(0.493, 0.458)
(0,0)	0.2	473	592	482	561	351	386	39	79	37	(0.501, 0.555)	(0.492, 0.451)
(-0.3,-0.3)	0.9	591	625	619	583	234	248	37	40	23	(0.501, 0.501)	(0.494, 0.471)
(-0.3,-0.3)	0.5	550	599	563	572	307	307	38	40	24	(0.499, 0.509)	(0.495, 0.461)
(-0.3,-0.3)	0.2	520	567	504	538	357	398	35	49	32	(0.500, 0.529)	(0.494, 0.453)
(0.3,0.3)	0.9	544	682	517	706	211	195	41	93	11	(0.497, 0.570)	(0.491, 0.470)
(0.3,0.3)	0.5	451	664	462	689	289	270	33	105	37	(0.504, 0.607)	(0.490, 0.454)
(0.3,0.3)	0.2	415	645	432	663	330	327	30	123	35	(0.505, 0.618)	(0.490, 0.447)

表 6. $n_{ig}=100, i=1,2,3, g=1,2$ ，機動檢定法的檢定力估計下，各個設置下

$(\hat{w}_{TOX}, \hat{w}_{REL})$ 個數分佈。

(μ_{12}, μ_{22})	(ρ_1, ρ_2)	$(\hat{w}_{TOX}, \hat{w}_{REL})$									平均	標準差
		(0,0)	(0,1)	(1,0)	(1,1)	(0,0~1)	(1,0~1)	(0~1,0)	(0~1,1)	(0~1,0~1)		
(0.5,0.5)	(0.9,0.9)	540	166	571	180	584	615	121	55	168	(0.515, 0.360)	(0.481, 0.388)
(0.5,0.5)	(0.5,0.5)	341	117	379	155	789	817	86	49	267	(0.519, 0.442)	(0.477, 0.365)
(0.5,0.5)	(0.9,0.5)	137	386	165	431	651	754	34	163	279	(0.530, 0.683)	(0.472, 0.350)
(0.5,0.5)	(0.2,0.2)	264	99	254	102	904	906	67	56	348	(0.498, 0.498)	(0.473, 0.344)
(0.5,0.5)	(0.9,0.2)	68	480	70	634	532	639	19	320	238	(0.546, 0.813)	(0.465, 0.283)
(0.5,0.2)	(0.9,0.9)	1031	77	930	61	291	277	236	20	77	(0.477, 0.144)	(0.480, 0.295)
(0.5,0.2)	(0.5,0.5)	680	56	736	74	550	538	181	20	165	(0.512, 0.235)	(0.480, 0.326)
(0.5,0.2)	(0.9,0.5)	326	218	383	248	704	727	89	83	222	(0.518, 0.481)	(0.478, 0.391)
(0.5,0.2)	(0.2,0.2)	538	56	527	65	691	745	138	16	224	(0.511, 0.305)	(0.478, 0.333)
(0.5,0.2)	(0.9,0.2)	112	380	115	475	676	752	35	179	276	(0.530, 0.713)	(0.471, 0.329)

表 7. 德國乳癌資料在各限制平均壽命的機動檢定與固定生活品質權重檢定統

計量，其對應的單尾 p 值列於括號中。

τ	500	750	1000	1250	1500	1750	2000
$Z(\hat{w}_{REL})$	2.257 (0.036)	2.449 (0.0175)	2.6042 (0.0095)	2.737 (0.008)	2.945 (0.004)	3.025 (0.0025)	3.121 (0.002)
\hat{w}_{REL}	0	0	0.15	0.04	0	0	0
\hat{w}_{REL} 顯著範圍	$\hat{w}_{REL} < 0.08$	$\hat{w}_{REL} < 0.55$	$\hat{w}_{REL} < 0.89$	$\hat{w}_{REL} < 0.92$	$\hat{w}_{REL} < 0.84$	$\hat{w}_{REL} < 0.85$	$\hat{w}_{REL} < 0.88$
$Z(0)$	2.257 (0.01200)	2.449 (0.00717)	2.597 (0.00470)	2.736 (0.00311)	2.945 (0.00162)	3.025 (0.00124)	3.121 (0.00090)
$Z(0.5)$	1.714 (0.04328)	2.2219 (0.01318)	2.530 (0.00570)	2.616 (0.00445)	2.643 (0.00411)	2.684 (0.00363)	2.762 (0.00287)
$Z(1)$	0.174 (0.43087)	1.342 (0.08987)	1.900 (0.02874)	1.943 (0.02603)	1.762 (0.03901)	1.785 (0.03712)	1.855 (0.03177)

表 8. 國際乳癌研究組織第五期資料在各限制平均壽命的機動檢定與固定生活

品質權重檢定統計量，其對應的單尾 p 值列於括號中。

τ	50	60	70	80	90	100
$Z(\hat{w}_{TOX}, \hat{w}_{REL})$	5.6967 (0)	5.7660 (0)	5.7707 (0)	5.8023 (0)	5.9093 (0)	5.9843 (0)
$(\hat{w}_{TOX}, \hat{w}_{REL})$	(1, 0)	(1, 0)	(1, 0)	(1,0)	(1,0)	(1,0)
$Z(0,0)$	0.9042 (0.1829)	1.9523 (0.02545)	2.6072 (0.004564)	3.0921 (0.000993)	3.5336. (0.000205)	3.8737 (5.36e-05)
$Z(0.5,0.5)$	1.1883 (0.1174)	2.0412 (0.02062)	2.6349 (0.004208)	3.1059 (0.000949)	3.4837 (0.000247)	3.7446 (9.03e-05)
$Z(1,1)$	1.3877 (0.08262)	1.8533 (0.03192)	2.3237 (0.01007)	2.7588 (0.002901)	3.0421 (0.001174)	3.1998 (0.000688)

表 9. 國際乳癌研究組織第五期資料在生活品質調整後存活時間差異的信賴域

中顯著差異 ($\hat{w}_{TOX}, \hat{w}_{REL}$) 範圍。

\hat{w}_{TOX}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
τ	\hat{w}_{REL}										
60		0	0~0.11	0~0.22	0~0.32	0~0.43	0~0.53	0~0.63	0~0.73	0~0.82	0~0.91
80	0~0.29	0~0.38	0~0.47	0~0.56	0~0.64	0~0.73	0~0.81	0~0.88	0~0.96	0~1	0~1
100	0~0.54	0~0.61	0~0.68	0~0.75	0~0.82	0~0.89	0~0.95	0~1	0~1	0~1	0~1