

東海大學統計研究所

碩士論文

加入短期反應訊息於多組治療存活時間比

● 較之適應性隨機指派設計

**Adaptive Randomization for Multiarm
Survival Clinical Trials Using Short-Term
Response Information**

指導教授：張玉媚 博士

研究生：涂瑞銓

中華民國一零八年七月

東海大學碩士班研究生

論文口試委員審定書

統計學系碩士班涂瑞銓君所提之論文

加入短期反應訊息於多組治療存活時間比較之適
應性隨機指派設計

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 張之媚
江存如

中華民國 108 年 07 月 04 日

致謝

首先要誠摯地感謝我的指導老師，張玉媚博士，從一開始的題目發想，提供符合我想研究的方向且非常有趣的題目，增進我想要完成的動力；在每一次的討論時，張老師總是不厭其煩地指導我，雖然我常常無法將我的想法表達地很完善，但張老師仍然能指引我，並提供我很多想法；在論文初稿完成後，仍細心地幫我校稿。這份論文的完成，要感謝張老師的付出與包容，協助我能順利地完成碩士論文。

此外，要感謝系上的沈葆聖博士，雖然沈老師不是我的指導老師，但依然主動地幫助我，提供多方面的想法；也感謝沈老師在繁忙之餘，擔任我的口試委員，幫我審核這份論文，在口試時，以豐富的學術涵養，提供我更多意見，使得這份論文更完全。同時也要感謝口試委員陳春樹博士對於這份論文的建言，感謝陳老師在百忙之中審核，以嚴謹的研究態度，給予這份論文精闢的見解與寶貴的建議，使得這份論文更為完整。

摘要

在第二期臨床試驗時，我們不希望花費太多的時間和資源，因此我們需要有效率的隨機指派方法，挑出可能有療效的治療方法，趕緊進到第三期試驗做藥效的確認。傳統上使用存活時間當作主要終點(primary endpoint)，將花費太長的時間，無法有效地在試驗期間適應性地隨機指派不同治療給病人，然而使用短期的療效(例如腫瘤反應率)當作隨機指派的依據，可能病人短期療效好，但存活時間卻不長。由於主要的目的是希望新的治療可以使病人存活時間延長，因此 Huang 等人(2009)等人使用短期反應率的訊息來加速存活時間適應性隨機指派臨床試驗。本論文將推廣 Huang 等人(2009)的方法由兩組治療的比較推廣至多組比較，並將存活時間為指數分配的假設推廣至韋伯分配。由於 Huang 等人(2009)提出的模型是每一個世代(cohort)一位病人進案時都進行期中分析，下一個進到試驗的病人則是根據修正的隨機指派之機率進行隨機指派，為節省時間，我們考慮使用每一個世代為五個人進到試驗後再進行期中分析，並比較每一個病人和每五個病人的模擬結果。

關鍵詞:適應性隨機指派、白血病、反應率、馬可夫鏈蒙地卡羅、替代終點

Abstract

We usually don't want to waste substantial amount of time and resources in phase two clinical trials, therefore we need effective methods of randomization, so that we can chose a potential treatment on survival, and confirm its benefit in phase three trial quickly. Traditionally, using survival time as a primary endpoint, takes a long period of time to assign patients to different groups ineffectively during the adaptive-randomization trails. However, using short-term efficacy (e.g. tumor response rate) as a primary endpoint, may show that patients have good efficacy shortly, but do not have long period of survival time. Owing to prolonging survival time is a common goal of a better treatment, Huang et al. (2009) proposed a model which used short-term response information to facility adaptive randomization for survival clinical trials. In this article, we extend the method which Huang et al. (2009) proposed to multi-arms clinical trials. Moreover, we assume that the survival time of patients is following the mixture Weibull distribution. In addition, since the model of Huang et al. (2009) proposed conduct interim analyses each cohort of one patient, and following patient randomize into the trial is based on the revised probability of adaptive randomization. In order to save time, we consider to conduct interim analyses each cohort of five patients, then compare the respective simulation results.

Key words: adaptively randomization, leukemia, response rate, Markov chain Monte Carlo, surrogate endpoint

目錄

一、研究背景和目的.....	1
二、文獻回顧.....	4
三、統計方法.....	8
四、模擬研究.....	12
五、結論與討論.....	26
參考文獻.....	27

一、研究背景和目的

臨床試驗通常可以分為四期，第一期主要是瞭解藥物的安全劑量，在人體所能夠承受之最大耐受劑量(Maximum Tolerance Dose, MTD)；第二期則是探索藥物的療效和安全性，為少數病人的臨床試驗，再根據該試驗結果來設計第三期臨床試驗的受試者人數；第三期受試者對象較多，為大型的臨床試驗，受試者人數可能由數百到數千人不等，在設計上一般以隨機分配、雙盲及對照試驗等方式進行，主要證實藥物療效及安全性，作為上市前的依據；第四期則是藥品上市後的追蹤，且與已核准的適應症相關，監測該藥物是否有先前幾個階段沒發現的副作用或是不良反應等情況，試驗設計較為簡單，但是需要大量的病人進行長時間的觀察。

傳統上第二期臨床試驗通常為單臂試驗(single-arm trial)，會將歷史性對照組(historical control)與實驗組比較。而常常在第二期試驗結果證明實驗組有希望具有療效後，第三期卻以失敗告終；這類和歷史性標準治療的比較，病人的母體可能不同，代表第二期可能有混淆因子(confounding)存在，導致第二期試驗結果的錯誤，也令研究人員漸漸重視第二期試驗的隨機化。

選擇主要終點為臨床試驗重要的一環；一個好的主要終點為，可以用以證明新藥是否有好的藥效為準則，例如：在病人接受治療後，觀察其反應是否良好，如腫瘤縮小。在白血病的第二期試驗中，常使用完全緩解(complete remission)作為終點；雖然完全緩解代表病人狀態短暫地顯示治療效果好，也可能是代表可以延長病人的存活時間，但是也有病人在完全緩解後，突然惡化。很多化學療法可以短暫地使病人有良好的治療效果，卻沒有真正地延長病人的存活時間，而我們最終的目的還是希望病人能有較長的存活時間，因此使用存活時間較適合當作主要終點。

存活時間可分為總存活期(overall survival)和無惡化存活時間(progression free survival)，個別代表觀察病人是否死亡和病人是否惡化、產生藥物不良反應，這些觀察時間可能是幾週多則數年。而使用存活時間當作主要終點，第一，會導致研究花費很長的時間才能做出結果。第二，在前面第二段提到的，為了避免病人母體不同，因此應採用雙臂(two-arms)或多臂(multi-arms)的試驗，代表可能使用更多的病人數量。因此在第二期試驗時，應使用療效-適應性隨

機指派(outcome-adaptive randomization)，根據前面病人的治療結果，來調整後續受試者進入各治療組的機率，使較多的受試者分派至治療效果較好的隨機指派方法，並進行期中分析，制定提早結束試驗的標準，在當前的訊息顯示藥物有顯著療效時，提早結束試驗，減少花費時間和病人數量。

由於在進行隨機指派時，還無法觀察到前面病人的完整的存活時間，因此，在新的病人進來要進行隨機指派時，可以使用跟存活時間有關係指標當作替代終點(surrogate endpoint)，例如：接受不同藥物治療之後，經過幾週的觀察病人膽固醇下降的幅度，來評斷不同藥物的療效。若使用替代終點，可以減少進行臨床試驗的時間，代表藥廠能更快的向市場揭露藥效的好壞，也代表可以減少藥廠的花費。重要的是，替代終點不一定總是能代表真正的臨床結果，且通常替代終點是唯一或者是主要影響疾病的主要療效。當研究人員在選用替代終點時，會根據已經證實某指標對臨床結果有顯著關係的證據來制定假設，但有可能，某疾病還未有這方面的證據，而研究人員卻制定指標跟臨床結果有關係的假設，可能使臨床試驗結果更差，也可能將病人指派到較差的治療，甚至造成極大的生命的喪失。葉嘉新等人(2008)提到，1980年代在美國曾上市 Flecainide 用以治療無症狀之心室早期收縮(VPC, Ventricular premature contraction)，當時廣為接受的醫學理論「控制 VPC，可以減少因突發性心律不整所導致的猝死」，結果在該藥上市後，美國國家衛生院執行一大型臨床試驗，試驗進行一半，資料安全監測委員會發現使用 Flecainide 的病人，VPC 固然大為減少，但突發性心律不整的病人反而增加，據估計，因此藥枉死的美國人可能高達五萬人。

為了要快速地建立隨機指派的機制，減少時間和病人數量的花費，又不能失去真正的臨床結果，我們考慮使用替代終點來加速使用存活時間當主要終點的臨床試驗，這裡我們假設腫瘤的反應率可以在接受治療後馬上得知，並使用腫瘤的反應率當作替代終點，在還未觀察到病人存活時間時，使用反應率和目前所觀察的存活時間做隨機指派，在觀察到之後，使用替代終點和完整的存活時間做隨機指派。

Huang 等人(2009)同時利用反應率和存活時間的訊息，根據得到的訊息更新後驗機率，而此後驗機率將作為適應性調整隨機指派的準則。雖然在還未觀察到主要療效指標之前，無法完全利用準確的病人存活時間，但可以知道病人至少活了多久，這些也是可以利用的訊息。在初期，由於各組病人還

未死亡，觀察到各組病人的存活時間訊息可能相近，因此較依賴短期的反應的訊息，去做適應性隨機指派，而隨著觀察時間越長，便可以得到的越多存活時間的訊息去更新後驗機率。因此，本論文推廣 Huang 等人(2009)的方法由兩組治療的比較推廣至多組比較，提出適應性調整隨機指派的準則。

本文在第二章的文獻回顧中介紹 Huang 等人(2009)使用短期反應訊息加速第二期適應性隨機指派臨床試驗的方法。第三章將原本兩組治療的方法推廣至三組，並將指數分配推廣至韋伯分配，看此方法是否穩健。第四章為模擬研究，將評估存活時間使用存活時間假設為韋伯分配時是否適用此方法，此外，我們將模擬每一人進到試驗就進行期中分析和每五人進到試驗才進行期中分析，比較兩者的差異，並將與不考慮短期反應的適應性隨機指派方法做比較。最後，在第五章根據模擬結果加以總結及討論未來進一步的研究方向。

二、文獻回顧

Huang 等人(2009)提出使用短期反應和存活時間兩者的訊息來加速急性骨髓性白血病(Acute myeloid leukemia)的第二期臨床試驗設計。他們認為短期反應是可以預測存活時間的，在期中分析時，還觀測不到正確的存活時間，因此使用短期反應當作替代終點，並將短期反應分為四類：(1) 產生抗藥性或死亡(resistance to treatment or death)、(2) 穩定(stable)、(3) 部分緩解(partial remission)、(4) 完全緩解(complete remission)。由於總存活期較難觀察，且若病人惡化，病人可能會去尋求其他的治療，所以使用的是無惡化存活期當作終點。令 $x = a$ 或 b ，個別代表接受 A 治療或 B 治療， n_x 代表接受 x 治療的病人個數。若在治療 x 中的第 i 個病人，發生第 k 類短期反應， $k=1, \dots, 4$ ，那麼定義 $S_{x,k,i} = 1$ 和 $S_{x,j,i} = 0$ ， $1 \leq j \leq 4, j \neq k$ ，假設向量 $(S_{x,1,i}, \dots, S_{x,4,i})$ ， $i=1, \dots, n_x$ 都是獨立且有相同的多項分配 $\text{Multi}(1, p_{x,1}, \dots, p_{x,4})$ ，其中 $p_{x,k}$ 代表接受 x 治療的病人分配到第 k 類短期反應的機率。此外，給定病人分配到第 k 類下， $T_{x,i}$ 代表接受 x 治療的第 i 個病人的無惡化存活期，且 $T_{x,i}$ 服從參數為 $\lambda_{x,k}$ (rate) 的指數分配；那麼病人還未分到某一個短期反應下的 $T_{x,i}$ ，服從混合(mixture)的指數分配。Huang 等人(2009)假設在接受治療後，可以馬上觀測到病人的短期反應，並根據歷史資料對參數假設先驗分配，提出的模型如下：

$$x = a, b \text{ (treatment indicator)}$$

$$(S_{x,1,i}, S_{x,2,i}, S_{x,3,i}, S_{x,4,i}) \sim \text{Multi}(1, p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4}), i = 1, \dots, n_x$$

$$T_{x,i} \sim \sum_{k=1}^4 p_{x,k} \text{Exp}(\lambda_{x,k}), i = 1, \dots, n_x$$

$$(p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4}) \sim \text{Dir}(\gamma_{x,1}, \gamma_{x,2}, \gamma_{x,3}, \gamma_{x,4})$$

$$\mu_{x,k} \equiv \frac{1}{\lambda_{x,k}} \sim \text{IG}(\alpha_{x,k}, \beta_{x,k}), k = 1, \dots, 4$$

根據指數分配和逆伽瑪分配的期望值分別為 $\mu_{x,k}$ 和 $\frac{\beta_{x,k}}{\alpha_{x,k}-1}$ ， $x=a, b, k=1, \dots, 4$ ，和無惡化存活期(週)的期望值歷史資料，對先驗分配 $\mu_{x,k}$ 的參數做假設： $\alpha_{x,k} = 11, x=a, b, \beta_{x,1} = 40, \beta_{x,2} = 300, \beta_{x,3} = 750, \beta_{x,4} = 1100, k=1, \dots, 4$ ，

即先驗分配 $\mu_{x,1}$ 的期望值為 $\frac{\beta_{x,1}}{\alpha_{x,1}-1} = 4$ ，相同地， $\mu_{x,2}$ 、 $\mu_{x,3}$ 、 $\mu_{x,4}$ 先驗分配的期望值個別為30、75、110，以上是根據無惡化存活期(週)的期望值歷史資料做的假設，這樣的先驗分配，是帶有訊息的先驗分配，但不會太嚴格，以至於可以在進行中的試驗，會根據新觀察到的資料而修正。且設定 $(p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4})$ 先驗分配為Dirichlet分配，參數 $\gamma_{x,k} = 0.5, x=a, b, k=1, \dots, 4$ ，為不帶有訊息的先驗分配。根據上述假設，可以簡單地得到後驗分配：

$$\begin{aligned} & \text{後驗機率 } p(p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4} | S_{x,1}, S_{x,2}, S_{x,3}, S_{x,4}) \propto \\ & L(p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4} | S_{x,1}, S_{x,2}, S_{x,3}, S_{x,4}) \times p(p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4}) \\ & = \prod_{i=1}^{n_x} \left(\frac{1!}{S_{x,1,i}! S_{x,2,i}! S_{x,3,i}! S_{x,4,i}!} \times p_{x,1}^{S_{x,1,i}} p_{x,2}^{S_{x,2,i}} p_{x,3}^{S_{x,3,i}} p_{x,4}^{S_{x,4,i}} \right) \\ & \quad \times \frac{\Gamma(\sum_{k=1}^4 \gamma_{x,k})}{\prod_{k=1}^4 \Gamma(\gamma_{x,k})} \times \prod_{k=1}^4 p_{x,k}^{\gamma_{x,k}-1} \\ & = p_{x,1}^{n_{x,1}} p_{x,2}^{n_{x,2}} p_{x,3}^{n_{x,3}} p_{x,4}^{n_{x,4}} \times \frac{\Gamma(\sum_{k=1}^4 \gamma_{x,k})}{\prod_{k=1}^4 \Gamma(\gamma_{x,k})} \times p_{x,1}^{\gamma_{x,1}-1} p_{x,2}^{\gamma_{x,2}-1} p_{x,3}^{\gamma_{x,3}-1} p_{x,4}^{\gamma_{x,4}-1} \\ & \propto \frac{\Gamma(\sum_{k=1}^4 (\gamma_{x,k} + n_{x,k}))}{\prod_{k=1}^4 \Gamma(\gamma_{x,k} + n_{x,k})} \times \prod_{k=1}^4 p_{x,k}^{\gamma_{x,k} + n_{x,k} - 1} \end{aligned}$$

所以 $(p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4})$ 後驗分配為

$$\text{Dir}(\gamma_{x,1} + n_{x,1}, \gamma_{x,2} + n_{x,2}, \gamma_{x,3} + n_{x,3}, \gamma_{x,4} + n_{x,4})。$$

假設接受 x 治療且發生第 k 類短期反應的病人數為 $n_{x,k}$ ，對於其觀察到或設

限的存活時間為 $t_{x,1}^{(k)}, \dots, t_{x,n_{x,k}}^{(k)}$ ，與之對應的指標為 $\delta_{x,1}^{(k)}, \dots, \delta_{x,n_{x,k}}^{(k)}$ (0代表設

限，1代表可觀察到惡化或死亡發生)。

$$\begin{aligned} & \text{後驗機率 } p(\mu_{x,k} | \mathbb{t}_x^{(k)}) \propto L(\mu_{x,k} | \mathbb{t}_x^{(k)}, \delta_x^{(k)}) \times p(\mu_{x,k}) \\ & = \prod_{i=1}^{n_{x,k}} [f(t_{x,i}^{(k)})^{\delta_{x,i}^{(k)}} \times S(t_{x,i}^{(k)})^{1-\delta_{x,i}^{(k)}}] \times \frac{\beta_{x,k}^{\alpha_{x,k}}}{\Gamma(\alpha_{x,k})} \times \mu_{x,k}^{-\alpha_{x,k}-1} \times \exp\{-\frac{\beta_{x,k}}{\mu_{x,k}}\} \\ & = \prod_{i=1}^{n_{x,k}} \left\{ \left[\frac{1}{\mu_{x,k}} e^{-t_{x,i}^{(k)}/\mu_{x,k}} \right]^{\delta_{x,i}^{(k)}} \left[e^{-t_{x,i}^{(k)}/\mu_{x,k}} \right]^{1-\delta_{x,i}^{(k)}} \right\} \times \frac{\beta_{x,k}^{\alpha_{x,k}}}{\Gamma(\alpha_{x,k})} \times \mu_{x,k}^{-\alpha_{x,k}-1} \times \exp\{-\frac{\beta_{x,k}}{\mu_{x,k}}\} \\ & = \mu_{x,k}^{-\sum \delta_{x,i}^{(k)}} \times e^{-\sum t_{x,i}^{(k)}/\mu_{x,k}} \times \frac{\beta_{x,k}^{\alpha_{x,k}}}{\Gamma(\alpha_{x,k})} \times \mu_{x,k}^{-\alpha_{x,k}-1} \times \exp\{-\frac{\beta_{x,k}}{\mu_{x,k}}\} \end{aligned}$$

$$\propto \mu_{x,k}^{-\sum \delta_{x,i}^{(k)} - \alpha_{x,k} - 1} \exp\left\{-\frac{\beta_{x,k} + \sum t_{x,i}^{(k)}}{\mu_{x,k}}\right\} \times \frac{(\beta_{x,k} + \sum t_{x,i}^{(k)})^{\alpha_{x,k} + \sum \delta_{x,i}^{(k)}}}{\Gamma(\alpha_{x,k} + \sum \delta_{x,i}^{(k)})}$$

那麼， $\mu_{x,k}$ 後驗分配為 $IG\left(\alpha_{x,k} + \sum_{i=1}^{n_{x,k}} \delta_{x,i}^{(k)}, \beta_{x,k} + \sum_{i=1}^{n_{x,k}} t_{x,i}^{(k)}\right)$ 。

以 μ_x 表示 $T_{x,i}$ 的期望值， $x = a, b$ 。給定 $(p_{x,1}, \dots, p_{x,4}, \mu_{x,1}, \dots, \mu_{x,4})$ 下，可以得到 $\mu_x = \sum_{k=1}^4 p_{x,k} \mu_{x,k}$ ，在試驗進行期間，不斷更新和比較 μ_a 和 μ_b ，用模擬的方式可以計算出 $p = \Pr(\mu_a > \mu_b | \text{data})$ ，並使用這個機率作為調整隨機指派的依據，即指派病人到 A 治療的機率為 p ，到 B 治療的機率為 $q = 1 - p$ 。如果在試驗進行的期間(包括試驗結束)，若 $p > p_U$ (或 $p < p_L = 1 - p_U$)，那麼就判斷 A 治療(或 B)為較好的治療。若在試驗結束後，還沒有選出哪一個為較好的治療，就判斷兩治療沒有顯著差異。模擬資料的病人個數為 120 人，假設招募率為每週一人，在觀察 120 週後，額外再觀察 40 週。在表一中第一種設計為假設兩種治療都相等(null hypothesis)，即發生四種短期反應的機率相同，且四種短期反應下的無惡化存活期有相同分配，選 $p_L = 0.025$ 當作決策標準，使得模擬結束後，選為 A 治療(B 治療)較好的機率為 0.046(0.048)，即如同頻率論(frequentist)的單尾型 I 誤差(type I error)；其他的模擬也使用相同的 p_L 當作決策標準，因此其他兩種模擬中被選為較好的治療的機率，如同頻率論中的 power，並控制在雙尾型 I 誤差約為 0.1。在不考慮短期反應的設計裡，將病人接受兩種治療個別的存活時間假設為指數分配且期望值為 μ_a, μ_b ，假設兩者的先驗分配為 $IG(\alpha, \beta)$ ， $\alpha = 2, \beta = 60$ ，那麼此先驗假設的期望值為 $\frac{\beta}{\alpha - 1} = 60$ ，此約為第一種模擬下存活時間參數的平均 $\left(\frac{4+30+75+110}{4} = 54.75\right)$ ，並選擇 $p_L = 0.007$ 判斷試驗提早停止的準則。

由表一可以看到三種模擬結果，考慮短期反應訊息的設計比起一般不考慮短期反應、只使用無惡化存活期做的設計，選擇出較好的治療的機率較大，平均也使用更少的病人。

表一、存活時間為混合指數分配，每一人進到試驗就進行期中分析，考慮短期反應和不考慮短期反應的設計

治療	反應率	四種反應下平均存活時間	考慮短期反應的設計		不考慮短期反應的設計	
	$p_1 p_2 p_3 p_4$	$\mu_1 \mu_2 \mu_3 \mu_4$	平均病人數	被選重的機率	平均病人數	被選重的機率
1. A 和 B 有相同的反應率和存活時間						
A	0.2 0.4 0.1 0.3	4 30 75 110	58(50%)	0.046	59(50%)	0.046
B	0.2 0.4 0.1 0.3	4 30 75 110	58(50%)	0.048	59(50%)	0.045
共			116(100%)		118(100%)	
2. B 有較好的反應率，AB 有相同存活時間						
A	0.2 0.4 0.1 0.3	4 30 75 110	16(18%)	0.0002	26(25%)	0.002
B	0.1 0.1 0.2 0.6	4 30 75 110	71(82%)	0.590	77(75%)	0.429
共			87(100%)		103(100%)	
3. B 有較好的反應率跟存活時間						
A	0.2 0.4 0.1 0.3	4 30 75 110	11(18%)	0.0002	21(29%)	0.001
B	0.1 0.1 0.2 0.6	6 45 112 165	51(82%)	0.976	72(71%)	0.648
共			62(100%)		93(100%)	

三、統計方法

假設招募率是每週一個病人，120 週後若試驗還未結束，額外觀察 40 週，即第一個病人觀察 159 週，第二個病人觀察 158 週，...，最後一個人最多觀察 40 週。因此我們假設設限時間為：

$$C_{t,i} \sim \text{Uniform}(0, 160 - i), i = 1, 2, \dots, 120,$$

我們假設存活時間 $T_{x,i}$ ， $i = 1, \dots, 120$ ，為混合韋伯分配，並將原本兩組治療推廣至多組，當有三組治療時， $x = a, b, c$ ，確認此方法是否穩健，以下是我們的模型：

$$x = a, b, c \text{ (treatment indicator)}$$

$$(S_{x,1,i}, S_{x,2,i}, S_{x,3,i}, S_{x,4,i}) \sim \text{Multi}(1, p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4}), i = 1, \dots, n_x$$

$$T_{x,i} \sim \sum_{k=1}^4 p_{x,k} \text{Wei}(\lambda_{x,k}, \theta_{x,k}), i = 1, \dots, n_x$$

$$(p_{x,1}, p_{x,2}, p_{x,3}, p_{x,4}) \sim \text{Dir}(\gamma_{x,1}, \gamma_{x,2}, \gamma_{x,3}, \gamma_{x,4})$$

$$\lambda_{x,k} \sim \Gamma(\alpha_{x,k}, \beta_{x,k})$$

$$\theta_{x,k} \sim \Gamma(\eta_{x,k}, \rho_{x,k}), k = 1, \dots, 4$$

存活時間 $T_{x,i}$ 服從韋伯分配參數為 $\lambda_{x,k}$ (scale) 和 $\theta_{x,k}$ (shape)，我們考慮四種短期反應下的病人之風險隨著時間分別為：快速上升、上升、風險不變、風險下降，假設先驗分配 $\theta_{x,k}$ 的期望值分別為：2、1.2、1、0.95，因為 Gamma 分配的期望值為 $\eta_{x,k} \times \rho_{x,k}$ ，因此我們假設先驗分配 $\theta_{x,k}$ 的參數為：
 $\eta_{x,k} = 300$ ， $x = a, b, c$ ， $k = 1, \dots, 4$ ， $\rho_{x,1} = 2/300$ ， $\rho_{x,2} = 1.2/300$ ， $\rho_{x,3} = 1/300$ ， $\rho_{x,4} = 0.95/300$ ， $x = a, b, c$ 。並假設先驗分配 $\lambda_{x,k}$ 的參數為： $\alpha_{x,k} = 5$ ， $\beta_{x,k} = 1/450$ ， $x = a, b, c$ ， $k = 1, \dots, 4$ 。

假設 $\lambda_{x,k}$ 、 $\theta_{x,k}$ 獨立，可得聯合後驗分配為：

$$P(\lambda_{x,k}, \theta_{x,k} | \mathbb{t}_x^{(k)}, \boldsymbol{\delta}_x^{(k)}) \propto L(\lambda_{x,k}, \theta_{x,k} | \mathbb{t}_x^{(k)}, \boldsymbol{\delta}_x^{(k)}) \times P(\lambda_{x,k}, \theta_{x,k})$$

$$= \prod_{i=1}^{n_{x,k}} \left\{ \lambda_{x,k} \theta_{x,k} t_{x,i}^{(k)\theta_{x,k}-1} \exp \left[-\lambda_{x,k} t_{x,i}^{(k)\theta_{x,k}} \right] \right\}^{\delta_{x,i}^{(k)}} \times \exp \left[-\lambda_{x,k} t_{x,i}^{(k)\theta_{x,k}} \right]^{1-\delta_{x,i}^{(k)}} \\ \times \frac{\lambda_{x,k}^{\alpha_{x,k}-1} \exp \left\{ -\frac{\lambda_{x,k}}{\beta_{x,k}} \right\}}{\Gamma(\alpha_{x,k}) \beta_{x,k}^{\alpha_{x,k}}} \times \frac{\theta_{x,k}^{\eta_{x,k}-1} \exp \left\{ -\frac{\theta_{x,k}}{\rho_{x,k}} \right\}}{\Gamma(\eta_{x,k}) \rho_{x,k}^{\eta_{x,k}}}$$

由此聯合後驗分配，我可以得到在給定另一個參數下的邊際後驗分配，分別為：

$$P(\lambda_{x,k} | \mathbb{t}_x^{(k)}, \boldsymbol{\delta}_x^{(k)}, \theta_{x,k}) \propto \\ \lambda_{x,k}^{\sum \delta_{x,i}^{(k)}} \times \exp \left[-\lambda_{x,k} \sum t_{x,i}^{(k)\theta_{x,k}} \right] \times \lambda_{x,k}^{\alpha_{x,k}-1} \exp \left\{ -\frac{\lambda_{x,k}}{\beta_{x,k}} \right\} \propto \\ \lambda_{x,k}^{\sum \delta_{x,i}^{(k)} + \alpha_{x,k} - 1} \times \exp \left[-\lambda_{x,k} \left(\sum t_{x,i}^{(k)\theta_{x,k}} + \frac{1}{\beta_{x,k}} \right) \right] \times \frac{[\sum t_{x,i}^{(k)\theta_{x,k}} + \frac{1}{\beta_{x,k}}]^{\sum \delta_{x,i}^{(k)} + \alpha_{x,k}}}{\Gamma(\sum \delta_{x,i}^{(k)} + \alpha_{x,k})}$$

所以 $\lambda_{x,k}$ 的邊際後驗分配為 $\Gamma(\sum \delta_{x,i}^{(k)} + \alpha_{x,k}, \frac{1}{\sum t_{x,i}^{(k)\theta_{x,k}} + \frac{1}{\beta_{x,k}}})$

$$P(\theta_{x,k} | \mathbb{t}_x^{(k)}, \boldsymbol{\delta}_x^{(k)}, \lambda_{x,k}) \propto \\ \theta_{x,k}^{\sum \delta_{x,i}^{(k)} + \eta_{x,k} - 1} \times \prod t_{x,i}^{(k)\delta_{x,i}^{(k)}(\theta_{x,k}-1)} \times \exp \left\{ -\lambda_{x,k} \sum t_{x,i}^{(k)\delta_{x,i}^{(k)}} - \frac{\theta_{x,k}}{\rho_{x,k}} \right\}$$

由上述式子可得知 $\theta_{x,k}$ 的邊際後驗分配沒有相近已知分配的型式。

同樣地，我們將 $T_{x,i}$ 的期望值表示為： $\mu_x = \sum_{k=1}^4 p_{x,k} \mu_{x,k}$ ， $x = a, b, c$ ， $k = 1, \dots, 4$ ，將指派病人到 A 治療的機率定為：

$$p_a = \Pr(\mu_a = \max\{\mu_x, x = a, b, c\} | data) \\ = \Pr(\mu_a > \mu_b > \mu_c | data) + \Pr(\mu_a > \mu_c > \mu_b | data)$$

指派病人到 B 治療的機率為：

$$p_b = \Pr(\mu_b = \max\{\mu_x, x = a, b, c\} | data) \\ p_b = \Pr(\mu_b > \mu_a > \mu_c | data) + \Pr(\mu_b > \mu_c > \mu_a | data)$$

指派病人到 C 治療的機率為：

$$p_c = \Pr(\mu_c = \max\{\mu_x, x = a, b, c\} | data)$$

$$= \Pr(\mu_c > \mu_b > \mu_a | data) + \Pr(\mu_c > \mu_a > \mu_b | data)$$

且當 p_x 足夠大時($p_x > p_U$)時，提前結束試驗，並說明 p_x 為較好的治療， $x = a, b, c$ 。

我們很難直接計算 p_x ， $x=a,b,c$ ，所以我們將對 $\mu_x = \sum_{k=1}^4 p_{x,k} \mu_{x,k}$ ，進行抽樣，再將三組樣本進行比較。對 μ_x 抽樣時，我們可以先分別對 $p_{x,k}$ 和 $\mu_{x,k}$ 抽樣，再將分別的樣本由 $\mu_x = \sum_{k=1}^4 p_{x,k} \mu_{x,k}$ 的定義，組合成 μ_x ； $\mu_{x,k}$ 為韋伯分配的期望值 $= \lambda_{x,k}^{-1/\theta_{x,k}} \times \Gamma(1 + 1/\theta_{x,k})$ ，所以我們要抽取 $\mu_{x,k}$ 的樣本又需要對 $(\lambda_{x,k}, \theta_{x,k})$ 抽樣。我們已經推導出 $(\lambda_{x,k}, \theta_{x,k})$ 給定另一個參數下的邊際後驗分配，我們利用兩個邊際後驗分配抽樣：給定一個起始值 $\theta_{x,k}^0$ ，對 $\lambda_{x,k}$ 抽取一個樣本， $\lambda_{x,k}^0$ ，再將其帶到 $\theta_{x,k}$ 的邊際後驗分配，即給定 $\lambda_{x,k}^0$ 下再對 $\theta_{x,k}$ 再抽取新的樣本 $\theta_{x,k}^1$ ，重複此動作即可得到 $(\lambda_{x,k}, \theta_{x,k})$ 的樣本，再將其組合成為 $\mu_{x,k}$ 的樣本。由於 $\theta_{x,k}$ 的邊際後驗分配沒有相近已知分配的形式，我們將利用 MCMC 方法中的 Metropolis-Hastings 演算法 (Metropolis 等人 1953; Hastings 1970)：

1. 給予一個容易抽取樣本的建議密度函數(proposal function)： $q(\cdot)$
2. 給定起始值： x_0 ，從 $q(x_0)$ 抽取一個候選點(candidate)： x_{cand} ，可得到 $q(x_{cand}|x_0)$ 、 $q(x_0|x_{cand})$ 和 $\pi(x_0)$ 、 $\pi(x_{cand})$ ，其中 $\pi(\cdot)$ 為我們 $\theta_{x,k}$ 的邊際後驗分配
3. 令一個接受函數(acceptance)：

$$\alpha(x_{cand}|x_0) = \min\left(1, \frac{\pi(x_{cand}) q(x_0|x_{cand})}{\pi(x_0) q(x_{cand}|x_0)}\right)$$

4. 從 uniform(0,1)生成一個值： u ，若 $u < \alpha$ ，則接受 x_{cand} 成為下一個樣本點， x_1 ，再從 $q(x_1)$ 抽取新的候選點；若 $u < \alpha$ ，則從 $q(x_0)$ 抽取新的候選點，重複此動作。

我們使用建議函數為常態分配，給定起始值 x_0 後，從 Normal(x_0, v_k)抽取候選點， v_k 為常態分配的變異數，考慮四種短期反應下 $\theta_{x,k}$ 先驗分配假設的不同，令 $v_1 = 0.42$ 、 $v_2 = 0.35$ 、 $v_3 = 0.24$ 、 $v_4 = 0.25$ ，能將更新率控制在 0.2~0.3 左右(Roberts 等人 1997)，由於我們的建議函數為對稱分配，我們的接受函數可以改成 $\alpha(x_{cand}|x_0) = \min\left(1, \frac{\pi(x_{cand})}{\pi(x_0)}\right)$ ；又因為將抽取的樣本帶入後驗分配之後，常發生 overflow 的問題，導致程式無法運算，因此我們將 $\pi(\cdot)$ 取對數 (Chen and Yang 2011)，並建立一個新接受函數：

$$\log(\alpha(x_{cand}|x_0)) = \min(0, \log(\pi(x_{cand})) - \log(\pi(x_0)))$$

若 $\log(u) < \log(\alpha(x_{cand}|x_0))$ ，則接受 x_{cand} 成為下一個樣本點。

四、模擬研究

我們模擬四種不同條件，第一種為三種治療有相同的反應率，即病人為產生抗藥性或死亡、穩定、部分緩解、完全緩解四種短期反應的機率都相同，且三種治療有相同的存活時間分配的參數；第二種為三種治療有相同的反應率，但存活時間分配的參數不同；第三種為三種治療有不同的反應率，但存活時間分配的參數相同；第四種為三種治療有不同的反應率，且存活時間分配的參數也不同。我們首先將存活時間假設為混合指數分配，先驗分配的假設與第二章的文獻相同，在這四種條件下，比較每一人進到試驗後就進行期中分析、依據分析結果改變隨機指派到各治療的機率，和每五人進到試驗後，才進行期中分析、改變隨機指派到各治療的機率。

每一種模擬執行 5000 次，當每一人進到試驗就進行期中分析，選擇 $p_U = 0.83$ 時，我們在第一種模擬時，假設三種治療相同，可以將每一種治療被選為較好的治療的機率(即犯錯的機率)控制在 0.05 左右。且考慮每五個人再進行期中分析、改變指派到各治療的機率，總共 24 次，招募率一樣維持每一週一個人，若 120 人進到試驗後，試驗還未停止，再額外觀察 40 週。選擇 $p_U = 0.81$ ，將第一種模擬(三組治療相等)選到任一組為最好的機率控制在約 0.05。(表二)

由四種模擬的結果可以得知，每一人進到試驗後就進行期中分析，和每五人進到試驗後才進行期中分析，兩者的結果不論是花費的病人個數，或是選到最好的治療的機率都差不多。

表二、存活時間為混合指數分配，比較每一人、五人進到試驗進行期中分析

	反應率	平均存活時間	每一人期中分析		每五人期中分析	
治療	$p_1 p_2 p_3 p_4$	$\mu_1 \mu_2 \mu_3 \mu_4$	平均病人數	被選的機率	平均病人數	被選的機率
1. A、B 和 C 有相同的反應率和存活時間						
A	0.4 0.3 0.1 0.2	4 30 75 110	35(33.33%)	0.0536	36(33.33%)	0.0521
B	0.4 0.3 0.1 0.2	4 30 75 110	35(33.33%)	0.0570	36(33.33%)	0.0500
C	0.4 0.3 0.1 0.2	4 30 75 110	35(33.33%)	0.0516	36(33.33%)	0.0456
共			105(100%)		108(100%)	
2. A、B 和 C 有相同的反應率，存活時間 C 優於 B，B 優於 A						
A	0.4 0.3 0.1 0.2	4 30 75 110	29(27.62%)	0.0346	29(28.16%)	0.0378
B	0.4 0.3 0.1 0.2	6 45 112 165	35(33.33%)	0.0532	34(33.01%)	0.0641
C	0.4 0.3 0.1 0.2	10 60 140 190	41(39.05%)	0.1338	40(38.83%)	0.1611
共			105		103(100%)	
3. A、B 和 C 有相同的存活時間，反應率 C 最好，B 次之，A 最差						
A	0.4 0.3 0.1 0.2	4 30 75 110	7(10.77%)	0.0006	8(11.94%)	0.0005
B	0.2 0.2 0.2 0.4	4 30 75 110	21(32.31%)	0.0332	22(32.84%)	0.0250
C	0.1 0.1 0.3 0.5	4 30 75 110	37(56.92%)	0.6674	37(55.22%)	0.6822
共			65(100%)		67(100%)	
4. 反應率 C 最好，B 次之，A 最差，存活時間 C 優於 B，B 優於 A						
A	0.4 0.3 0.1 0.2	4 30 75 110	5(10%)	0.0002	6(11.11%)	0.0003
B	0.2 0.2 0.2 0.4	6 45 112 165	16(32%)	0.0342	18(33.33%)	0.0258
C	0.1 0.1 0.3 0.5	10 60 140 190	29(58%)	0.8687	30(55.56%)	0.8773
共			50(100%)		54(100%)	

接下來，我們將使用上述每五人進行期中分析的模擬結果，和不考慮短期反應的試驗，即一般只使用存活時間做適應性隨機指派的準則的試驗作比較，看有多組治療下，使用短期反應的訊息，是否也能加速使用存活時間當作主要終點的

臨床試驗。在不考慮短期反應的設計裡，將病人接受三種治療個別的存活時間假設為指數分配且期望值為 μ_a 、 μ_b 、 μ_c ，假設三者的先驗分配為 $IG(\alpha, \beta)$ ， $\alpha = 11$ 、 $\beta = 547.5$ ，即先驗分配的期望值為 $\frac{\beta}{\alpha-1} = 54.75$ ，為上述第一種模擬四種狀態下存活時間參數(4、30、75、110)的平均，且我們選擇 $P_U = 0.77$ ，將第一種模擬(三組治療相等)選到任何一組為最好的機率控制在約 0.05。

在表三中我們可以看到除了在第二種條件下的結果，其他兩種(第三、四條件下)皆是使用短期反應的試驗平均使用較少的病人完成試驗，選到為最佳治療的機率也較高。

表三、存活時間為混合指數分配，每五人進到試驗進行期中分析

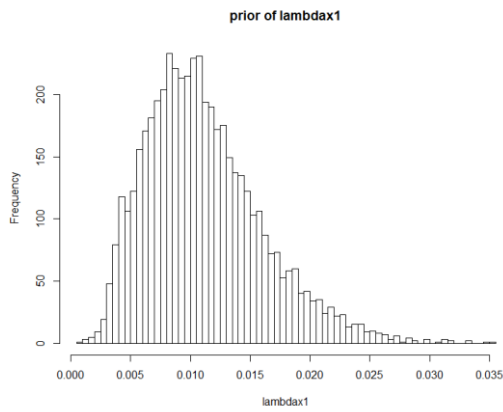
治療	反應率				平均存活時間				考慮短期反應		不考慮短期反應	
	p_1	p_2	p_3	p_4	μ_1	μ_2	μ_3	μ_4	平均病人數	被選的機率	平均病人數	被選的機率
1.A、B 和 C 有相同的反應率和存活時間												
A	0.4	0.3	0.1	0.2	4	30	75	110	36(33.33%)	0.0521	38(33.33%)	0.0522
B	0.4	0.3	0.1	0.2	4	30	75	110	36(33.33%)	0.0500	38(33.33%)	0.0560
C	0.4	0.3	0.1	0.2	4	30	75	110	36(33.33%)	0.0456	38(33.33%)	0.0482
共									108		114(100%)	
2.A、B 和 C 有相同的反應率，存活時間 C 優於 B，B 優於 A												
A	0.4	0.3	0.1	0.2	4	30	75	110	29(28.16%)	0.0378	26(24.76%)	0.0037
B	0.4	0.3	0.1	0.2	6	45	112	165	34(33.01%)	0.0641	35(33.33%)	0.0630
C	0.4	0.3	0.1	0.2	10	60	140	190	40(38.83%)	0.1611	44(41.91%)	0.3361
共									103(100%)		105(100%)	
3.A、B 和 C 有相同的存活時間，反應率 C 最好，B 次之，A 最差												
A	0.4	0.3	0.1	0.2	4	30	75	110	8(11.94%)	0.0005	16(18.83%)	0.0008
B	0.2	0.2	0.2	0.4	4	30	75	110	22(32.84%)	0.0250	30(35.29%)	0.0638
C	0.1	0.1	0.3	0.5	4	30	75	110	37(55.22%)	0.6822	39(45.88%)	0.6704
共									67(100%)		85(100%)	
4.反應率 C 最好，B 次之，A 最差，存活時間 C 優於 B，B 優於 A												
A	0.4	0.3	0.1	0.2	4	30	75	110	6(11.11%)	0.0003	11(15.72%)	0.0003
B	0.2	0.2	0.2	0.4	6	45	112	165	18(33.33%)	0.0258	25(35.71%)	0.0387
C	0.1	0.1	0.3	0.5	10	60	140	190	30(55.56%)	0.8773	34(48.57%)	0.8641
共									54(100%)		70(100%)	

在存活時間假設為韋伯分配的模擬之前，我們生成 10 筆來自韋伯分配的資料，分別在產生抗藥性或死亡($k=1$)和完全緩解($k=4$)的病人的先驗分配下，看此演算法是否能收斂，且將後驗分配和先驗分配比較，並看假設資料分配和先驗分

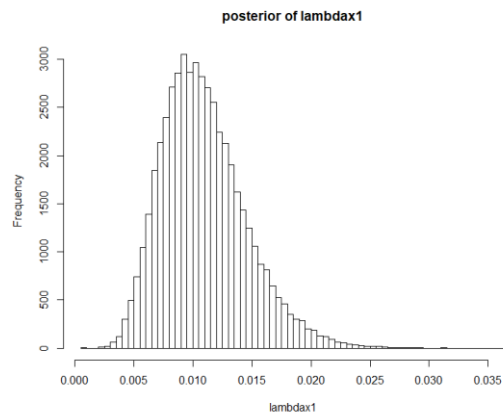
配有落差時是否能在後驗分配中得到修正。

由圖一、三我們可以發現在同一筆資料下，雖然先驗分配假設不同，但在後驗分配可以得到修正，使得兩者後驗分配結果相近。此外，由圖二、四我們可以看到 $\mu_{x,k}$ 的後驗分配沒有很完美地收斂且有一些離群值，由於我們不是要估計 $\mu_{x,k}$ 的期望值，而是要抽取 μ_x 的樣本在和其他的 $\mu_{x'}$ 的樣本進行比較，因此這不會影響到我們在計算 p_x ， $x=a, b, c$ 。

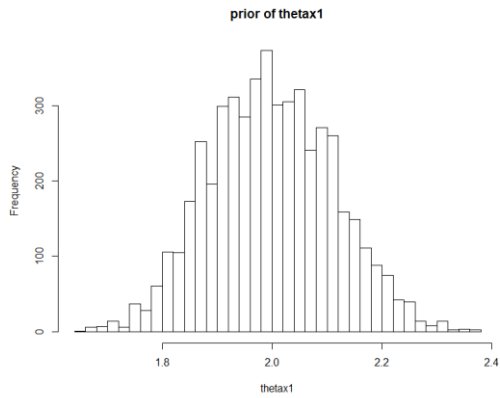
(1) $\lambda_{x,1}$ 先驗分配



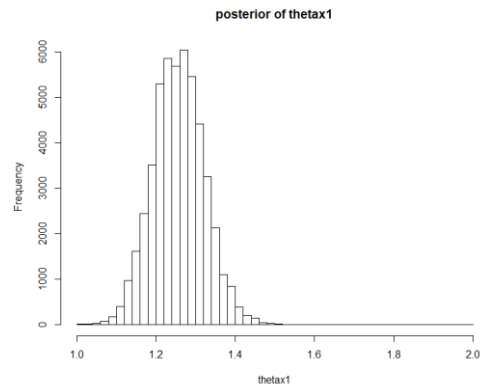
(2) $\lambda_{x,1}$ 後驗分配



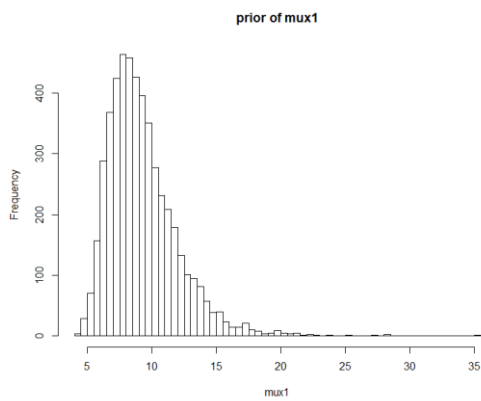
(3) $\theta_{x,1}$ 先驗分配



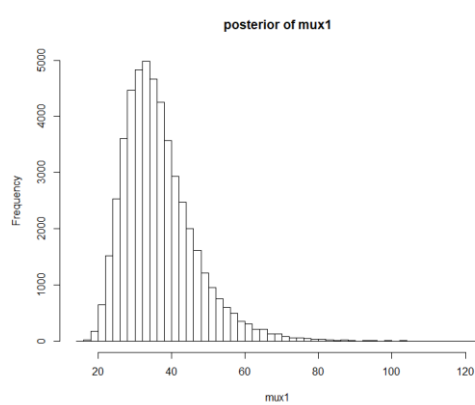
(4) $\theta_{x,1}$ 後驗分配



(5) $\mu_{x,1}$ 先驗分配

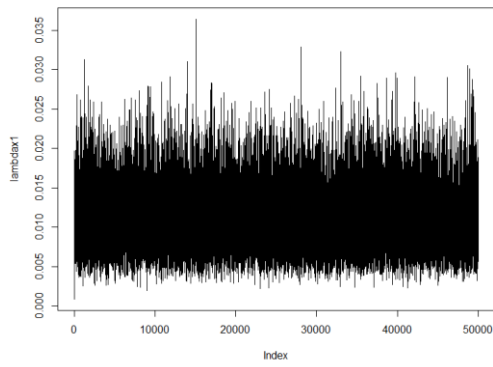


(6) $\mu_{x,1}$ 後驗分配

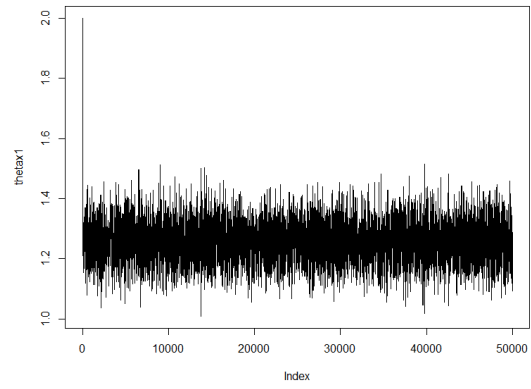


圖一、 $k=1$ 時， $\lambda_{x,k}, \theta_{x,k}$ 先驗期望值 = $1/90, 2$ 且 10 筆資料來自 Weibull(1/120, 1.2)，各參數先驗分配和後驗分配

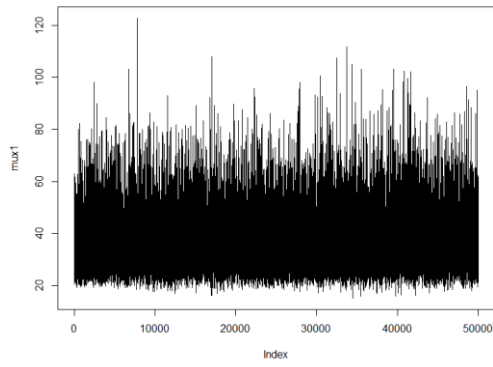
(1) $\lambda_{x,1}$



(2) $\theta_{x,1}$

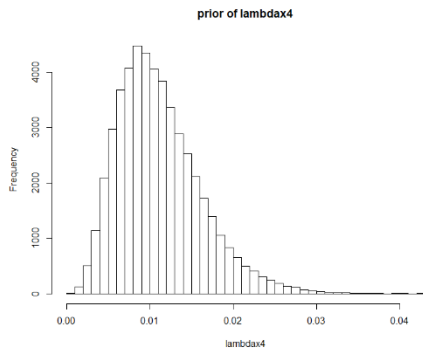


(3) $\mu_{x,1}$

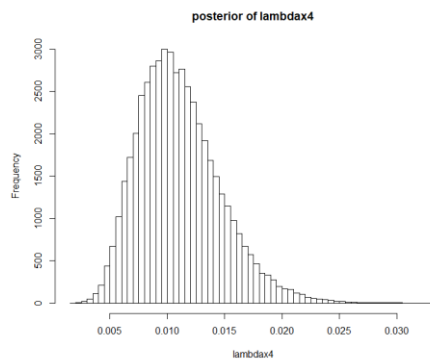


圖二、 $k = 1$ 時，給定 10 筆資料後，各參數的追蹤圖

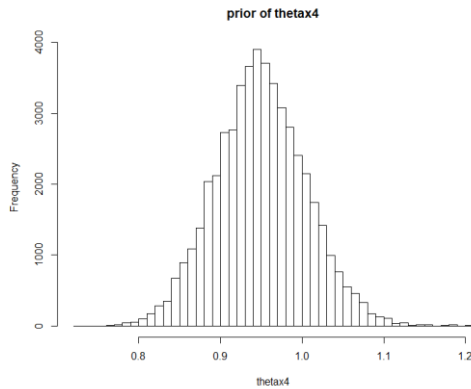
(1) $\lambda_{x,4}$ 先驗分配



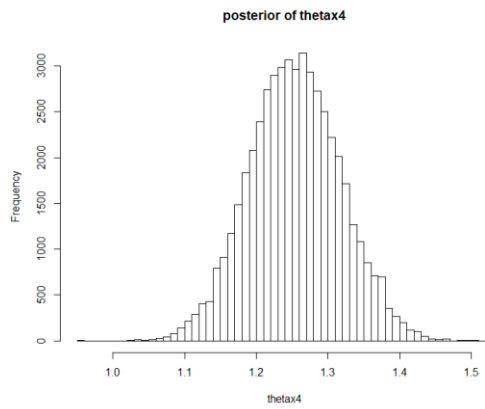
(2) $\lambda_{x,4}$ 後驗分配



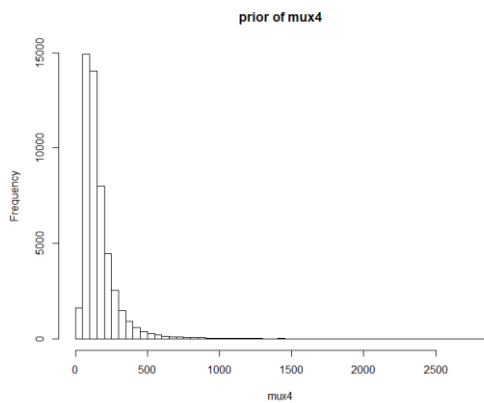
(3) $\theta_{x,4}$ 先驗分配



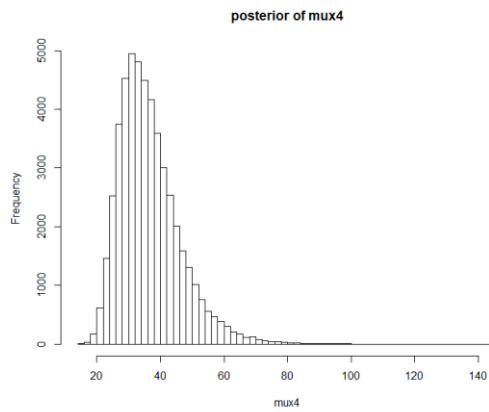
(4) $\theta_{x,4}$ 後驗分配



(4) $\mu_{x,4}$ 先驗分配



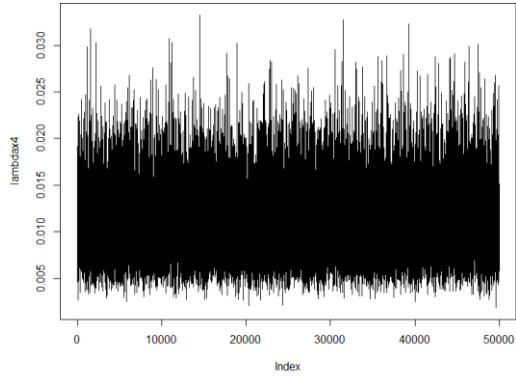
(5) $\mu_{x,4}$ 後驗分配



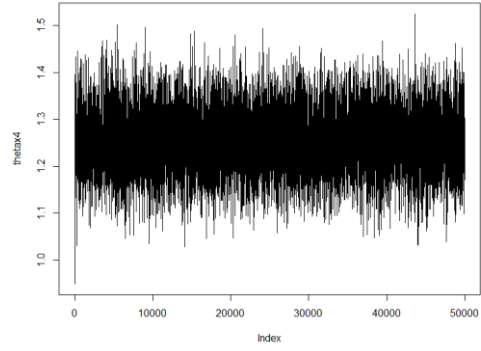
圖三、當 $k=4$ ($\lambda_{x,k}, \theta_{x,k}$ 先驗期望值 = 1/90, 0.95)

假設 10 筆資料來自 Weibull(1/120, 1.2) (和上個例子使用相同筆資料)

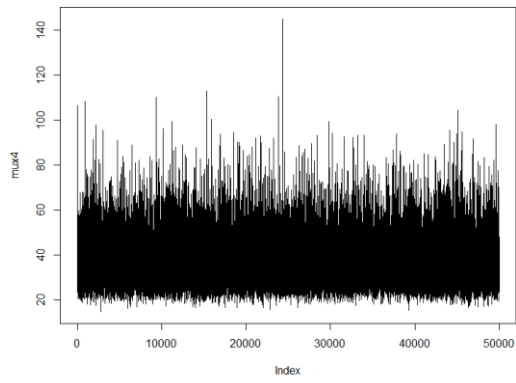
(1) $\lambda_{x,4}$



(2) $\theta_{x,4}$



(3) $\mu_{x,4}$



圖四、 $k = 4$ 時，給定10筆資料，各參數的追蹤圖

使用 MCMC 方法我們可以得到 $\mu_x, x=a,b,c$, 近似的分配, 即可以計算 p_a, p_b, p_c 。我們選擇 $P_U = 0.755$, 可以將第一種條件(三組治療相等)選到任一組為最好的機率控制在約 0.05。在第二種條件時, 我們將控制三組治療的反應率相同, 但存活時間的尺度參數($\lambda_{x,k}$)不同($\lambda_{a,k} = \frac{1}{90}, \lambda_{b,k} = \frac{1}{120}, \lambda_{c,k} = \frac{1}{140}$), 形狀參數($\theta_{x,k}$)相同, 第三種條件時, 我們控制三組治療有不同的反應率, 但存活時間的參數皆相同; 第四種條件時, 我們控制三組治療有不同的反應率, 且存活時間的尺度參數也不同, 形狀參數相同。

我們由表二的結果可以發現使用每五人進到試驗再進行一次期中分析, 和每一人就做一次的試驗所使用的病人數, 和停止試驗的機率都差不多, 於是以下改為韋伯分配時的模擬也使用每五人進到試驗再進行一次期中分析。相同地, 我們將使用短期反應訊息的設計和一般不使用短期反應訊息的設計進行比較。

在不使用短期反應訊息的模擬中, 我們有參數 $\lambda_x, \theta_x, x = a, b, c$, 將 λ_x 的先驗分配假設為 Gamma 分配, 參數 $\alpha_x = 5, \beta_x = 1/450$; θ_x 的先驗分配假設為 Gamma 分配, 參數為 $\eta_x = 300, \rho_x = 1.2875/300$, ρ_x 的假設為考慮短期反應的 $\rho_{x,k}, k = 1, \dots, 4$, 之平均。在 MCMC 方法抽取後驗樣本時, 選擇建議函數為常態分配, 變異數選擇 0.31, 使樣本更新率為 0.2 左右。我們選擇 $P_U = 0.905$, 可以將第一種條件(三組治療相等)選到任一組為最好的機率控制在約 0.05。

在第三、四種條件下, 我們可以看到考慮短期反應訊息的設計, 使用更少的病人, 且選到最佳治療的機率也較高。相同地, 我們發現, 在第二種條件時(治療間反應率相同), 考慮短期反應的設計, 並無法有效地加速試驗。(表四)

表四、存活時間為混合韋伯分配，每五人進到試驗進行期中分析

治療	反應率	四種反應下，存活時間參數	考慮短期反應		不考慮短期反應	
	p_1 p_2 p_3 p_4	θ_1 θ_2 θ_3 θ_4 λ_1 λ_2 λ_3 λ_4	平均病人數	被選的機率	平均病人數	被選的機率
1.A、B 和 C 有相同的反應率和存活時間						
A	0.4 0.3 0.1 0.2	2 1.2 1 0.95 $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$	37(33.64%)	0.0496	37(32.46%)	0.0507
B	0.4 0.3 0.1 0.2	2 1.2 1 0.95 $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$	37(33.64%)	0.0529	38(33.33%)	0.0543
C	0.4 0.3 0.1 0.2	2 1.2 1 0.95 $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$	36(32.72%) 110(100%)	0.0575	39(34.21%) 114(100%)	0.0503
2.A、B 和 C 有相同的反應率，存活時間 C 優於 B，B 優於 A						
A	0.4 0.3 0.1 0.2	2 1.2 1 0.95 $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$	33(30.28%)	0.0323	31(27.43%)	0.0173
B	0.4 0.3 0.1 0.2	2 1.2 1 0.95 $\frac{1}{120}$ $\frac{1}{120}$ $\frac{1}{120}$ $\frac{1}{120}$	37(33.94%)	0.064	39(34.51%)	0.0650
C	0.4 0.3 0.1 0.2	2 1.2 1 0.95 $\frac{1}{140}$ $\frac{1}{140}$ $\frac{1}{140}$ $\frac{1}{140}$	39(35.78%) 109(100%)	0.0924	43(38.06%) 113(100%)	0.1306
3.A、B 和 C 有相同的存活時間，反應率 C 最好，B 次之，A 最差						
A	0.4 0.3 0.1 0.2	2 1.2 1 0.95 $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$	18(20.45%)	0.0013	23(21.70%)	0.0027
B	0.2 0.2 0.2 0.4	2 1.2 1 0.95 $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$ $\frac{1}{90}$	31(35.23%)	0.0677	37(34.90%)	0.0687

C	0.1 0.1 0.3 0.5	2 1.2 1 0.95 $\frac{1}{90} \frac{1}{90} \frac{1}{90} \frac{1}{90}$	39(44.32%) 88(100%)	0.4207	46(43.40%) 106(100%)	0.3100
4.反應率 C 最好，B 次之，A 最差，存活時間 C 優於 B，B 優於 A						
A	0.4 0.3 0.1 0.2	2 1.2 1 0.95 $\frac{1}{90} \frac{1}{90} \frac{1}{90} \frac{1}{90}$	14(17.28%)	0.0013	18(17.65%)	0.0017
B	0.2 0.2 0.2 0.4	2 1.2 1 0.95 $\frac{1}{120} \frac{1}{120} \frac{1}{120} \frac{1}{120}$	29(35.80%)	0.0650	36(35.29%)	0.0650
C	0.1 0.1 0.3 0.5	2 1.2 1 0.95 $\frac{1}{140} \frac{1}{140} \frac{1}{140} \frac{1}{140}$	38(46.92%) 81(100%)	0.5500	48(47.06%) 102(100%)	0.4147

表五、5000 次模擬的後驗期望值的平均、百分位數

	考慮短期反應的設計			不考慮短期反應的設計		
	平均	2.5 百分位數	97.5 百分位數	平均	2.5 百分位數	97.5 百分位數
1.A、B 和 C 有相同的反應率和存活時間						
A	41.74705	31.57922	77.45577	25.34448	16.15812	36.71803
B	41.79852	31.37095	73.84238	25.62223	15.86766	36.95708
C	41.96657	31.40397	77.14116	25.54407	16.24830	36.41792
2.A、B 和 C 有相同的反應率，存活時間 C 優於 B，B 優於 A						
A	42.82906	32.19153	66.71503	26.45117	16.42410	38.30879
B	45.69974	33.65254	79.95770	29.20876	17.17451	42.22532
C	47.16077	34.19827	82.20658	30.93308	18.26473	45.22091
3.A、B 和 C 有相同的存活時間，反應率 C 最好，B 次之，A 最差						
A	48.45912	34.68900	66.7106	27.89638	17.11118	41.45424
B	58.56267	39.70437	98.65897	34.40368	19.24491	54.14661
C	76.16325	41.16624	144.2020	40.8694	21.16875	65.06252
4.反應率 C 最好，B 次之，A 最差，存活時間 C 優於 B，B 優於 A						
A	51.03337	36.24932	71.36858	29.13448	17.48630	44.89513
B	65.20808	43.92969	110.3935	39.4457	20.66136	63.52197
C	89.11768	47.68427	151.4555	49.36726	23.60047	77.34483

表六、混合韋伯分配在不同參數下的期望值

1. A、B 和 C 有相同的反應率和存活時間
$\mu_a = \mu_b = \mu_c = 47.70278$
2. A、B 和 C 有相同的反應率，存活時間 C 優於 B，B 優於 A
$\mu_a = 47.70278, \mu_b = 62.73107, \mu_c = 72.6993$
3. A、B 和 C 有相同的存活時間，反應率 C 最好，B 次之，A 最差
$\mu_a = 47.70278, \mu_b = 74.36453, \mu_c = 90.19583$
4. 反應率 C 最好，B 次之，A 最差，存活時間 C 優於 B，B 優於 A
$\mu_a = 47.70278, \mu_b = 99.30705, \mu_c = 141.7458$

由表五、六可以發現，使用短期反應的設計，所得到的 $\mu_x, x = a, b, c$ ，的後驗期望值，比不使用短期反應之設計的結果更相近於各條件下混合韋伯分配的期望值。

五、結論與討論

有很多新藥試驗為了要趕快通過第三期試驗，在第二期只使用短期反應跟非常少的病人跟歷史性對照組比較，以快速地進到第三期試驗，甚至不做第二期試驗，直接進到第三期試驗，導致現今第三期試驗的失敗率非常高；為了要有效地利用資源，又要謹慎地選擇對存活時間有益的治療方法，進到第三期試驗進行確認，在第二期臨床試驗使用存活時間當作主要終點是必須的。

我們推廣 Huang 等人(2009)提出的方法，存活時間為混合韋伯分配時，且將兩組治療推廣至多組，使用短期反應訊息也能加速適應性隨機指派臨床試驗，要注意的是，在第二種條件時(治療間的反應率相同)，考慮短期反應的設計，無法有效地加速試驗，此外，我們也考慮 Huang 等人(2009)的方法為每一人進到試驗就進行期中分析，只有一個樣本可能造成前期病人指派到較差的治療機率變高，且花費較多人力，由表二我們發現使用每五人進到試驗才進行期中分析，結果是差不多的，但我們可以減少人力花費，更有效地完成試驗。

只考慮反應率進行隨機指派，可能造成隨機性不夠，導致實驗組跟對照組的預後因子的分布不均，產生評估偏差，例如：急性骨隨性白血病在臨床上的預後因子包括：年紀、病人活動度、是否有接觸過放射線或細胞毒性藥物等等。雖然我們提出的方法在後期有更多的存活時間訊息可以參考，但可能在前期就已造成組間共變數分布不均，使得結果產生誤差，未來，可以加入調整共變異數的方法，來使組間預後因子平均分布。

參考文獻

- 葉嘉新、林志六編，林志六等著(2008)。《新藥開發與臨床試驗》。台灣：秀威代理
- Agarwal, S. K., Mangal, N., Menon, R. M., Freise, K. J. and Salem, A. H. (2017). Response Rates as Predictors of Overall Survival: A Meta-Analysis of Acute Myeloid Leukemia Trials. *J Cancer* 8(9): 1562-1567
- Aronson, J. K.. (2005). Biomarkers and surrogate endpoints. *Br J Clin Pharmacol*, 59(5): 491-494
- Chen, C. S. and Yang, H. D. (2011) A joint modeling approach for spatial earthquake risk variations. *Journal of Applied Statistics*, 38(8): 1733-1741
- Hastings, W. K. (1970) Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57(1): 97-109.
- Huang, X., Ning, J., Li, Y., Estey, E., Issa, Jean-Pierre and Berry, A. D. (2009). Using short-term response information to facilitate adaptive randomization for survival clinical trials. *Statistics in Medicine*, 28(12): 1680-1689
- Korn, E. L. and Freidlin, B. (2011). Outcome-Adaptive Randomization: Is It Useful? *J Clin Oncol*. 29(6): 771–776.
- Lin, J., Lin, L. A. and Sankoh, S. (2016). A General Overview of Adaptive Randomization Design for Clinical Trials. *Journal of Biometrics & Biostatistics*, 7(2):294
- Nowacki, A. S., Zhao, W., Palesch, Y. Y. (2017) A surrogate-primary replacement algorithm for response-adaptive randomization in stroke clinical trials. *Stat Methods Med Res*. 26(3):1078-1092.
- Ning, J. and Huang, X. (2010). Response-adaptive Randomization for Clinical Trials with Adjustment for Covariate Imbalance. *Stat Med*. 29(17): 1761-1768.

Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, and Augusta H. Teller. (1953) Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* 21: 1087.

Pallmann, P., Bedding, A. W., Choodari-Oskooei, B. et al. (2018) Adaptive designs in clinical trials: why use them, and how to run and report them *BMC Med.* 16: 29.

Roberts, G. O., Gelman, A., and Gilks, W. R. (1997) weak convergence and optimal scaling of random walk metropolis algorithms. *The Annals of Applied probability*, 7: 110-120

Stallard, N., Whitehead, J., Todd, S., and Anne W. (2001) Stopping rules for phase II studies. *Br J Clin Pharmacol*, 51(6): 523–529.

Yildirim, I. (2012). Bayesian Inference: Metropolis-Hastings Sampling.