

東 海 大 學

工業工程與經營資訊學系

碩士論文

應用資料探勘於肝癌治療傳統栓塞
及微球栓塞療效評估之研究

研 究 生：郭承定
指 導 教 授：洪堯勳 教授

中 華 民 國 一 〇 七 年 七 月

**Applying Data Mining to Efficacy Assessment of
Conventional Transcatheter Arterial Chemoembolization
and Drug-eluting Micro-bead Embolization for Hepatoma
Treatment**

By
Cheng-Ding Kuo

Advisor : Prof. Jau-Shin Hon

A Thesis
Submitted to the Institute of Industrial Engineering and
Enterprise Information at Tunghai University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in
Industrial Engineering and Enterprise Information

July 2018

Taichung , Taiwan

應用資料探勘於肝癌治療傳統栓塞及微球栓塞療效評估之研究

學生：郭承定

指導教授：洪堯勳 教授

東海大學工業工程與經營資訊學系

摘 要

肝癌是亞洲地區重要面對的議題，台灣肝癌罹患人數有逐年下降的趨勢，但仍舊維持癌症之首，不易被察覺的特質導致發現時往往已是中晚末期。目前療法之中以肝癌動脈栓塞較為平和，且適用較廣，一直是不斷研究改良的重點標的，栓塞目前應用於臨床醫學有兩種療法，分別為傳統栓塞及微球栓塞，有部分研究認為兩種療法各有所適用之情況，故本研究期望在現今具備較大量完整資料的背景下採用資料採礦技術對肝癌栓塞兩種療法做更進一步的分類探討。

本研究以台中榮總接受肝動脈栓塞之肝癌病患資料為研究對象。以卡方檢定比較療法療效的差異，邏輯斯迴歸計算勝算比比較差異程度，再分別對兩種療法以決策樹 C4.5 更進一步分類各自適用領域，最後以卡方及邏輯斯迴歸驗證分類結果。

研究結果顯示，卡方檢定顯示傳統療法顯著優於微球療法，平衡兩方法比較基礎，僅保留癌症分級 A、B 級患者再次比較，卡方結果差異不大，以計算勝算比得知傳統療法有效的機率是微球療法的 2.57 倍。進一步以決策樹 C4.5 做為分類器，結果發現，傳統療法受限於腫瘤大小，小於等於 9.3 公分有效，大於則無效，均值檢定有療效的腫瘤大小平均顯著較小，卡方檢定腫瘤小於等於 9.3 公分顯著有效於大於 9.3 公分的，勝算比顯示腫瘤小治療有效機率是腫瘤大的 8.475 倍。微球療法受制於腫瘤數量，多顆無效，單顆腫瘤 100-300(μm)有效，卡方檢定單顆腫瘤顯著有效於多顆腫瘤，勝算比發現使用微球療法治療單顆腫瘤成功機率會比多顆腫瘤高 3.39 倍的。單顆腫瘤中，卡方檢定 100-300(μm)顯著有效於 300-500(μm)，勝算比結果為 100-300(μm)治療成功機率是 300-500(μm)的 3.412 倍。以此研究結果提供肝癌栓塞選擇療法上的一項參考。

關鍵字詞：肝癌、肝動脈栓塞、栓塞預後、資料採礦、決策樹、邏輯斯迴歸

Applying Data Mining to Efficacy Assessment of Conventional Transcatheter Arterial Chemoembolization and Drug-eluting Micro-bead Embolization for Hepatoma Treatment

Student : Cheng-Ding Kuo

Advisor : Prof. Jau-Shin Hon

Department of Industrial Engineering and Enterprise Information
Tunghai University

ABSTRACT

Liver cancer is an important issue of cancer death in the Asian region. The number of liver cancer patients in Taiwan has been declining year by year, but it still maintains the top of cancer problem. The symptoms of liver cancer are not easy to detect, so it may be found in intermediate staged or late staged cancer. At present, trans-arterial chemoembolization (TACE) is relatively calm and widely used, and has become a key target for continuous research and improvement. Clinical medicine currently uses two embolization therapies, traditional embolization and drug eluting microsphere embolization. Some studies have shown that each of the two therapies has their own suitable situation, so our study hopes to use the current large amount of complete data, using data mining technology to further classify the two treatment methods for liver cancer embolization.

Our study collected data of liver cancer patients at Taichung Veterans General Hospital. Using chi-square test to determine the difference in the efficacy of the comparison therapy. Using logistic regression to calculate the odds ratio. The two treatments were further classified into the applicable fields by decision tree C4.5. Finally, the classification results were verified by chi-square and logistic regression.

The results of our study showed that the chi-square test showed that traditional therapy was significantly better than microsphere therapy. To balance the basis of the two methods, the patients who retained the cancer grades A and B were compared again. The chi-square results were not significantly different from the original ones. The effective rate of calculating traditional treatments was 2.57 times than microsphere therapy. Furthermore, using decision tree C4.5 as a classifier, it was found that conventional treatment is limited by tumor size,

tumors less than or equal to 9.3 cm were effective, but if they were larger than 9.3 cm, they were not effective. The result of the T test, the tumor size with a curative effect was significantly smaller on average. The chi-square test results showed that the tumor was less than or equal to 9.3 cm was significantly effective at more than 9.3 cm. The odds ratio shows that the effective rate of small tumor treatment was 8.475 times larger than that of tumor. Microsphere therapy was subject to the number of tumors, multiple ineffective, single tumor 100-300 (μm) was effective. The chi-square test for a single tumor was significantly more effective than multiple tumors. The odds ratio found that the use of microsphere therapy, the probability of success in treating a single tumor was 3.39 times higher than that of multiple tumors. In a single tumor, the chi-square test bead size 100-300 (μm) was significantly more effective than size of 300-500 (μm). The odds ratio showed the success rate of bead size 100-300 (μm) treatment was 3.412 times than the size of 300-500 (μm). This study provides a reference for the selection of liver cancer embolization therapy.

Keywords: Liver Cancer, Hepatic Artery Embolization, Embolization Prognosis, Data Mining, Decision Tree, Logistic Regression.

誌謝

本研究能順利進行並完成，首先要感謝指導教授 洪堯勳 博士，兩年來不遺餘力的指導，以各種方式培養我在邏輯思考上更加有效率也更加準確，同時也在論文架構上給予極大的幫忙，真的於碩士班這兩年學習到非常多。執行研究的過程中，非常感謝台中榮民總醫院放射線部 張碧倚 主治醫師，在排除萬難撥空協助我突破研究上的問題，同時給予醫學領域上專業的指導，讓我能更了解醫療專業領域以輔助加深研究的深度。也非常感謝臺中榮總臨床資訊研究發展中心(Clinical Informatics Research & Development Center of Taichung Veterans General Hospital)所提供完整的資料，使本研究能順利完成。

感謝碩士班已畢業的宗佑學長給予我許多的幫助與建議，在我遇到瓶頸時給予方向的建議。也感謝研究室的同學富群、至翔在研究期間給我的幫助與鼓勵。

感謝清大工工的朋友韋智、在大葉資訊工程的珈誠於研究方法上都給我許多的幫助以及經驗交流。

最後要感謝我最親愛的家人，支持我繼續深耕工工專業，並時常關心我的學習情況，如果沒有家人們的支持我將無法完成碩士學業。

在此奉上最至誠的感謝，謝謝所有幫助過我的人，謝謝你們！

郭承定 謹誌於

東海大學工業工程與經營資訊研究所

民國 107 年 7 月

目錄

摘要.....	i
ABSTRACT.....	ii
誌謝.....	iv
目錄.....	v
表目錄.....	vii
圖目錄.....	viii
第一章 緒論	1
1.1 研究背景與動機	1
1.2 研究目的	4
1.3 研究方法	4
1.4 研究範圍與限制	5
1.5 研究架構	6
第二章 文獻探討	8
2.1 肝癌治療	8
2.2 藥物比較	12
2.3 資料探勘	17
第三章 研究材料與方法	19
3.1 研究材料	19
3.2 資料探勘技術	22
3.3 評估模式	28
第四章 實證分析	33
4.1 傳統療法及微球療法療效比較	34
4.2 傳統療法限制與適用細分	40
4.3 微球療法限制與適用細分	44
第五章 結論	52
5.1 研究發現與討論	52

5.2 研究結論及未來研究建議	53
參考文獻.....	56

表目錄

表 2.1 歷年資料探勘技術應用於癌症預後文獻整理.....	18
表 3.1 資料項目表.....	21
表 3.2 邏輯斯迴歸演算邏輯表.....	24
表 3.3 決策樹種類比較表.....	25
表 3.4 決策樹演算邏輯表.....	27
表 3.5 混淆矩陣表.....	30

圖目錄

圖 1.1 世界男性最高癌發分佈圖.....	1
圖 1.2 民國 105 年台灣 10 大死因圖.....	2
圖 1.3 民國 85 至民國 105 年台灣各種癌症趨勢圖.....	2
圖 1.4 研究流程圖.....	7
圖 2.1 肝腫瘤診斷流程圖.....	8
圖 2.2 Child-Pugh 分級計分制圖.....	9
圖 2.3 巴塞隆那臨床肝癌分類(Barcelona Clinic Liver Cancer, BCLC).....	10
圖 2.4 肝癌肝動脈栓塞示意圖.....	11
圖 2.5 多柔比心藥物動力學.....	13
圖 2.6 肝癌動脈栓塞療效與重大不良反應比較圖.....	15
圖 2.7 肝功能曲線圖及新舊療法副作用比較圖.....	16
圖 2.8 資料採礦技術流程圖.....	17
圖 3.1 研究架構圖.....	20
圖 3.2 邏輯斯迴歸方程式圖示.....	24
圖 3.3 十摺交叉驗證運行圖.....	29
圖 3.4 概似比對應實驗通過機率圖.....	31
圖 3.5 ROC 曲線與 PRC 曲線比較圖.....	32
圖 4.1 卡方分配產出.....	34
圖 4.2 去除 C、D 級病患卡方分配產出.....	35
圖 4.3 邏輯斯迴歸產出結果圖.....	37
圖 4.4 邏輯斯迴歸模型適用性圖.....	38
圖 4.5 去除 C、D 級病患邏輯斯迴歸模型.....	39
圖 4.6 傳統療法決策樹分類.....	40
圖 4.7 傳統療法決策樹分類模型結果.....	40
圖 4.8 cTACE 有效無效及腫瘤大小 T 檢定產出.....	41
圖 4.9 cTACE 以 9.3 公分劃分卡方檢定產出.....	42

圖 4.10 cTACE 以 9.3 公分劃分邏輯斯迴歸結果	43
圖 4.11 微球療法決策樹分類.....	44
圖 4.12 微球療法決策樹模型情形.....	44
圖 4.13 微球療法去除混球後決策樹分類.....	45
圖 4.14 微球療法去除混球後決策樹模型情形.....	46
圖 4.15 微球療法去除混球去除 500-700 μ m 後決策樹分類.....	46
圖 4.16 微球療法去除混球去除 500-700 μ m 後決策樹模型情形.....	47
圖 4.17 微球療法去除混球去除 500-700 μ m 後腫瘤數量卡方.....	48
圖 4.18 微球療法去除混球去除 500-700 μ m 後腫瘤數量邏輯斯迴歸.....	49
圖 4.19 微球療法去除混球去除 500-700 μ m 後單腫瘤微球大小卡方.....	50
圖 4.20 微球療法去除混球去除 500-700 μ m 後單腫瘤微球大小邏輯斯迴歸.....	51

第一章 緒論

1.1 研究背景與動機

依據 2014 世界衛生組織(WHO)全球癌症報告，提出全球最高癌發分佈圖，圖 1.1，亞洲男性最高癌發的癌症為肺癌和肝癌，而台灣亦是肺癌和肝癌高癌發地區。衛生福利部於民國 105 年台灣 10 大死因統計中，如圖 1.2 所示，癌症為台灣最嚴重的死因，而從衛生福利部長期對各種癌症的監測，由圖 1.3 可以看出來肺癌及肝癌一直保持著癌症中最多人罹患的癌症，也吻合了世界衛生組織所提出最高癌發分佈圖亞洲的現況。雖然隨著時代的進步、安定，營養更能滿足以及衛生福利的推行，肝癌有下降的趨勢，但肝癌仍然與肺癌名列前茅。

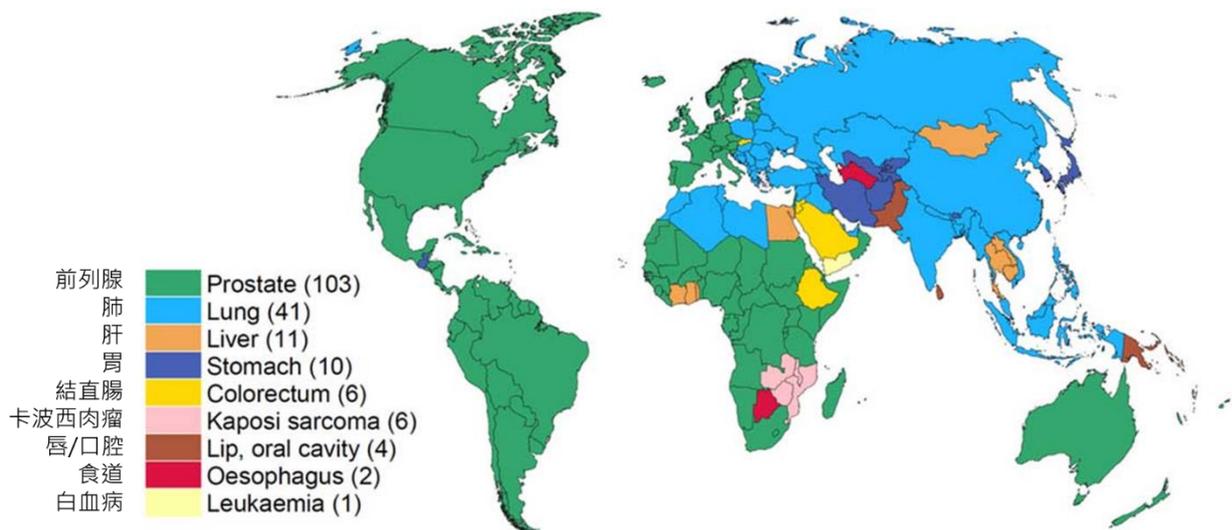


圖 1.1 世界男性最高癌發分佈圖

資料來源:世界衛生組織(WHO)2014 全球癌症報告

105年台灣10大死因

所有死亡原因	標準化死亡率 (每十萬人口)	
	順位	較上年 增減%
所有死亡原因	439.4	1.8
惡性腫瘤	1	126.8 -0.9
心臟疾病(高血壓性疾病除外)	2	50.3 4.7
肺炎	4	26.9 9.3
腦血管疾病	3	28.6 2.4
糖尿病	5	24.5 0.8
事故傷害	6	23.1 1.2
慢性下呼吸道疾病	7	15.1 3.5
高血壓性疾病	8	13.5 2.3
腎炎、腎病症候群及腎病變	10	12.4 5.4
慢性肝病及肝硬化	9	13.4 -1.8

$$\text{標準化死亡率} = \frac{\sum(\text{年齡別死亡率} \times \text{標準組年齡別人口數})}{\text{標準組總人口數}}$$

圖 1.2 民國 105 年台灣 10 大死因圖

資料來源: 衛生福利部統計處

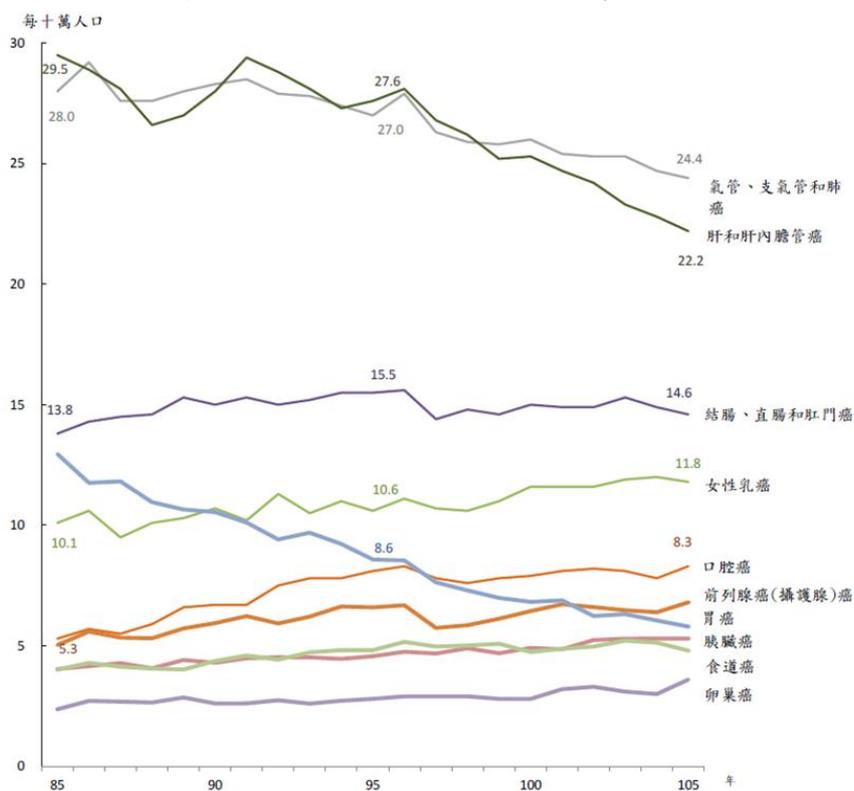


圖 1.3 民國 85 至民國 105 年台灣各種癌症趨勢圖

資料來源: 衛生福利部統計處

絕大部分的民眾發現身體有問題的時候已經是癌症的中期或末期了，而大部分的癌症在初期發現其存活機率是很高的(李宗佑，2015。)因此癌症的預防及提早診斷上對於尚未發病的人是一項重要的課題；但對於發病的患者，挽救治療上更為重要，故治療與預後監測的議題更是我們不容忽視的重要問題。

其中又以肝癌最為麻煩，因為肝癌早期無痛感的關係是一種較難發現癌症，通常發現時已經是中期、晚期，而中期、晚期的治療方式依據劉鐘軒、蔡正中、陳海雄於2013提出的「肝癌的診斷及治療最新發展」提到巴塞隆那臨床肝癌分類(Barcelona Clinic Liver Cancer, BCLC)中肝癌中期以肝動脈栓塞治療為主，晚期則是化療，而栓塞治療從Llovet JM, Bruix J等人於2003以隨機對照試驗的薈萃分析證實，肝動脈化療栓塞術(TACE)有效成為無法手術肝癌患者於肝臟分離大或多結節非侵襲性肝癌細胞(HCC)並保留肝功能的標準護理治療方式，除此之外，還可以阻止癌細胞藉由血液流動的方式擴散至其他器官(Sutter health CPMC，2010。)也因此肝癌肝動脈栓塞治療方式的改善或提升也將會是一項重要的課題。

目前肝動脈栓塞治療方式主要有兩種，新舊療法，差異在於攜藥載體不同，舊療法為傳統經導管動脈化學藥物栓塞治療(Conventional Transcatheter arterial chemoembolization, cTACE)，油水包覆(碘油包覆著化療藥物)的概念將藥物帶入腫瘤血管進行栓塞，栓塞藥物為碘化油(lipiodol)，是一種含有碘的油性物質，除了有栓塞效果之外，也能攜帶化療藥物，經由微小血管進入腫瘤內。然而，lipiodol並不是永久性的栓塞物質，有一部分可能會隨時間被血流帶離肝臟。新療法為微球載藥動脈栓塞(Drug-eluting micro-bead embolization, DC Bead)則是以微球載藥，微珠藥物傳遞栓塞系統是高分子聚合物所製成的微小球體，這些微小球體除了會永久的將血管栓塞住，而且這些小球體可以長期留在微小肝動脈的末端，以穩定的速率，釋放出化學治療藥物，並長期維持化學治療藥物的局部濃度，作用時間可持續一個月以上，延長有效的治療時間而取得較佳的治療效果。(癌症防治季刊第九期 肝癌，2010，陳嘉邦。)

Song MJ 等人於2011、2012、2016研究分別都顯示，與cTACE相比，使用DC珠的TACE顯示出更好的治療響應和延遲的腫瘤進展。但肝臟治

療相關毒性無顯著差異，於文獻探討中諸多文獻採用幾個個案統計方式得出結果也支持 DC Bead 有較優的表現。從 Gao S 等人於 2013 年的研究中提出目前的證據表明，DC Bead 能夠完成與 cTACE 相同的腫瘤反應，雖然此分析綜合了 DC Bead 與 cTACE 相比的臨床療效資料，但仍需進一步大規模的隨機對照研究。

但 Do Seon Song 等人在肝癌栓塞比較上依據臨床實例提出結論，表示針對常規 TACE 難治的 HCC，使用 DC 珠的 TACE 是較有效的，且這種治療有較好的反應，尤其是當腫瘤很小並且在血管造顯影延遲或增強時；此結論指出常規 TACE 仍有適用空間並非為全能被取代。而 Lammer J 2010 年的研究顯示副作用較輕微，療效與傳統的碘油栓塞相近，且依據卡方檢定產出的結果發現 DC bead 僅在於某些條件下具有顯著優勢。由前人文獻能推論在肝癌某些情況下 DC bead 可能有較佳的表現，也就是目前 DC Bead 有證據顯示能部份取代 cTACE 的能力，但還無法證實是否在療效上能完全追上 cTACE，cTACE 是否還具有在某些情形下的優勢，並且這些情況是什麼，而進一步新療法具有什麼進步空間，這些成為本研究想探究的問題，也因此衍伸出本研究的研究目的。

1.2 研究目的

- 1、利用病歷資料及追蹤預後資料驗證肝癌栓塞新舊療法對病人治療成效是否有所不同。
- 2、找出影響肝癌栓塞新舊療法預後各自的限制，輔助醫生在療法選擇時的參考。
- 3、進一步探討限制下各自療效有效成功機率的差異，以供醫療風險上更進一步參考。

1.3 研究方法

前人研究多以統計方式對療效進行評估，在統計上較容易出現大樹法則並較難理出資料中規則，且因前人研究較無法蒐集較大量的病歷、治療資料，本研究配合完善資料保存之「醫學中心」級的公立醫院以取得更完整及大量的病歷、治療資料，故本研究將採用資料探勘技術期望理出具有價值的規則以提供醫師參考。

資料探勘(Data Mining)的目的是從大量的資料中找出關聯、規則，並經由轉化解讀成為可實際使用的知識，本研究參考前人使用於醫療資料的研究，先以卡方分配檢視蒐集的資料基本療效概況，進一步採用資料探勘技術中的邏輯斯迴歸方法對新舊療法進行有效程度上差異的比較，本研究期望能更進一步探討兩種療法各自存在的適用情況及限制，因此將進一步以資料探勘中決策樹方法做為分類器更進一步對資料進行劃分，而劃分後的結果再以統計卡方分配方式進行驗證支持，最後以邏輯斯迴歸再次比較分類後療效上的差異程度，以以上方式試圖找出新舊療法在肝癌栓塞治療預後成果上所受到的限制以及適用領域，提供醫療選擇療法上一項參考。

1.4 研究範圍與限制

1.4.1 研究範圍

本研究蒐集臺中榮總臨床資訊研發中心(Clinical Informatics Research & Development Center of Taichung Veterans General Hospital)，所提供民國 99 年至民國 106 年期間的接受肝動脈栓塞之肝癌病患資料進行研究分析。

本研究主要蒐集新療法(DC Bead TACE / 鉍 90)微球載體攜藥之療法，以及使用舊療法(cTACE)油水載體攜藥之療法，接受兩種療法任一療法進行治療之肝癌患者。

1.4.2 研究限制

1、因本研究所蒐集之資料僅取得自臺中榮民總醫院，故所產出之結果並無法代表所有接受肝動脈栓塞之肝癌患者皆適用所產出之結果，本研究所產出規則僅供未來醫師於治療上參考使用，以及所比較之模型適用程度可供後續肝動脈栓塞預後資料研究者參考。

2、因本研究與醫師討論及礙於資料取得有限，所選用之屬性資料未將肝動脈栓塞接受治療之患者生活飲食習慣、完整病例等等全數匯入，故具有更顯著之因子存在的可能性，以待後續研究者更完整完備進行探討。

1.5 研究架構

本研究將分作以下五個章節進行論，可參考圖 1.4 以輔助了解本研究之研究架構：

第一章: 緒論

本研究首先簡介目前肝癌治療的概況及背景，並從背景中衍生出研究的動機及目的，再進一步概述所需使用的方法及研究的範圍與限制。

第二章: 文獻探討

本論文將對三大類相關文獻進行探討，首先，是肝癌的部分，先針對肝癌現今檢查的流程，進一步如何劃分級數，再進一步探討現今主要治療的方法，瞭解現今肝癌治療方法後對其中肝癌動脈栓塞進行探討，以了解整個肝癌治療從發現到以肝癌動脈栓塞治療的過程。

了解完肝癌及該動脈栓塞的基本情形後，本研究藥進行新療法及舊療法的比較，以前人的研究經驗，主要分做三部分進行探討，首先是新舊療效的差異及原理，其次是藥物動力學簡介及前人應用藥物動力學於比較肝動脈栓塞新舊療法之文獻，最後是療效評估目前應用於癌症治療相關文獻。

最後，因應前面研究方法所提及本研究採用的資料探勘技術，文獻探討將簡介資料探勘技術以及尋找前人以資料探勘技術應用於醫療、癌症或預後追蹤的相關研究，進行整理與分析，加以改良或應用，以支持本研究。

第三章: 研究材料與方法

本章節將說明資料如何蒐集及選擇因子作詳細的說明，並對所使用資料探勘之技術的原理及數學模型加以論述及探討。

第四章: 實證分析

將資料餵入選擇後之資料探勘之模型進行演算，並探討其產出之結果，進一步找出有價值的法則以供醫師治療參考。

第五章: 結論

以實證分析中整理出結果，提出結論、建議以及可供未來研究的方向及議題。

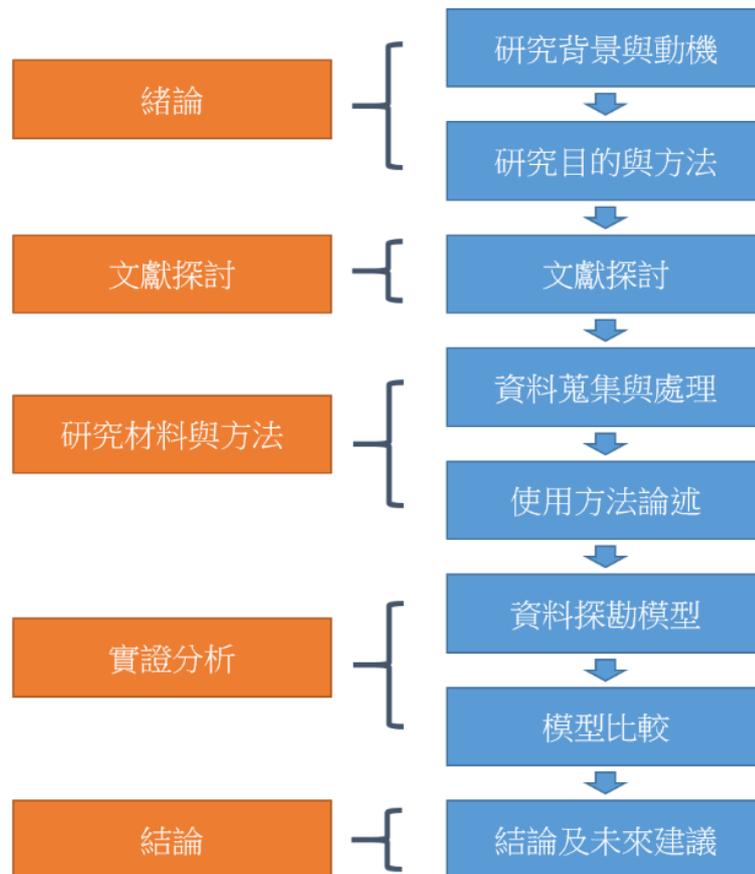


圖 1.4 研究流程圖

資料來源:本研究整理

第二章 文獻探討

2.1 肝癌治療

2.1.1 肝癌檢查流程

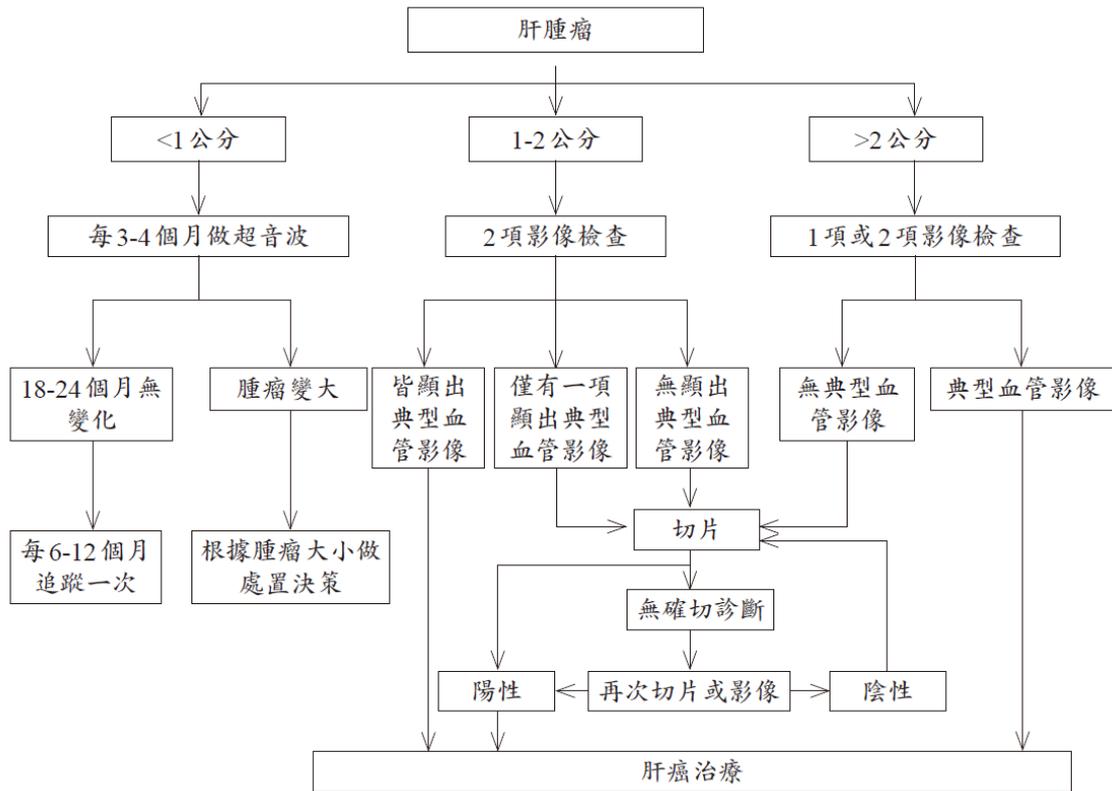


圖 2.1 肝腫瘤診斷流程圖

資料來源:劉鐘軒、蔡正中、陳海雄。2013。

目前肝癌癌症醫療診斷流程如圖 2.1，在發現肝腫瘤後先要劃分腫瘤大小，大小小於 1 公分先採用追蹤還不用治療，1-2 公分，須進行 2 項影像顯影，若兩顯影都出現典型血管影像，則須進行肝癌治療，若 1 項或無典型血管影像則須以切片再次檢驗，陽性則進入肝癌治療，大於 2 公分只要有血管影像就要進入肝癌治療，若無影像，切片觀察。

2.1.2 肝癌分級

臨床指標	分數		
	1	2	3
肝性腦病變 Encephalopathy	無	輕度至中度	嚴重
腹水 Ascites	無	輕度至中度 (對利尿劑有反應)	嚴重 (對利尿劑反應差)
膽紅素 (mg/dl) Total bilirubin	< 2	2-3	> 3
白蛋白 (g/dl) Albumin	> 3.5	2.8-3.5	< 2.8
凝血酶原時間 Prothrombin time 國際標準化比值 International Normalised Ratio	< 1.7	1.7-2.3	> 2.3

肝功能嚴重程度
 Class A = 分數總和5到6分(輕度的肝硬化)
 Class B = 分數總和7到9分(中度的肝硬化)
 Class C = 分數總和10到15分(嚴重的肝硬化)

圖 2.2 Child-Pugh 分級計分制圖

資料來源:劉鐘軒、蔡正中、陳海雄。2013。

當確定進入肝癌治療時，會先依據 5 項臨床指標區分肝硬化程度，如圖 2.2，肝性腦病變從無到嚴重分成 1-3 級；腹水從無到嚴重分成 1-3 級；膽紅數 <2 為 1 分，2-3 為 2 分，>3 為 3 分；白蛋白 >3.5 為 1 分，2.8-3.5 為 2 分，<2.8 為 3 分；凝血酶原時間國際標準化比值 <1.7 為 1 分，1.7-2.3 為 2 分，>2.3 為 3 分。再將分數加總，5-6 分為 A 級輕度肝硬化，7-9 分為 B 級中度肝硬化，10-15 分 C 級重度肝硬化，產出肝硬化指標為後續肝癌治療方法選擇的一項重要指標。

2.1.3 肝癌治療

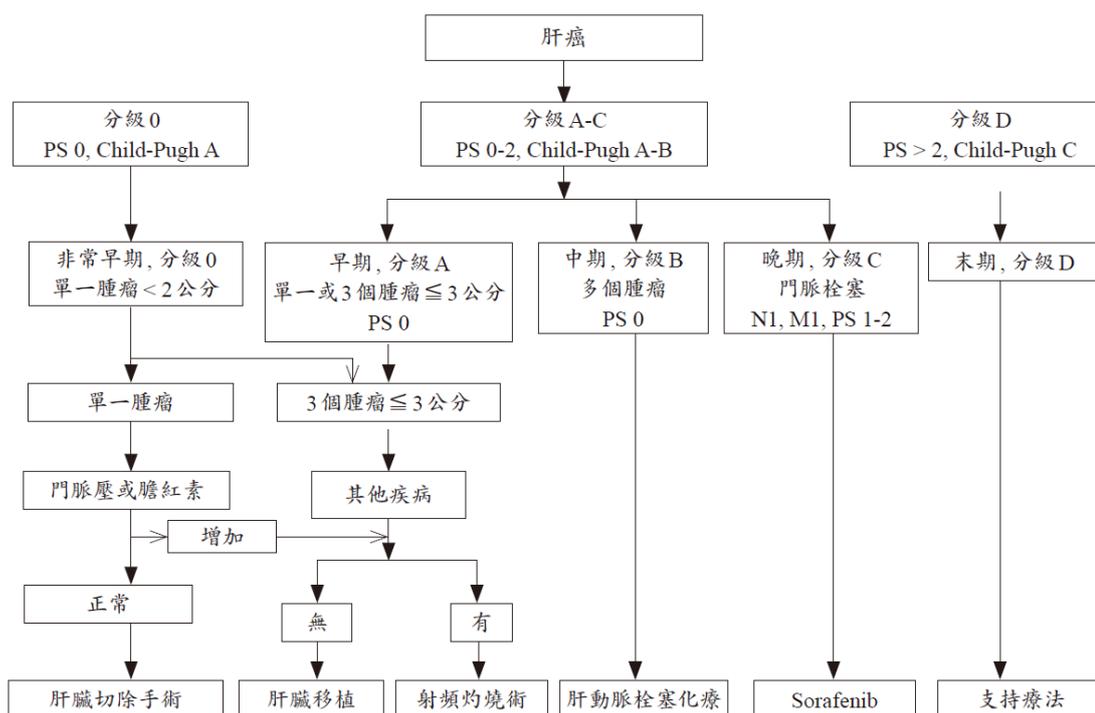


圖 2.3 巴塞隆那臨床肝癌分類(Barcelona Clinic Liver Cancer, BCLC)

資料來源:劉鐘軒、蔡正中、陳海雄。2013。

依據肝硬化程度(Child-Pugh)以及日常體能狀態 PS (Performance Status, PS)來決定癌症分級；PS 0：無症狀，PS 1：有症狀但對生活無影響，PS 2：躺在床上的時間<50%，PS 3：躺在床上的時間>50%，PS 4：長期完全臥床；癌症分級: PS 0 ,Child-Pugh A 為 0 級，PS 0-2 ,Child-Pugh A-B 為 A-C 級，PS >2 ,Child-Pugh C 為 D 級。圖中 N1：淋巴轉移，M1：遠處轉移。

分級 0，單一<2 公分腫瘤為癌症非常早期，主要以切除腫瘤為主；分級 A，3 顆內≤3 公分腫瘤，PS 0 為癌症早期，無其他疾病以肝臟移植為主，有的話以射頻灼燒處理；分級 B，多腫瘤 PS 0 為癌症中期以肝動脈栓塞化療為主；分級 C，有淋巴轉移、遠處轉移的狀況，PS 1-2 有病患會以肝動脈栓塞化療拖緩癌症擴散，但主要會以標靶治療或化療的方式延遲惡化；當情況更加惡化，分級到 D 時為癌末，此時能以嗎啡等藥物減緩疼痛，或進入安寧治療，輔助生命安寧直到死亡。

2.1.4 肝癌栓塞

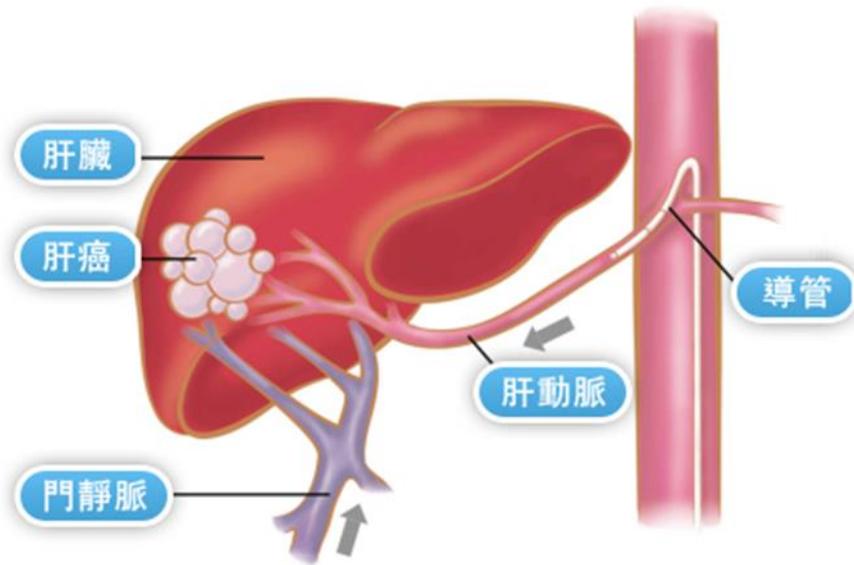


圖 2.4 肝癌肝動脈栓塞示意圖

資料來源: 肝癌防護學術基金會

本研究主要研究對象為肝癌肝動脈栓塞，治療的邏輯主要是針對肝癌所需的血液及營養 90~95% 是由肝動脈所供應的特性，故利用栓塞物把肝動脈阻塞，使肝癌細胞因沒有血流供應而缺氧缺營養萎縮。而正常的肝細胞則由門靜脈供應血液，仍可保持機能。這就是動脈栓塞療法治療肝癌的基本原理，可參考圖 2.4 示意圖。栓塞並不是根除性治療，肝腫瘤有可能由不同的動脈供應血流，故可能會重覆栓塞或以局部酒精注射治療殘存腫瘤。通常醫師會為您在鼠蹊部，實行局部麻醉再由股動脈插入導管放置到欲栓塞之肝動脈，再由導管注入栓塞物質，或加入抗癌藥物及油性溶劑，引起栓塞部位肝動脈血管阻塞，造成腫瘤細胞得不到血液營養供應而萎縮。(肝動脈血管栓塞術，趙美施，2018)

而肝癌動脈栓塞治療目前有許多文獻都認為此療法是一種無法手術病患的一項福音，且相較化療能讓在體內殘留化學治療藥物更少，也因此肝癌動脈栓塞治療將是未來更加需要被重視的課題。

Llovet JM, Bruix J 等人於 2003 以隨機對照試驗的薈萃分析證實，肝動脈化療栓塞術 (TACE) 有效成為無法手術肝癌患者於肝臟分離大或多結節非侵襲性肝癌細胞 (HCC) 並保留肝功能的標準護理治療方式。

加利福尼亞州太平洋醫療中心(Sutter health CPMC)發表於 2010 肝癌栓塞治療優點:

1. 肝臟腫瘤的血液供給通常來自肝動脈，因此使用顆粒切斷腫瘤的供給血液就可以延緩腫瘤的生長。當阻斷了腫瘤的血液供應時就形成了栓塞。
2. 顆粒還可以阻斷血液流向身體其他部位，讓化學治療藥物留在腫瘤內。這樣能夠循環到機體健康細胞處的藥物較少，就可以使用更高劑量的藥物。

2.2 藥物比較

2.2.1 藥物動力學

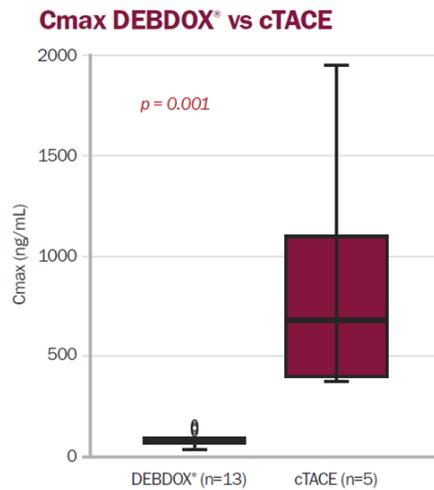
藥物代謝動力學 (Pharmacokinetics, PK)，簡稱藥代動力學或藥動學，主要是研究藥物在生物體內的變化及其規律，藥物包括藥劑、激素、營養素和毒素等。將藥物在體內過程（從吸收、分布、到代謝、排泄）以數學邏輯及模型來展示，輔助醫療研究了解藥物反應。藥代動力學輸出主要受給藥方式和劑量影響，同時也會影響吸收率的狀況。(Pharmacokinetics. In Mosby's Dictionary of Medicine, Nursing & Health Professions.)

台灣生物精神醫學暨神經精神藥理學學會表示，通常比較藥物優劣，會以兩大參數進行比較，Cmax 以及 AUC，Cmax 為最高血中藥物濃度(峰濃度)，Cmax 越高，表示藥物吸收速率越快；AUC (即 area under curve)越高，表示進入人體的藥量越多，對於藥物的清除率則相對降低。

Varela M, Real MI, Burrel M et al. 以 27 位 Child-Pugh A 肝硬化 (76% male, 59% HCV) 未經治療過的大型/多顆腫瘤肝癌患者進行新藥栓塞治療，抽樣 13 位病患檢視藥物殘留並與傳統治療方式比較得出藥物殘留顯著差異，如圖 2.5。

本研究希望能取得接受肝動脈栓塞之肝癌患者治療後的多柔比星的峰濃度 Cmax 及多柔比星的 AUC，以此能更進一步檢測肝動脈栓塞新療法與舊療法，藥物在身體中的影響程度，但目前追蹤資料尚未有此項目，故本研究抽樣未納入此屬性資料，提出以做未來研究可納入之預後追蹤指標。

多柔比星Cmax



多柔比星AUC

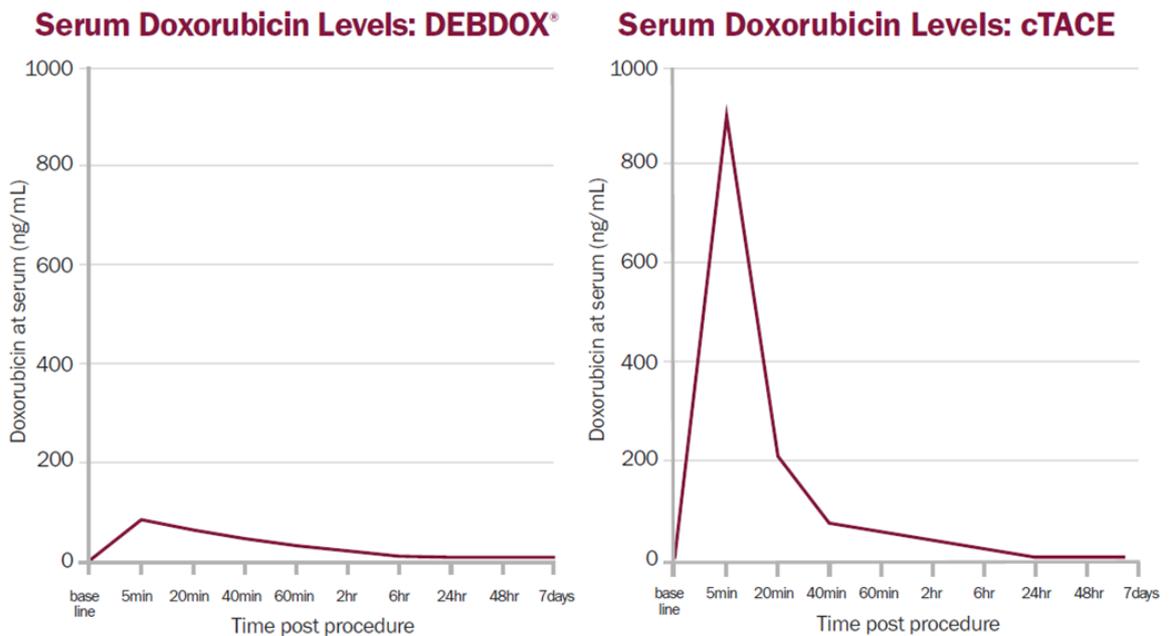


圖 2.5 多柔比星藥物動力學

資料來源: Biocompatibles UK Ltd is a BTG International group company

2.2.2 藥物療效

治療效果主要指標為:完全有效(Complete Response, CR)、部分有效(Partial Response, PR)、無差別(Stable disease, SD)、惡化(Progressive Disease, PD)。將針對四項指標如何判斷進行探討,依據彰化基督教醫院教學資料整

理出，治療後身體內如果沒有再發現癌細胞，可謂『完全有效』，一般治療追蹤 5 年，都沒有復發的跡象，即可謂『痊癒』。治療後腫瘤如果縮小一半以上，算是『部份有效』。如果病人的腫瘤既無變大，也沒變小的，則視為『無差別』，而『惡化』就是腫瘤成長變大。

雖然接受化療，但是腫瘤仍然持續變大，就視為『病程惡化』，代表癌細胞對該類治療產生抗藥性，雖然如此，其他種類的化療藥物仍然可能有效，此時要更改選擇何種藥品或治療方式，應該進一步與您的醫師商量。

《实用癌症杂志》2011 年 第 6 期，龚巧英等人以生存率及緩解情況，完全緩解(CR)率、部分緩解(PR)率及 PR+CR 比率對肝癌栓塞及 3D 適型放射治療合併栓塞兩種治療方式進行比較。

Franco M. Muggia 等針對 3、4 期卵巢癌患者進行藥物比較，在接受六週療程，每三週測量其 CR、PR 進行比較驗證藥效。

Do Seon Song 等人在肝癌栓塞比較上依據臨床實例提出結論，對於常規 TACE 難治的 HCC，使用 DC 珠的 TACE 是有效的，且這種治療有較好的反應，尤其是當腫瘤很小並且在血管造顯影延遲或增強時。在肝癌某些情況下 DC 珠可能有較佳的表現。

此屬性廣應用於癌症治療療效評估，且前人研究多採用此屬性為輸出指標，故本研究也是選擇以此屬性為輸出項進行研究。

Lammer J 等人 2010 年研究中進行抽樣統計將 212 名患有 Child-Pugh A /B 肝硬化和大型和/或多結節等因子並不可手術之癌症病患，產出圖 2.6 的結果，結果指出在四種情形疾病控制上 DC bead 優於 cTACE，且藥物不良反應也是 cTACE 比較嚴重，研究顯示出 DC bead 在某些情形下具有優勢，而 cTACE 仍有優勢空間。

本研究希望在數據量更加龐大的現在，能加入臨床使用指標再次將此議題重新檢視，以目前累積起來的肝癌動脈栓塞臨床病例資料進行驗證，期望得到更接近病患接受治療後的情形，以利於往後醫療上參考應用。

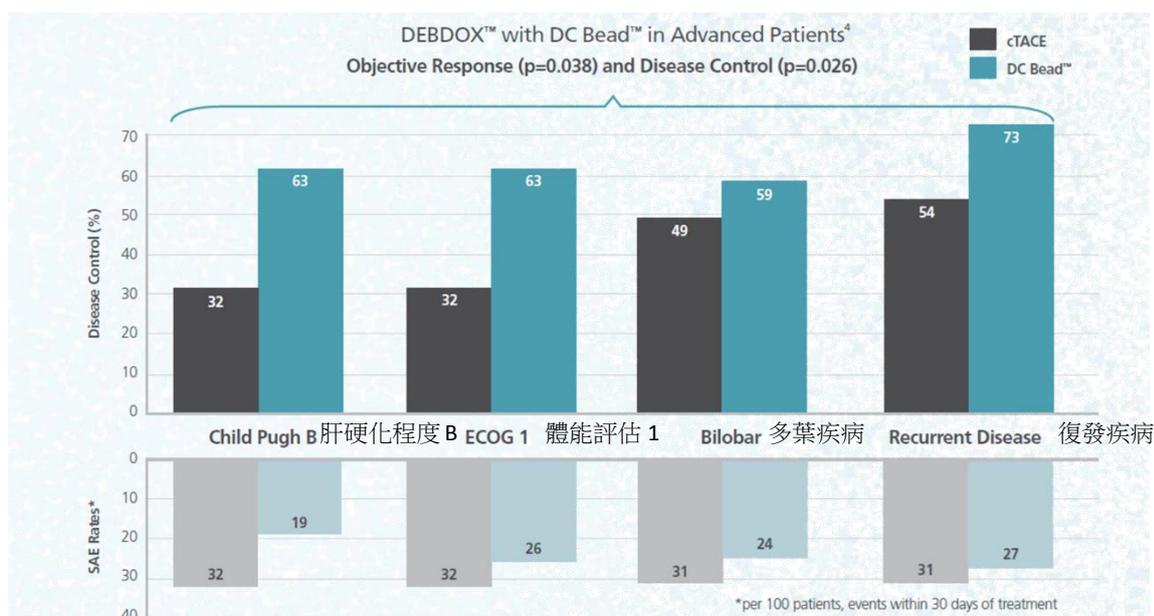


圖 2.6 肝癌動脈栓塞療效與重大不良反應比較圖

資料來源: Lammer J, Malagari K, Vogl T et.2010.

2.2.3 其他

本研究參考前人研究，應用於療效評估上還有使用其他的指標，主要還有兩項指標，第一種是肝功能狀況，另一種是副作用情形以及重大副作用事件數目，以下將針對這兩項指標進行探究。

肝功能狀況，主要是看檢查標穀草轉氨酶(AST)、谷丙轉氨酶(ALT)，ALT和AST都屬於轉氨酶。轉氨酶是人體代謝過程中必不可少的「催化劑」，主要存在於肝細胞內，在肝細胞發生炎症、壞死、中毒等，造成肝細胞受損時發揮作用。ALT主要存在於肝細胞中，它是肝細胞損傷最敏感和特異的指標。AST在心肌細胞、肝細胞和骨骼肌細胞中都有存在，如果有上述組織的損傷，它都會升高。(醫學界消化頻道，2016)。

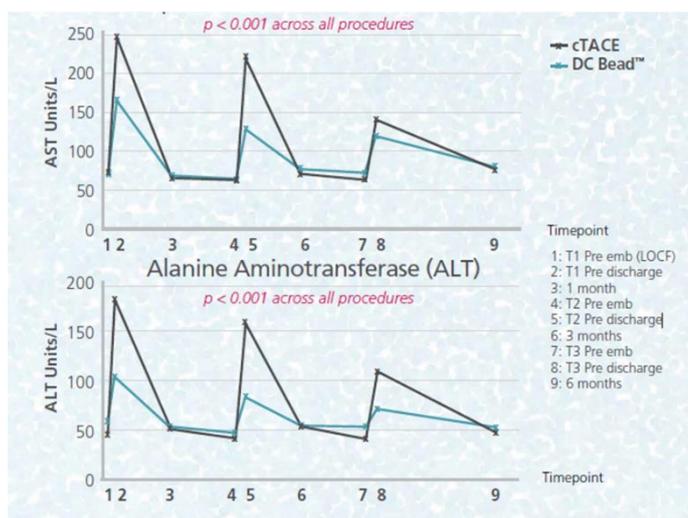
副作用(side effect)主要是人體對治療藥物產生出不良的反應，而肝癌動脈栓塞所使用的藥物多柔比星主要出現的急性副作用包括噁心、嘔吐和心律不整。它也可能引起嗜中性白血球缺乏症(neutropenia)，以及完全的脫髮。比較溫和的副作用是給藥後48小時會將尿液變成鮮豔的紅色。當多柔比星累積劑量達到500-550mg/m²，副作用包括充血性心衰竭(CHF)、擴張型心肌病甚至死亡的發生風險都顯著提高。部分患者可能發展為「肢端紅腫症」

(PPE)，特點為手掌或腳掌上的皮膚出疹、腫脹、疼痛和紅斑。由於這些副作用和紅色的症狀。(節錄自維基百科)

Lammer J 等人於 2010 年研究中將肝癌栓塞新療法及舊療法進行以上兩種指標比對產出結果如圖 2.7，結果發現新療法在兩指標上皆具有優勢，長期觀察肝功能狀態顯示肝功能狀態皆受到藥物影響，故數值皆有超出標準值之情形，但相較之下新療法照成之影響來的較低。而在副作用的方面新藥產生副作用的案例也相較來的低，以此推斷在這方面新療法確實能降低肝臟損害並較少有副作用產生。

此項屬性需要病人長期體檢累積監控，故整理出此屬性以供後續研究可再次參考使用，以台灣醫療案例進行驗證。

預後肝功能時間軸比較比例



AST和ALT的正常水平

- 血清AST的正常範圍值大約為：5-40 U/L。
- 血清ALT的正常範圍值大約為：7-56 U/L。

出現多柔比星相關副作用比例

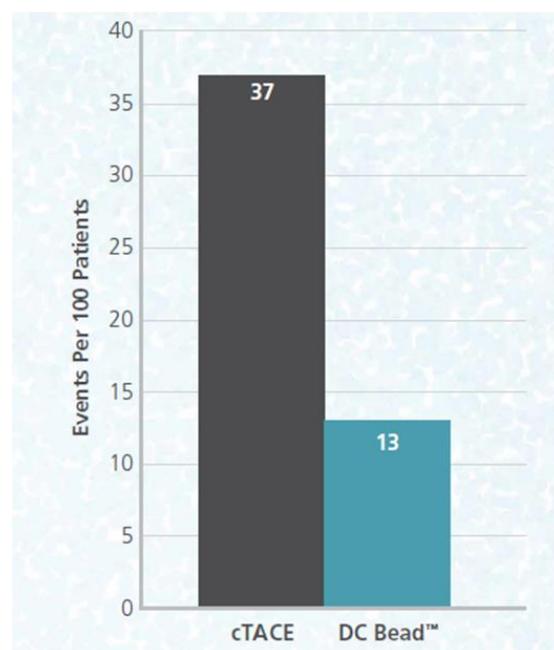


圖 2.7 肝功能曲線圖及新舊療法副作用比較圖

資料來源: Lammer J, Malagari K, Vogl T et.2010.

2.3 資料探勘

2.3.1 資料探勘定義與流程

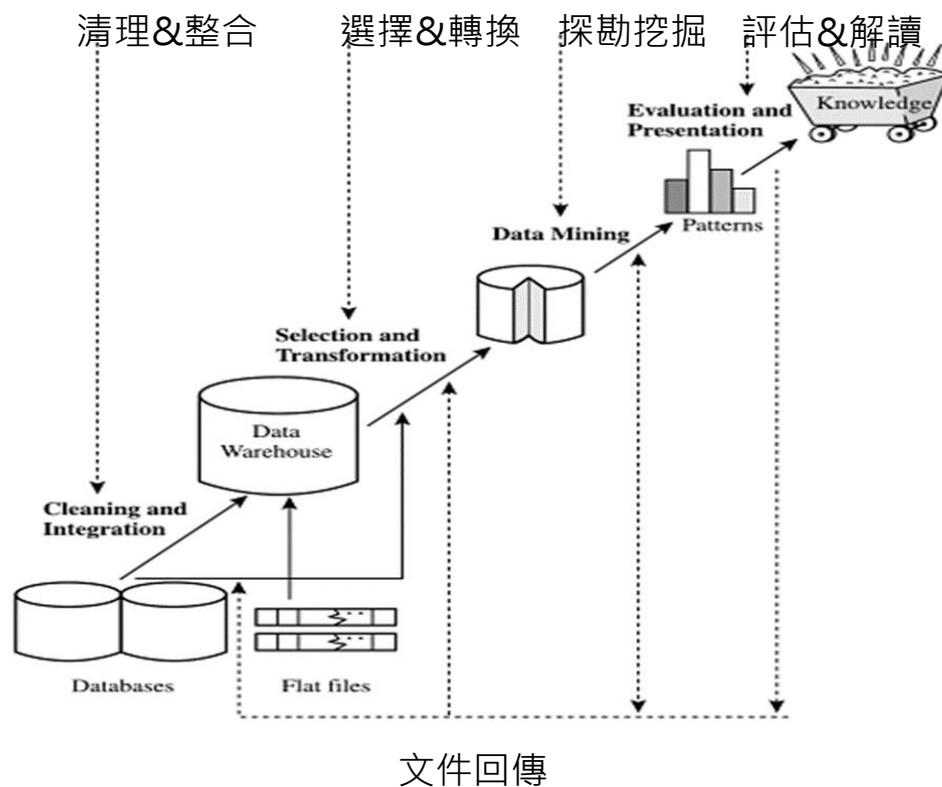


圖 2.8 資料採礦技術流程圖

資料來源:Han, J., & Kamber, M. (2006).

最早由 Fayyad(1992)所提出，其目的是從龐大的資料中找出規則，而 Han and Kamber(2000)也提到資料採礦是從大量資料中挖掘或萃取知識的一個程序，並利用這程序發掘出有用的資訊。

資料探勘與統計最大的差異在於資料處理的量，分別追求的點也略有不同，資料探勘主要在計算資料與資料間的距離藉由距離找出法則，而統計是藉由大數法則推論出一項規則，再探討推估與實際的誤差，雖然做法略有不同但兩者相輔相成。

2.3.2 資料探勘技術應用於疾病預後相關研究

表 2.1 歷年資料探勘技術應用於癌症預後文獻整理

作者	論文名稱	資料採礦技術	預測疾病	收集數據	結論
何子銘 (2006)	運用三種資料探勘方法預測子宮頸癌存活情形之比較	類神經網路、決策樹、邏輯斯迴歸	乳癌患者存活情況	433,272 筆資料記錄及 72 個變項進行資料分析	預測能力上決策樹普遍較邏輯斯迴歸高，但相差不大。但兩者皆高於類神經網路，且類神經預測較兩者來的不穩定。
陳麗帆 (2009).	使用比較資料探勘演算法預測非小細胞肺癌患者預後因子、存活情形及其效能	決策樹(DT)、類神經網路(ANN)、邏輯斯迴歸(LR)	肺癌	131,257 位美國癌症登記資料檔中診斷為 NSCLC 患者	預測力以 ANN 表現較好，外推能力以 LR 表現較好，LR 模型最易受小樣本影響；DT 則受到所提供訊息之不足而無法成樹。
胡嘉仁 (2011).	運用三種資料探勘方法預測原發性肝癌患者存活情形之比較	神經網路、決策樹、邏輯斯迴歸	肝癌	原發性肝癌共 65,088 筆個案資料記錄及 72 個變項	發現預測準確度整體平均以決策樹分析表現較好，優於類神經網路以及邏輯斯迴歸，顯示決策樹分析在內部學習效能及穩定度上比另外兩種模式好。

資料來源: 本研究整理

本研究參照前人應用資料探勘於癌症預後上的文獻，主要皆是以決策樹、類神經網路及邏輯斯迴歸為主進行預測及比較，同時本研究想探討之問題也與其中幾個技術對應，進一步胡嘉仁(2011)也有以資料探勘來預測肝癌預後之情形，同本研究也是以肝癌為研究對象，更加強支持這資料探勘方法適用於本研究。

第三章 研究材料與方法

3.1 研究材料

3.1.1 資料來源:

本研究蒐集資料之病患資料來自臺中榮總臨床資訊研發中心(Clinical Informatics Research & Development Center of Taichung Veterans General Hospital)。

3.1.2 資料期間:

民國 99 年至民國 106 年。

3.1.3 研究對象:

接受肝癌動脈栓塞治療之肝癌患者，本研究主要蒐集新療法(DC Bead TACE / 釷 90)微球載體攜藥之療法，以及使用舊療法(cTACE)油水載體攜藥之療法，接受兩種療法任一療法進行治療之肝癌患者。

本研究研究架構如圖 3.1，首先會先與專業醫師討論並結合前人研究文獻提出預其需求的資料，並由專業醫生以研究名義向醫院提出申請，申請核准後進行資料收集，將所需屬性的資料完成收集後進行資料的清洗與轉換；本研究資料清洗主要將愛滋病患加以排除以及產出遺漏之資料也先排除，進一步進行資料轉換，主要將療效依據客觀反應(CR + PR)進行劃分，以及將紀錄混亂的腫瘤數量劃分為多顆或單顆；完成以上後的資料開始進行實證研究，分別對應研究目的使用不同的方法進行分析，因所選用的三種模型皆可獨立預測治療有無效的能力，故最後對三種模型進行比較，以供未來研究參考使用。

本研究蒐集接受肝癌動脈栓塞治療之肝癌患者，排除接受肝癌動脈栓塞治療之肝癌患者中愛滋病患者，輸入項共蒐集接受肝癌動脈栓塞治療新療法之肝癌患者 171 位，接受肝癌動脈栓塞治療舊療法之肝癌患者 201 位，共計 372 位；共產出肝癌動脈栓塞治療具成效病患 215 位，肝癌動脈栓塞治療未具成效或惡化病患 157 位。

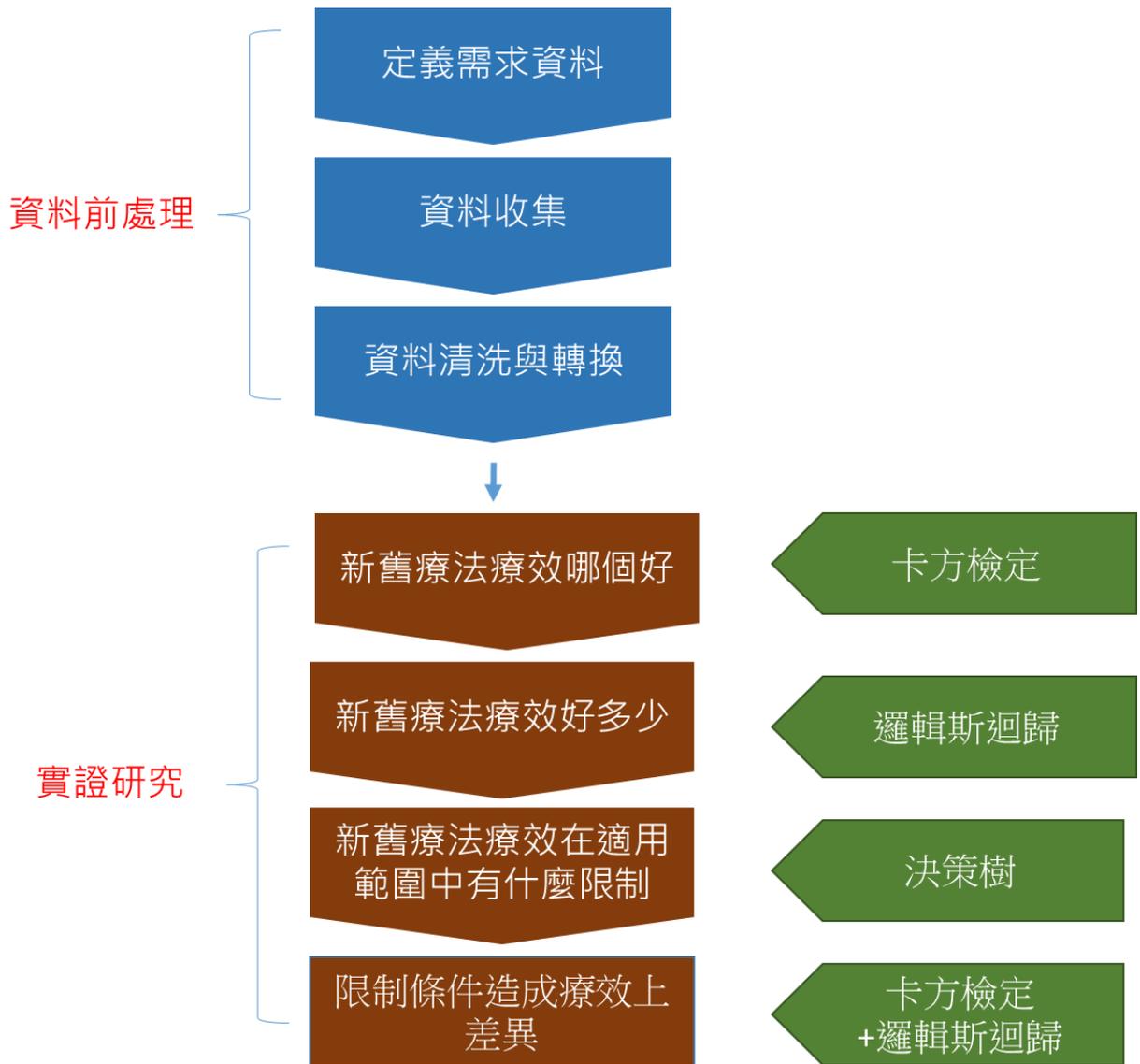


圖 3.1 研究架構圖
資料來源:本研究整理

3.1.4 資料項目:

表 3.1 資料項目表

治療評估
肝癌分期 (肝癌的診斷及治療最新發展，劉鐘軒，2013)
腫瘤大小 (肝癌的診斷及治療最新發展，劉鐘軒，2013)
腫瘤數量 (肝癌的診斷及治療最新發展，劉鐘軒，2013)
新藥舊藥 (Varela M, Real MI, Burrel M et al. J Hepatol 46 (2007): 474-481)
微球大小 (Koorosh Ashrafi et al. , Characterization of a novel intrinsically radiopaque Drug-eluting Bead for image-guided therapy: DC Bead LUMI,2017)
肝炎類型 (肝癌資料庫之建立及特徵分析，魯惠雲，2004)
預後追蹤
癌症治療效果 (Lawrence H. Schwartz et al. , RECIST 1.1 – Update and Clarification: From the RECIST Committee,2016)

資料來源:本研究整理

本研究參照前人研究文獻採用屬性資料如表 3.1，輸入項為肝癌分期(單位:期)、腫瘤大小(單位:公分 cm)、腫瘤數量(單位:顆)、新藥舊藥、微球大小(單位:微米 μm)、肝炎類型；輸出項為肝癌預後指標，癌症治療效果。

3.1.5 資料處理與轉換說明

劉怡萱在 2011 研究中提出，在 HIV/AIDS 患者罹患癌症方面，本研究確實看到一些與病毒感染相關之癌症其發生密度在 HIV/AIDS 族群當中有

較高的現象，如：口腔癌、鼻咽癌、肛門癌及肝癌。而且本研究之結果與國外許多研究結果結論相近，由此可知罹患 HIV/AIDS 確實會影響癌症的發生。並對於罹患癌症的 HIV/AIDS 患者存活具有很大的影響，其平均存活年數比較起沒有罹患癌症之 HIV/AIDS 患者少了 3 至 4 年的存活時間，而且具有統計上的顯著差異。

本研究參照前人研究，將嚴重影響預後的愛滋病患者先行替除，在進行資料的轉換，剔除完後所蒐集資料為 372 筆。

其中對兩項屬性進行調整，腫瘤數量因為病例紀錄詳細程度不同，有些有紀錄確切腫瘤數量，但多數只記錄多顆或單顆，故本研究為維持資料的一致性，採用多顆與少顆來做為該屬性的因子。

癌症療效結果，治療效果主要指標應為：完全有效(Complete Response, CR)、部分有效(Partial Response, PR)、無差別(Stable disease, SD)、惡化(Progressive Disease, PD)。本研究期望判斷治療有效還是無效，故採用 Lammer J 等人 2010 年的研究，客觀反應 (OR) 定義為 CR + PR，先做第一部分劃分，將具備效果(CR + PR)轉化為有效；將無成效(SD)及惡化(PD)轉化為無效，因子由 4 個轉成 2 項因子。

3.2 資料探勘技術

本研究參究前人研究，於醫療上最常應用之三種資料探勘模型工具進行研究比較，先以邏輯斯迴歸計算出勝算比，以此方法驗證新舊療法是否對肝癌栓塞預後療效有所差異；接著再以決策樹試圖找出影響肝癌栓塞預後療效的因子及法則；最後由類神經網路建立一組倒傳遞神經網路預測模型，以供肝癌栓塞療效預測；完成三組模型後進入評估步驟，進一步比較三種模型的適用情形。

3.2.1 邏輯斯迴歸：

一般的迴歸分析，應變項(Dependent Variable ,DV)為連續變項(continuous variable)；當 DV 不是連續變項，而是二分變項(dichotomous variable)實，就必須使用邏輯斯迴歸(logistic regression)或是以最小平方法迴歸(OLS regression)。最小平方法迴歸雖然能找到自變項(Independent variables)的係數(coefficient)，但通常會低估 IV 對 DV 的影響。

邏輯斯迴歸，算式為公式(3.2.1.1)，可由圖 3.2 看出來，將輸入之變數由無限制之數值轉換成 0 到 1 之間之數值，也就轉換成機率輸出，其運算邏輯整理如表 3.2，第一步先以訓練資料計算出迴歸係數($\beta_1, \beta_2, \beta_3, \dots$)，再將驗證資料匯入具備係數的模型輸出預測出的機率，再由預測出的機率換算勝算(odds)，即一件事情發生的機率與一件事情沒發生機率的比值，再進一步計算勝算比(odds ratio)，拿兩件事情的勝算來進行比較，計算比值。

舉例來說，假設有兩個屬性 X_1 、 X_2 ，產出成功 Y，藉由資料迴歸運算得出 $\beta_0 = -2.1464$ ， $\beta_1 = 0.3416$ ， $\beta_2 = 1.0987$ ，產出

$$Y = \frac{e^{-2.1464+0.3416X_1+1.0987X_2}}{1-e^{-2.1464+0.3416X_1+1.0987X_2}} \text{， 假設 } X_1 = 2 \text{、} X_2 = 0 \text{ 帶入}$$

$$Y = \frac{e^{-2.1464+0.3416(2)+1.0987(0)}}{1-e^{-2.1464+0.3416(2)+1.0987(0)}} = \frac{e^{-1.4632}}{1-e^{-1.4632}} = 0.188$$

此時意義代表當 $X_1 = 2$ 、 $X_2 = 0$ 時，能成功的機率為 0.188。

假設統計肝癌栓塞患者採用新療法具有療效的機率為 0.8 則勝算為 $0.8/0.2=4$ ；而肝癌栓塞患者採用舊療法具有療效的機率為 0.4 則勝算為 $0.4/0.6=2/3$ ；則勝算比為 $4/(2/3)=6$ ，則可說肝癌栓塞患者採用新療法的有療效機率是肝癌栓塞患者採用舊療法的 6 倍。

$$E(y) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2}} \quad (3.2.1.1)$$

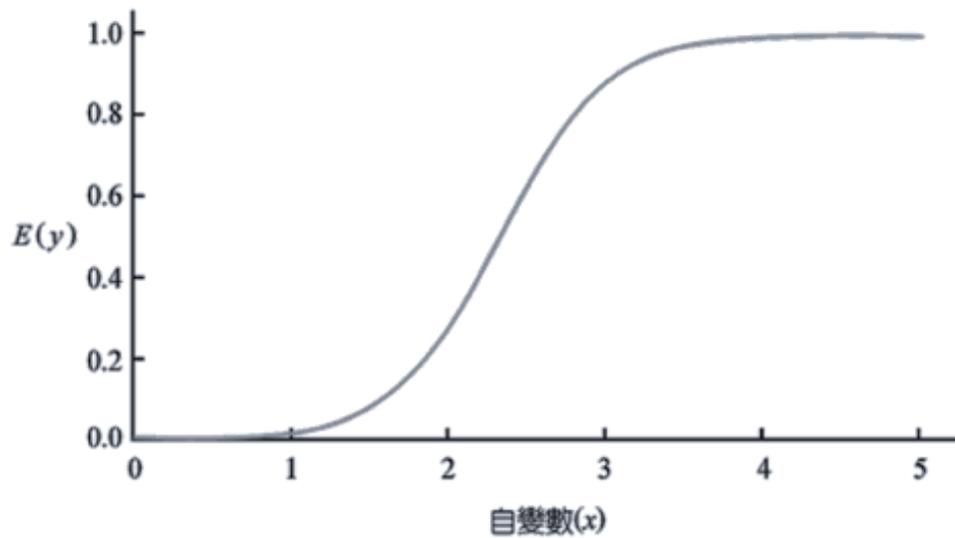


圖 3.2 邏輯斯迴歸方程式圖示

資料來源:滄海書局

表 3.2 邏輯斯迴歸演算邏輯表

- 1 • 以訓練資料找出回歸之係數(β_1 、 2 、 3)
- 2 • 驗證資料輸入得出預測結果是與否的機率
- 3 • 運用產出的機率進行勝算計算
- 4 • 計算勝算比

資料來源:本研究整理

3.2.2 決策樹:

表 3.3 決策樹種類比較表

決策樹	作者	分割依據	優點	缺點
分類與回歸樹(Classification And Regression Tree, CART)	Briemen (1984)	Gini Index	<ul style="list-style-type: none">• 完整的體系，包含枝生長到修剪整體過程。• 可處理迴歸問題	<ul style="list-style-type: none">• 屬性要求只能有兩個子節點• 類別過多時，錯誤增加較快
ID3	Quinlan (1979)	Entropy 、 Gain	<ul style="list-style-type: none">• 最早提出的決策樹演算法	<ul style="list-style-type: none">• 無法處理連自變項的問題
C4.5 (J48)	Quinlan (1993)	Gain Ratio	<ul style="list-style-type: none">• 改良 ID3，以信息增益率選擇特徵節點，避開偏袒的情形。• 改善自變項受限問題。	<ul style="list-style-type: none">• 效率比較低
C5.0	Quinlan	Gain Ratio	<ul style="list-style-type: none">• 改善 C4.5 效率問題	<ul style="list-style-type: none">• 僅能處理分類問題

資料來源:本研究整理

由李治、李國琳於《電子技術與軟件工程》2013 年第 10 期發表，針對醫學診斷資料挖掘的應用進行決策樹 C4.5 與 CART 的比較，C4.5 算法與 CART 算法相比，二者的分類精確度接近，但 C4.5 算法可以較好地控制決策樹的規模，生成的規則可理解性較高。本研究將決策樹 CART 及決策樹 ID3 體系進行整理於表 3.3，同時因本研究研究對象為醫學診斷資料挖掘的應用，故參考前人研究採用 ID3 體系的 C4.5 (J48)決策樹做為本研究之研究方法。

本研究將以舉例概述 ID3 決策樹以及 C4.5 (J48)決策樹演算法，以應用於研究使用。

● ID3 演算法:

核心概念:信息熵越大代表資料越混濁

$$\text{原信息熵} : I(p,n) = -\left(\frac{p}{n+p}\right) \times \log_2\left(\frac{p}{n+p}\right) + \left(\frac{n}{n+p}\right) \times \log_2\left(\frac{n}{n+p}\right) \quad (3.2.2.1)$$

$$\text{預期處理後信息熵} E(x) = \sum_{i=1}^n \left(\frac{n_i+p_i}{n+p}\right) \times I(n_i, p_i) \quad (3.2.2.2)$$

$$\text{資料增益} \text{Gain}(x) = I(p,n) - E(x) \quad (3.2.2.3)$$

各屬性之資料增益量即為決策樹分支決策參數，也就是哪一個屬性的增益量最大就是分支的節點，其意義為以此屬性劃分可最大化降低資料的混濁程度。舉例來說：

假設：

類別 P: 療效=“yes” 9 人 ; 類別 N: 療效=“no” 5 人

則資料不純度 : $I(p, n) = I(9, 5) = 0.940$ (帶入 3.2.2.1 公式)

計算經過年齡分類後的 $I(p, n)$

年齡	p_i	n_i	$I(p_i, n_i)$
≤ 30	2	3	0.971
30~40	4	0	0
> 40	3	2	0.971

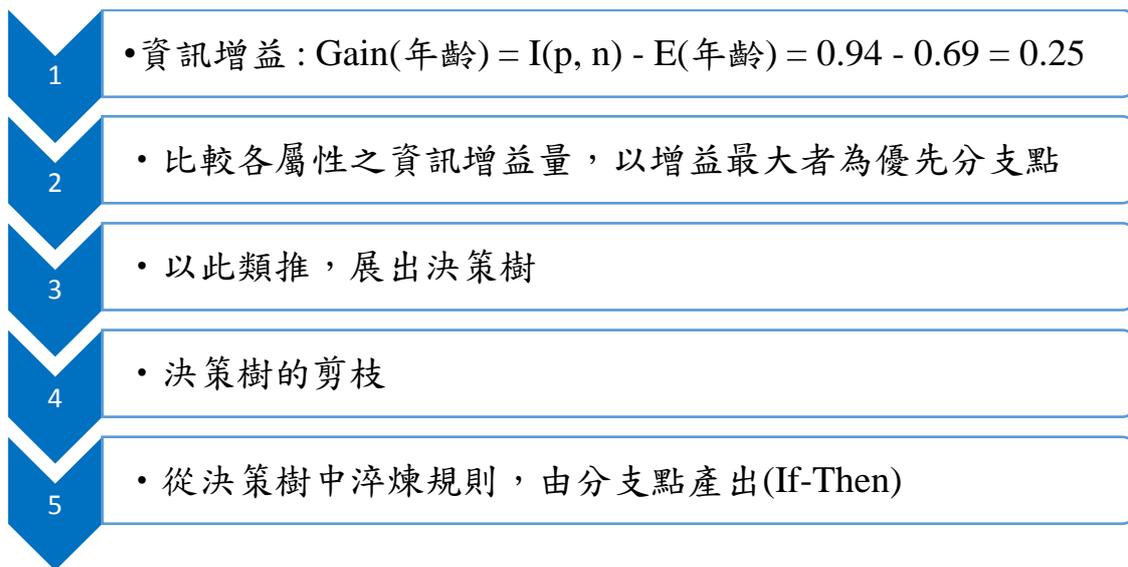
預期處理後信息熵：

$$\begin{aligned} E(\text{年齡}) &= \frac{5}{14} I(2,3) + \frac{4}{14} I(4,0) \\ &\quad + \frac{5}{14} I(3,2) = 0.69 \end{aligned} \quad (\text{帶入 3.2.2.2 公式})$$

資料增益量：

$$\text{Gain}(\text{年齡}) = I(p, n) - E(\text{年齡}) = 0.94 - 0.69 = 0.25 \quad (\text{帶入 3.2.2.3 公式})$$

表 3.4 決策樹演算邏輯表



資料來源:本研就整理

算出所有屬性之增益量後進行比較，後續處理可參照本研就整理之表 3.4。

● **C4.5 (J48)演算法:**

ID3 具有偏袒問題：如 ID 號，每資料一個 ID，以它為分支點，Gain 會最大化；又如一種自變項於資料中每筆皆相同，Gain 會最小，為防止這狀況出現，C4.5 因此改變使用增益率(Gain Ratio)來進行分支決策，增益量/自變項本身信息熵，即考慮自變項本身體質的問題，則可避免偏袒的情況發生。

核心概念:同 ID3 核心概念，加入考量自身信息熵，排除偏袒問題。

$$\text{自變項信息熵: } SI(X) = -\sum_{j=1}^v \left(\left(\frac{N_j}{N} \right) \times \log_2 \left(\frac{N_j}{N} \right) \right) \quad (3.2.2.4)$$

$$\text{資訊增益率: } GR(X) = \frac{\text{Gain}(X)}{SI(X)} \quad (3.2.2.5)$$

各屬性之資料增益率即為決策樹分支決策參數，也就是哪一個屬性的增益率最大就是分支的節點，其意義為以此屬性劃分可最快速度降低資料的混濁程度。延續 ID3 舉例來說:

年齡	p_i	n_i	$p_i + n_i = N_j$	$I(p_i, n_i)$
<=30	2	3	5	0.971
30~40	4	0	4	0
>40	3	2	5	0.971

N=14

資訊增益:Gain(年齡) = I(p, n) - E(年齡) = 0.94 - 0.69 = 0.25

(帶入 3.2.2.3 公式)

自變項信息熵:SI(年齡) = $-\left(\frac{5}{14} \times \log_2\left(\frac{5}{14}\right) + \frac{4}{14} \times \log_2\left(\frac{4}{14}\right) + \frac{5}{14} \times \log_2\left(\frac{5}{14}\right)\right) = 1.5774$

(帶入 3.2.2.4 公式)

資訊增益率:GR(年齡) = $\frac{\text{Gain(年齡)}}{\text{SI(年齡)}} = \frac{0.25}{1.5774} = 0.1585$

(帶入 3.2.2.5 公式)

後續處理同 ID3，可參照表 3.2.2，依據增益率做節點劃分參數。

3.3 評估模式

於本章節主要概述本研究訓練資料及驗證資料的劃分依據，以及最後評估模型適用程度的評估方法。本研究參照前人研究提出最常用的方法將採用十摺交叉驗證的方式進行訓練資料及驗證資料的劃分方式，並採用混淆矩陣及 ROC 曲線下面積及 PRC 曲線下面積進行隊本研究使用的 3 種模型進行綜合比較。

3.3.1 十摺交叉驗證

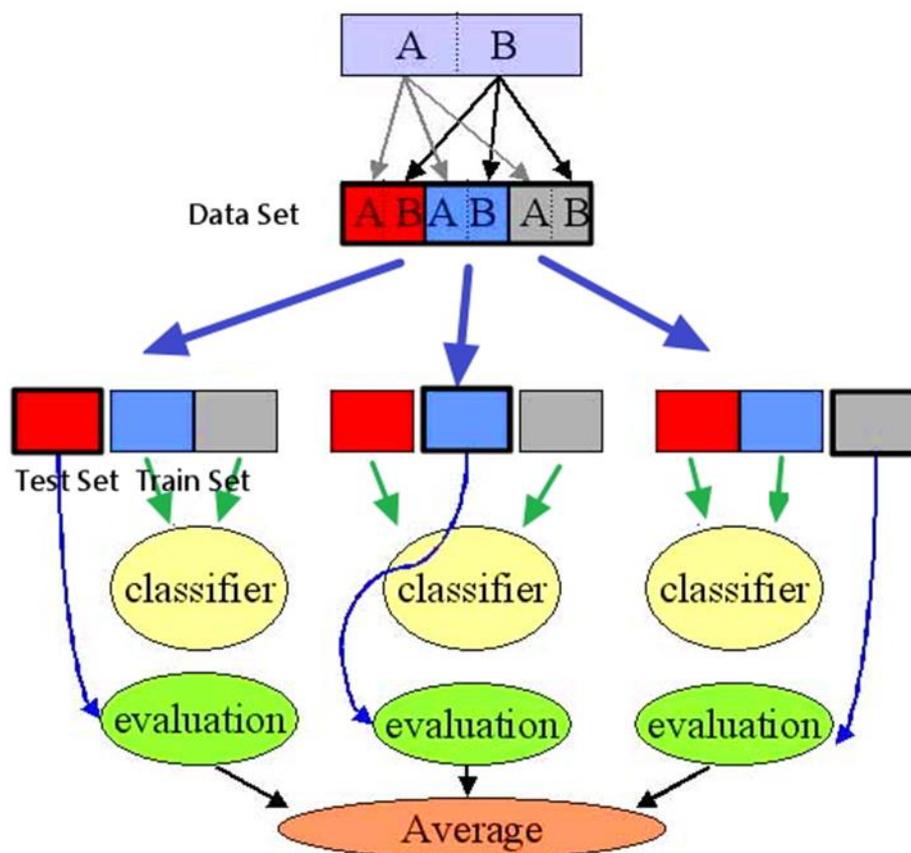


圖 3.3 十摺交叉驗證運行圖

資料來源: 逍遙文工作室。2012。

本研究參考逍遙文工作室(2012)並整理出交叉驗證概述，交叉驗證(Cross Validation)，又稱做循環估計，為統計學上將數據樣本切割成較小子集的實用方法。一開始的子集被稱為訓練集(training set)。而其他的子集則被稱為驗證集(validation set)或測試集(testing set)。K 摺交叉驗證即為將資料區分 K 等份，取其中一份當驗證資料，其餘 K-1 份當訓練資料，以此方式重複 K 次取其平均值，以此方式提高準確率，運作方式如圖 3.6。

最常見的交叉驗證方法有二折交叉驗證(2-fold cross-validation)，十折交叉驗證(10-fold cross-validation)，而以十折交叉驗證較常被運用。(Kohavi, Ron, 1995.)，故本研究採用十摺交叉驗證的方式進行研究。

3.3.2 混淆矩陣:

表 3.5 混淆矩陣表

		預測		
		正確	錯誤	
真實	正 確	True Positive (TP)真陽	False Negative (FN)假陰	TP+FN=A
	錯 誤	False Positive (FP)假陽	True Negative (TN)真陰	FP+TN=B
		TP+FP=C	FN+TN=D	A+B=C+D=E

資料來源:本研究整理

由表 3.5 可以了解整體混淆矩陣的運作模式，接下來是由混淆矩陣延伸出的判斷指標，本研究將依據表中所使用的代號，逐一簡介各項指標。

首先從預測準確面介紹(C、D 面):

精確率(precision) = $\frac{TP}{C} = \frac{TP}{TP+FP}$ ，也就是預測正確中正中的。

召回率(recall) = $\frac{TN}{D} = \frac{TN}{TP+FN}$ ，也就是預測錯誤中正中的。

從真實面介紹(A、B 面):

真陽性率 (敏感度)(true positive rate, TPR) = $\frac{TP}{A} = \frac{TP}{TP+FN}$

假陰性率(不一致性) (false negative rate, FNR) = $\frac{FN}{A} = \frac{FN}{TP+FN}$

真陰性率 (特異性)(true negative rate, TNR) = $\frac{TN}{B} = \frac{TN}{FP+TN}$

假陽性率(不一致性)(false positive rate, FPR) = $\frac{FP}{B} = \frac{FP}{FP+TN}$

整體模型準確成度(E 面):

$$\text{正確率} = \frac{TP+TN}{E} = \frac{TP+TN}{TP+FN+TN+FP}$$

進一步，由真陽性率、假陰性率、真陰性率、假陽性率衍伸出的概似比(Likelihood ratio)，可依據圖 3.7 對應 LR 值算出的數據判斷實驗通過的機率。

$$LR(+)= \text{真陽性率} / \text{假陽性率} = TP/(TP+FN) / FP/(FP+TN)$$

$$LR(-)= \text{假陰性率} / \text{真陰性率} = FN/(TP+FN) / TN/(FP+TN)$$

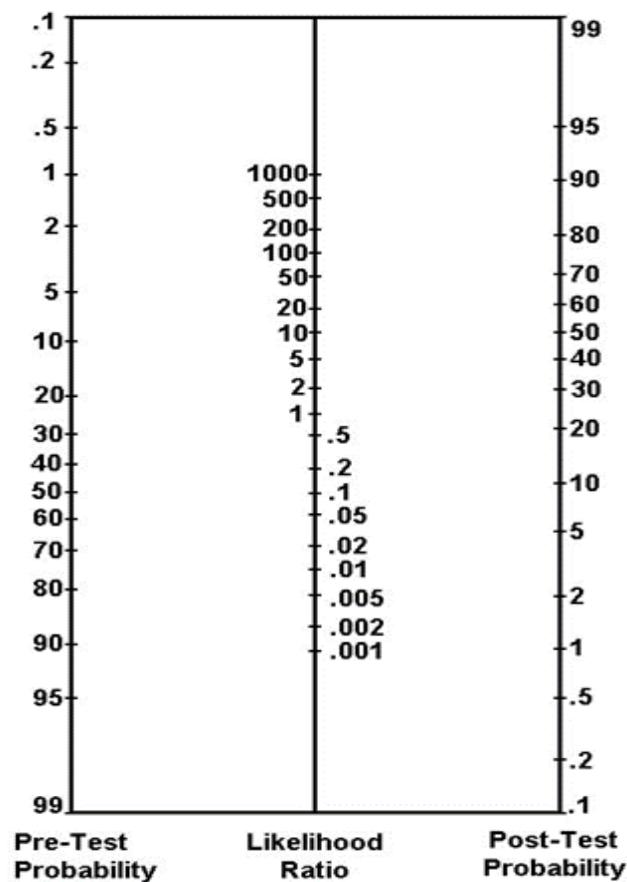


圖 3.4 概似比對應實驗通過機率圖

資料來源:CodeHealth

3.3.1 ROC 曲線及 PRC 曲線:

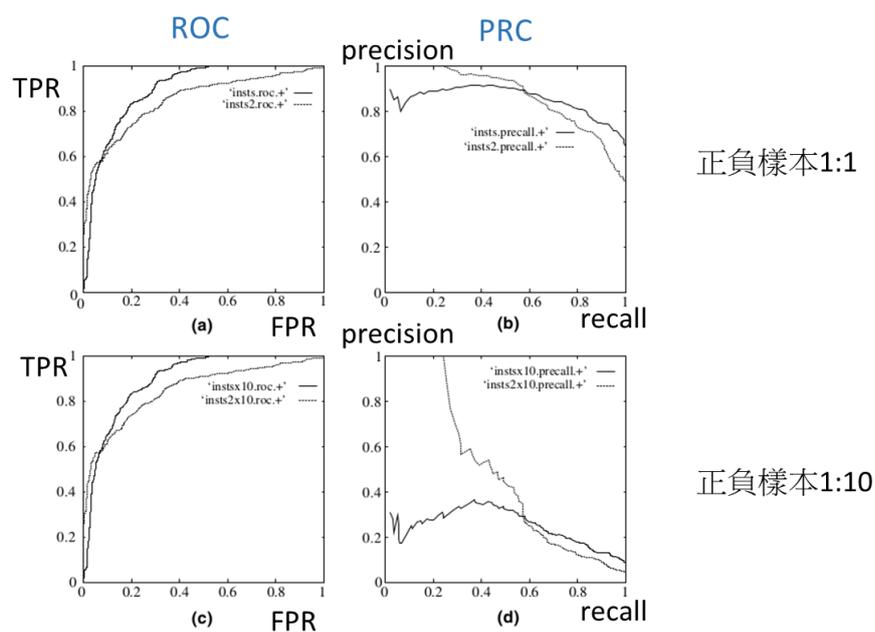


圖 3.5 ROC 曲線與 PRC 曲線比較圖

資料來源: 皮皮 BLOG(2016)、本研究整理

ROC、PRC 曲線下面積都是判斷模型準確程度的重要依據，若面積為 0.5，同等用猜的機率，低於 0.5 則比猜的命中率還低；兩曲線最大差異在於對正負樣本的敏感程度，由圖 3.8 可看出來當正負樣本比例調大時，ROC 曲顯無明顯變動，而 PRC 曲顯嚴重變化，故當正負樣本差異大時較適用 PRC 曲線。

第四章 實證分析

本研究採用 IBM SPSS Statistics 19 對資料進行卡方檢定，再採用 Weka 處理資料探勘分析需求，先將資料分作使用新療法的病患以及使用舊療法的病患，以卡方分配檢視蒐集來的肝癌栓塞全體病患資料基本療效概況，進一步採用資料探勘技術中的邏輯斯迴歸方法對新舊療法進行有效程度上差異的比較；接著，本研究期望能更進一步探討兩種療法各自存在的適用情況及限制，因此將進一步以資料探勘中決策樹方法做為分類器更進一步對資料進行劃分，而劃分後的結果再以統計卡方分配方式進行驗證支持，最後以邏輯斯迴歸再次比較各個分類後療效上的差異程度，以以上方式試圖找出新舊療法在肝癌栓塞治療預後成果上所受到的限制以及適用領域，提供醫療選擇療法上一項參考。

4.1 傳統療法及微球療法療效比較

療效 * 新舊療法 交叉表

			新舊療法		總和
			c T A C E	D C b e a d	
療效	有效	個數	137	78	215
		在療效之內的	63.7%	36.3%	100.0%
		在新舊療法之內的	68.2%	45.6%	57.8%
	無效	個數	64	93	157
		在療效之內的	40.8%	59.2%	100.0%
		在新舊療法之內的	31.8%	54.4%	42.2%
總和	個數	201	171	372	
	在療效之內的	54.0%	46.0%	100.0%	
	在新舊療法之內的	100.0%	100.0%	100.0%	
	整體的 %	54.0%	46.0%	100.0%	

卡方檢定

	數值	自由度	漸近顯著性 (雙尾)	精確顯著性 (雙尾)	精確顯著性 (單尾)
Pearson卡方	19.253 ^a	1	.000		
連續性校正 ^b	18.340	1	.000		
概似比	19.366	1	.000		
Fisher's精確檢定				.000	.000
線性對線性的關連	19.201	1	.000		
有效觀察值的個數	372				

a. 0格 (0%) 的預期個數少於 5。最小的預期個數為 72.17。
b. 只能計算 2x2 表格

圖 4.1 卡方分配產出

資料來源:本研究產出

有無效 * 新舊藥 交叉表

			新舊藥		總和
			cTACE	DC	
有無效	無效	個數	64	78	142
		在 有無效 之內的	45.1%	54.9%	100.0%
		在 新舊藥 之內的	31.8%	54.5%	41.3%
		整體的 %	18.6%	22.7%	41.3%
有效	有效	個數	137	65	202
		在 有無效 之內的	67.8%	32.2%	100.0%
		在 新舊藥 之內的	68.2%	45.5%	58.7%
		整體的 %	39.8%	18.9%	58.7%
總和	總和	個數	201	143	344
		在 有無效 之內的	58.4%	41.6%	100.0%
		在 新舊藥 之內的	100.0%	100.0%	100.0%
		整體的 %	58.4%	41.6%	100.0%

卡方檢定

	數值	自由度	漸近顯著性 (雙尾)	精確顯著性 (雙尾)	精確顯著性 (單尾)
Pearson卡方	17.770 ^a	1	.000		
連續性校正 ^b	16.845	1	.000		
概似比	17.793	1	.000		
Fisher's精確檢定				.000	.000
線性對線性的關連	17.718	1	.000		
有效觀察值的個數	344				

a. 0格 (0%) 的預期個數少於 5。最小的預期個數為 59.03。

b. 只能計算 2x2 表格

圖 4.2 去除 C、D 級病患卡方分配產出

資料來源:本研究產出

本研究首先針對新舊療法療效進行統計分析，以卡方分配來檢視新舊療法哪個有效的比例比較多，從統計結果圖 4.1，從中可以發現傳統療法(cTACE)有效比例為 68.2%，而無效的比例為 31.8%；新療法(DC bead)有效比例為 45.6%，而無效的比例為 54.4%。且雙尾檢定 <0.05 ，故卡方檢定具有顯著差異，也就是說目前以資料來看，傳統療法具有顯著較好的療效。比較前人文獻，本研究資料顯示出不同結果，兩種療法有效比例差約 22%，傳統療法有較好療效，但我們回推資料本身並與專業醫師探討，目前新療法廣泛運用癌症各期，但舊療法僅限用在癌症 B 級及比 B 級更早期的級數，故此統計結果僅能證實舊療法在總體未考慮癌症分級下具備療效優勢。

進一步，本研究將癌症分級 C、D 級患者先抽離再次比較兩療法，結果如圖 4.2，可發現傳統療法資料數目沒變，代表 C、D 級病患皆未採用傳統栓塞療法，同樣的有效比例也不會有所變動，但新療法病患數量由 171 變成 143，其中有 28 位癌症分級 C、D 的病患，再看有效比例，由 45.6%變成 45.5%，基本上無大的變化，可知 C、D 級病患有效比例類似總體，同時也代表在癌症分級 A、B 級傳統療法具有顯著優勢。

由目前分析結果來說，具有以下兩項意義，第一，癌症分級 A、B 級的患者使用傳統療法具有較高的成功控制癌症機率，但目前尚無傳統療法應用於癌症分級 C、D 級的治療資料，因此無法比較癌症分級 C、D 級病患兩種療法的優劣，可做為未來傳統療法加強深入研究的部分；第二，微球療法雖有擴大肝癌栓塞治療領域的功效，但療效有效比率僅 45.6%，治療成效偏低，且去廚較難治療先天情況較差的癌症分級 C、D 疾病患後，仍是偏低的治療成效，故未來發展可加強對微球治療無效原因的研究及探討。

Variable	Class
	有效
肝炎=B	-0.1327
肝炎=C	0.0991
肝炎=non-B non-C	-0.0975
肝炎=B+C	0.3167
分級=B	-0.063
分級=C	0.2176
分級=A	-0.0334
分級=D	0.5472
數量=單	1.055
大小	-0.0903
微球=300-500	-1.2534
微球=300-500, 500-700	0.4087
微球=100-300	-0.6229
微球=500-700	-0.8346
微球=100-300, 500-700	46.0508
微球=100-300, 300-500	17.6603
微球=0	0.0648
新舊藥=CTACE	0.0648
Intercept	0.7307
肝炎=C	1.1042
肝炎=non-B non-C	0.9071
肝炎=B+C	1.3726
分級=B	0.9389
分級=C	1.2431
分級=A	0.9672
分級=D	1.7284
數量=單	2.8719
大小	0.9137
微球=300-500	0.2855
微球=300-500, 500-700	1.5049
微球=100-300	0.5364
微球=500-700	0.4341
微球=100-300, 500-700	9.990488157293326E19
微球=100-300, 300-500	46749492.226
微球=0	1.0669
新舊藥=CTACE	1.0669

圖 4.3 邏輯斯迴歸產出結果圖

資料來源:本研究產出

=== Summary ===

Correctly Classified Instances	240	64.5161 %
Incorrectly Classified Instances	132	35.4839 %
Kappa statistic	0.2547	
Mean absolute error	0.4312	
Root mean squared error	0.4777	
Relative absolute error	88.3782 %	
Root relative squared error	96.7225 %	
Total Number of Instances	372	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.758	0.510	0.671	0.758	0.712	0.258	0.676	0.728	有效
	0.490	0.242	0.597	0.490	0.538	0.258	0.676	0.562	無效
Weighted Avg.	0.645	0.397	0.640	0.645	0.639	0.258	0.676	0.658	

圖 4.4 邏輯斯迴歸模型適用性圖

資料來源:本研究產出

邏輯斯迴歸產出結果如圖 4.3，邏輯斯迴歸的回歸係數以及各組的勝算比，從迴歸係數可以發現微球混球 100-300,300-500 μm 以及 100-300,500-700 μm 的係數都比較高，勘查其資料，發現有較高機率產出有效的成果，但臨床應用資料較少，故未來可能可以多朝混球方面累積資料進形更伸一步得探討，同時也顯示出腫瘤數量也是較有影響力的因子。

從產出的勝算比來看，具有最大勝算的是微球=100-300,300-500 μm ，此性具有超高的勝算比，將資料展開來看醫療資料中只有 4 筆且每筆皆產出治療有效的結果，故其勝算會變得非常大，同時也影響勝算比產出極大，此預測可能要具備更大量使用微球=100-300,300-500 μm 的資料以確保預測準確；由產出也發現肝動脈栓塞對於肝炎 B 或肝炎 B+C 的患者相較其他肝炎情形患者有較好的療效；肝動脈栓塞也對分級 C 或 D 較其他分級有較好的療效，由文獻探討中我們能發現分級 C 為肝晚期 D 為肝癌末期，對應治療方法建議皆是延遲癌症擴散及惡化，本研究產出之結果，因晚期及末期治療上皆是採姑息治療的方式，只要有拖延惡化皆會列入治療有效故在 C 級、D 級，相較 A、B 級要求標的會有所不同，故可能較有較好的療效成果；在數量上明顯吻合前人研究，單顆腫有較高的有效療效機率，單顆腫瘤有效療效機率是多顆腫瘤的 2.8719 倍；微球大小 300-500,500-700 μm 也具有較高的有效療效機率。

而最重要的，以全體病患計算出的勝算比結果發現傳統療法成功的機率是 1.0669 倍，差異並不大，此結果因納入癌症分級 C、D 疾病患一起觀察，所以可能因為傳統療法在於癌症分級 A、B 級病患具有強大優勢，故綜觀來說療效成果差異並不太大，本研究也將更進一步將癌症分級 C、D 級病患加以排除，再次進行比較。

而模型的適用狀況如圖 4.4，模型的準確度為 64.5%，假陽率為 0.397，精確率為 0.64，召回率為 0.39，預測能力看 ROC 曲線下面積及 PRC 曲線下面積，分別為 0.676 及 0.658。

```

Coefficients...
      Class
Variable 有效
=====
新舊藥=CTACE    0.9434
Intercept    -0.1823

Odds Ratios...
      Class
Variable 有效
=====
新舊藥=CTACE    2.5687

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      215          62.5 %
Incorrectly Classified Instances    129          37.5 %
Kappa statistic                    0.2273
Mean absolute error                0.4609
Root mean squared error            0.4807
Relative absolute error            95.0467 %
Root relative squared error        97.6357 %
Total Number of Instances         344

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.678    0.451    0.682    0.678    0.680    0.227    0.572    0.629    有效
0.549    0.322    0.545    0.549    0.547    0.227    0.572    0.462    無效
Weighted Avg.    0.625    0.397    0.625    0.625    0.625    0.227    0.572    0.560

```

圖 4.5 去除 C、D 級病患邏輯斯迴歸模型

資料來源:本研究產出

去除癌症分級 C、D 級病患後，再次以邏輯斯迴歸對資料進行分析，結果發現在於 A、B 級，確實傳統療法具有較強的優勢，以計算出的勝算比如圖 4.5 可以發現，傳統療法有效的機率是微球療法的 2.57 倍，在次佐證在癌症分級 A、B 級患者的治療上選擇傳統療法會有較大的成功機率，此結果提供醫療選擇療法上的參考。

而進一步，本研究希望探討兩種療法是否存在各自適合的應用情況，且是否具有各自的限制，也因此本研究將採用資料探勘的決策樹方式做為分類器，對資料更進一步切分。

4.2 傳統療法限制與適用細分

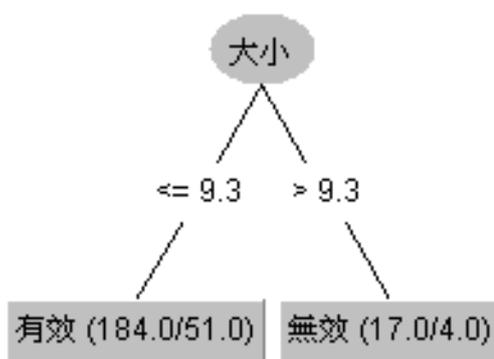


圖 4.6 傳統療法決策樹分類

資料來源:本研究產出

```

=== Summary ===

Correctly Classified Instances      140          69.6517 %
Incorrectly Classified Instances    61           30.3483 %
Kappa statistic                    0.1216
Mean absolute error                 0.4088
Root mean squared error             0.465
Relative absolute error             93.9964 %
Root relative squared error         99.7812 %
Total Number of Instances          201

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                0.956   0.859   0.704     0.956   0.811     0.172   0.533    0.698    有效
                0.141   0.044   0.600     0.141   0.228     0.172   0.533    0.387    無效
Weighted Avg.   0.697   0.600   0.671     0.697   0.625     0.172   0.533    0.599
  
```

圖 4.7 傳統療法決策樹分類模型結果

資料來源:本研究產出

本研究以 C4.8 決策樹產出的結果如圖 4.6，傳統療法節點為腫瘤大小，依照 9.3 為分水嶺，大於 9.3 公分則治療無效，此結果有正呼應了前面統計所產出的結果，造成無效的腫瘤平均大小較大；本研究依據真實病歷資料得出 9.3 為其有無療效的腫瘤大小分水嶺。

模型的適用狀況如圖 4.7，模型的準確度為 69.7%，假陽率為 0.6，精確率為 0.671，召回率為 0.637 預測能力看 ROC 曲線下面積及 PRC 曲面下面積，分別為 0.533 及 0.599。

F 檢定：兩個常態母體變異數的檢定		
	無效	有效
平均數	5.69375	4.438321
變異數	8.584722222	6.259936
觀察值個數	64	137
自由度	63	136
F	1.371375369	
P(F<=f) 單尾	0.065049369	
臨界值：單尾	1.409443063	
t 檢定：兩個母體平均數差的檢定，假設變異數相等		
	無效	有效
平均數	5.69375	4.438321
變異數	8.584722222	6.259936
觀察值個數	64	137
Pooled 變異數	6.995923688	
假設的均數差	0	
自由度	199	
t 統計	3.134886405	
P(T<=t) 單尾	0.00098968	
臨界值：單尾	1.652546746	
P(T<=t) 雙尾	0.001979359	
臨界值：雙尾	1.971956544	

圖 4.8 cTACE 有效無效及腫瘤大小 T 檢定產出

資料來源:本研究產出

本研究以統計均值 T 檢定方式對以腫瘤大小劃分加以檢視，將療效有效的腫瘤大小平均值與療效無效的腫瘤大小平均值進行比較，結果產生如

圖 4.8，療效無效的腫瘤大小平均為 5.69 公分，而療效有效的腫瘤大小平均 4.44 公分，檢定結果為顯著，確實療效無效的腫瘤大小平均值較大，但均值檢定很難找出劃分界限，也因此本研究以 C4.5 決策樹對資料進行劃分，劃分界限為 9.3 公分。

觀察值處理摘要

	觀察值					
	有效的		遺漏值		總和	
	個數	百分比	個數	百分比	個數	百分比
有效無效 * 腫瘤大小9.3劃分	201	100.0%	0	.0%	201	100.0%

有效無效 * 腫瘤大小9.3劃分 交叉表

			腫瘤大小9.3劃分		總和
			<= 9.3	> 9.3	
有效無效	無效	個數	51	13	64
		在有效無效之內的	79.7%	20.3%	100.0%
		在腫瘤大小9.3劃分之內的	27.7%	76.5%	31.8%
		整體的 %	25.4%	6.5%	31.8%
	有效	個數	133	4	137
		在有效無效之內的	97.1%	2.9%	100.0%
在腫瘤大小9.3劃分之內的		72.3%	23.5%	68.2%	
	整體的 %	66.2%	2.0%	68.2%	
總和	個數	184	17	201	
	在有效無效之內的	91.5%	8.5%	100.0%	
	在腫瘤大小9.3劃分之內的	100.0%	100.0%	100.0%	
	整體的 %	91.5%	8.5%	100.0%	

卡方檢定

	數值	自由度	漸近顯著性 (雙尾)	精確顯著性 (雙尾)	精確顯著性 (單尾)
Pearson卡方	17.044 ^a	1	.000		
連續性校正 ^b	14.871	1	.000		
概似比	15.749	1	.000		
Fisher's精確檢定				.000	.000
線性對線性的關連	16.959	1	.000		
有效觀察值的個數	201				

a. 0格 (0%) 的預期個數少於 5。最小的預期個數為 5.41。
b. 只能計算 2x2 表格

圖 4.9 cTACE 以 9.3 公分劃分卡方檢定產出

資料來源:本研究產出

進一步本研究以卡方分配再次檢驗以 9.3 公分劃分後有效無效的比例是否有顯著的差異，卡方檢定結果如圖 4.9，小於等於 9.3 公分腫瘤大小的無療效比例為 27.7%，而有療效比例為 72.3%；大於 9.3 公分腫瘤大小的無療效比例為 76.5%，而有療效比例為 23.5%；且卡方檢定為顯著差異，故可以此推斷確實傳統療法以 9.3 公分劃分有效無效有一定參考價值，也因此

建議治療上若癌症分級 A、B 級病患若腫瘤大小小於等於 9.3 公分時可選用傳統療法，當若大於 9.3 公分則建議考慮其他療法。雖然知道 9.3 公分為分水嶺，但本研究想進一步了解有效程度上的差異，因此本研究將使用邏輯斯迴歸來比較分水嶺兩邊療效差異的程度。

```

Coefficients...
      Class
Variable      有效
=====
公分劃分2=大於    -2.1372
Intercept      0.9585

Odds Ratios...
      Class
Variable      有效
=====
公分劃分2=大於    0.118

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      146      72.6368 %
Incorrectly Classified Instances    55      27.3632 %
Kappa statistic                     0.2162
Mean absolute error                 0.3987
Root mean squared error             0.4481
Relative absolute error             91.6894 %
Root relative squared error         96.1656 %
Total Number of Instances          201

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.971   0.797   0.723     0.971   0.829     0.291   0.552    0.700    有效
0.203   0.029   0.765     0.203   0.321     0.291   0.552    0.402    無效
Weighted Avg.  0.726   0.552   0.736     0.726   0.667     0.291   0.552    0.605

```

圖 4.10 cTACE 以 9.3 公分劃分邏輯斯迴歸結果

資料來源:本研究產出

由邏輯斯迴歸產出的勝算比如圖 4.10 可以發現腫瘤大小大於 9.3 公分有效的機率是腫瘤大小小於等於 9.3 公分的 0.118 倍，反過來說比較好理解，也就是腫瘤較小治療有效機率是腫瘤大的治療有效機率的 8.475 倍。而此模型的準確度為 72.6%，假陽率為 0.552，精確率為 0.736，召回率為 0.726 預測能力看 ROC 曲線下面積及 PRC 曲面下面積，分別為 0.552 及 0.605。

整理以上產出結果，以決策樹分類發現傳統療法受限於腫瘤大小劃分界限為 9.3 公分，以均值檢定亦發現相同結果，治療無效的腫瘤大小均值顯著較大，又以卡方檢定對 9.3 公分劃分後的數據進行檢定，發現療效上小於等於 9.3 公分的腫瘤大小具有顯著成功療效，而進一步比對兩種機率差異，由邏輯斯迴歸計算顯示小於等於 9.3 公分治療成功機率是大於 9.3 公分腫瘤

大小的 8.475 倍，故建議若肝癌癌症分級 A、B 及的病患，若腫瘤大小小於等於 9.3 公分可使用傳統療法，可取得不錯的療效，若腫瘤大於 9.3 公分以傳統療法治療成功機率僅有小於等於 9.3 公分治療成功機率的 0.118，此結果提供醫生選用療法上的參考。

4.3 微球療法限制與適用細分

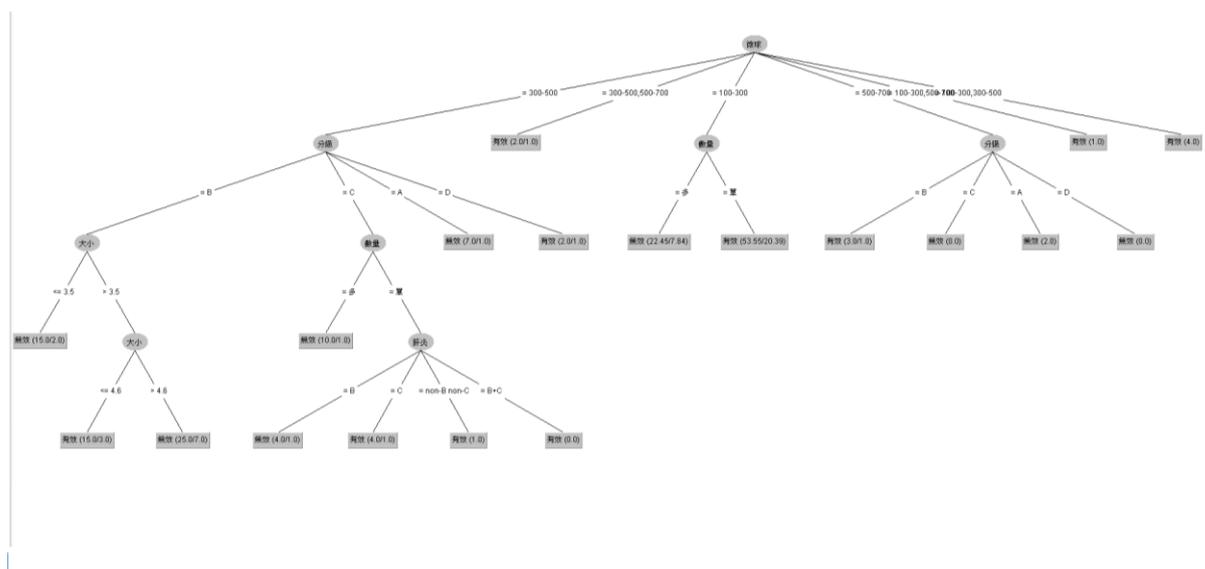


圖 4.11 微球療法決策樹分類

資料來源:本研究產出

=== Summary ===

Correctly Classified Instances	101	59.0643 %
Incorrectly Classified Instances	70	40.9357 %
Kappa statistic	0.18	
Mean absolute error	0.4625	
Root mean squared error	0.5107	
Relative absolute error	93.193 %	
Root relative squared error	102.512 %	
Total Number of Instances	171	
Ignored Class Unknown Instances	201	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.590	0.409	0.548	0.590	0.568	0.180	0.617	0.441	有效
	0.591	0.410	0.632	0.591	0.611	0.180	0.586	0.435	無效
Weighted Avg.	0.591	0.410	0.594	0.591	0.591	0.180	0.600	0.438	

圖 4.12 微球療法決策樹模型情形

資料來源:本研究產出

接下來本研究以 4.5 決策樹對微球療法進行分類產出結果如圖 4.11，從決策樹中可以發現微球療法採用混用不同大小微球的效果都是有成效的，從資料中去探討，發現使用混球的資料數量都非常少，如前面對全體病患劃分新舊療法使用邏輯斯迴歸比較時所發現的一樣，從資料中看，本研究所蒐集之資料，使用混球 100-300,300-500(μm)的資料僅有 4 筆且皆為有效，而使用 100-300,500-700(μm)只有一筆且有效，300-500,500-700(μm)只有兩筆一有效一無效，依據決策樹判斷結果此三種混球皆為有效，但礙於使用混球的病例實在有限，因此是否混用不同大小微球有更好的療效需要再進行長期觀察需要更多的臨床資料再進行分析，故本研究進一步先將混球加以排除，再次以決策樹做為分類器進行分類。

模型的適用狀況如圖 4.12，模型的準確度為 59.1%，假陽率為 0.41，精確率為 0.594，召回率為 0.591 預測能力看 ROC 曲線下面積及 PRC 曲面下面積，分別為 0.6 及 0.438。準確度偏低，預測能力也不太理想，原因也可能跟混球資料量太少有關聯，將混球先行排除後再來做更進一步比較。

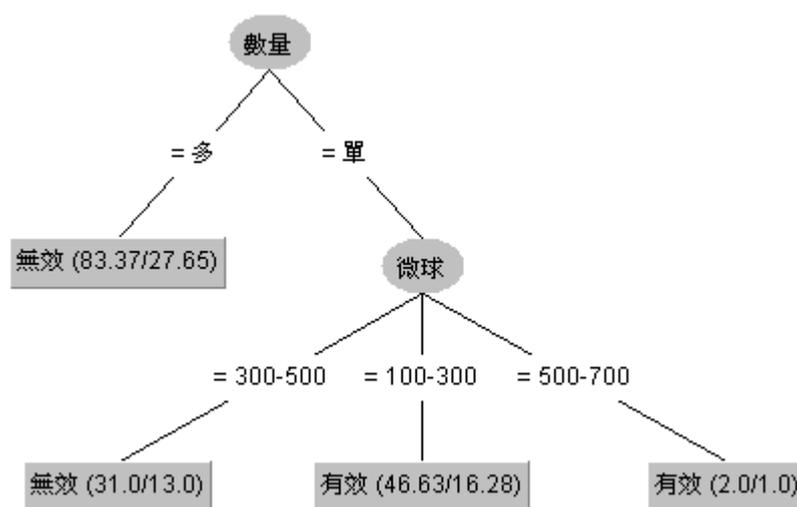


圖 4.13 微球療法去除混球後決策樹分類

資料來源:本研究產出

=== Summary ===

Correctly Classified Instances	106	65.0307 %
Incorrectly Classified Instances	57	34.9693 %
Kappa statistic	0.2576	
Mean absolute error	0.4692	
Root mean squared error	0.4938	
Relative absolute error	95.0973 %	
Root relative squared error	99.4221 %	
Total Number of Instances	163	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.389	0.143	0.683	0.389	0.496	0.282	0.586	0.522	有效
	0.857	0.611	0.639	0.857	0.732	0.282	0.586	0.609	無效
Weighted Avg.	0.650	0.404	0.659	0.650	0.628	0.282	0.586	0.570	

圖 4.14 微球療法去除混球後決策樹模型情形

資料來源:本研究產出

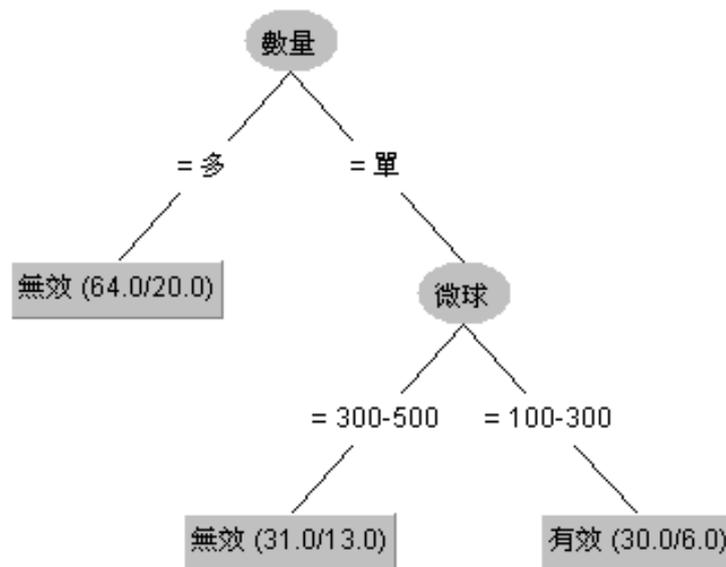


圖 4.15 微球療法去除混球去除 500-700 μ m 後決策樹分類

資料來源:本研究產出

```

=== Summary ===

Correctly Classified Instances      88          70.4 %
Incorrectly Classified Instances    37          29.6 %
Kappa statistic                    0.3868
Mean absolute error                 0.414
Root mean squared error             0.4753
Relative absolute error             83.4028 %
Root relative squared error         95.3878 %
Total Number of Instances          125

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                -----  -----  -
                0.509   0.132   0.763     0.509   0.611     0.408   0.639   0.585   有效
                0.868   0.491   0.678     0.868   0.761     0.408   0.639   0.636   無效
Weighted Avg.   0.704   0.328   0.717     0.704   0.693     0.408   0.639   0.613

```

圖 4.16 微球療法去除混球去除 500-700 μm 後決策樹模型情形

資料來源:本研究產出

排除混球後，以決策樹再次分類，結果如圖 4.13，可以發現模型歸納分做兩層節點，第一層節點是以腫瘤數量多寡來劃分，而第二節點為為球大小來劃分，相較前面混球存在的模型樹的規模縮減很多，但是從模型情況來看，比較圖 4.12 與圖 4.14 可以發現模型準確度由 59.1% 變為 65%，各項模型指標都有部分的優化，以此推測可能受到部分化分因子資料量過少導致預測能力下降的情況。

本研究再次檢視資料數量上的問題，發現微球大小 500-700(μm) 資料量也偏低，僅 5 筆臨床案例資料，故本研究由原本 130 筆資料中排除微球大小 500-700(μm) 5 筆資料再次以決策樹進行分類得到結果如圖 4.15，分類未受排除的影響，仍以腫瘤數量為第一決策分支點，而第二分支點也維持以微球大小進行劃分，影響最大的是模型適用情形，比較圖 4.14 及圖 4.16 可以發現模型準確度再次上升，由 65% 上升為 70.4%，而整體其他指標也大幅上升，再次驗證屬性資料多寡確實影響模型準確程度。

本研究將以最後去蕪存菁後的 125 筆資料進行更進一步的驗證，首先針對腫瘤多寡對療效上的差異進行卡方檢定，並已邏輯斯迴歸計算差異程度，接著在依照決策樹分類器分類的結果對微球大小進行卡方檢定以及邏輯斯迴歸，以提供更完整的結果於醫療上參考。

有效無效 * 腫瘤數量 交叉表

			腫瘤數量		總和
			單顆	多顆	
有效無效	無效	個數	24	44	68
		在有效無效之內的	35.3%	64.7%	100.0%
		在腫瘤數量之內的	39.3%	68.8%	54.4%
	整體的 %	19.2%	35.2%	54.4%	
	有效	個數	37	20	57
		在有效無效之內的	64.9%	35.1%	100.0%
在腫瘤數量之內的		60.7%	31.3%	45.6%	
整體的 %	29.6%	16.0%	45.6%		
總和	個數	61	64	125	
	在有效無效之內的	48.8%	51.2%	100.0%	
	在腫瘤數量之內的	100.0%	100.0%	100.0%	
	整體的 %	48.8%	51.2%	100.0%	

卡方檢定

	數值	自由度	漸近顯著性 (雙尾)	精確顯著性 (雙尾)	精確顯著性 (單尾)
Pearson卡方	10.887 ^a	1	.001		
連續性校正 ^b	9.734	1	.002		
概似比	11.046	1	.001		
Fisher's精確檢定				.001	.001
線性對線性的關連	10.800	1	.001		
有效觀察值的個數	125				

a. 0格 (0%) 的預期個數少於 5。最小的預期個數為 27.82。

b. 只能計算 2x2 表格

圖 4.17 微球療法去除混球去除 500-700 μ m 後腫瘤數量卡方

資料來源:本研究產出

本研究將以最後去蕪存菁後的 125 筆資料以卡方檢定的方式驗證是否差異的存在，卡方檢定的結果如圖 4.17，單顆腫瘤治療無效比例為 39.3%，而有效比例為 60.7%；多顆腫瘤無效比例為 68.8%，而有效比例為 31.3%，且由卡方檢定為 0.01 具有顯著差異，驗證微球療法對單顆腫瘤有比較保握的優勢，若要針對多顆腫瘤以微球療法治療，本賢就將繼續探討會有多少風險的產生同時，腫瘤多寡對於治療成效上有多大的有效機率差異。

```

Coefficients...
      Class
Variable 有效
=====
數量=單   1.2213
Intercept -0.7885

Odds Ratios...
      Class
Variable 有效
=====
數量=單   3.3918

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      81          64.8  %
Incorrectly Classified Instances    44          35.2  %
Kappa statistic                    0.2945
Mean absolute error                0.456
Root mean squared error            0.4797
Relative absolute error            91.8693  %
Root relative squared error        96.272  %
Total Number of Instances         125

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.649   0.353   0.607     0.649   0.627     0.295   0.579    0.502    有效
0.647   0.351   0.688     0.647   0.667     0.295   0.579    0.594    無效
Weighted Avg.  0.648   0.352   0.651     0.648   0.649     0.295   0.579    0.552

```

圖 4.18 微球療法去除混球去除 500-700 μ m 後腫瘤數量邏輯斯迴歸

資料來源:本研究產出

本研究將以邏輯斯迴歸計算腫瘤數量對治療有效無效的勝算比，結果如圖 4.18 可以發現單顆腫瘤治療會比多顆腫瘤治療多 3.39 倍的成功機率，也因此提供醫師在治療多顆腫瘤評估風險參考，但可能因為數據量仍有所限制，雖然模型的精準度約 64.8%，但預測能力並不是非常理想，可能待未來具備更加豐富之臨床資料以提高模型預測的情形。

進一步，本研究做第二分支點的驗證，先將資料中腫瘤數量為多的資料加以排除，剩下腫瘤數量為單顆的資料 61 筆，在進行不同微球大小的比較。

微球大小 * 有效無效 交叉表

			有效無效		總和
			無效	有效	
微球大小 100-300	個數		6	24	30
	在微球大小之內的		20.0%	80.0%	100.0%
	在有效無效之內的		25.0%	64.9%	49.2%
	整體的 %		9.8%	39.3%	49.2%
300-500	個數		18	13	31
	在微球大小之內的		58.1%	41.9%	100.0%
	在有效無效之內的		75.0%	35.1%	50.8%
	整體的 %		29.5%	21.3%	50.8%
總和	個數		24	37	61
	在微球大小之內的		39.3%	60.7%	100.0%
	在有效無效之內的		100.0%	100.0%	100.0%
	整體的 %		39.3%	60.7%	100.0%

卡方檢定

	數值	自由度	漸近顯著性 (雙尾)	精確顯著性 (雙尾)	精確顯著性 (單尾)
Pearson卡方	9.256 ^a	1	.002		
連續性校正 ^b	7.730	1	.005		
概似比	9.583	1	.002		
Fisher's精確檢定				.004	.002
線性對線性的關連	9.105	1	.003		
有效觀察值的個數	61				

a. 0格 (0%) 的預期個數少於 5。最小的預期個數為 11.80。
b. 只能計算 2x2 表格

圖 4.19 微球療法去除混球去除 500-700 μ m 後單腫瘤微球大小卡方

資料來源:本研究產出

資料進行劃分，劃分為不同微球大小，100-300(μ m)30 筆，300-500(μ m)31 筆，共計 61 筆資料，進行卡方檢定，結果如圖 4.19，100-300(μ m)治療無效比例為 20%，而有效比例為 80%；300-500(μ m)治療無效比例為 58.1%，而有效比例為 41.9%，其卡方檢定結果為 0.02，具顯著差異，也就是確實 100-300(μ m)在單顆腫瘤中的治療效果較 300-500(μ m)來的好。進一步也探討療效的差異程度。

```

Coefficients...
      Class
Variable 有效
=====
微球=100-300  1.2272
Intercept    -0.603

Odds Ratios...
      Class
Variable 有效
=====
微球=100-300  3.4115

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      81          64.8 %
Incorrectly Classified Instances    44          35.2 %
Kappa statistic                    0.2761
Mean absolute error                 0.4624
Root mean squared error             0.4844
Relative absolute error             93.1613 %
Root relative squared error         97.2293 %
Total Number of Instances          125

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
      0.491   0.221   0.651     0.491   0.560     0.284   0.549    0.490    有效
      0.779   0.509   0.646     0.779   0.707     0.284   0.549    0.563    無效
Weighted Avg.  0.648   0.377   0.649     0.648   0.640     0.284   0.549    0.529

```

圖 4.20 微球療法去除混球去除 500-700 μm 後單腫瘤微球大小邏輯斯迴歸

資料來源:本研究產出

以邏輯斯迴歸對微球大小 100-300(μm)、300-500(μm)進行勝算比的計算，結果如圖 4.20，100-300(μm)治療成功機率是 300-500(μm)的 3.412 倍，同樣此模型精準度約為 64.8%，但其預測能力仍有待更加強，以上研究結果提供未來研究參考使用。

第五章 結論

5.1 研究發現與討論

- 1、本研究使用邏輯斯迴歸發現傳統療法療效有效機率是微球療法的1.0669倍，但有另一項指標更為明顯就是微球100-300,300-500(μm)，有效機率是其他微球的4千多萬倍，進一步探討，發現其中資料只有4筆且全部都是有效，因此造成勝算變得超級大，因此造成此超大勝算比產出，微球100-300,300-500(μm)是否確實具有非常高的有效機率，待未來更多使用此治療的病患以進行驗證。同時也發現腫瘤數量單顆治療有效機率是多顆的2.8719倍，此項結果也呼應前人研究。
- 2、本研究使用決策樹發現，其中使用混球栓塞皆產出有效的結果，但是，情況如同邏輯斯迴歸的結果，因使用混球的病患資料都太少，故這個點都將待未來資料量更多時進一步分析。
- 3、傳統方法在對治癌症分級A、B級的病患上具有較強的優勢，但對於腫瘤較大等情形時較難對付，此結果需進一步取得C、D級病患以傳統栓塞治療的資料進行比對。
- 4、癌症分級A、B可使用比較有優勢的傳統栓塞，而出現轉移或較大顆的癌症分級C、D級病患，可使用微球栓塞，而微球療法小型微球對單顆腫瘤具有較大的優勢，但目前微球栓塞以卡方統計結果來看成功比率約為45%，其原因有待未來研究發現，是否是醫療手法技術上問題，又或者是微球本身的問題，有待進一步提升加強。
- 5、本研究發現整體影響最大的主要是腫瘤多寡以及腫瘤大小，且資料中發現有部分資料會採離建議值差異較大的治療方式，導致產出變異加大，與醫師討論，認為可能是未考慮到病人已發生之先天條件，故也建議進行栓塞前可以多從病人先天條件下進行考量，以達到更好的治療效果。
- 6、最後我們從微球療法以決策樹當作分類器中可以發現，模型的準確度以及預測能力會隨著被劃分出來的決策分支點所使用屬性資料的資料量而變化，當資料量較不足時會造成準確度及預測能力都下滑，故未來研究應注意決策點屬性資料數量問題，若不足數量可能無法初顯良好的預測能力，本研究所遇到之混球及500-700(μm)微球大小就是遇到這種狀

況，期望未來若這些臨床應用更加普及能取得更大量的資料，可在針對上述所說標的做更深入的研究及探討。

5.2 研究結論及未來研究建議

5.2.1 結論

本研究對應研究目標，將結論分做四部分近行探討，本研究結果發現依據台中榮總肝癌病患真實病歷資料，得出不同於前人研究之結果，以統計方面已顯著顯示傳統動脈栓塞療法有較優的療效表現，進一步，本研究探尋資料本身及與專業醫師討論得知目前傳統栓塞僅用於癌症分級 A、B 級病患，故將癌症分級 C、D 及病患暫時排除再次進行統計，所得結果如總體結果，故得知傳統栓塞於癌症分級 A、B 級病患治療上具較大優勢，而癌症分級 C、D 級病患則可選擇微球療法，但療效仍有待加強提升。

本研究進一步以邏輯斯迴歸探討兩種藥效的差異程度，結果顯示傳統療法成功療效機率是微球療法的 1.0669 倍，探究原因可能因為傳統療法在於癌症分級 A、B 級患者治療上具有較強的優勢，而微球療法是適用領域較廣的關係，故造成療效程度上未有明顯差異；進一步排除癌症分級 C、D 級患者，針對癌症分級 A、B 級患者以邏輯斯迴歸計算出的結果得知傳統療法有效的機率是微球療法的 2.57 倍，因此建議癌症分級 A、B 級患者以傳統療法為主。

本研究分別探討兩種療法所受到的限制，以決策樹輔助做為分類器，先針對傳統療法進行分類，結果發現，傳統療法受限於腫瘤大小，以 9.3 公分為分水嶺，小於等於 9.3 公分有效，大於則無效，本研究以統計均值檢定檢視，確實有療效的腫瘤大小平均顯著較小，進一步以卡方檢定對 9.3 公分分水嶺對兩邊進行檢驗，小於等於 9.3 公分腫瘤大小的有療效比例為 72.3%，大於 9.3 公分腫瘤大小的有療效比例為 23.5%，卡方檢定為顯著差異；以邏輯斯迴歸檢視差異程度，腫瘤較小治療有效機率是腫瘤大的治療有效機率的 8.475 倍，以此結論建議癌症分級 A、B 級患者當腫瘤大於 9.3 公分時不建議使用傳統療法。

對於微球療法，採用決策樹分類後發現混球都歸類為有效但從資料去看卻發現資料量都非常少，也因此先將混球加以排除再次以決策樹進行分

類，進一步，本研究為確保資料皆有一定數量因此也再次檢查分類劃分初的支點，發現微球 500-700(μm)也只有 5 筆資料，故也先加以排除，再次以決策樹劃分，而從模型適用情形來看去除這些因子後模型的準確度以及模型預測能力都有所提升，而決策樹展出來顯示，微球療法受制於腫瘤的數量對於多顆腫瘤無效，而對於單顆腫瘤，100-300(μm)有效，進一步，本研究以卡方檢定分別對腫瘤數量對療效的影響，結果顯示確實單顆腫瘤有效比例為 60.7%有顯著高於多顆腫瘤有效的比例 31.3%，以邏輯斯迴歸比較差異程度發現使用微球療法單顆腫瘤治療會比多顆腫瘤治療多 3.39 倍的成功機率。單顆腫瘤中在以卡方檢定驗證 100-300(μm) 有效比例為 80%顯著高於 300-500(μm)治療有效的比例 41.9%，以邏輯斯迴歸比較差異程度，結果為 100-300(μm)治療成功機率是 300-500(μm)的 3.412 倍。以此研究結果提供肝癌栓塞選擇療法上的一項參考，期望能取得成功機率更大的療法來進行治療，以提供更可靠的醫療使病患增加康復機率。

5.2.2 未來建議

- 1、本研究利用台中榮總 372 位肝癌病患病歷資料，故產出之結果僅能代表台中榮總目前肝癌動脈栓塞之情形，並對此領域研究提供一個可參考的模型，未來學者可跨足多家代表性醫療機構，取得更完整資料，提出代表範圍更大之研究成果，以利於提高肝癌治療上成功的機率。
- 2、本研究資料上取得有所受限，且肝癌患者數量雖然較其他癌症多，但資料量仍有瓶頸，未來學者可累積更大的資料量，期望能更加提高模型預測能力，輔助醫療應用更加精準。
- 3、本研究所蒐集之屬性因子，主要是以前人文獻認為最大影響因子，同時參考主治醫師給予之專業建議，以及醫療機構方便追蹤之因子，未來研究者可更多元蒐集因子，包括生理、作息等因子，從中找到更全面的法則，以利於預防。
- 4、對於藥物優劣的比較還有其他的指標，如藥物動力學的 AUC 跟峰濃度，副作用情形，還有用藥後肝功能情形，存活情況，但這些預後指標需要更長的時間追蹤，以及目前追蹤設備及制度皆有待發展有待更加完備，待未來資料更加完備本研究提出之項目可供研究參考使用，但長期追蹤

要注意其他環境可控不可控等因素介入的問題，以免產生非因果的假關係出現。

參考文獻

中文文獻

1. 醫學界消化頻道(民 95 年 03 月 04 日)。健康每日頭條【三院科普·讀懂肝功能 ALT，AST——人體肝臟功能的晴雨表】。取自 <https://kknews.cc/zh-tw/health/4nm8mg.html>。
2. 世界衛生組織(WHO)(2014)。全球癌症報告。
3. 加利福尼亞太平洋醫療中心 Sutter health CPMC. (2010)。使用藥物緩釋顆粒的經導管動脈化學藥物栓塞治療 (TACE)。
4. 李治、李國琳(2013)。C4.5 和 CART 算法在醫學數據挖掘中的對比研究。電子技術與軟件工程 *Electronic Technology & Software Engineering*，**10**，47-48。
5. 李宗佑(2015)。應用資料採礦於肝膿瘍患者罹患癌症之研究。(碩士論文)。東海大學工業工程與經營資訊研究所，台中市。
6. 何明霖。彰化基督教醫院網頁【治療效果判定】。取自 <http://www2.cch.org.tw/lungcancer/rescist.htm>
7. 何子銘等(2006)。運用三種資料探勘方法預測子宮頸癌存活情形之比較。台灣家庭醫學雜誌 **16:3 期(2006/09)**: 192-203。
8. 周辰熹(2013)。臨床藥物動力學試驗之基本概念。台灣生物精神醫學暨神經精神藥理學學會，台北市。
9. 胡嘉仁(2011)。運用三種資料探勘方法預測原發性肝癌患者存活情形之比較。(碩士論文)。嘉南藥理科技大學醫療資訊管理研究所，台南市。
10. 施品光(2009)。類神經網路與癌症基因統計資訊應用於白血病醫學文件分類之研究。(碩士論文)。國立高雄應用科技大學電機工程系，高雄市。
11. 陳麗帆等(2009)。使用比較資料探勘演算法預測非小細胞肺癌患者預後因子、存活情形及其效能。腫瘤護理雜誌 **9 卷 2 期(2009/12)**: 33-47。
12. 陳嘉邦(2010)。肝癌的治療方式~肝癌肝動脈栓塞治療(DC Bead TACE / 鉍 90)。癌症防治季刊 **2010:第九期 肝癌**。彰化基督教醫院影像醫學部，彰化縣。
13. 徐仁徽(2013)。植基於資料探勘演算法在人類大腸癌資料集標靶基因篩選研究中採用倒傳遞類神經網路進行研究。(碩士論文)。中興大學基因體暨生物資訊學研究所，台中市。
14. 逍遙文(民 101 年 10 月 22 日)。逍遙文工作【室交叉驗證 (Cross Validation)】。取自 <https://cg2010studio.com/2012/10/22/%E4%BA%A4%E5%8F%89%E9%A9%97%E8>

%AD%89-cross-validation/。

15. 滄海書局(2006)。PPT 教材【15.9 羅吉斯迴歸】。取自 stat.thu.edu.tw/stat-member/teacher/chlin-file/regression/logit%20regression.ppt。
16. 張斐章等(2001)。回饋式類神經網路於河川流量推估之應用。《農業工程學報》47(2):32-39。
17. 魯惠雲(2004)。肝癌資料庫之建立及特徵分析。(碩士論文)。國立臺北大學企業管理學系，新北市。
18. 蔡蕙如、柯明中、張偉斌、劉德明(2007)。應用類神經網路與分類迴歸樹於肝癌分類模式。《北市醫學雜誌》4卷8期:658-667。
19. 衛生福利部統計處(2016)。民國 105 年台灣 10 大死因統計。
20. 維基百科【阿黴素】。取自 <https://zh.wikipedia.org/wiki/%E9%98%BF%E9%BB%B4%E7%B4%A0>。
21. 趙美施(2018)。肝動脈血管栓塞術。彰化基督教醫院內科衛教護理，彰化縣。
22. 龔巧英等(2011)。單純肝動脈栓塞化療和三維適形放療聯合肝動脈栓塞化療對原發性肝癌的療效比較。《實用癌症雜誌》，2011年26卷06期:634-636頁。
23. 劉鐘軒、蔡正中、陳海雄(2013)。肝癌的診斷及治療最新發展。《內科學誌》2013:24:85-94。健仁醫院胃腸肝膽科，高雄榮民總醫院一般外科、胃腸肝膽科，高雄市。
24. 劉怡萱(2011)。利用全民健康保險資料庫分析台灣愛滋病患者癌症及自體免疫疾病發生率之世代研究。(碩士論文)。國立陽明大學 International Health Program，台北市。
25. 博客專家柚子皮(民 105 年 06 月 30 日)。皮皮 BLOG【分類模型評估之 ROC-AUC 曲線和 PRC 曲線】。取自 <https://blog.csdn.net/pipisorry/article/details/51788927>。

英文文獻

1. Do Seon Song, Jong Young Choi, corresponding author Sun Hong Yoo et al. (2013). DC Bead Transarterial Chemoembolization Is Effective in Hepatocellular Carcinoma Refractory to Conventional Transarterial Chemoembolization: A Pilot Study. *Gut Liver*. 2013 Jan, 7(1), 89–95.
2. Franco M. Muggia, Patricia S. Braly, Mark F. Brady et al. (2000) Phase III Randomized Study of Cisplatin Versus Paclitaxel Versus Cisplatin and Paclitaxel in Patients With Suboptimal Stage III or IV Ovarian Cancer: A Gynecologic Oncology Group Study. *J Clin Oncol*. 2000 Jan, 18(1), 106-115.
3. Han, J., & Kamber, M.(2006). *Data Mining, Southeast Asia Edition: Concepts and Techniques*. San Mateo, Calif :Morgan Kaufmann .
4. Hee Chul Nam, Bohyun Jang, and Myeong Jun Song. (2016). Transarterial chemoembolization with drug-eluting beads in hepatocellular carcinoma. *World J*

Gastroenterol. 2016 Oct 28, 22(40), 8853–8861.

5. Kohavi, Ron (1995). A Study of CrossValidation and Bootstrap for Accuracy Estimation and Model Selecti. *IJCAI'95 Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2* (pp.1137-1143). San Francisco, CA, USA. Morgan Kaufmann. ISBN:1-55860-363-8
6. Koorosh Ashrafi et al. (2017). Characterization of a novel intrinsically radiopaque Drug-eluting Bead for image-guided therapy: DC Bead LUMI. *J Control Release.* 2017 Mar 28, 250, 36-47.
7. Llovet JM1, Burroughs A, Bruix J. (2003). Hepatocellular carcinoma. *Lancet.* 2003 Dec 6, 362(9399), 1907-1917.
8. Lammer J, Malagari K, Vogl T (2010) Prospective randomized study of doxorubicin-eluting-bead embolization in the treatment of hepatocellular carcinoma: results of the PRECISION V study. *Cardiovasc Intervent Radiol* 33 (2010): 41-52.
9. Lawrence H. Schwartz,1 Saskia Litière,2 Elisabeth de Vries et al. (2016). RECIST 1.1 – Update and Clarification: From the RECIST Committee. *Eur J Cancer.* 2016 Jul, 62, 132–137.
10. Llovet JM1, Bruix J. (2003). Systematic review of randomized trials for unresectable hepatocellular carcinoma: Chemoembolization improves survival. *Hepatology.* 2003 Feb, 37(2), 429-442.
11. Mosby. (2016). *Mosby's Dictionary of Medicine, Nursing & Health Professions 10th Edition*. Philadelphia, United States: Elsevier - Health Sciences Division.
12. Song MJ, Park C-H, Kim JD. (2011). Drug-eluting bead loaded with doxorubicin versus conventional Lipiodol-based transarterial chemoembolization in the treatment of hepatocellular carcinoma: a case-control study of Asian patients. *Eur J Gastroenterol Hepatol* 23. Advance online publication. doi: 10.1097/MEG.0b013e328346d505.
13. Song MJ1, Chun HJ, Song DS, Kim HY et al. (2012). Comparative study between doxorubicin-eluting beads and conventional transarterial chemoembolization for treatment of hepatocellular carcinoma. *J Hepatol.* 2012 Dec, 57(6), 1244-1250.
14. Varela M, Real MI, Burrel M et al. (2007). Chemoembolization of hepatocellular carcinoma with drug eluting beads: efficacy and doxorubicin pharmacokinetics. *J Hepatol.* 46(3). Advance online publication. doi: <https://doi.org/10.1016/j.jhep.2006.10.020>.