

# 機器學習模型應用於重症肝硬化病人之死亡預測 - 長期追蹤報告

蕭嘉瑩<sup>1</sup> 馬瑞菊<sup>2</sup> 鄭群翰<sup>1</sup> 曾葭茵<sup>3</sup> 李佳欣<sup>2</sup>  
鄭婉如<sup>2</sup> 郭豐吉<sup>1</sup> 蘇珉一<sup>1</sup>

台東馬偕紀念醫院 <sup>1</sup> 內科部 <sup>2</sup> 護理部 <sup>3</sup> 醫學教育研究部

## 摘 要

風險預測管理在臨床研究和醫療照護中意義重大，尤其肝硬化病人與一般重症病人相比預後較差，但目前尚無針對肝硬化病人長期死亡風險的預測指標，因此本研究試圖使用機器學習方法：Bagging、Adaboost、Support Vector Machine (SVM)、Random forest 建構預測模型。研究納入 2013 年 7 月 1 日至 2018 年 12 月 31 日就醫的 528 名個案並追蹤至 2020 年 6 月，並與目前臨床肝硬化嚴重度量表：Child-Pugh Score (CTP)、MELD Score (Original)、MELDNa Score、MELD Score 以及重症單位常用的 The Acute Physiology and Chronic Health Evaluation (APACHE II) 進行比較，結果顯示機器學習模型皆優於現有的臨床評估量表。預測 1 個月死亡模型以 Random forest 表現最佳 (準確率：0.836；AUC: 0.844)；預測 12 個月內死亡模型以 SVM 表現最佳 (準確率：0.826；AUC: 0.869)；預測 12 個月後死亡模型以 Adaboost 表現最佳 (準確率：0.759；AUC: 0.760)，充分證實機器學習模型在預測肝硬化病人的長期死亡率方面具有應用價值。

**關鍵詞：**機器學習 (Machine learning)  
肝硬化 (Cirrhosis)  
死亡預測 (Predictors of death)  
長期追蹤 (Long-Term Follow up)

## 前 言

肝硬化是常見的慢性疾病，2019 年在全球常見死因排行第 11 名，造成 131.5 萬人死亡<sup>1</sup>。當肝臟受到病毒、藥物、酒精等原因反覆傷害，炎症反應超過肝臟修復能力肝臟就會逐漸纖維化，嚴重纖維化末期即是肝硬化<sup>2</sup>。肝硬化產生急性惡化容易導致靜脈曲張出血、腹水、

肝腎症候群、自發性細菌性腹膜炎和敗血症等併發症，需要重症照護的機率極高<sup>3</sup>。且隨著肝硬化病程進展至末期，完全根治的治療方法僅有肝臟移植<sup>4-5</sup>。

社會逐漸朝向精密醫學時代發展，風險預測管理在臨床研究和醫療照護中逐漸成為主流，尤其與一般重症病人相比肝硬化病人預後較差<sup>6</sup>，肝硬化病人因併發症而在 12 個月後死

亡的機率高達 57%<sup>7</sup>。因此若能夠提前預測肝硬化病人的長期死亡率，對輔助醫療決策將具有臨床價值<sup>8</sup>。雖然目前臨床已有數個用於預測肝硬化風險的傳統評分工具，例如：Child-Pugh Score (CTP)、MELD Score (Original)、MELDNa Score、MELD Score，但多用於短期死亡率預測<sup>9</sup>。而在重症使用最廣泛的 The Acute Physiology and Chronic Health Evaluation (APACHE II)，其預測死亡率可能比上述肝硬化風險傳統評分工具更準確，但實際上其並非來自肝硬化病人群體<sup>10</sup>。故預測程度有限<sup>11-12</sup>。

機器學習 (Machine learning, ML) 是自我學習的演算方法<sup>13-14</sup>，近年醫療資訊被整合為大數據，使 ML 更容易被應用在醫療領域，而且比傳統評分工具及傳統回歸方法更具優勢<sup>15</sup>，但目前關於其在重症肝硬化病人的長期預後預測能力研究極少且未得驗證。因此本研究追蹤肝硬化病人長期存活情形，用於建構 ML 模型進行預測，並與傳統評分系統比較預測長期結果的能力。

## 研究材料與方法

### 一、研究對象及場所

採電子病歷回溯性研究，收案時間從 2013 年 7 月 1 日至 2018 年 12 月 31 日，使用來自東

部某區域教學醫院內科加護病房主診斷包含肝硬化診斷之個案，一共收集 528 筆資料並長期追蹤其死亡與否 (追蹤至 2020 年 6 月)，收集流程如圖一所示。

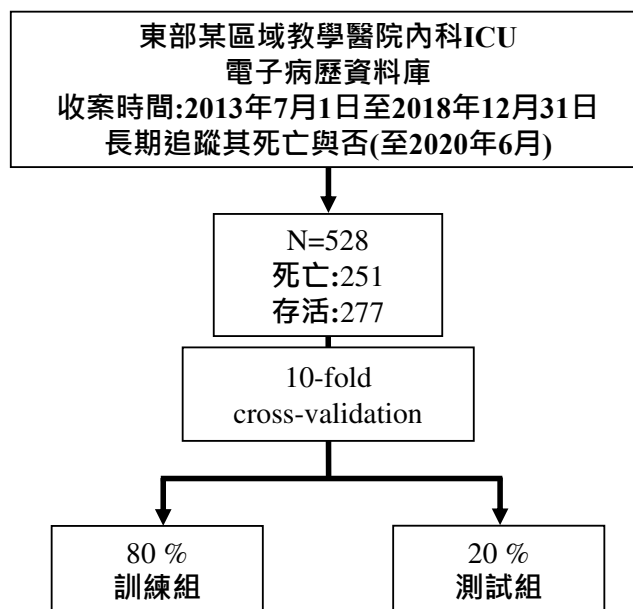
### 二、資料處理與分析

使用 R Studio Cloud (版本 3.6.1) 軟件進行統計分析。資料庫包含 34 個變項，生理數據變項擷取病人入院 24 小時內第一筆。

首先將數據透過聚類熱圖 (Cluster heat map) 以圖形化形式轉換為資訊並將之用於決策<sup>16</sup>。熱圖使數據可視化並結合多種圖形，能夠以顏色差異呈現數值落差變化，助於觀察隱藏在數據集中的趨勢；相似的行和列會以聚類方式進行換位顯示結構；垂直和水平邊緣上是分層聚類樹，表達數據間的相似性<sup>17-18</sup>。

接著將資料庫隨機亂數切割成 10 等份，使用十折交叉驗證 (Cross-validation, CV) 來評估預測性能，其中 80% 作為訓練組建構模型，剩餘 20% 為測試組用以驗證模型效果，並選擇十折中結果最佳的一次來呈現。

機器學習模式包含監督式學習、非監督式學習以及強化學習，本研究使用監督式學習的 Bagging、Adaboost、Support Vector Machine (SVM)、Random forest 四個方法，分別建構三



圖一：數據收集及分析流程圖。

個不同時間線的死亡預測模型：1 個月內死亡、12 個月內死亡、12 個月後死亡。最終再與傳統評分工具：APACHE II、CTP、MELD Score (Original)、MELDNa Score、MELD score 進行比較。

模型建構完成後，計算各模型的準確率 (Accuracy)、靈敏度 (Sensitivity)、特異度 (Specificity)、精密度 (precision)、F1 分數 (F1 Score)、ROC 曲線下面積 (Area under the ROC curve, AUC) 評估效果，AUC 是用於評估診斷分類模型性最常用的工具<sup>28</sup>。由於傳統指標在臨床上已被充分驗證及廣泛使用，因此採用 AUC 與機器學習模型進行比較。

### 三、統計方法

#### (一) Bagging

Bagging 用於使用相同基礎學習器生成多個結果，平均各個分類器結果再以投票法進行最終預測<sup>19-20</sup>。

#### (二) AdaBoost

Boosting 其思想是通過結合許多弱分類器 (weak classifier) 以建立預測能力高的分類器，其應用十分廣泛，AdaBoost 便是 Boosting 算法的延伸<sup>20</sup>，該算法根據每次疊代中弱分類器的分類結果改變訓練樣本權重，理念為加強學習弱分類器分類錯誤的樣本，從而將弱分類器提升為強分類器<sup>21-22</sup>。

#### (三) SVM

SVM 是解決最佳化線性約束條件問題的算法，通過建立數據之間的最佳分隔邊界來解決此數據集中的二次優化問題<sup>23-24</sup>。意即將資料樣本化為空間中的點，從空間平面中找出一條線，將資料樣本分開以進行分類<sup>25</sup>。

#### (四) Random forest

利用隨機子空間方法構建多棵決策樹<sup>26</sup>，最終使用集成獲得預測結果。由於組成的預測變量是樹狀結構，並且每棵樹都是隨機切割，因此稱作隨機森林<sup>27</sup>。

## 結 果

### 一、描述性統計

數據詳見表一，資料庫中病人一個月內死亡率 28.4% (150/528)；1 個月以後至 12 個月內死亡率 10.2% (54/528)，12 個月後整體死亡率為 47.5% (277/528)。

### 二、聚類熱圖

將數據繪製成帶有圖例顏色條、列側註釋和行側註釋的熱圖 (圖二)。顏色越紅值越大，顏色越黃值越小。由圖可以得知死亡的族群多集中在上半部，檢驗數值白血球、動脈血氧分壓、肌酸酐越大者越容易死亡，有吸菸病史、肝硬化病史、酒精病史的族群雖然死亡率也極高，但在該群人中年紀較輕的病人相對存活可能性較高。

### 三、機器學習模型

透過準確率比較肝硬化病人 1 個月內、12 個月內、12 個月後模型預測死亡率能力的表現，比較統評分工具則選用 AUC 效果如表二。在預測 1 個月內死亡模型 Random forest 表現最佳 (準確率：0.836；AUC: 0.844)；預測 12 個月死亡模型 SVM 表現最佳 (準確率：0.826；AUC: 0.869)；預測 12 個月後死亡模型 AdaBoost 表現最佳準確率：0.759；AUC: 0.760)。在所有模型中機器學習模型效果皆比現有臨床評分工具來得高。

圖三 - 五為機器學習模型的重要性變量圖，分別列出三個模型的重要性變量圖，圖列出各模型中平均值最高的 10 個 (總共 34 個變量)，其中有八個變量在三個模型的重要性變量排行中的皆排在前十，表示其對大多數模型分類器重要性較高，分別是白血球、血中尿素氮、尿酸、總膽紅素、直接膽紅素、丙氨酸轉胺酶、年齡及肝性腦病變。

將三個模型中預測效果最佳的機器學習模型與傳統風險評分工具進行比較 (圖六 - 八)，ROC 曲線圖顯示機器學習模型效果更佳 (1 個月內死亡模型：Random forest (AUC=0.844))

表一：資料庫的描述性統計 (n=528)

變項	合計 (n=528)
性別 (Sex)	
男 n(%)	381(72.2%)
年齡 (Age) mean (±SD)	56.45(±13.967)
肝硬化病史 (Cirrhosis history)	
有 n(%)	370(70.1%)
吸菸病史 (Smoke)	
戒菸 n(%)	69(13.0%)
有 n(%)	266(50.4%)
無 n(%)	193(36.6%)
酒精病史 (Alcohol)	
戒酒 n(%)	123(23.3%)
有 n(%)	317(60.0%)
無 n(%)	88(16.7%)
酒精性肝炎 (Alcoholic hepatitis)	
有 n(%)	287(54.4%)
病毒性肝炎 (Viral hepatitis)	
B 型 +C 型 n(%)	6(1.1%)
C 型 n(%)	66(12.5%)
B 型 n(%)	100(18.9%)
無 n(%)	357(67.5%)
食道靜脈曲張 (Esophageal varice)	179(33.9%)
有 n(%)	
腹水 (Ascites)	
多量 (massive) n(%)	32(6.1%)
中量 (moderate) n(%)	85(16.1%)
少量 (slight) n(%)	107(20.3%)
無 (absent) n(%)	304(57.6%)
肝性腦病變 (Hepatic encephalopathy)	
昏迷 (coma) n(%)	63(11.9%)
半昏迷 (light-coma) n(%)	38(7.2%)
嗜睡 (drowsy) n(%)	82(15.5%)
混亂 (confused) n(%)	91(17.2%)
無 n(%)	254(48.1%)
是否使用呼吸器 (Ventilator)	
有 n(%)	269(50.9%)
是否洗腎 (H/D+CVVH< 2week)	
有 n(%)	51(9.7%)
是否使用血管升壓素 (Vasopressors)	137(25.9%)
有 n(%)	
心率 (HR)(Bpm) mean(±SD)	101.52(±25.317)
血色素 (Hb)(g/dL) mean(±SD)	9.54(±2.525)
白血球 (WBC)(10 <sup>3</sup> /uL) mean(±SD)	10.72(±7.541)
嗜性白血球 (Seg) mean(±SD)	75.85(±14.961)
帶狀型嗜中性白血球 (Band) mean(±SD)	2.09(±5.50)
血小板 (PLT)(10 <sup>3</sup> /uL) mean(±SD)	118.58(±83.836)

表一：資料庫的描述性統計 (n=528) (續)

變項	合計 (n=528)
凝血時間 (PT international normalized ratio)(mg/dL) mean(±SD)	1.49(±1.144)
直接膽紅素 (Direct bilirubin)(mg/dL) mean(±SD)	2.40(±435.064)
總膽紅素 (Total bilirubin)(mg/dL) mean(±SD)	4.58(±6.222)
血清麩胺草醋酸轉胺酶 (SGOT)(U/L) mean(±SD)	311.18(±1418.497)
丙氨酸轉胺酶 (SGPT)(U/L) mean(±SD)	99.03(±468.809)
白蛋白 (Albumin)(mg/dL) mean(±SD)	2.76(±0.865)
尿酸 (Uric acid)(mg/dl) mean(±SD)	26.08(±434.856)
血中尿素氮 (BUN)(mg/dL) mean(±SD)	52.23±(4.4.620)
肌酸酐 (Cr)(mg/dL) mean(±SD)	2.45(±2.649)
腎絲球過濾率 (eGFR) mean(±SD)	53.39(±42.901)
鈉 (Na)(mEq/L) mean(±SD)	137.67(±6.526)
動脈血氧分壓 (PaO <sub>2</sub> )(mmHg) mean(±SD)	131.23(±84.501)
血氧飽和度 (SpO <sub>2</sub> )(%) mean(±SD)	97.15(±3.849)
氣氧分壓 (FiO <sub>2</sub> )(%) mean(±SD)	30.64(±35.304)
APACHE II <sup>a</sup> mean(±SD)	19.82(±8.564)
CTP <sup>b</sup> mean(±SD)	8.98(±2.305)
MELD(Original) <sup>c</sup> mean(±SD)	19.99(±9.119)
MELDNa <sup>d</sup> mean(±SD)	21.78(±8.659)
MELD <sup>e</sup> mean(±SD)	20.72(±8.992)
死亡 (Death)	
1 個月內死亡 n(%)	150(28.4%)
1 個月以後至 12 個月內死亡 n(%)	54(10.2%)
12 個月以後死亡 n(%)	47(8.9%)
存活 n(%)	277(52.5%)

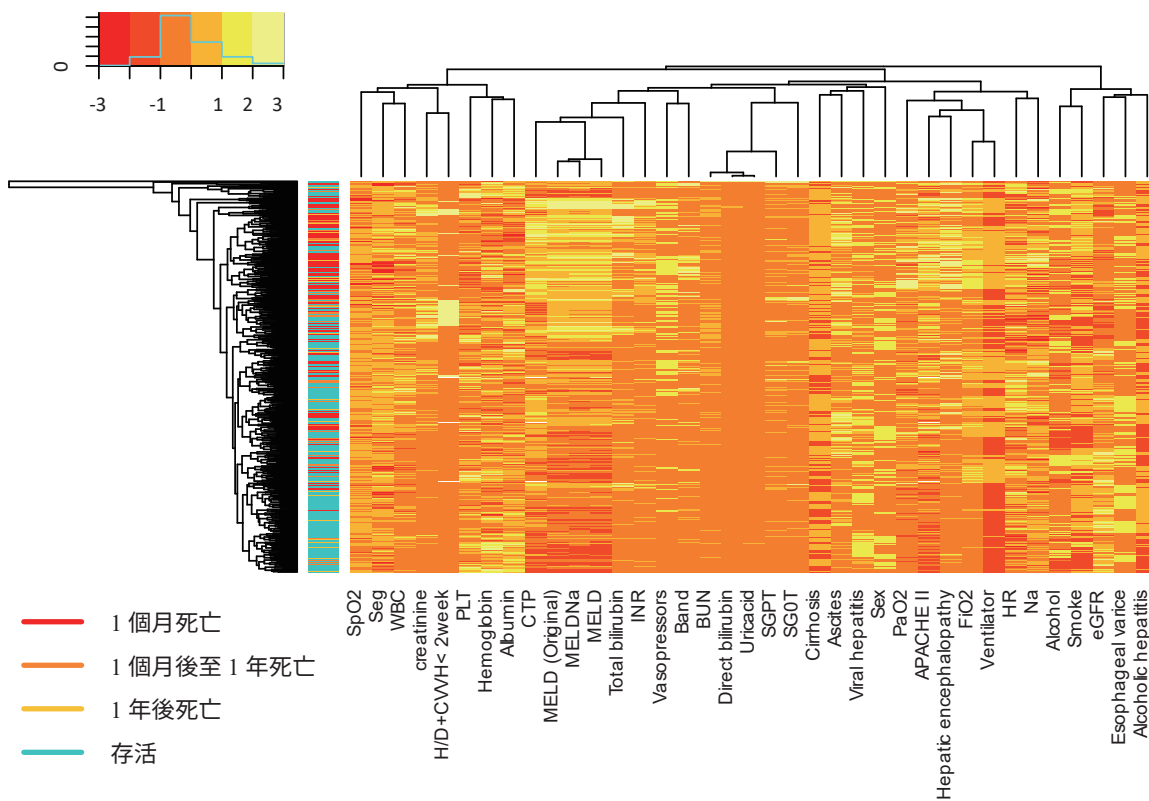
備註：

<sup>a</sup>：急性生理性健康評估 II (Acute physiology and chronic health evaluation II)。<sup>b</sup>：Child-Turcotte-Pugh。<sup>c</sup>：原版終末期肝病模型 (Original, Pre-2016, Model for End-Stage Liver Disease)。<sup>d</sup>：MELDNa/MELD-Na Score。<sup>e</sup>：終末期肝病模型 (Model for end-stage liver disease)。

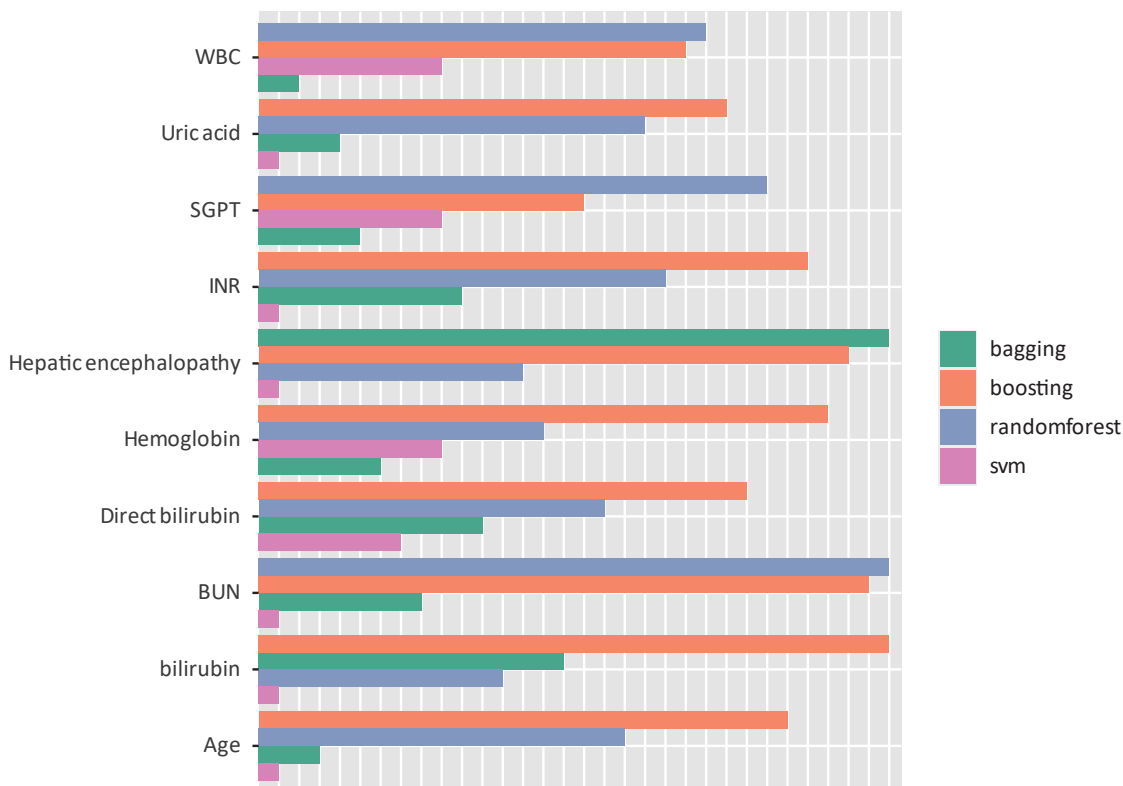
v.s CTP (AUC=0.785)(如圖六)、12 個月內死亡模型：SVM (AUC=0.869) v.s MELDNa (AUC=0.797) (如圖七)、12 個月後死亡模型：Adaboost (AUC=0.760) v.s CTP (AUC=0.746) (如圖八)。

最後為證實本研究機器學習模型方法優於臨床評分工具，將三個時間線模型中準確率最佳的機器學習模型和現有臨床評分工具做比較 (圖九～十一)，1 個月內死亡預測模型

最佳的模型：Random forest (準確率=0.836；AUC=0.844) (如圖九)。優於 APACHE II (AUC=0.752)、CTP (AUC=0.785)、MELD (Original) (AUC=0.778)、MELDNa (AUC=0.780)、MELD (AUC=0.771)；12 個月內死亡預測模型最佳的模型：SVM (準確率=0.826；AUC=0.869)。優於 APACHE II (AUC=0.715)、CTP (AUC=0.796)、MELD (Original) (AUC=0.786)、MELDNa (AUC=0.797)、MELD (AUC=0.785)



圖二：聚類熱圖，數據經過標準化為 Z 分數以顏色表示差異，變項行列都已分別聚類



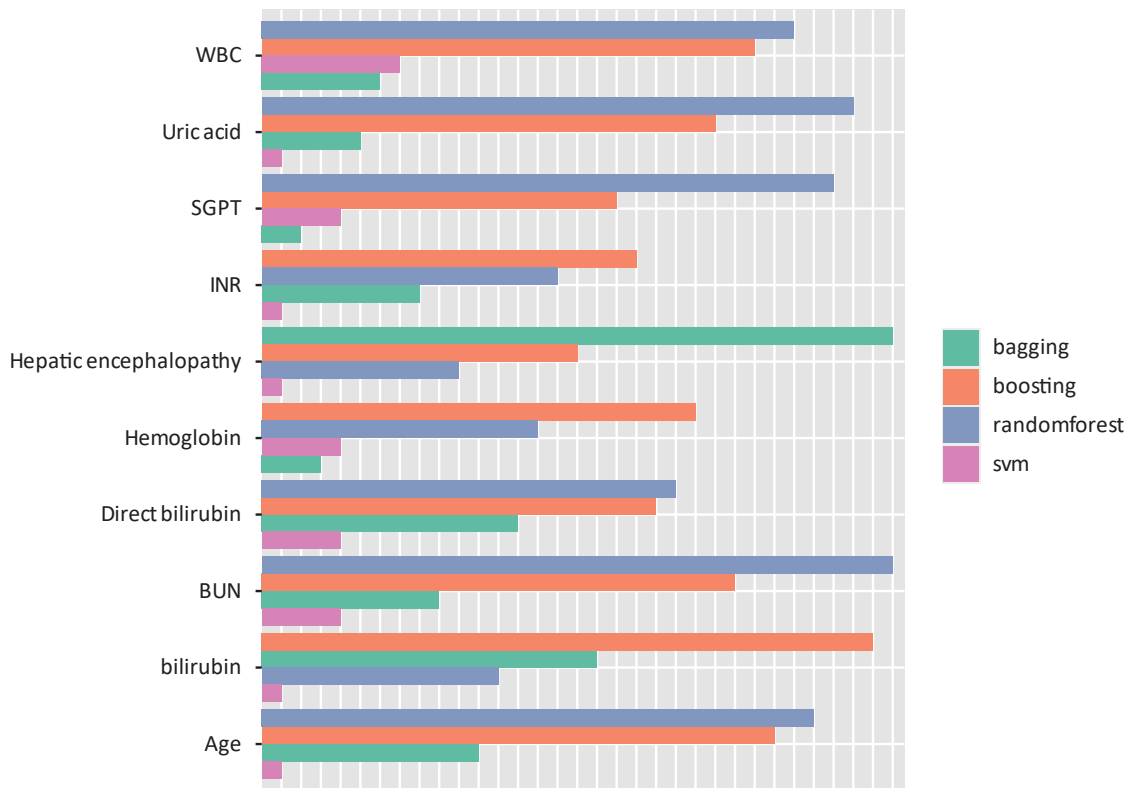
圖三：1 個月內死亡模型之重要性變量圖。

表二、模型的 Accuracy、Sensitivity、Specificity、precision、F1 Score、和曲線下面積 (AUC) 值

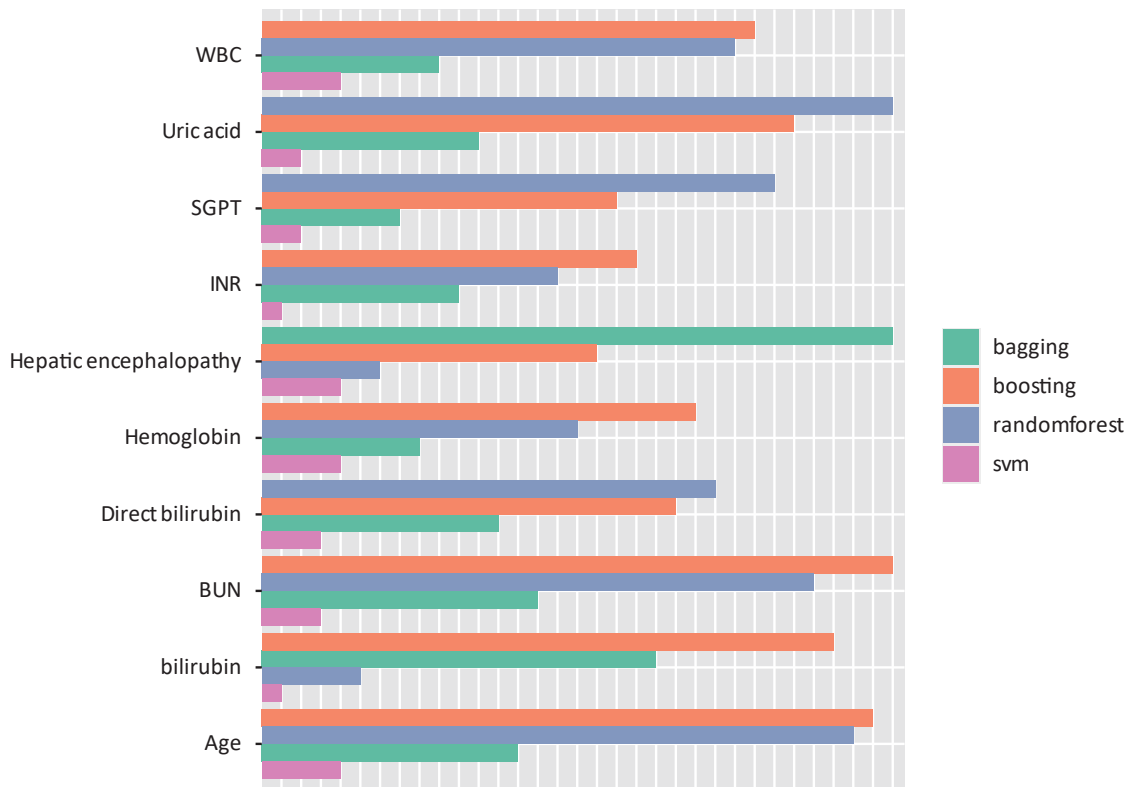
模型	Accuracy	Sensitivity	Specificity	precision	F1 Score	AUC
1 個月內死亡						
Bagging <sup>a</sup>	0.778	0.881	0.500	0.827	0.853	0.797
Adaboost <sup>b</sup>	0.826	0.881	0.678	0.881	0.881	0.838
Random forest <sup>c</sup>	0.836	0.934	0.571	0.855	0.893	0.844
SVM <sup>d</sup>	0.826	0.921	0.571	0.853	0.886	0.833
APACHE II						0.752
CTP						0.785
MELD (Original)						0.778
MELDNa						0.780
MELD						0.771
12 個月內死亡						
Bagging <sup>a</sup>	0.750	0.782	0.685	0.830	0.805	0.825
Adaboost <sup>b</sup>	0.807	0.855	0.714	0.855	0.855	0.847
Random forest <sup>c</sup>	0.778	0.797	0.742	0.859	0.827	0.853
SVM <sup>d</sup>	0.826	0.855	0.771	0.880	0.867	0.869
APACHE II						0.715
CTP						0.796
MELD (Original)						0.786
MELDNa						0.797
MELD						0.785
12 個月以後死亡						
Bagging <sup>a</sup>	0.692	0.724	0.652	0.724	0.724	0.747
Adaboost <sup>b</sup>	0.759	0.775	0.739	0.789	0.782	0.760
Random forest <sup>c</sup>	0.740	0.758	0.717	0.771	0.765	0.784
SVM <sup>d</sup>	0.740	0.758	0.717	0.771	0.765	0.806
APACHE II						0.701
CTP						0.746
MELD (Original)						0.711
MELDNa						0.711
MELD						0.701

備註：

<sup>a</sup>：引導聚集算法 (Bootstrap aggregating)。<sup>b</sup>：自適應增強 (Adaptive boosting)。<sup>c</sup>：隨機森林。<sup>d</sup>：支持向量機 (Support vector machine)。

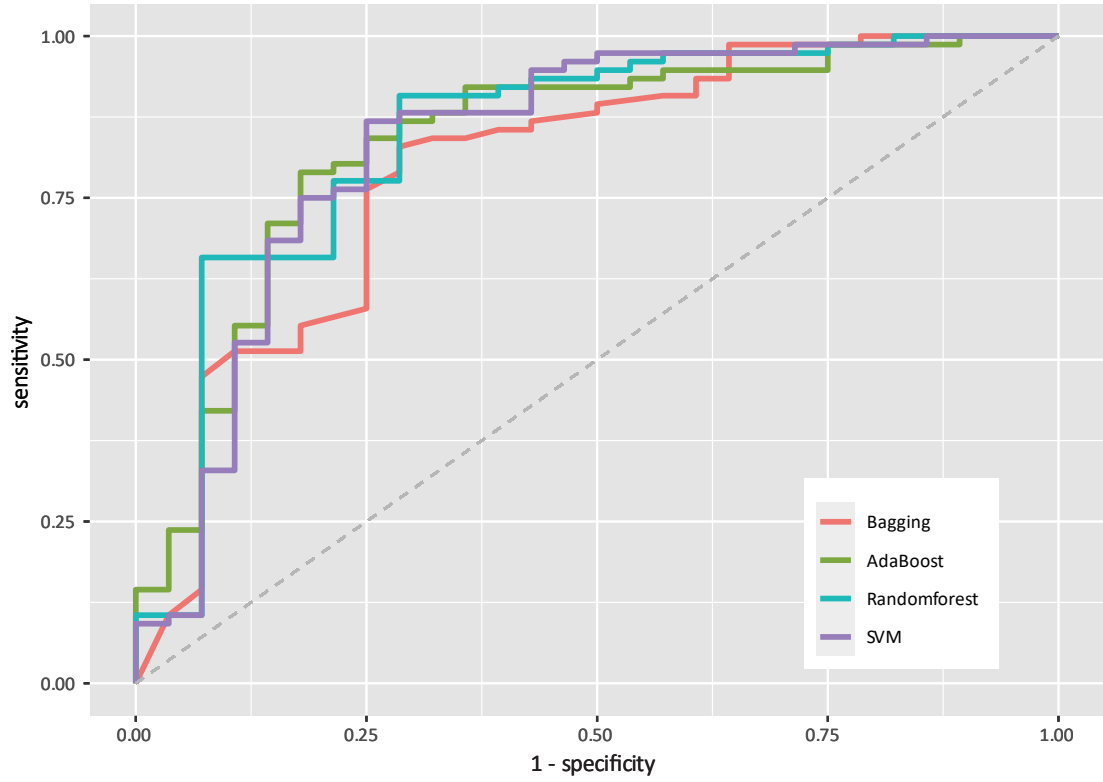


圖四：12個月內死亡模型之重要性變量圖。

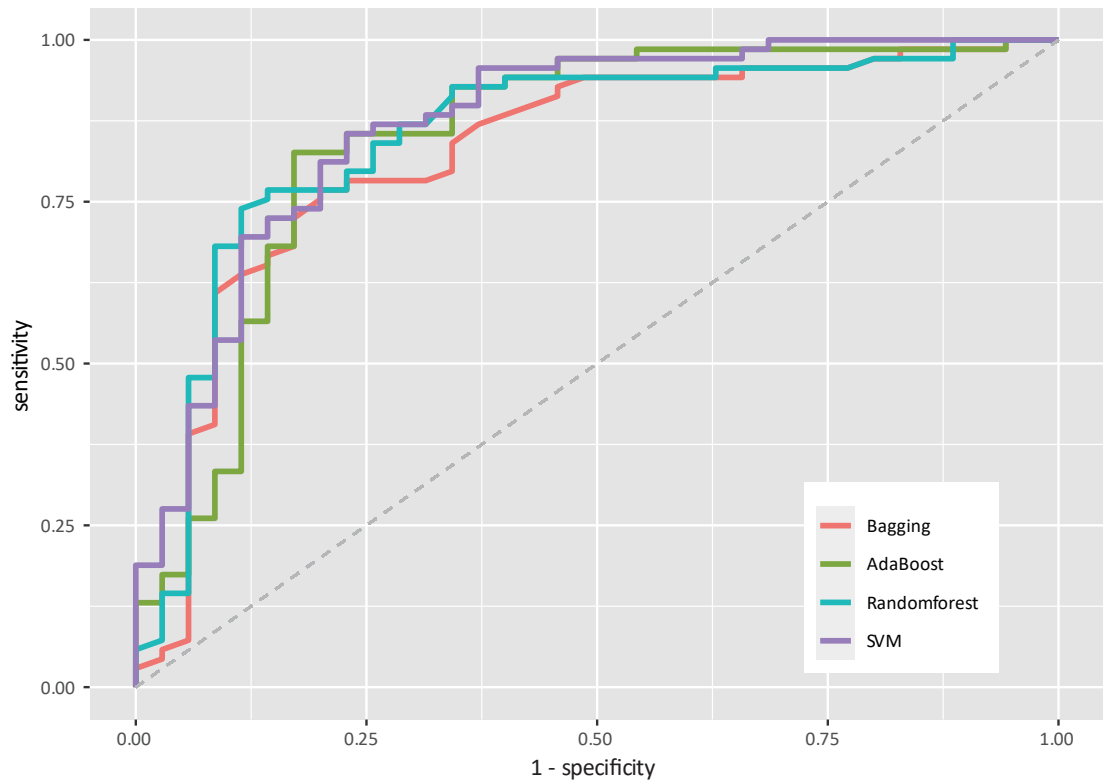


圖五：12個月後死亡模型之重要性變量圖。

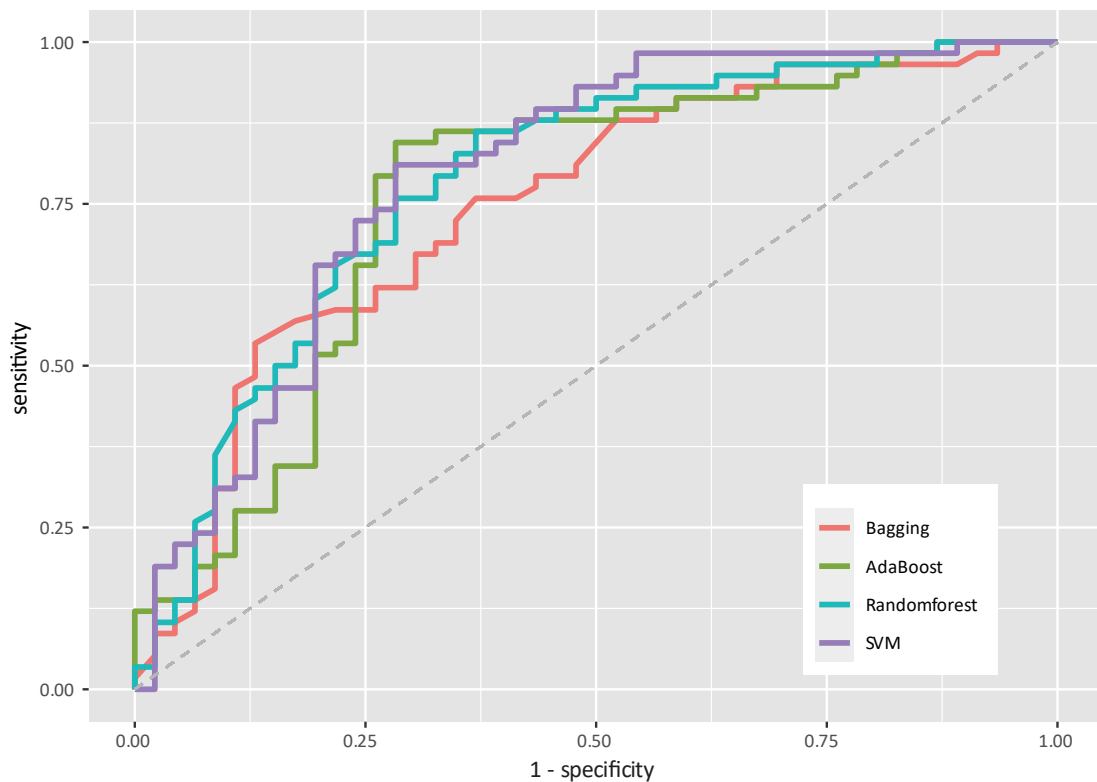




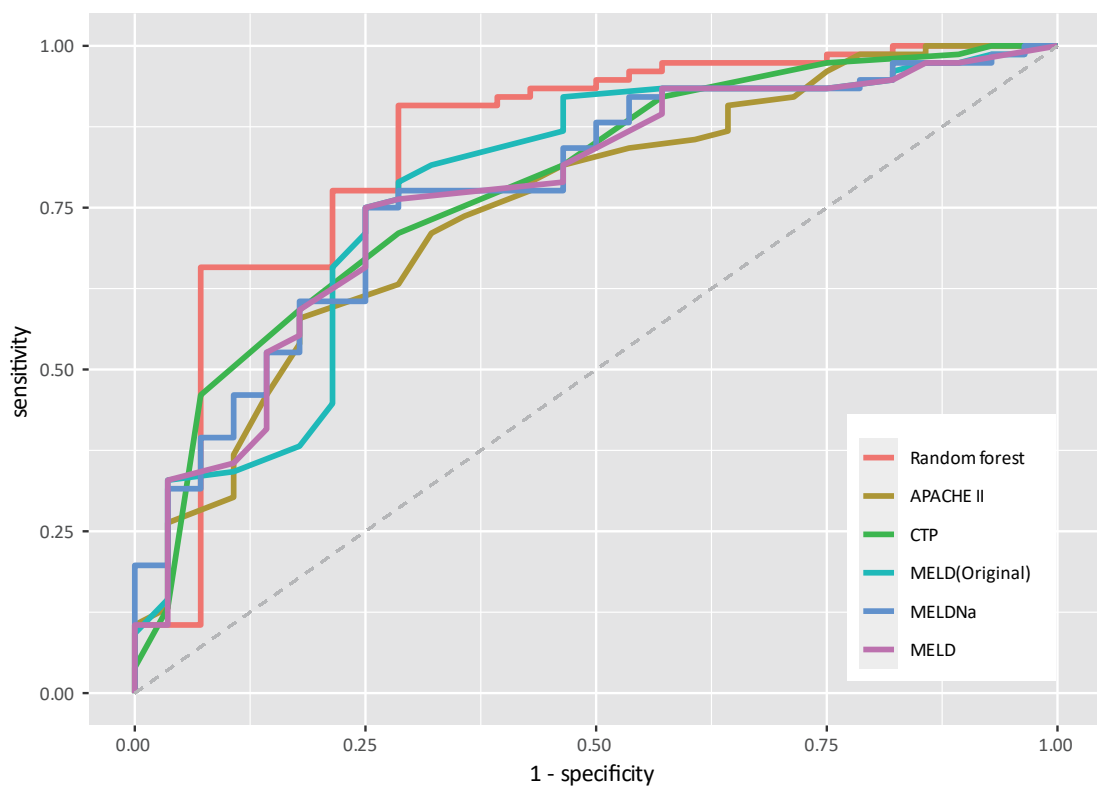
圖六：1 個月內死亡模型之 ROC 曲線圖。



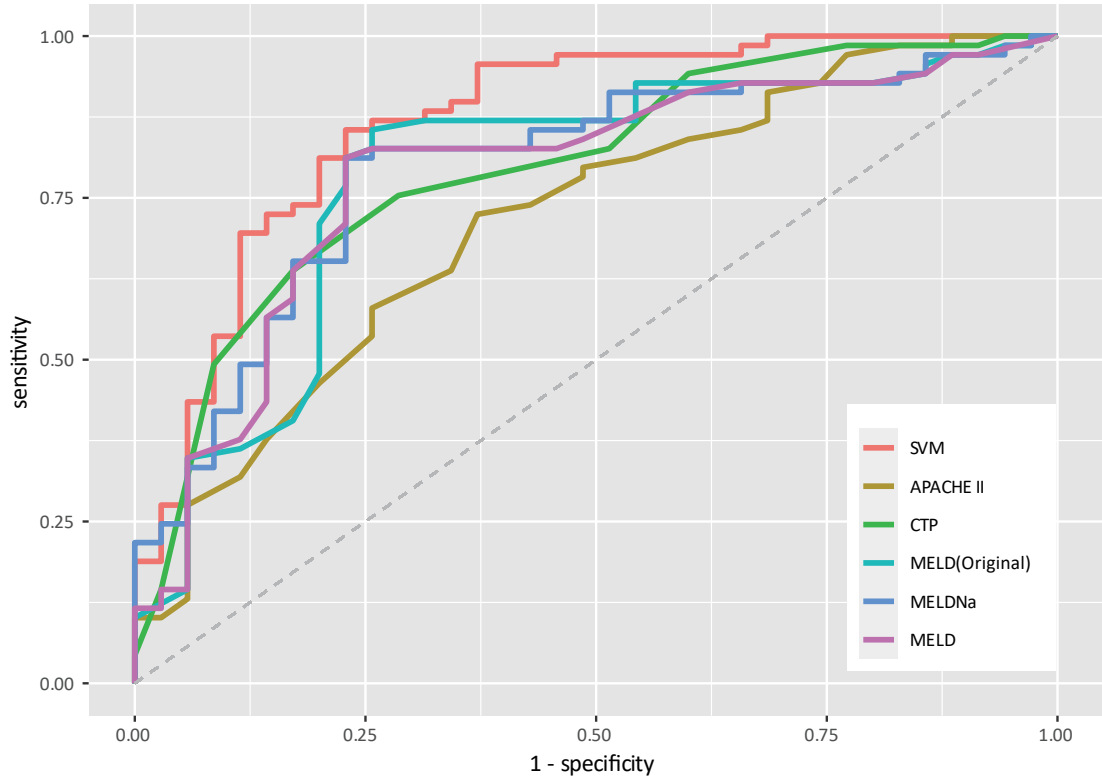
圖七：12 個月內死亡模型之 ROC 曲線圖。



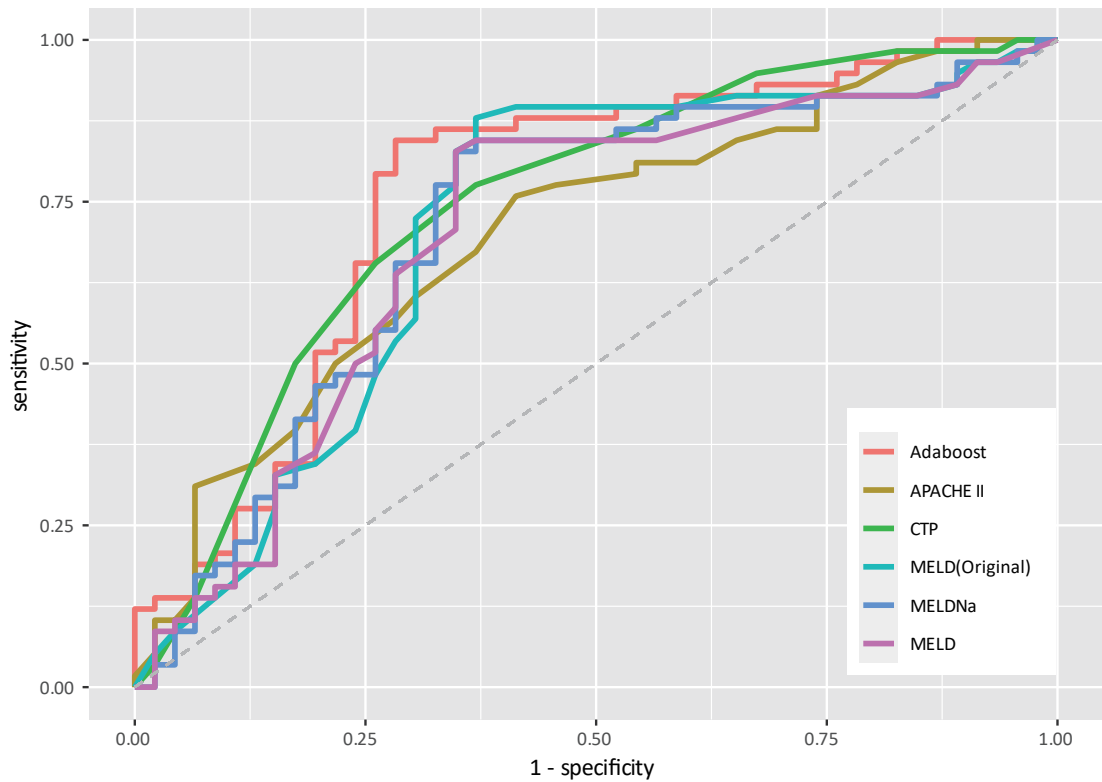
圖八：12 個月後死亡模型之 ROC 曲線圖。



圖九：1 個月內死亡率預測效果最佳的方法 (Random forest) 與傳統評分工具 (APACHE II、CTP、MELD (Original)、MELDNa、MELD) 之 ROC 曲線圖。



圖十：12 個月內死亡率預測效果最佳的方法 (SVM) 與傳統評分工具 (APACHE II、CTP、MELD (Original)、MELDNa、MELD) 之 ROC 曲線圖。



圖十一：12 個月後死亡率預測效果最佳的方法 (Adaboost) 與傳統評分工具 (APACHE II、CTP、MELD (Original)、MELDNa、MELD) 之 ROC 曲線圖。

(如圖十)；12個月後死亡預測模型最佳的模型：Adaboost (準確率=0.759；AUC=0.760)。優於 APACHE II (AUC=0.701)、CTP (AUC=0.746)、MELD (Original) (AUC=0.711)、MELDNa (AUC=0.711)、MELD (AUC=0.701) (如圖十一)。

## 討 論

肝硬化病人出院後的死亡率普遍偏高，且重症病人常合併使用氣管內管、呼吸器、血管加壓素、中心靜脈導管和腎臟替代療法等。由於重症病人資源耗用高，相對付出的醫療成本也較高。若能及早識別高風險病人，並根據病人病況利用標準化措施進行干預 (例如門靜脈高壓相關出血的管理)，將有助於提高生存率<sup>29</sup>。故對病人進行風險及長期預測有利於管理與資源分配。

本研究建構之模型除及時識別嚴重程度較高的病人外，還能對長期死亡風險進行預測。收案追蹤 528 名病人，入住內科加護病房於 1 個月內死亡率為 28%，另外 19% 病人平均追蹤 17.05 月後死亡。研究結果顯示 ML 模型優於傳統評分工具，且模型準確率相近，在四種機器學習方法 (Bagging、Adaboost、SVM 和 Random forest) 預測 1 個月內死亡模型以 Random forest 表現最佳 (Accuracy: 0.836；AUC: 0.844)；預測 12 個月內死亡模型以 SVM 表現最佳 (Accuracy: 0.826；AUC: 0.869)；預測 12 個月後死亡模型以 Adaboost 表現最佳 (Accuracy: 0.759；AUC: 0.760)。

機器學習根據變量選擇時的差異會產生不同結果，亦可藉此確認影響生存率之重要變量<sup>30</sup>，然在本研究中三個模型中皆被視為最重要的變量是白血球、血中尿素氮、尿酸、總膽紅素、直接膽紅素、丙氨酸轉胺酵素、年齡及肝性腦病變。其中最主要的因素是肝性腦病變，肝性腦病變是肝硬化的主要神經類併發症，早年就有多數研究指出肝性腦病變與肝硬化病人的短期死亡有關<sup>31</sup>。其次則是血中尿素氮，這與最近的多項研究結果一致<sup>12,32</sup>，Arihan 等人指出血中尿素氮與死亡率不僅密切相關，

且對肌酸酐有疊加作用，使的入住重症加護病房的病人構成異質群體，因此血中尿素氮可用來識別重症病人<sup>32</sup>。

關於機器學習應用在肝硬化病人長期分析模型近年已被提出，多和傳統肝硬化評分工具進行比較，但仍缺乏以台灣病人為母體的研究，現有研究結果未必能外推至其他族群<sup>33-34</sup>。由於目前已有許多研究指出肝硬化重症病人死亡率可透過 CTP、APACHE II 及 MELD score 等評估工具進行預測<sup>35-37</sup>，因此本研究使用數個與肝硬化評估相關的傳統評分工具藉以增加可信度。

多數風險預測模型尚未完全適用於臨床環境，故其性能受到限制<sup>38</sup>。雖機器學習領域目前蓬勃發展中，但仍有許多細節未被充分研究<sup>39</sup>。因此本研究旨在探討機器學習應用在肝硬化病人長期死亡風險預測的可行性，期望結果能協助臨床及早識別並介入高風險病人療程，以期達到降低死亡率俾使醫療資源充分被利用。

## 研究限制

本研究係屬病歷回溯性研究故存在以下幾項限制：(1) 研究樣本受限：本研究僅限於東部某區域教學醫院，與不同層別及不同地域之醫院結果可能有出入。(2) 研究資料範圍受限：取得資訊係以加護病房常用的生理項目與相關評分為主，未能依肝硬化病人常用的項目如慢性肝衰竭急性發作 (Acute-on-chronic liver failure, ACLF) 與 failure-SOFA (CLIF-SOFA) scores 等評分工具且病歷回溯缺乏當下判斷，亦難以事後追查當下疾病狀況及主要死因，未來可考慮和其他重症評分進行比較，增加研究結果的可信度。(3) 缺乏肝臟移植資料：本院為區域教學醫院，無法施行肝移植手術，因此收案病人皆未進行肝臟移植，故研究結果僅應用於預測未接受肝臟移植之肝硬化病人的長期存活。

## 致 謝

感謝內科加護病房醫療團隊及病人使本研究得以順利完成。

## 參考文獻

- World Health Organization. Cause-specific mortality 2000-2019. 2020.
- Wang X, Wu B. Critical issues in the diagnosis and treatment of liver cirrhosis. *Gastroenterol Rep* 2019;7(4):227-30.
- Levesque E, Saliba F, Ichai P, Samuel D. Outcome of patients with cirrhosis requiring mechanical ventilation in ICU. *J Hepatol* 2014;60(3):570-8.
- Rudow DL, Goldstein MJ. Critical care management of the liver transplant recipient. *Crit Care Nurs Q* 2008;31(3):232-43.
- Sargent S, Wainwright SP. A qualitative study exploring patients perceived quality of life following an emergency liver transplant for acute liver failure. *Intensive Crit Care Nurs* 2007;23(5):272-80.
- Vaz A, Eusebio M, Antunes A, et al. Mortality in patients with cirrhosis admitted to the ICU: Time to rethink strategies? *Crit Care* 2015;19(Suppl 1):1-201.
- D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: A systematic review of 118 studies. *J Hepatol* 2006;44(1):217-31.
- Deelen J, Kettunen J, Fischer K, et al. A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. *Nat Commun* 2019;10(1):1-8.
- Tsai YW, Tzeng IS, Chen YC, Hsieh TH, Chang SS. Survival prediction among patients with non-cancer-related end-stage liver disease. *PloS One* 2018;13(9):e0202692.
- Theocharidou E, Pieri G, Mohammad AO, et al. The Royal Free Hospital score: a calibrated prognostic model for patients with cirrhosis admitted to intensive care unit. Comparison with current models and CLIF-SOFA score. *Am J Gastroenterol* 2014;109(4):554.
- Chen RC, Cai YJ, Wu JM, et al. Usefulness of albumin-bilirubin grade for evaluation of long-term prognosis for hepatitis B-related cirrhosis. *J Viral Hepat* 2017;24(3):238-45.
- Lin YJ, Chen RJ, Tang JH, et al. Machine-learning monitoring system for predicting mortality among patients with noncancer end-stage liver disease: Retrospective study. *JMIR Med Inform* 2020;8(10):e24305.
- Michie D, Spiegelhalter DJ, Taylor CC. Machine learning, neural and statistical classification. USA. Prentice Hall 1994; 1-290.
- Allyn J, Allou N, Augustin P, et al. A comparison of a machine learning model with EuroSCORE II in predicting mortality after elective cardiac surgery: A decision curve analysis. *PloS One* 2017;12(1):e0169772.
- Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J* 2017;38(23):1805-14.
- Dupin-Bryant PA, Olsen DH. Business intelligence, analytics and data visualization: A heat map project tutorial. *J Manag Inf Syst* 2014;18(3):185-200.
- Wilkinson L, Friendly M. The history of the cluster heat map. *J Am Stat Assoc* 2009;63(2):179-84.
- Babicki S, Arndt D, Marcu A, et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* 2016; 44(W1):W147-53.
- Breiman L. Bagging predictors. *Machine Learning* 1996; 24(2):123-40.
- Schapire RE. Explaining adaboost. In: *Empirical Inference*. Springer 2013;37-52.
- Freund Y, Schapire RE. Experiments with a new boosting algorithm. *International Conference on Machine Learning* 1996; 96:148-56.
- Meir R, Rätsch G. An introduction to boosting and leveraging. In *Advanced lectures on machine learning*. Springer 2003; 118-83.
- Ganggayah MD, Taib NA, Har YC, Lio P, Dhillon SK. Predicting factors for survival of breast cancer patients using machine learning techniques. *BMC Medical Inform Decis Mak* 2019;19(1):1-17.
- Joachims T. Making large-scale SVM learning practical. *JFE Tech* 1998;28.
- Cortes C, Vapnik V. Support-vector networks. *Machine Learning* 1995;20(3): 273-97.
- Ho TK. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*. IEEE 1995; 1:278-82
- Biau G. Analysis of a random forests model. *J Mach Learn Res* 2012;13(1):1063-95.
- Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* 2007; 115(5):654-7.
- Weil D, Levesque E, McPhail M, et al. Prognosis of cirrhotic patients admitted to intensive care unit: a meta-analysis. *Ann Intensive Care* 2017;7(1):1-14.
- Ganggayah MD, Taib NA, Har YC, Lio P, Dhillon SK. Predicting factors for survival of breast cancer patients using machine learning techniques. *BMC Medical Inform Decis Mak* 2019;19(1):1-17.
- Bustamante J, Rimola A, Ventura PJ, et al. Prognostic significance of hepatic encephalopathy in patients with cirrhosis. *J Hepatol* 1999;30(5):890-5.
- Arihan O, Wernly B, Lichtenauer M, et al. Blood Urea Nitrogen (BUN) is independently associated with mortality in critically ill patients admitted to ICU. *PloS One* 2018;13(1): e0191697.
- Kanwal F, Taylor TJ, Kramer JR, et al. Development, validation, and evaluation of a simple machine learning model to predict cirrhosis mortality. *JAMA Network Open* 2020; 3(11):e2023780.
- Hu C, Anjur V, Saboo K, et al. Low Predictability of Readmissions and Death Using Machine Learning in Cirrhosis. *Am J Gastroenterol*, 2021;116(2):336-46.
- Karvellas CJ, Bagshaw SM. Advances in management and prognostication in critically ill cirrhotic patients. *Curr Opin Crit Care* 2014;20(2):210-7.
- Bao Q, Wang B, Yu L, Weng H, Ge J, Li L. A modified prognostic score for critically ill patients with cirrhosis: an observational study. *Gastroenterol Hepatol* 2016;31(2):450-8.

37. Chen RC, Cai YJ, Wu JM, et al. Usefulness of albumin-bilirubin grade for evaluation of long-term prognosis for hepatitis B-related cirrhosis. *J Viral Hepat* 2017;24(3):238-45.
38. Mpanya D, Celik T, Klug E, Ntsinjana H. Machine learning and statistical methods for predicting mortality in heart failure. *Heart Fail Rev* 2021;24(3):238-45.
39. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. *Science* 2015;349(6245):255-60.

## Machine Learning to Predict the Mortality in Intensive Care Unit - Long-term Tracking Report

Chia-Ying Hsiao<sup>1</sup>, Jui-Chu Ma<sup>2</sup>, Cheng Chun Han<sup>1</sup>, Lang-Yin Zeng<sup>3</sup>,  
Chin-Hsin Li<sup>2</sup>, Wan-Ju Cheng<sup>2</sup>, Feng Chi Kuo<sup>1</sup>, and Min-I Su<sup>1</sup>

<sup>1</sup>Department of Internal Medicine, <sup>2</sup>Department of Nursing, <sup>3</sup>Department of Medical Research,  
Taitung MacKay Memorial Hospital, Taitung, Taiwan

Risk prediction management is of great significance in clinical research and medical care. The prognosis of patients with liver cirrhosis is worse than compared with other groups of critically ill patients. However, there are no scoring systems design of long-term mortality risk for patients with cirrhosis. The study included 528 patients who received medical treatment from July 1, 2013 to December 31, 2018 and was followed up to June 2020. This research attempts to use machine learning methods (Bagging, Adaboost, Support Vector Machine (SVM) and Random forest) to predicted mortality of patients with cirrhosis were performed and compared with CTP, MELD Score (Original), MELDNa Score, MELD Score, and The Acute Physiology and Chronic Health Evaluation (APACHE II). The result pointed out that compared with the existing scoring systems, machine learning method model has better prediction effect. Random forest is the best model for predicting death within 1 month (accuracy rate: 0.836; AUC: 0.844); SVM is the best model for predict death within 12 months (accuracy rate: 0.826; AUC: 0.869); Adaboost is the best model for predict death after 12 months (accuracy rate: 0.759; AUC: 0.760). It fully proves that the machine learning model has application value in predicting the long-term mortality rate of patients with cirrhosis. (*J Intern Med Taiwan* 2022; 33: 141-154)