以類神經網路模型預測大腸直腸癌病人之存活率

楊靜如 [a,b*]，梁子豪 [a],劉立 [a]

[a] *臺北醫學大學醫學資訊研究所*

[b] *臺北市立聯合醫院仁愛院區放射診斷科*

E-mail: A0226@tpech.gov.tw

摘要

類神經網路因為有分類及預測功能，是早期常被應用在臨床決策支援的一種系統。本文將計劃如何以類神經網路，設計一個預測大腸直腸癌病人存活率的電腦系統，並就其準確度及解決方法作探討。資料來源是台北國泰醫學中心，癌症防治中心屬下癌症登錄小組的「大腸直腸癌資料庫」，此資料庫登錄了自民國 91 年 1 月開始，在國泰醫院被診斷出大腸癌或直腸癌等六種癌症的病人資料，資料庫應申報欄位共計 65 個，主要分為身分證字號及人口學特徵、癌症診斷、申報醫院特別資料、癌症分期/預後因子、首次癌症的治療、追蹤等六大類，本研究收集個案總數共計 1193 名，去除 55 名追蹤期間為 0 年者後剩 1138 名，追蹤期間平均 2.5 年，最長 8.6 年，最短 0.1 年，參數取決採取排除法，運用 STATISTICA Neural Networks (StatSoft Inc.)，建構類神經網路模型，過程中測試超過 100 個模型，保留其中測試錯誤(test error)最小的 5 組，結果顯示預測能力最好的類神經網路模型是一個 feedforward, fully-connected, MLP network，它是由反饋監督式學習(back-propagation supervised training)演算法所構成。以現在的類神經網路模型，加上癌症登錄資料庫的適當欄位作為輸入變項，可以獲得更好的預測結果。測試此訓練完成的類神經網路的敏感度高達 94.1%、AUC0.854、特異性 71.7%，在這個研究裏面，發展出針對病人本身的存活率預測模型，而其準確度比以往的統計方法都要來得高。

關鍵字：類神經網路、大腸直腸癌、反饋監督式學習

Abstract

Cancer death proceeded to the leading cause of death since 1982 and thetrend of both mortality rate and prevalence are increasing as time. Colorectalcancer is not only has an increasing mortality rate but also be the third leadingcause of cancer death recently. The percentage of people in a study or treatmentgroup who are alive for a certain period of time after they were diagnosed withor treated for a disease, such as cancer. The survival rate is often stated as afive-year survival rate, which is the percentage of people in a study or treatmentgroup who are alive five years after diagnosis or treatment. But it is only anumber of percentage of people survived not the probability of survival of an individual. Clinically, it is very difficult to predict the survival rate of anindividual with cancer because there are many influencing factors and thefactors are interacting to each other. In this study, I try to establish an artificialneural network model to predict the survival rate of an individual with colorectalcancer based on a cancer registry database.