B 型肝炎病毒序列突變率預測之研究

孫光天[a]*，劉威良[a]，劉紋君[b]，張定宗[c]

國立臺南大學 資訊教育研究所[a]

成功大學基礎醫學研究所[b]　成功大學分子醫學研究所[c]

* ktsun@mail.nutn.edu.tw

摘要

B 型肝炎病毒序列突變率對於患者癒後有很大影響，當病毒序列突變率產生異常偏高時，患者在未來病情轉變為肝硬化及肝癌的機率也可能隨之升高。因此本研究實驗樣本係以成功大學醫學院所提供之B 型肝炎患者臨床資料。並選擇對時間序列較常應用之類神經網路模型稍作改良，期望能以患者過去各月的樣本特徵，透過類神經網路推論，預測患者下個月病毒序列突變率，甚至預測至更久的時間。經本研究實驗結果證實，本研究之類神經網路模型不論在病毒序列突變率短期預測或長期預測上皆有不錯的預測能力。

關鍵詞：類神經網路、B 型肝炎病毒、突變率預測

壹、緒論

台灣位於 B 型肝炎高流行區，B 型肝炎研究學者 Beasley 等人也曾經在台灣平均追蹤了 11.25 年，結果發現 B 型肝炎帶原者發生肝癌的機率為一般人的 100 倍。日後研究也被証實 B 型肝炎和台灣地區居民慢性肝炎、肝硬化有密切的關係，同時也是肝細胞癌症之發生率和死亡率偏高的主要原因[4,7,9]。B 型肝炎病毒基因體是不完全雙股的 DNA（partly double-strand circular DNA），全長約為 3.2kb（3200 個鹼基）[9]，目前已被証實 B 型肝炎的嚴重程度、致癌機制以及用藥療效好壞，主要與 HBV 基因前 C 區（Precore region）、基礎啟動子（Basic core promoter）以及表面抗原基因(preS1、preS2、S gene)的序列突變有關[8]。但在過去的 B 型肝炎研究中，大部份皆以病毒序列突變點的方式做為分析，甚少會以病毒序列突變率做進一步分析。但事實上，病毒突變率和病毒演化有著密切的關係，這也是病毒對人體產生威脅的重要因素之一。原因在於病毒本身不一定會致病，但它可能會經過多次突變後而產生另一種致病的變種病毒。在各項研究中也證實，當病毒序列突變率產生異常偏高時，對人體的生理機能可能存在著威脅。因此若能預測病毒序列的突變演化，了解病毒週期動態，未來或許可以提前預防疾病及製藥。

在過去資訊技術的預測模型方面，類神經網路具有良好之非線性模式對應能力，且不須於建構時如傳統統計模式般設定許多假設條件，近年來，類神經網路已廣泛使用於商業、工程及社會科學等方面。其相關應用有工程製程管制、商業股票預測、教育選題策略以及資料分析等[1,2,3]。同時在醫學上亦有優異的表現，常被廣泛地應用於診斷與分類問題。如慢性活動性肝炎之診斷、老年癡呆症診斷分析和乳癌腫瘤分類診斷等，而且都有很高的預測正確率[5,6]。

因此本研究中亦使用類神經網路做為病毒突變率預測的分析模型。但由於所使用的研究樣本皆有時間性的關係，因此本研參考過去常用於時間序列分析之類神經網路模型，其中較有名的例子為太陽黑子預測[10]。本研究針對突變率預測將模型稍作改良以提高預測準確率。其預測的主要目的為：以 B 型肝炎患者過去各月之血液抽血參數及病毒突變率，透過類神經網路推論預測下個月的病毒突變率及六個月後的病毒突變率。

貳、實驗方法

一、研究樣本說明:

本研究以成功大學醫學院張定宗醫師研究團隊所提供之 B 肝病患長期追蹤之病毒檢驗數據及病毒序列，作為實驗分析的原始資料集。其樣本特徵如下。

(一) 病患數目：5位

(二) 總樣本數：64筆，樣本資料為5位病患連續12個月的血液參數及病毒序列（共60筆資料），其中4位病患增加第18個月血液參數及病毒序列，可作