

行政院國家科學委員會補助專題研究計畫 成果報告
期中進度報告

提昇私立大學校院研發能量專案計畫

「先進醫學研究資料庫之建立」

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 91-2745-P-038-006

執行期間：89 年 12 月 01 日至 92 年 11 月 30 日

計畫主持人：李友專

共同主持人：潘憲(醫學系)、陳品玲(護理系)、邱弘毅(公衛系)、趙振瑞(保健系)、徐建業(醫資所)、邱泓文(醫資所)、蔣以仁(醫資所)、劉立(醫資所)

計畫參與人員：吳惠弘、陳庚賢、萬序恬

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、
列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：臺北醫學大學醫學資訊研究所

中 華 民 國 九 十 三 年 二 月 二 十 七 日

一、中文摘要

隨著科學研究的演進，資訊越來越多；資訊科技的產生，使得一切的資訊都數位化，也增加了流通性。在此同時也造成資訊過多或是無法分類的窘況。本計劃希望建置整合的資料庫系統，將學校以及附醫萬芳兩家醫院所產生的大量醫學學術資源，再加上健保資料庫、基因資料等之生物資訊資料庫，並提供相關工具如資料探勘(Data mining)及類神經網路(Neural Networks)，生物統計工具以及知識發掘(Knowledge Discovery)工具等，藉以提昇校內的其他研發計劃之研發能量。

本計畫已完成之成果：a.收集並建立生物及醫學資訊的資料，開發生物醫學資訊相關分析軟體工具，應用於醫學文獻中基因與疾病的關聯性分析。b.建立多媒體醫學資料庫及開發儲存、展示相關技術，支援研究、教學及臨床應用。c.建立台北醫學大學線上生物統計中心，提供線上統計諮詢服務以及預約，支援校內其他研究計劃之統計需求。

關鍵詞：資料探勘、知識發掘、多媒體、資料庫、醫學文獻、生物統計、生物資訊、生物醫學

二、英文摘要 Abstract：

In the health environment, a huge amount of data are generated. For example, there are more than 300 million hospital visits in Taiwan each year, which translate into 400GB of medical data each year. This project will build an information infrastructure to store, integrate, manage, and retrieve these medical data, including text data and image data from systems such as PACS(Picture Archiving and Communication System).

There are 3 component projects in this program projects, namely, Biomedical Computing Lab, Hypermedia Medical Data Warehousing, and Online Biostatistical Center. They represent an important infrastructure that will facilitate future medical research in Taipei Medical University.

Keywords: Data mining, knowledge discovery , ultimedia , medical database , medical literature biostatistics, bioinformatics

目 錄

一、中文摘要.....	I
二、英文摘要 Abstract :	I
壹、前言.....	3
貳、研究目的.....	4
參、文獻探討.....	5
肆、研究方法.....	11
(1) 總計畫	12
(2) 醫學及生物資訊運算實驗室.....	12
(3) 超媒體醫學資源中心.....	13
(4) 網路統計運算資源中心.....	17
伍、結果與討論.....	18
附件一：生物統計資源現況及需求問卷.....	39
附件二：生物統計資源現況及需求問卷調查結果.....	41
附件三：統計諮詢線上預約系統網頁範例.....	43
附件四：資料分析建議系統網頁範例.....	47
附件五：統計諮詢線上預約系統使用滿意度問卷.....	53
附件六：統計諮詢線上預約系統使用滿意度調查結果.....	54
附件七：資料分析建議系統使用滿意度問卷.....	55
附件八：資料分析建議系統使用滿意度調查結果.....	56
附件九、發表論文清單.....	57

壹、前言

資訊科技的演進已經影響到醫療的品質，越來越多的現象可以發現民眾、醫事人員、醫學研究人員對於醫學資訊的急切需求比率正在上升中。再加上基因科技的進步，生物資訊與醫學資訊的連接與整合已經成為全世界相關領域的重要研究議題之一。

台北醫學大學醫學資訊研究所長期在醫學資訊領域耕耘，除了已經在台灣的醫學資訊界執牛耳，由於近年來生物資訊的發展，因此本所的未來發展計畫中也增列了這一項，於民國 89 年獲得通過「提昇私校研發能量計畫 建立先進之醫學研究資料庫」的三年期整合型計畫之後。就是以建立多種醫學資訊與生物資訊資料庫以提供研究者作進一步的研究與使用作為目標。

本整合型計畫除了總計畫外，另外分為三項子計畫：醫學及生物資訊運算實驗室、超媒體醫學資源中心、網路統計運算中心。子計畫名稱及主持人詳列如下表：

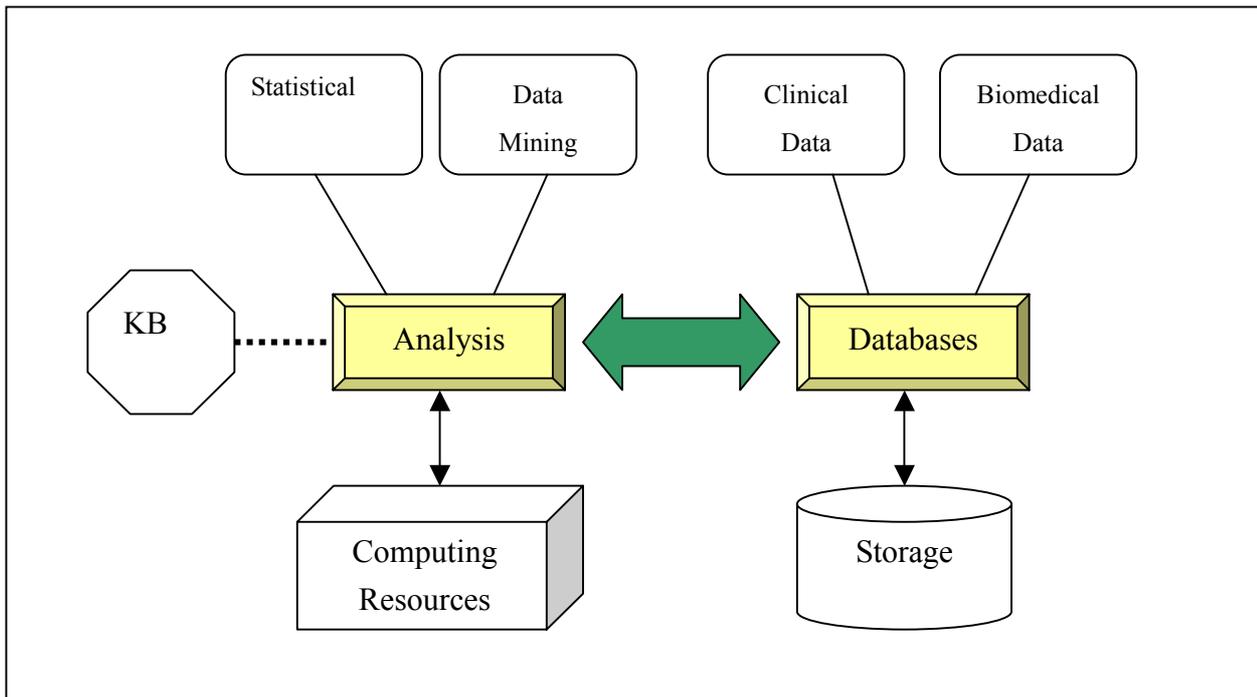
計畫項目	主持人	服務單位 系 所	職稱	計畫名稱
總計畫	李友專	醫學資訊研究所	教授兼所長	先進醫學研究資料庫之建立 (Advanced Research Medical Databases)
子計畫一	徐建業	醫學資訊研究所	副教授	醫學及生物資訊運算實驗室 (Biomedical Computing Lab)
子計畫二	潘憲	北醫附設醫院	教授兼院長	超媒體醫學資源中心 (Hypermedica Medical Data Warehousing)
子計畫三	邱弘毅	公共衛生學系	教授	網路統計運算資源中心 (Online Biostatistical Center)

計畫參與人員：

計畫項目	主持人	共同主持人	協同人員
總計畫	李友專教授	邱文達教授、劉建財副教授	萬序恬、林明錦、吳惠弘、陳博楊、顏永泰
子計畫一	徐建業副教授	邱泓文助理教授、蔣以仁助理教授	楊騰芳、陳庚賢
子計畫二	潘憲教授	劉立副教授、彭建綸副教授	曾續萍
子計畫三	邱弘毅教授	陳品玲副教授、趙振瑞副教授	許雅

貳、研究目的

由於醫學資訊技術的進步，各式各樣的醫學媒體日新月異。涵蓋文字格式之診斷、處



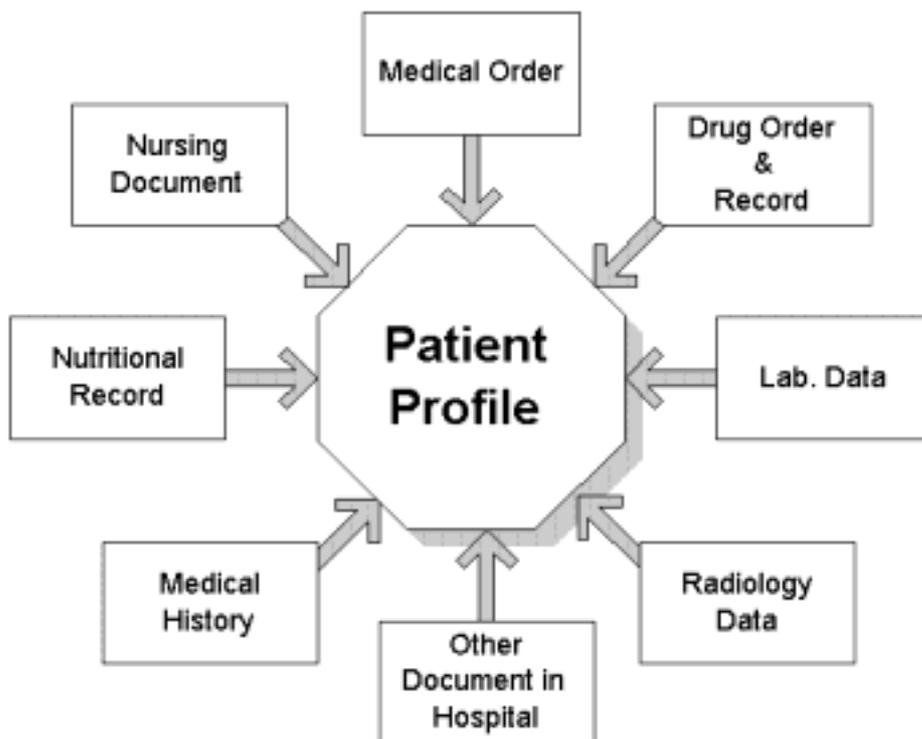
方、醫囑，2D 之 X-RAY 影像，3D 之 CT、MRI 及 Video 之內視鏡檢查。無論是提供醫學研究或是醫師診斷之依據，都是不可或缺之資源。迄今，臺北醫學院已經累積了超過 100 Giga Byte 的看診資訊，數十萬張醫學影像以及難以數計的醫學報告，如此龐大的醫學資源，若能善加運用，相信對於醫學研究之進步與病患之治療都有強大的助益。

如上頁圖所示，醫學資訊資料庫(Database)必須要有大量的儲存設備來儲存(Storage)，這些資料內容可以包括臨床資料(Clinical Data)，如醫學影像、醫學診斷、或是檢驗資料等。以及生物醫學資料(Biomedical Data)，如個人基因譜、健康紀錄等。

有了資料，必須再加上分析(Analysis)才能有效的利用這麼大量且異質的資料，要分析這麼大量的資料，勢必需要高速的運算資源(Computing Resources)。用來分析的技術包括資料探勘、統計等工具。建立了這種研究架構以及流程之後，就可以架構一個完整的醫學資訊知識庫(Knowledge Base, KB)。

然而醫學媒體格式複雜，光是以 Image 為例就有 GIF、JPEG、BMP、DCM 等格式造成研究者處理之困難。此外，儲存空間的需求龐大，研究者必須絞盡腦汁的儲存與備份資料，為了提昇研究能量，本計劃將建構一個先進的醫學研究資料庫基礎建設：其中包含醫學及生物資訊運算實驗室、超媒體醫學資源中心以及網路統計運算中心。希望藉由這些中心的建立以及工具與資料庫的開發，使得各種醫學媒體皆能在這樣的架構下可以被統一的搜尋、存放、備份、保全、分類、展示、傳播、研究。

臨床醫學資訊資料庫是以病患為中心整合資料，以分散式架構建立伺服器端之機制，並以瀏覽器做為檢視之工具，提供預防醫學、治療醫學與復健醫學之照護需求。



以目前的病歷之記載方式係以發生之先後順序為基礎，此乃由於傳統之紙張記錄並無法對已經產生之資料進一步重組，而今雖然電子化，然而並未有任何改變。除了未能改變之外，更因資訊化的過程將原有之資料分散在各處，因此在利用上更容易因疏忽而失去資訊的完整性。以病歷而言，完整性乃是最不容許忽視之一環，失去了完整性可能導致醫師判斷病情上之困擾，因此本系統之首要目的即是依據醫療上之需求整合資訊化的過程所分散之資料，並以病患為中心重新組合以提供除了發生之先後順序以外之其他觀點，輔助醫師進行照護。

由於一份醫療記錄的內容可能同時包含各種不同來源或格式的媒體，為了達到資料整合之目的，同時又為了空間上之考量，因此目前以資料導向取得來自不同資料表的文字及不同電腦中的影像之醫療記錄具有相當多的困難，且無法滿足因應不同醫學資料的種類區分來取得所需醫學資料的集合，只能取得單一的醫學資料紀錄，更不能以物件導向的方式將所需的醫學資料包裝以提供動態方式提供使用者取得所需的醫學資料，本系統之另一個目的乃是以分散式架構建立伺服器端之機制，改變傳統以物件(Object)為基礎之存取提昇為以醫療記錄物件集(Set of Object)為基礎之存取方式，以符合醫療上之需求。

除了伺服器端之機制，為醫師存取之方便性，本系統採取瀏覽器作為在客戶端之工具，由於瀏覽器與伺服器之間傳統上是以 URL 作為彼此溝通存取資訊之方式，而 URL 的方式僅能描述實際提供資料的位置而取得單一的項目資料，因此本系統之希望透過以醫學資料名稱描述方式制定適合描述醫療資料之通訊協定的方式，達到傳遞醫療資料於瀏覽器與伺服器之間的目的

參、文獻探討

在以往，由於醫學資料散佈在 Internet/Intranet 醫學資料庫中，為了可以讓使用者可以透過一個共通介面查詢到所需的資料，最常採用的是共通閘道介面(Common Gateway Interface, CGI)的方式，使用者透過這介面取得所需的資料。然而 CGI 本身有其本身的限制，如會造成主機相當大的負載，且其 CGI 本身的設計並不適合於彈性的需求，如常因後端主機的改變或者前端使用者需求的改變而時常需改變其程式的結構。

在傳統的以靜態及資料導向取得醫學資料的方式無法滿足因應不同醫學資料的種類區分來取得所需醫學資料的集合，更不能以物件導向的方式將所需的醫學資料包裝以提供動態方式提供使用者取得所需的醫學資料。因此傳統的以靜態及資料導向取得醫學資料的方式在因應不同的檢查、檢驗的系統不同型態的資料種類上，只能取得單一的醫學資料紀錄，而在醫學上的單一類別資料不足以表示一個完整的醫學資料，如一個腸胃切片的檢查，其只有文字檢查報告而無影像上面的資料加以輔證是難以成為一份有意義的醫學上的資料。所謂一份有意義的醫學資料的集合定義即是應是一份可因應實際上的需求取得一段有效期間下的醫學資料集合，並不是侷限於所取得的資料僅是於單一項目的醫學資料，如一個肝癌病患接受其 MRI、CT、腹部超音波及血液方面的檢驗、檢查，在臨床的觀點一份有意義的醫學資料必須如同一份完整的病歷，其內容應該是包含下列幾個項目的集合：

1. 醫學資料提供者：提供資料的醫療院所。
2. 檢驗、檢查資料：在醫學資料提供者下所做的所有的檢驗、檢查資料，而這些資料可應映實際的需求的集合。如所需的資為一段期間內這並換所做過的各類檢驗、檢查資料的集合。

這樣資料的集合對一臨床醫師才是一份具有實質醫學資料集合。

為了增加散佈在特定區域內的伺服器效能、克服網路的問題及伺服器問題而無法提供服務。為了解決這些問題目前已有許多的為了增加 Web Server 的存取量既可適性的並且降低前端的存取的延遲時間的解決方案已被提出。而這些方案大致可分為伺服器端的解決方式(Server-side solution)及前端的解決方式(Client-side solution)。在伺服器端的解決方式中可分為兩種方法第一、使用區域網路中的叢集技術，並且分散負載在這些伺服器上。第二、使用區域性分布的伺服器的集合，在這區域性的分布的伺服器上的資料是互相被映射(Mirror)或者互相被快取(Cache)著，而前端需求的處理是被導引至一個區域上最接近前端的伺服器上。而前端解決方式例如快取(Cache)、代理快取(Proxy cache)、事先擷取(Pre-fetch)或者多重擷取(multi-fetching)等都能減少其前端存取的時間。其中事先擷取(pre-fetch)的方式由伺服器方式起始所提供的服務是目前較為使用的，但前端解決方式仍然是被考慮使用的，因為前端解決方式是使用來避免由前端需求處理所導致的網路延遲。在我們提出的系統中考慮一個為了提高伺服器的輸出量(Throughput)，可適應(Scaleable)及高容錯性(Fault-tolerant)伺服器解決方式的 Web Server。

近兩年來，Java 的發展速度十分驚人，其跨平台的特性，以及優異的物件導向語法，使得 Java 非常適合用來發展系統、開發應用程式。Java 虛擬機器利用軟體模擬電腦的特色，讓 Java 不只是一種新的語言而已。廣義的看，Java 已經是一種新的作業平台了。Java 虛擬機器可以在現有的電腦硬體和許多作業系統上執行，也可在專為 Java 設計的硬體上執行。

Java 把軟體抽線的層次往上推一層，在虛擬機器之下，甚麼都是抽象的，在設計之初，考慮的便是與作業系統或 CPU 無關的環境，不管底層的軟硬體如何替換，上面的 Java 程式永遠執行的安然無恙。

因此，我們所提供的發展平台，為了能適應各種環境的需求，一律使用 Java 為主要的發展與執行環境。藉助於 Java 優良的物件導向語法及跨平台特性，使的系統發展十分容易架構，也很容易在各平台上執行。

而在 WWW 的伺服器方面，SUN 改善原先 Web Server 所提供 CGI 效能不佳的缺陷而推出了 Servlet 的標準，配合 JSP (Java Server Page) 只要是以 Java 所寫出的應用程式就可以在任何提供 Servlet Engine 的 Web Server 來產生動態的網頁；第二，推出了 JDBC (Java DataBase Connectivity) 標準，使得 Java 應用程式沿襲“Write Once, Run Anywhere”的特性，而可以透過一定的方式連至各廠商的關連式資料庫系統中；再者，JavaBean 標準的制訂可以使得元件的定製標準化，而使得該元件可以在不同的開發環境中 (IDE Tools) 被再利用 (Reusable)。

- 作業系統的獨立

EJB 是以 Java 平台為基礎的規格，所謂的 Java 平台便是以「Virtual Machine」的型態獨立於各作業系統，故以 Java 程式碼所 Implement 的企業邏輯便能在各作業系統平台下的 JVM 來編譯執行了。

- 中介軟體層的獨立

既然 EJB 規格制訂了 EJB Server 所應具備的責任與義務，提供了標準的 API，所以由 IT 人員所開發的企業邏輯便可依其預算、功能及實際的需求來選擇廠商所提供的 EJB Server，爾後隨著時間的遷移及現實環境作業的需求，可以在不改寫應用程式的狀況下而可以無縫式(Seamless)的移植到其它廠商的 EJB Server 上。

- 資料庫系統的獨立

將企業邏輯 Component 化需要有一個永續性的機制，而現階段最佳的儲存機制仍然是關連式資料庫系統。

透過 JDBC(Java Database Connectivity)，使得元件對應關連式資料庫能標準化(Object-Relational Standardize)，換言之，不再需要為 Sybase、Oracle、Informix 等各種資料庫撰寫不同的 SQL 程式碼。

在醫療影像方面，PACS (影像的傳輸及擷取系統) 包括經由整合各色各樣數位化網路 (digital network) 來獲得 (acquisition) 影像與資料、加以貯存 (storage) 並顯示影像 (display)，並且利用 data management system 與院內其他資訊系統，例如醫院資訊系統 hospital information system、放射線部門資訊系統 radiology information system 等互相整合，作為醫療決策的根據，而形成所謂的證據醫學 evidence-based medicine。它的規模小自將 film digitizer 連接到顯示影像的工作站，也可以大到包含全院性的影像管理 (image management)，也有學者因而稱為 Image Management And Communication System (IMACS)。

PACS 的發展開始於八零年代的晚期,剛開始的設計目的在於針對許多放射線部門的某些特定儀器。每一種 PACS 的模組(module),功能上均為自主而無法與其他模組互相連繫。當日益增多的放射線儀器加入醫院的網路時,這項弱點也就越來越明顯了。系統的維護 (maintenance) 決定影像的發出路線 (routing decision) 機器間的協調 (coordination of machines) 容錯性 (fault tolerance) 系統的擴充性 (expandability of the system), 成為越來越難以解決的難題。這有部分是因為早期在設計的觀念上,並沒有想到一個大規模的 PACS,在實施時的複雜性,另一方面也因為在當時某些與 PACS 有關的資訊科技尚未成熟 (比如說如何及時並可令人相信的,將影像與相關的病患資料自影像檢查的機器傳送到 PACS controller 等等), 使得 PACS 的發展遇到了瓶頸。

在美國,由於軍方對於 teleradiology 的高度興趣,大力贊助與 PACS 有關的研究計劃,成立了所謂 Installation Site for Digital Imaging Network and Pictures Archiving and Communication System (DIN/PACS)。之後, Philip Medical System、AT& T 的參與,並選擇在 The University of Washington in Seattle, and Georgetown University /George Washington University Consortium in Washington ,D.C.來安裝 PACS。National Cancer Institute 也以 Multiple viewing Stations for Diagnostic Radiology 的名義,給 UCLA 一筆基金來發展第一個與 PACS 有關的研究計劃。

在產官學的通力合作之下,發現以往的問題關鍵在於儀器廠商並不打算遵循 ACR-NEMA 的工業標準,以致於放射線部門的各色各樣儀器之間,無法互相連繫並分享醫學影像資訊。所以陸續出現了 1982 年版的 ACR-NEMA,重點在於點對點訊息傳送 (point-to-point message transmission) 資料 (data formatting) 及呈現 (presentation), 還包括了初期有關指令及 data format 的一套字典。第二版的 ACR-NEMA2.0 出版於 1988 年,內容包括了對於硬體的定義 hardware definition、軟體的規範 software protocol、標準的數據辭典 standard data dictionary。

但是有關連接網路的議題,卻一直缺乏適宜的討論。也因為如此,在 1992 年版的便針對以連接網路的規範,作為中心議題,又因為其中的內容有相當的修改與增補,乾脆改一個新的名稱 Digital Imaging and Communication in Medicine (DICOM 3.0)。而 DICOM 3.0 最大的特色在於它採取物件導向的模式來交換訊息,以及利用現有標準的網路溝通規範。等到 DICOM 3.0 的規格 specification 越來越被廣為接受,廠商的態度趨於合作,同時也開始根據 DICOM 3.0 的規格要求,發展出新版本的軟體與設備。

在日本,第一個有關 PACS & PHD (Personal Health Data) 的國際會議,在 Japan Association of Medical Imaging Technology (JAMIT) 的贊助之下,在 1982 年的七月舉行。往後,這個會議便與 Medical Imaging Technology 的會議合併舉行,成為一年一度的盛事。在歐洲,自 1983 年後,每年一度的 EuroPACS 會議成為歐洲 PACS 訊息交換的主導力量。

PACS 的硬體需求包含完全數位化的影像系統 full digital system、影像存檔 image archiving、image communication and network infrastructure、呈現影像 displays 及動力學 ergonomics 等五大類。

PACS 的軟體需求包含標準化之影像聯繫 standardized image communication and

DICOM、整合放射線部門資訊系統與醫院資訊系統 RIS/HIS integration、讓 PACS 環境下之工作流程更加順暢 optimizing the workflow in a PACS enviroment。

影像數位化是 PACS 的基本前提，近年來幾乎所有的影像設備（斷層影像 cross-section imaging、透視攝影 fluoroscopy、血管攝影 angiography）都已經完全數位化，即使因為經濟因素的考量，而一時無法將所有的影像設備數位化，也可以用數位掃描的方式來完成數位化的要求。其次談到影像存檔 image archiving，對於一些大型醫學中心而言，放射線部門每年所產生的影像資料，高達數個 terabyte (TB, 10¹² Byte)。乍看之下，這麼大量的影像資料似乎是推動 PACS 最大的困難，但是如果考慮到醫學領域外，例如金融機構、保險業等，就不難發現其實這些實在是不成問題。

典型的 PACS 存檔方式，根據所需重複叫出的頻率以及所需貯存的時間長短，而有所不同。若是對於住院只有數天到數周的病患，則快速而昂貴的 RAID (Redundant Array of Inexpensive discs) 系統，因為它可提供 data mirroring，增加了資訊保存的安全性，是很理想的選擇。至於光碟櫃 (Optical disc jukebox) 的貯存量在 0.5~1.0 TB，最多只能提供線上 (on line) 貯存 1~2 年的資訊。因此現階段 PACS 的觀念，對於長期離線 (off-line) 的資訊保存，仍然必須依賴人力，將所要的光碟放入 PACS。如果資料量超過 20 TB，則需要用到磁碟機 (Tape-based storage system)。除了安全性高，一般認為目前現代的磁帶，可以用到 30 年之久。甚至於有學者主張用 WORM (write-once-read-multiple) 的光碟，取代 MO (magneto-optical) 光碟。不管用甚麼方法貯存，資料的完整性以及避免被盜用，才是最重要的。

影像資料在貯存時，有可回復式壓縮 (reversible compression)，可以到 2:1 或是 3:1 的比例。經常用在長期的貯存，而並不是那麼計較把影像資料重新叫出的時間長短。不可回復式壓縮 (irreversible compression)，則是以超過 10:1 的比例來壓縮影像。雖然有研究顯示即使壓縮比例到 40:1，臨床上仍舊看不出影像品質變差的跡象。但是理論上，不可回復式壓縮，仍舊存在失去重要影像資料的危機，所以不應該在放射線醫師作初步診斷之前使用。

目前世界上較為著名使用超過 10:1 的比例來壓縮影像，作為長期的貯存方式有倫敦的 Hammersmith Hospital 和美國軍方的 Medical Diagnostic Imaging Support System (MDIS)。或許未來的另外解決方案是許多醫院合作，在醫院外共同組成一個影像資訊貯存中心。第三是影像的連接與網路的基礎建設 image communication and network infrastructure。PACS 的有效運作，網路的架構是非常重要的決定因素。一般若要是能讓放射線醫師在專業的工作站有效率的作診斷 Soft-copy reading，影像調閱的時間，不能超過 3 秒鐘。標準的乙太網路 Ethernet，平均的傳輸速度約在 3 Mbit/s (雖然學理上的最大頻寬為 10 Mbit/s) 對於一張未經壓縮的 10 Mbyte (= 80 Mbit) 的影像，傳輸的時間超過 25 秒，這在日常工作中是不可能被接受的。

因此工作站與網路骨幹 backbone 的連接，頻寬至少要 100 Mbit/s，而能夠符合要求的網路規範 network protocol 有 FDDI (Fiber Distributed Data Interface) 和 fast Ethernet 頻寬為 100Mbit/s，ATM (Asynchronous Transfer Mode) 頻寬為 155Mbit/s。除此之外，有兩個重要的觀念分別是 autorouting 以及 prefetching。所謂 autorouting，是指新得到的影像，自動傳

輸到需要閱片的工作站。所謂事先調閱 prefetching，是指事先在網路傳輸較不壅塞時，將相關已經獲得的影像下載到需要閱片的工作站。這兩個觀念，對於影像管理（image management）非常重要。

接下來要討論的影像的顯示 image display，也許是目前整個 PACS 硬體中最弱的一環。因為除了最近上市的 2x2.5-k 高解析度監視器 high resolution monitor 外，一般的 video monitor 若是與傳統的看片箱 light box 比較起來，不論是空間解析度，或者是最大亮度，都要稍差一截。可是昂貴的價錢，迫使一般醫院將 2x2.5-k 高解析度監視器局限於放射線部使用。另一方面要注意的是任何陰極射線管的監視器，在最大亮度及解析度方面，每一台的表現都有所不同，而且會隨時間而品質更糟。

最後不能不討論的是 PACS 動力學 Ergonomics，這是長久以來一直被忽視的議題。多數的 PACS 設施中，嘈雜的電腦工作站及硬碟，就緊接著放射線醫師長時間專心致志打報告的閱片室。有些空調又不足以配合電腦或監視器的散熱，甚至於室內燈光的位置、亮度都會影響到醫師使用 PACS 的意願。所以 PACS 工作場所的動力學，應該在一開始規劃閱片室時，就應從長計議。

早期在開發 PACS 的過程中，有一大缺失，在於 PACS 的安裝，以 PACS 的各成份而，影像與資料的交通，純粹是站在以賣方為主的設計，有時甚至於同一公司的儀器，也無法相互交通。1994 年，Peters 和 Imhof 在一篇專欄中，寫到“如果電子醫學產業能夠支持一個更開放，而且與賣方無關的 PACS 結構，那麼 PACS 就能成真”。但是經由 the American College of Radiology（ACR）和 the National Electronic Manufacturers' Association（NEMA）不斷的努力，終於發展到今天有 1993 年 DICOM,3.0 版的規範的出現，使標準化影像的交流能存在於不同廠商所生產的 PACS 的各種成份的美夢得以成真。因為廠商並不一定支援每一種 DICOM 標準中的要素（elements），所以如果預期合乎 DICOM 標準（DICOM compliant），那麼製造業者就應該提出 DICOM conformance statement 來作為證明。另一方面，仍然有所謂 shadow groups in DICOM 的存在，這是為了要因應廠商要貯存它們所獨有的資訊，因為在 DICOM 的規格下，讀寫影像的速度相對較慢而不得不如此。18 如果資訊是與檢查結果之後的繼續處理有關，比如 CT 或 MRI 的切面位置，如果這些資訊被貯存在未被規範的 shadow groups，一旦影像被傳送到不同廠商的廠商儀器，則這些資訊將不復存在。

另外一個大問題在於 RIS/HIS 的整合。以往人們將 PACS 和 RIS 視為兩個截然不同的個體，結果 PACS 和 RIS 中的數據 database 無法相互配合。PACS 不知道 RIS 的檢查排程，當然無從作有效的預先調閱（prefetching），或是進一步的影像管理。而僅僅 PACS 和 RIS 整合仍然是不夠的，比如一位病患如果每次到醫院的病歷號碼都不相同，那麼 PACS 就無法找出病患之前的檢查記錄。在這種情形下，用人工的方式來合併病患的病歷，不但耗費時間，另一種潛在的危機則是錯把馮京當馬涼，把影像歸納到另一位完全不相關的病患。所以經由適當的 HIS-RIS-PACS 的界面，使 PACS 能夠得知病患在院內目前的位置，進而得以將影像發送到門診或是病房。

最後要討論到的是如何讓 PACS 的工作流程更為順暢。多年來，許多 PACS 的製造廠

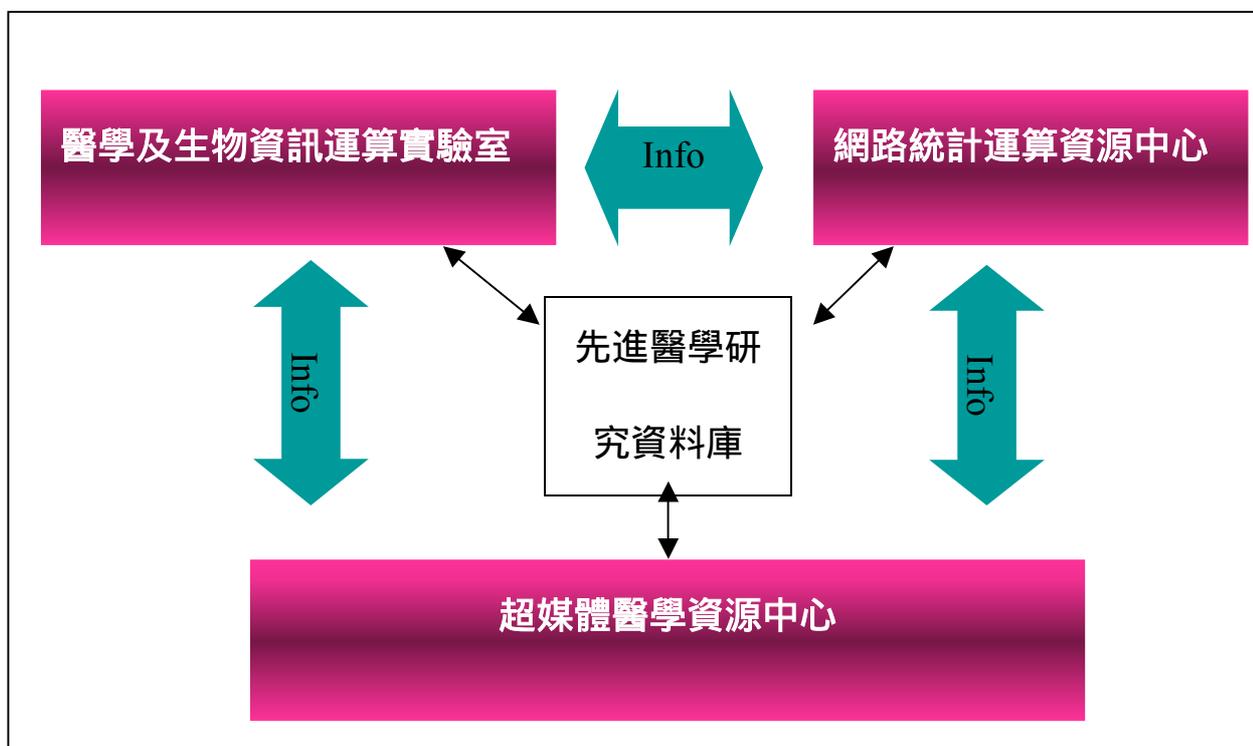
商或是開發軟體的工程師，都忽略了 PACS 的成功與否，在於 PACS 能否讓放射線醫師的生活過得更快樂，至少不是讓他們更難過日子。近來的 PACS 發展，已經能夠作到智慧型的預先調閱 intelligent prefetching, default display arrangement, 以及 one-button image loading, 使得 soft-copy 的閱片方式比起以往更為有效率。除此之外，工作站中的各種軟體，比如 3D 後處理時的多重平面重組 multiplanar reconstruction volume rendering segmentation 等，對於病灶本身得以有更進一步的觀察，而增加診斷的正確性。此外如與專家系統 expert system 的結合，或是電腦輔助診斷 computer aided diagnosis scheme, 對於間質性肺病 interstitial lung disease、血管狹窄或腫瘤體積的自動定量 automatic quantification of vascular stenosis or tumor volume, 電腦輔助發現肺部結節、乳房腫塊，使放射線醫師的工作更加能勝任愉快。

PACS 對醫療所產生的影響，可以分為下列幾點來作說明：首先，因為全面實施 PACS，而走向無片化 filmless。無片化，則病患做完檢查不需在放射線科等候，縮短了看門診所需時間。放射線科的技術人員則因為不需拍片子，而省下更多的時間來作更多的檢查，或是把省下的時間來作品管。這樣一來，放射線部門的營收增加，影像的品質也大幅提升。對於急診或是加護病房的患者而言，PACS 的閱片監視器 viewing monitor 可以調整視窗範圍 window level、中心值 center 的大小，減低了因攝影條件不佳，而必需重照的比率 reuptake rate。對於放射線醫師而言，在判讀時可以自行調閱既往的舊片，而不要假手檔案室管理員，減少了等待的時間，也提高了診斷的正確性。臨床醫師在門診看預約的病人，可以預先調閱，又可參考以前的檢查影像或記錄，很快便能掌握病情，而作迅速的決定，提高了醫療照護的品質。更值得一提的是，PACS 剛好創造了醫療資源共享的環境，多位不同科系的專家，可以分享影像資源，而不會像以往因為片子在其他醫師手中，結果會診或是其他醫療作業無法繼續進行。有些調查發現，在 PACS 實施之後，臨床醫師對於 CT、MRI 之類的斷層影像，較以往更常以電話的方式，來徵詢放射線醫師的意見，因而加強了與放射線醫師的互動。

肆、研究方法

資訊科技的演進已經影響到醫療的品質，越來越多的現象可以發現民眾、醫事人員、醫學研究人員對於醫學資訊的急切需求比率正在上升中。再加上基因科技的進步，生物資訊與醫學資訊的連接與整合已經成為全世界相關領域的重要研究議題之一。本計畫及希望能建立醫學資訊與生物資訊資料庫以提供研究者作進一步的研究與使用。因此本計畫分為三項子計畫：醫學及生物資訊運算實驗室、超媒體醫學資源中心、網路統計運算中心。

本計畫中的三個中心與核心設施的關係圖如下所示：



各計畫分別敘述如下：

(1) 總計畫

協助資源規劃與分配協助各項子計畫達到下列目標：

1. 收集並建立生物及醫學資訊的資料，開發生物醫學資訊相關分析軟體工具，應用於如生理信號分析、microarray 資料分析、基因序列資料探勘與蛋白質交互作用預測。
2. 建立多媒體醫學資料庫及開發儲存、展示相關技術，支援研究、教學及臨床應用。
3. 建立台北醫學大學線上生物統計分析之環境，支援校內其他研究計劃之統計需求。

並可以與校內其他計畫相互支援或是提供資訊服務，例如：

1. 做為校內另一提昇能量計畫 " MOLECULAR GENETIC AND BIOLOGIC STUDIES OF ATOPIC DERMATITIS " 在 microarray data analysis 部分的技術支援，主要運用 Data mining(e.g. Clustering and Classification)與 Neural Networks 加以分析基因表現之差異，並藉以驗證與開發相關應用。
2. 內試鏡影像教學系統之建立：醫學影像的教學系統將擁有為數眾多的醫療影像。此計畫可以提供研究數據整理、統計以及計劃事項之管理工具。以提昇這些醫療影像的用途。
3. 醫學診斷決策支援系統開發：藉由本資料庫的提供，可以挖掘出更多決策支援可以應用的文獻以及臨床實驗數據，增進決策支援系統的可信度以及可行性。
4. 證據醫學的實作模型：證據醫學需要大量的醫學資訊來佐證其立論，因此建置本資料庫將有助於資料的蒐集與分享。

總計畫之執行時程如下：

第一年：規劃硬體架構、採購及安裝，各子計畫方向整合

第二年：規劃資料庫建置，各子計畫進度監督與整合

第三年：持續資料庫與相關軟硬體建置，協助各子計畫推廣項下所提供的醫學資訊服務

(2) 醫學及生物資訊運算實驗室

在本計畫中，首先目標是建立一個醫學及生物資訊運算實驗室，並收集相關之資料庫，接下來我們計劃建立模型、工具及運算方法用來做醫療資訊探勘並評估其功效。為了要建構一個這樣的模型，我們將利用目前已知的技術包括：決策樹，類神經網路，原始運算法則，貝式網(Bayesian nets)，差分函數等。這些技術可以應用在影像辨識，病歷本文資料探勘等。

數位生醫訊號處理要研發實驗室工具，我們預計將研究、引進及開發相關數位生醫訊號處理技術，以全面提升數位生醫訊號處理應用於醫學教育及研究之能力。預期進行：(1)測試應用分析各種基礎及臨床醫學訊號的可行性及可信度。(2)根據我們應用的結果，發展、設計、測試其相關軟體工具，鼓勵及推廣，讓研究人員可以用全新的角度去詮釋他們記錄的訊號。

本實驗室期望利用三年的時間，從基礎建設到分析方法的開發最後到提供服務。

第一年：基礎建設：場地規劃、設備購買、資料庫收集。

第二年：持續進行資料庫之收集基礎建設資料庫之分析、資料探勘及統計分析方法及工具之收集及研究開發、校內服務規劃。

第三年：同時對資料庫繼續充實、資料探勘及統計分析方法及工具之測試及實際運作、開始進行校內服務的工作。

(3) 超媒體醫學資源中心

本計畫希望藉由完成醫學研究超媒體資源儲存中心，讓每個研究者或研究計畫既是超媒體資源的使用者又是資源的提供者。並且宛如擁有一群專業的資訊人員為其研究資訊的儲存、備份、安全管理與使用效能並控制媒體格式之複雜度，使其能安全舒適地專心於研究的方向，運用資訊而非受制於資訊，提升研究能量。

近年來，Java 的發展速度十分快速，其跨平台的特性以及優異的物件導向語法，使得 Java 非常適合用來發展系統與應用程式。Java 虛擬機器利用軟體模擬電腦的特色，讓 Java 不只是一種新的語言而已。廣義的看，Java 已經是一種新的作業平台了。Java 虛擬機器可以在現有的電腦硬體和許多作業系統上執行，也可在專為 Java 設計的硬體上執行。因此，為了能適應各種環境的需求，我們的發展平台，一律使用 Java 為主要的發展與執行環境。藉助於 Java 優良的物件導向語法及跨平台特性，使得系統發展十分容易架構，也很容易在各平台上執行。

為求達成使用者與系統之間以名稱方式溝通，必須先行定義雙方的通訊規範，此通訊規範包含了下列項目：

1. Protocol：存取該醫療資源使用之通訊協定。
2. SocietyID：醫療資源擁有者之身分證號。
3. Provider Descriptor：醫療資源之提供者。
4. Category Descriptor：醫療資源歸屬之類別。
5. Date Descriptor：醫療資源發生之日期區間。
6. Physician：醫療資源之著作醫師。
7. Item Descriptor：醫療資源隸屬之項目。

8. Information Descriptor：儲存醫療資源之資料型態。

該通訊規範在本系統稱為 UMRL(Uniform Medical Resource Locator)通訊協定。當系統接受 UMRL Protocol 的需求，系統先行確認使用者的安全性並且以邏輯的方式提供使用者以醫學資料名稱的方式取得所需的醫學資料。透過以醫學資料名稱描述方式所表現對於醫學資料的需求，如下列的描述：

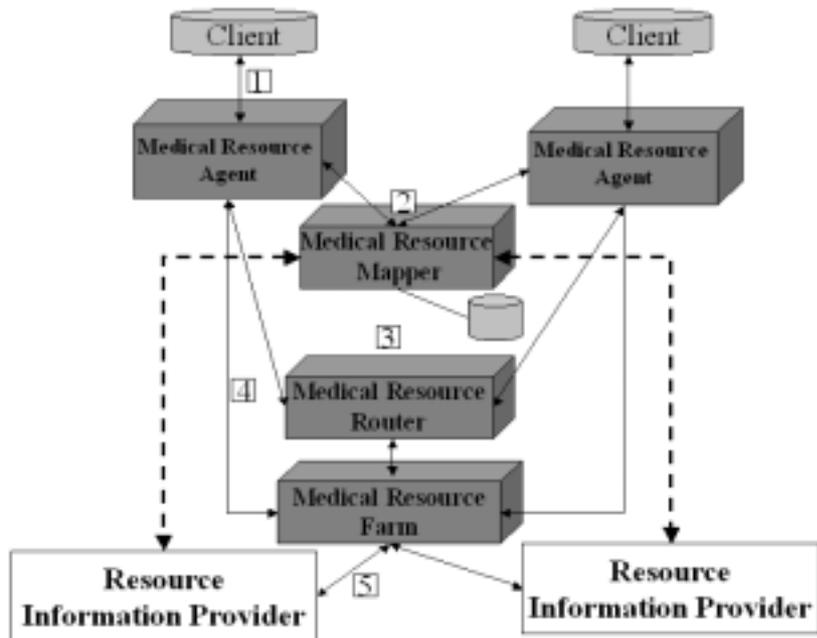
<Protocol/SocietyID/Provider/Category/Item/Physician/Date/Information>

此種以醫學資料的名稱表現方式完整的描述所需要資訊種類、型態及邏輯上儲存位置的集合。並將此位置的集合轉換為具體醫學資料集合。透過 UMRL 與傳統 URL 取得資料方式的差異性比較如下表：

UMRL	URL
Logical	Physical
Medical Record	Medical Item
Resource Name	Resource Address
Meaningful Address	Meaningless Address
Category Based	Directory Based
Object Oriented	Data Oriented

使用者透過 Internet 的方式以資料名稱解譯取得所需要的醫學資料集合，以此資料集合產生相對應的具體醫學資料，主要分為六個模組：

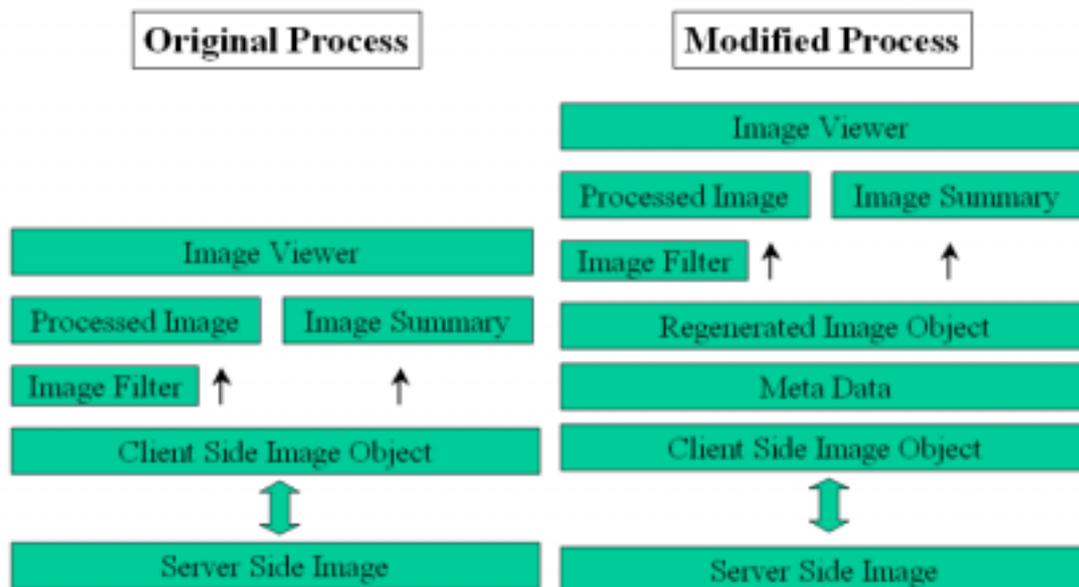
1. Medical Resource Agent (MRA): 接受使用者透過 URML 通訊的查詢，MRA 程式產生相對應於醫學資料名稱予以解譯成所需的 Logical Medical Name。
2. Medical Resource Mapper (MRM): 將 MRA 所得到的 Logical Medical Name 轉為資料實際儲存位置，MRM 的資料庫能夠解譯所有 Logical Medical Name 及其所對應的實際位置，並將結果傳回 MRA。
3. Medical Resource Router (MRR)：依照目前網路及 MRF 中的伺服器的狀況分散來自於經過 MRM 解譯過所產生的 MRA 對 MRF 的需求。
4. Medical Resource Farm (MRF)：提供 Fail Over 及 Load Balance 的功能來接受由 MRR 機制判斷 MRF 中的伺服器能夠接受其 MRA 所發出的 Request。
5. Resource Information Provider：提供醫院各檢驗部門上傳資料的轉換介面，並且透過網路定時將所獲得資料轉為對 MRM 所需要的邏輯名稱與實際相對應位置上傳至 MRM。



除了伺服器端之機制，為醫師存取之方便性，本系統採取瀏覽器作為在客戶端之工具，而在前端之影像處理其一般流程如下：

1. 利用前端之影像物件儲存伺服器端之影像
2. 由影像過濾器在改變影像之色盤以達到改變亮度及對比之效果
3. 將調整色盤之結果輸出到影像檢視器





然而經過測試以後，雖然 Client side Image object 已經將後端之資料取至客戶端，但是只要以影像過濾器改變影像之色盤時就會自動到伺服器端重新取得影像，經過了許多測試並將製程修改，將影像資料還原成 RGB 的資料轉為 Meta Data 才能擺脫重新存取網路之現象，提高影像處理之效能。

第一年

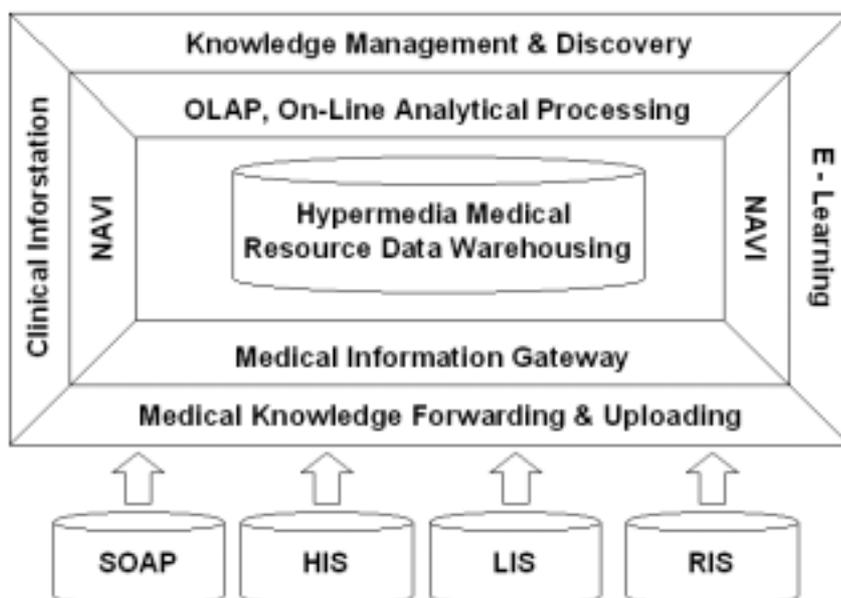
由於醫學資源之主要來源乃是病歷，因此病歷電子化乃是充實醫學資源之最主要工作，病歷電子化後才能進一步提供搜尋及統計之用。然而，各種醫學資源之型態因各臨床科別特性不同而異，因此整理並訂定資源格式方得以收集醫學資源。一旦資源格式訂定後，實作收集介面以累積醫學資源亦為一項重大工程。本年度之重點即是將不同來源之資料整合，並提供介面以資查詢。

第二年

基於第一年之基礎，本計劃將持續病歷電子化及收集介面以累積醫學資源，並進一步建構資料倉儲，將第一年實作收集介面累積之醫學資源作多維度之轉換，管理大量多媒體醫學資源，並利用線上分析處理之功能，提供醫學資訊服務。本年度之重點即是將整合之資料提供得以直接運用於統計與資料探勘使用之資料集。

第三年

當大量資料之資料倉儲建構完成，由資料倉儲本身或統計與資料探勘產生之知識必須有效管理，因此建立知識庫之需求更加殷切。本年度之重點即是結合知識管理與資料倉儲，並提供生物資訊運算與臨床資訊之整合服務，達成提昇研究能量之目標。



(4) 網路統計運算資源中心

台北醫學大學研究團隊人員近三年來接受國科會、衛生署、國家衛生研究院及其他公私機構委託執行的研究計劃案約一百二十三至一百二十九件。大多數計劃主持人或共同主持人並未具備統計背景，以致資料統合、分析時常面臨困境，造成研究進行之瓶頸。目前處理的方式多為計劃主持人個自摸索或個別向本校目前具有統計背景的教師請求支援，缺乏全面性、系統性地規劃與整合。本計劃係整合台北醫學大學之統計資源。藉由結合具統計專長人員的諮商服務及發展智慧型生物統計專家輔助系統等相關工具，設置網路統計運算資源中心，輔助基礎與臨床醫學研究人員研究設計、數據統計與結果分析，以提昇台北醫學大學研究團隊之資料分析能力，進而凝聚研究能量，促進研究風氣。

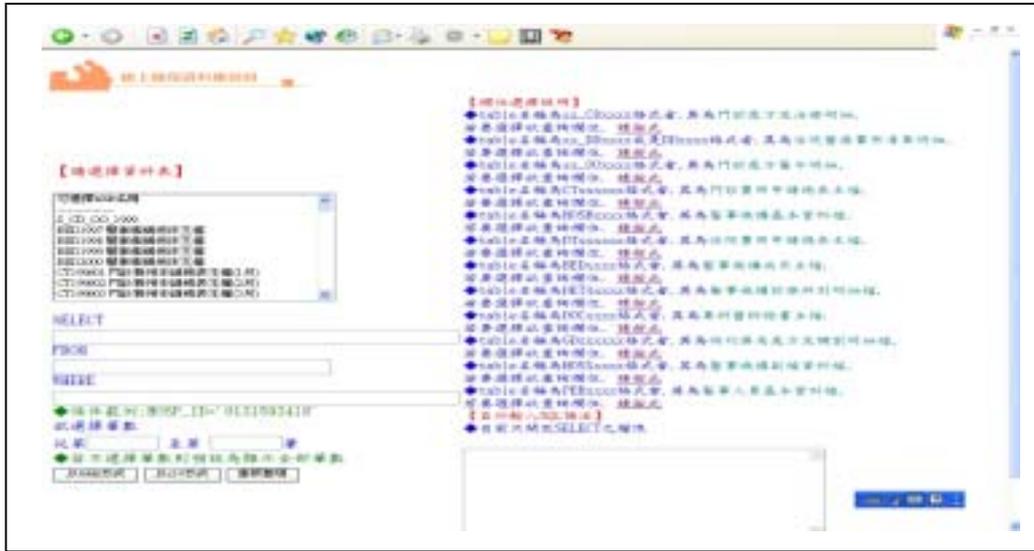
伍、結果與討論

(一) 健保資料庫：根據中央健保局所提供之全民健保資料，範圍為民國 85 年至 89 年之間的就診資料，目前共計有 221,359,487 筆資料，資料總容量為 38,378 MB。資料庫的系統操作介面為 RedHat Linux7.3 及 MySQL，共建有 453 個資料表，包括各年的特定主題分檔：

1. 癌症門診處方及治療明細與住院醫療費用清單明細
2. 重大傷病醫療使用資源門診處方及治療明細與住院醫療費用清單明細
3. 醫學中心門診處方及治療明細與住院醫療費用清單明細
4. 職業病與職業傷害門診處方及治療明細與住院醫療費用清單明細
5. 外傷門診處方及治療明細與住院醫療費用清單明細
6. 交通意外事故住院醫療費用清單明細
7. 復健治療門診處方及治療明細與住院醫療費用清單明細
8. 論病例計酬門診處方及治療明細
9. 糖尿病門診處方及治療明細與住院醫療費用清單明細

由這個資料庫的建立可分析及推斷出國民所罹患的疾病以及醫療行為，加上因為全民健保資料庫是一個全國性的健康資料庫，因此可以經由資料查詢與統計的方法來得知國民最易罹患的疾病及照成國民十大死因的因素等資料，進而可以推廣國民預防醫學的概念，由改變國民的飲食等方法來減少國民致病的因素進而達到延長國民平均壽命的目的。本計畫除了將此資料庫建立之外，並建有線上查詢之介面，以方便計畫內的研究人員使用。範例使用畫面如下：





(二) 醫院醫療資料庫

前項所述的為健保局所提供五百分之一的抽樣檔資料，本計畫除了建立抽樣健保資料庫之外，也建有數個醫院的完整健保資料資料庫可供作進一步的分析與查詢。為了保障病患隱私，這些資料皆有加密處理。

(三) 醫學及生物資訊運算實驗室資料庫：

1. 已建立一個醫學及生物資訊運算實驗室，目前積極收集資料庫包括：GenBank：為美國 NCBI 所公布的基因資料庫，收集所有 public domain 的 DNA sequences，目前大約有 18,197,000 sequence records 中有 22,617,000,000 個 bases。此資料庫每兩個月會更新一次，在北醫有一個 mirror site，可供有興趣的人下載使用。北醫生物資訊資料庫 <http://srs.tmu.edu.tw/>
2. NCBI Toolbox：為美國 NCBI 針對 GenBank, Entrez, BLAST, Sequin, OMIM, RefSeq 等等所發展的基因資料庫工具。
3. Cn3D：為美國 NCBI 所發展透過 Entrez 服務來觀看分子結構的 3D 立體瀏覽工具，他可以同時秀出 structure、sequence、以及 alignment。現在它更可以做到 annotation 及編輯 alignment 的功能。
4. SNP Databank：目前已下載人體二十三條染色體之原始資料，可以供研究人員研究分析。
5. 生物醫學訊號資料庫：包含多種心音及心電圖資料庫。設置新型數位訊號處理裝置包括：Stellate System: Harmonic, Digital EEG and Long-term Monitoring System. 和 BIOPAC System, MP150 Data Acquisition and Analysis System. 利用來收集生理訊號並建立資料庫 <http://203.71.92.116/BMS/index.php>
6. 在 ECG database 方面，收集有以下資料：
 - (1) MIT-BIH Arrhythmia Database: 48 個病人各 3 小時的心律不整心電圖資料，已經過專家確認並標示每一心跳之屬性，可作為心電圖教學訓練與設計心電圖心律不整分析方法時測試使用。資料庫大小約 400MB。
 - (2) European ST-T Database: 48 個病人各 2 小時的疑似心肌缺血之心電圖資料，已經過專家確認並標示異常 ST 段發生之時間，可作為心電圖教學訓練與設計心肌缺血分析方法時測試使用。資料庫大小約 300MB。
 - (3) Apnea-ECG Database: 70 個病人各 8 小時的之心電圖資料，期間病人有 Apnea 發生，ECG 可用於幫助預測 Apnea 之發生，並加上標示。此資料庫可用以心肺關聯之教學研究與預測 Apnea 之方法測試。資料庫大小約 600MB。

(4) PAF Prediction Challenge Database: 100 個病人各 30 分鐘的心電圖資料，期間病人有 AF(心房顫動)發生，並加上標示。此資料庫可用以研究 AF 發生之預測，也包含教學價值。資料庫大小約 400MB。

7. 在心音資料的收集方面，我們採用 BIOPAC 數位生理監測儀器以及一個與胸部接觸的麥克風 TSD108 來截取心音。我們分別對其主動脈區、肺動脈區、三尖瓣區及僧帽瓣區等四個位置做心音量測，每個位置的量測時間皆為二分鐘。而透過心臟專科醫師的協助，我們已分別收集到皆由臨床醫師判別有心雜音病患的心音 31 例以及一般人的心音 20 例，總計為 204 個檔案。大約 1.4GB。

8. 在生物資訊資料庫方面目前主要收集核甘酸與蛋白質序列資料庫包含:

(1) Entrez: 主要來自 GenBank, RefSeq 和 PDB, 包含有 Human, Mouse, HIV, Hepatitis C Virus, E. coli 等物種之基因序列資料。也包含如 dbEST, dbGSS 與 dbSTS 等子目錄。

(2) dbSNP: single nucleotide polymorphism 之資料庫，研究個體差異。

(3) OMIM: *Online Mendelian Inheritance in Man* 人類相關基因疾病資料庫。

以上 3 資料庫大小總合約 7GB。

9. 建立一個醫學及生物資訊運算(Biomedical Informatics Computing)實驗室：生物資訊(Bioinformatics)正在扮演非常重要的角色於人類疾病及基因關聯性的發現和描述，其中包括了許多分子生物學研究。我們將發展而且整合一組新的工具，利用語言分析及資料探勘，並且使用這些工具來分析多種資料庫資源訊息。

(1) 生物資訊運算高效能工作站

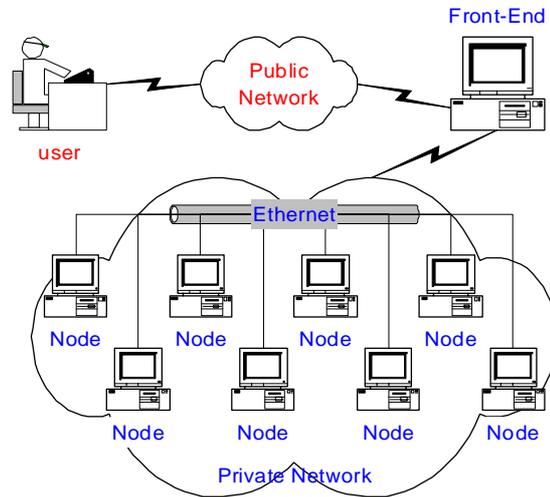
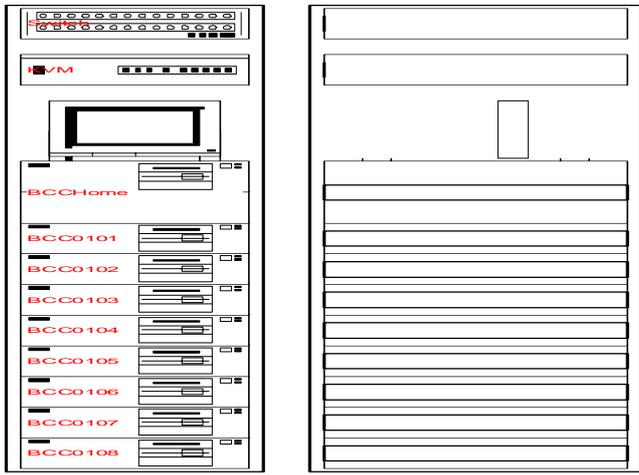
本計畫補助了部分的經費於生物資訊運算高效能工作站。目前系統已完成各項測試，置放於本所生物資訊運算中心，由一位老師與兩名研究生管理，並定期與廠商做技術交流，以維護系統的穩定與更新。在研究與教學方面，除前述之資料大部分建構收集於此外，目前開始研究生物資訊的工具，如序列比對用的 Blast 等如何使用 PC-cluster 來完成；在醫學影像處理上，對於高解析度的影像如數位 X 光影像等，需應用複雜影像處理方式，如 image registration 時，如何使用本系統來加快處理速度；在教學上，我們搭配本系統從事包含平行處理程式設計、生物資訊、醫學影像處理等相關課程的教學。本系統對於需高效能計算能力之教學研究上有相當的助益，並在培育相關技術人才上有所貢獻。

A. 系統架構

(a) 系統邏輯架構

利用兩部以上的個人電腦與電腦網路通訊設備等硬體設備，互相的串接後，就構成一個獨立 (standard-alone) 的區域網路 (local area network, LAN) 結構，搭配叢集技術所需要的軟體系統 (例如：作業系統 (operating system)、佇列系統 (queuing system)、平行計算函式庫 (parallel library) 發展工具 (develop tools)、編譯器 (compiler))，以及相關的演算法 (algorithms)，於是形成了一組具有平行與分散的計算平台 (computing platform)。

(b) 系統功能架構



(c) 硬體架構

在前述的概念中，整個硬體系統的架構基本上主要是由兩部以上的個人電腦與電腦網路通訊設備所完成的；前者又可分成：前端控制伺服器（front - end computer），以及運算節點（computational node computer）等兩種不同的角色；後者，則使用交換器（switch）來完成。

i. 電腦網路通訊設備 - 交換器

如何地使兩部以上的個人電腦能同時的一起工作（目前先排除軟體的部分），首先，必須利用實際線路串連在一起，故我們就必須使用到電腦網路通訊設備，其中則包括有：集線器（HUB）、橋接器（bridge）、交換器，以及路由器（router）等不同的設備。在整個叢集系統中，網路延滯時間（network latency）- 從甲台電腦傳輸到乙台電腦的反應時間最短，以及輸出入效能（I/O performance）- 在同一時點內網路所能傳輸的資料量相較於大 - 為最主要的考量兩大因素，再加上考慮價格/效能比（performance - cost ratio）的問題之下，並根據上述的討論，實為選擇使用交換器的主要原因。

ii. 個人電腦 - 前端控制伺服器

所有的個人電腦經過實體線路串接後，則必須有一個進入點（entry point），進入該系統，也就是所謂的主控伺服器（一般使用者與系統管理員皆利用此台電腦，進行工作的指派與日常或緊急的維護與管理）在此我們特稱作：前端控制伺服器。其主要負責的工作任務計有：管理使用者的帳號/群組（Network Information System, NIS）、分享儲存空間（Network File System, NFS）、工作分派（Queuing System）、檔案備份與還原（Backup and Restore）、其他運算節點的管理，以及其他等項。

iii. 個人電腦 - 運算節點

整個平行或分散系統的概念，則是指必須有兩台以上的個人電腦設備在分別同時地執行必要的運算作業，這些僅負責計算工作，稱為運算節點。我們常聽到「大量」的這個概念，實指此一部份。在接受到前端控制伺服器來的任何命令（包括有 command script file，以及執行檔等不同的類型）後，忠實不二地執行該項指令，待完成後，則將輸出結果送回。

iv. 多台電腦監控器（keyboard, video & mouse console switch）

若每部個人電腦都必須使用一組的鍵盤、滑鼠，以及顯示器的話，其成本則為大幅地上揚，此外，在整個系統中，僅有前端控制伺服器需要使用到該組設備，以及，簡化兩部以上的個人電腦的操作方式與節約成本等的考量下，在此使用了該項裝置。使用者若必須至本系統實體放置地點操作本系統，則可以在同一個位置上，利用該項設備，分別地監控每一個人電腦目前的狀況，而不須要每次針對單一個人電腦進行螢幕、鍵盤，以及滑鼠插拔的動作。

(d) 軟體架構

i. 作業系統 (operating system)

本系統以 Red Hat Linux 6.2 (English) 做為基礎的運作核心系統，由於其具有高度的彈性，亦即，我們可視實際的需求修改作業系統的任一部份，以達成我們所需要的系統環境，故採用之。該系統目前由美國 Red Hat 公司負責維護。所有的個人電腦系統中，我們均安裝相同 (包括 distributed OS version、Kernel version、language、software packages) 的作業軟體，以求全部的統一與將來維護工作的簡化。

ii. 佇列系統 (queuing system)

分派工作 (dispatch)，以及監控運算節點 (computational node computer) 負載的狀況為佇列系統 (queuing system) 的最主要工作。當使用者欲 submit 一個工作時，佇列系統會比對事先定義好的佇列規則與使用方式來判斷接受或不接受該項工作，以及工作的執行方式。

本專案系統採用 DQS (Distributed Queuing System)。該系統目前由美國 Florida State University 的 Computational Science and Information Technology 負責維護。

iii. 平行計算函式庫 (parallel library)

要使循序處理的程式達成平行處理的功能，則原程式碼必須加入 PVM (parallel virtual machine) 或 MPICH (message passing interface) 等兩大函式庫，再配合訊息傳遞 (message - passing) 或鎖定 (lock) 等的機制，才能達成所謂的平行運算方式。

iv. 發展工具 (develop tool, language) 與編輯器 (compiler)

主要的發展工具 (程式語言) 仍以 C、C++，以及 Fortran 77 為主。編輯器則使用 gnu 所提供的版本。

10. EMBOSS：我們利用運算中心的 PC-cluster 架構了一套生物資訊分析應用軟體 EMBOSS (European Molecular Biology Open Software Suite)。

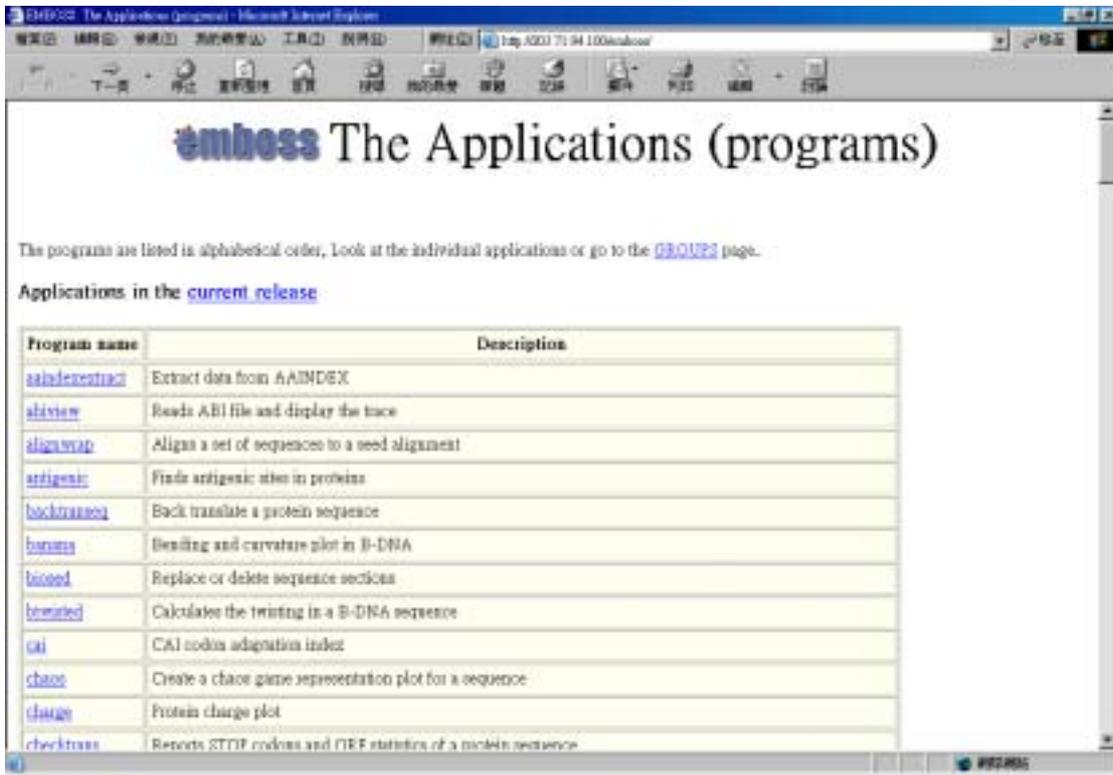
EMBOSS 是由歐洲多個學術單位共同開發出的分子生物學程式組合，免費提供全球下載使用，並且是開放原始碼的軟體。整個程式組合已包含超過一百個程式以上，分析功能相當完整，加上沒有分析序列的大小限制，以及具有自動辨認各種序列格式等功能，已能取代價格越來越貴的 GCG 套組，更適用於基因體層次上的分析工作。在程式操作方面，EMBOSS 除了可讓使用者在命令列模式下執行程式外，更有多種圖型化及網頁型式的使用者介面可供安裝使用，如 Jemboss 使用 Java；Pise 使用網頁及 XML 等。配合完整的使用說明文件，能大幅地降低初學者的進入門檻。整個 EMBOSS 根據其功能分為以下幾組：

Group	Description
<u>Alignment consensus</u>	Alignment consensus
<u>Alignment differences</u>	Alignment differences
<u>Alignment dot plots</u>	Alignment dot plots
<u>Alignment global</u>	Alignment global
<u>Alignment local</u>	Alignment local
<u>Alignment multiple</u>	Alignment multiple
<u>Display</u>	Display
<u>Edit</u>	Edit
<u>Enzyme kinetics</u>	Enzyme kinetics
<u>Feature tables</u>	Feature tables
<u>Information</u>	Information

<u>Nucleic 2d structure</u>	Nucleic 2d structure
<u>Nucleic codon usage</u>	Nucleic codon usage
<u>Nucleic composition</u>	Nucleic composition
<u>Nucleic cpg islands</u>	Nucleic cpg islands
<u>Nucleic gene finding</u>	Nucleic gene finding
<u>Nucleic motifs</u>	Nucleic motifs
<u>Nucleic mutation</u>	Nucleic mutation
<u>Nucleic primers</u>	Nucleic primers
<u>Nucleic profiles</u>	Nucleic profiles
<u>Nucleic repeats</u>	Nucleic repeats
<u>Nucleic restriction</u>	Nucleic restriction
<u>Nucleic transcription</u>	Nucleic transcription
<u>Nucleic translation</u>	Nucleic translation
<u>Phylogeny</u>	Phylogeny
<u>Protein 2d structure</u>	Protein 2d structure
<u>Protein 3d structure</u>	Protein 3d structure
<u>Protein composition</u>	Protein composition
<u>Protein motifs</u>	Protein motifs
<u>Protein mutation</u>	Protein mutation
<u>Protein profiles</u>	Protein profiles
<u>Utils database creation</u>	Utils database creation
<u>Utils database indexing</u>	Utils database indexing
<u>Utils misc</u>	Utils misc

EMBOSS 目前架構在生物運算中心的一組 Linux Cluster 上。使用者可以透過瀏覽器連線到：<http://203.71.94.100/emboss/>網址使用

連結後畫面如下：



目前已規劃應用於本校生物資訊學相關之研究教學活動。未來並將安裝 Jemboss 讓使用者有更簡易的圖型介面。

11. 利用文字語言方法分析生物資訊文獻資料庫

在現今生物資訊爆發的時代，生物基因工程是很多人想去研究的計畫。但是如何去跨出第一步，或是在搜尋到很多的相關文章時，卻發現說無法在為量眾多的基因資料庫裡，找到符合自己所需要的資料。本計劃目前提供了一個生物基因文獻的資料庫，可提供部分基因與相關疾病文章供使用者下載，另外也提供一個運算環境(利用醫學及生物資訊運算實驗室的設備)，可提供使用者快速的找出符合自己需求的資料，該運算環境的規格配備如下：

1. P4 2.53G 的高速 CPU 核心，及 512MB 容量的記憶體。
2. 以效能較高的 UNIX 作業系統(採 FreeBSD)來作為核心。
3. 可使用 Perl 作文句的快速處理。
4. 數十 GB 的硬碟容量，在運算時提供所需的足夠容量環境。
5. 允許使用者遠端登入操作，方便使用者使用。
6. 北醫提供服務，可以利用帳號來登入：

Login name : user1~user9(提供 9 個 user name，擇一即可)
 Password ; user1~user9(注意大小寫，要與 uesr name 一樣)
 登入後便可使用該運算環境。

我們應用資訊檢索(Information Retrieval)和自然語言處理(Natural Language Processing)的方法，發展一套改進 Entrez 系統搜尋結果的演算法。我們已經建立了一個簡易的查詢系統，Biomedical Literature Searching System (BLSS)，其所搜尋出來的文獻和搜尋條件的相關性，比 Entrez 系統更高。未來，我們會以此系統開發應用程式，幫助 MEDLINE 的使用者更有效率地搜尋出其所感興趣的文獻。(Chien-Yeh Hsu and Yung-Tai Yen, Developing an IR and NLP-Based Algorithm for Biomedical Literature Searching, APAMI & CJKMI-KOSMI Conference 2003.)

而如果需要尋找相關的文章，也可以到下列的網址來尋找：

1. MeSH term : <ftp://203.71.92.119:877/MeSH/mtrees2002.bin>

2. Biomedical literature database

- New 70 萬篇含 MeSH term: All “Neoplasms” MsSH term articles →
<ftp://203.71.92.119:877/Biomedical%20Database/Neoplasms%20articles/>
- Asthma : <ftp://203.71.92.119:877/Biomedical%20Database/asthma/>
- Breast cancer : <ftp://203.71.92.119:877/Biomedical%20Database/breast/>

3. Controlled vocabulary

- (Without alias)Pure HUGO:
ftp://203.71.92.119:877/controlled_vocabulary/Gene%20name/
- Original HUGO
ftp://203.71.92.119:877/controlled_vocabulary/HUGO_nomenclature.txt
- Gene ontology: ftp://203.71.92.119:877/controlled_vocabulary/Gene_ontology/

4. Reference articles

<ftp://203.71.92.119:877/Reference%20PDF/>

How to capture Gene symbol in Breast cancer

ftp://203.71.92.119:877/Reference%20PDF/gene_name/Find_gene_Breast.pdf

以下是一個在該運算環境中所做的實例。先在上述的網址中先抓取相關的基因文章，再將其與基因資料庫作比較，最後把文章中需要的資訊標誌出來，以下便是實際的做法及流程。

如何在文章中標誌基因，我們應該考慮：

- 一、基因本身是否有別名及基因名稱是否重複？
- 二、基因在文章中的位置為何？
- 三、基因在文章中是否為獨立元件？
- 四、基因名稱是否內含其他基因名稱？
- 五、當基因名稱如果有空白符號時，如何避免與在文章中的空白符號混淆？
- 六、文章中所標誌的基因是否有其他層面的涵義？

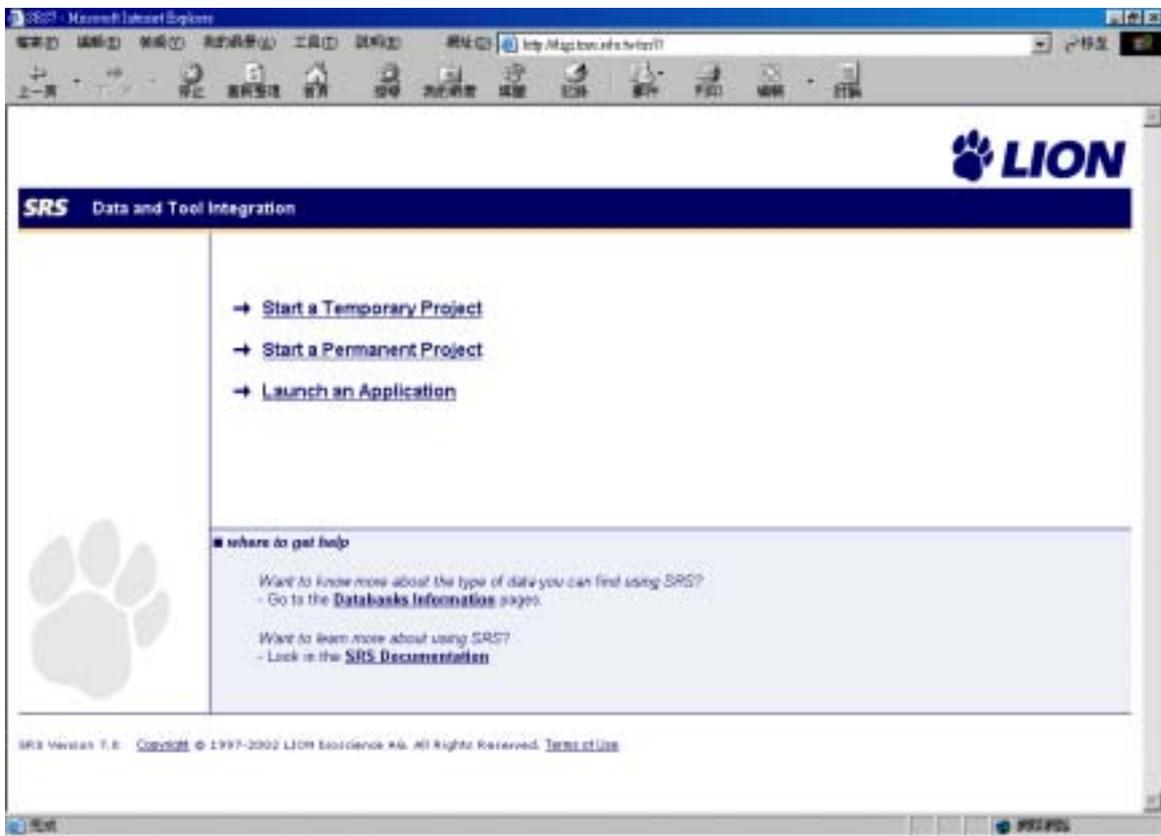
在上述五大項的流程完成後，已經算是把文章中所有以基因名稱相同的字都做上標誌，然而在標誌的字詞裡，並無法確定是否該標記名稱是否一定是基因。舉個例來說，“IV”如果是一個基因名稱，它同時有像“靜脈注射(IV)”的涵義，或是數字中的“四(IV)”，所以必須要建立一套 RULE 或詞性字典，慢慢的去將可能標錯的非基因名稱去除，因為目前 RULE 尚在建構中，所以並無法呈現完整的程序。

我們應用資訊檢索(Information Retrieval)和自然語言處理(Natural Language Processing)的方法，發展一套改進 Entrez 系統搜尋結果的演算法。我們已經建立了一個簡易的查詢系統，Biomedical Literature Searching System (BLSS)，其所搜尋出來的文獻和搜尋條件的相關性，比 Entrez 系統更高。未來，我們會以此系統開發應用程式，幫助 MEDLINE 的使用者更有效率地搜尋出其所感興趣的文獻。(Chien-Yeh Hsu and Yung-Tai Yen, Developing an IR and NLP-Based Algorithm for Biomedical Literature Searching, APAMI & CJKMI-KOSMI Conference 2003.)

12. SRS web sit

建置於 <http://digi.tmu.edu.tw/srs7> 網頁如下圖所示：

此網頁與資料庫主要在整合大量不同性質的生物資訊資料庫，使能透過同一查詢介面獲取資料。目前由歐洲生物資訊研究院(European Bioinformatics Institute, EBI)所建置與維護，使用界面則為 Lion Bioscience 公司所開發的使用介面。目前支援的資料庫有一百多種，包含蛋白質序列、核酸序列，代謝路徑、3D 結構、基因組功能及疾病與性狀等資訊。此資料庫介面之最大優點為可使用邏輯搜尋於不同資料庫；缺點是要對資料庫內有一定之認識，才能獲得較好的結果。



13. 生物醫學訊號資料庫網站

本計劃也初步建立了一個以網際網路危機處理之生物醫學訊號處理網站 (黃芷瑩, 徐建業, Building a Web-Based Environment for Biomedical Signal Processing, International Medical Informatics Symposium and Show in Taiwan 2003, Oct. 2003.)。其網址為:

<http://203.71.92.116/BMS/index.php>, 目前已收集到的資料:

- (1) 心音: 31 位有射出雜音的臨床病患, 以及 18 位無心雜音的正常人的心音訊號
 - (2) 腦波: 18 位進行音樂治療實驗志願者的腦波訊號
 - (3) 心電圖: 分別進行兩種音樂治療實驗之志願者的心電圖訊號, 總共 86 位。
 - (4) SPECT 影像
 - (5) 研究生之研究資料和參考文獻: 由 4 位醫學資訊研究所二年級學生提供
- 生醫訊號資料庫網站之初步網站架構如下圖所示:



a. 本網站提供之資料和功能包括：

- (1) 網站簡介
- (2) 聯絡方式
- (3) 相關連結收集
- (4) 簡易搜尋系統
- (5) 成員登入系統
- (6) 成員管理系統
- (7) 生醫訊號資料上傳 / 下載
- (8) 生醫訊號資料編輯 / 刪除
- (9) 研究資料和參考文獻上傳 / 下載
- (10) 研究資料和參考文獻編輯 / 刪除
- (11) 生醫訊號處理系統

b. 系統平台：

- (1) 作業系統： Microsoft Windows 2000 Server
- (2) WWW 伺服器： Apache
- (3) 動態網頁語言： PHP
- (4) 資料庫伺服器： MySQL
- (5) 訊號處理程式語言： MATLAB 和 MATLAB Web Server

c. 使用者設定：

- (1) 訪客：可瀏覽資訊
- (2) 實驗室成員：可下載資料檔，並可使用訊號處理功能
- (3) 資料所有人：可編輯資料
- (4) 管理者：成員管理、資料管理

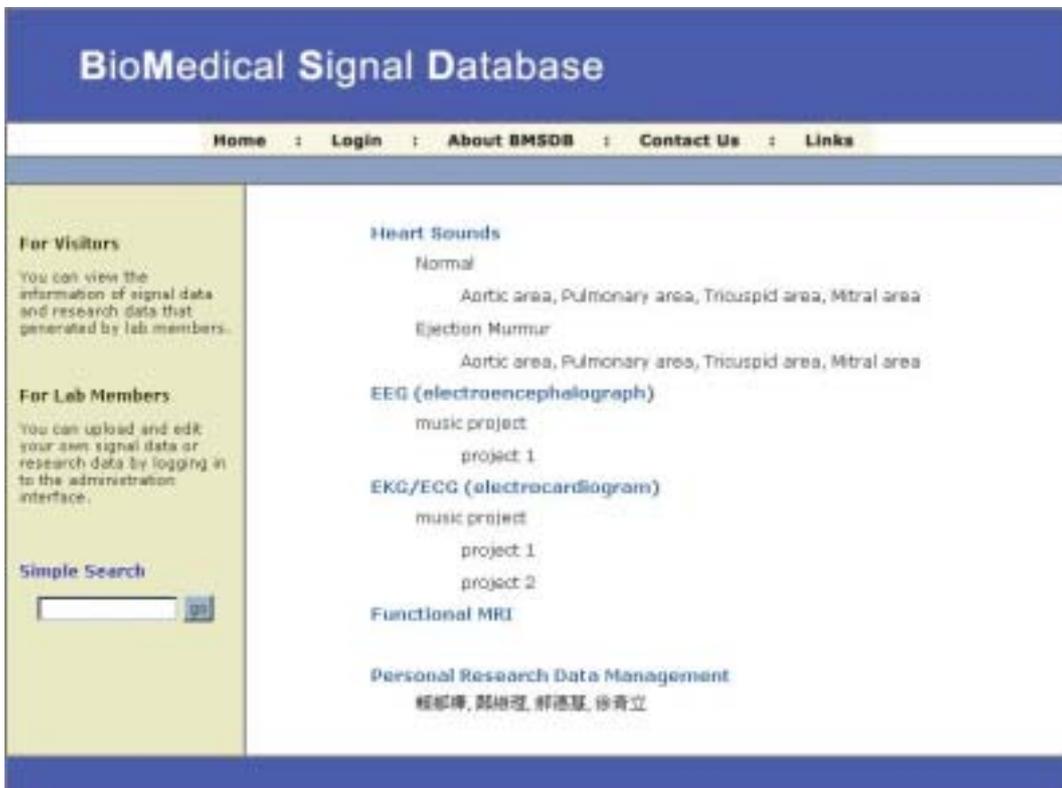
d. 生醫訊號處理系統：

利用 MATLAB 和 MATLAB Web Server 兩項技術，撰寫 MATLAB 程式來對生醫訊號檔案進行計算，並將計算結果透過 MATLAB Web Server 在網頁介面上呈現。已經完成的功能包括：

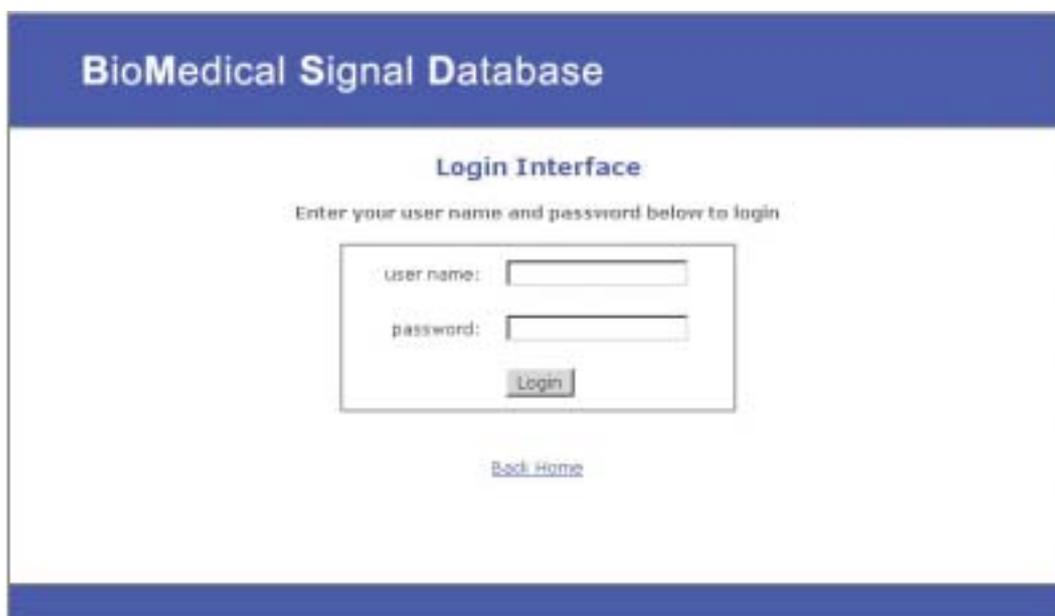
- (1) Time domain plot
- (2) Time domain interested area plot
- (3) Frequency domain plot
- (4) Frequency domain interested area plot
- (5) Lowpass filter analysis
- (6) Highpass filter analysis
- (7) Bandpass filter analysis
- (8) Bandstop filter analysis

e. 網站頁面範例：

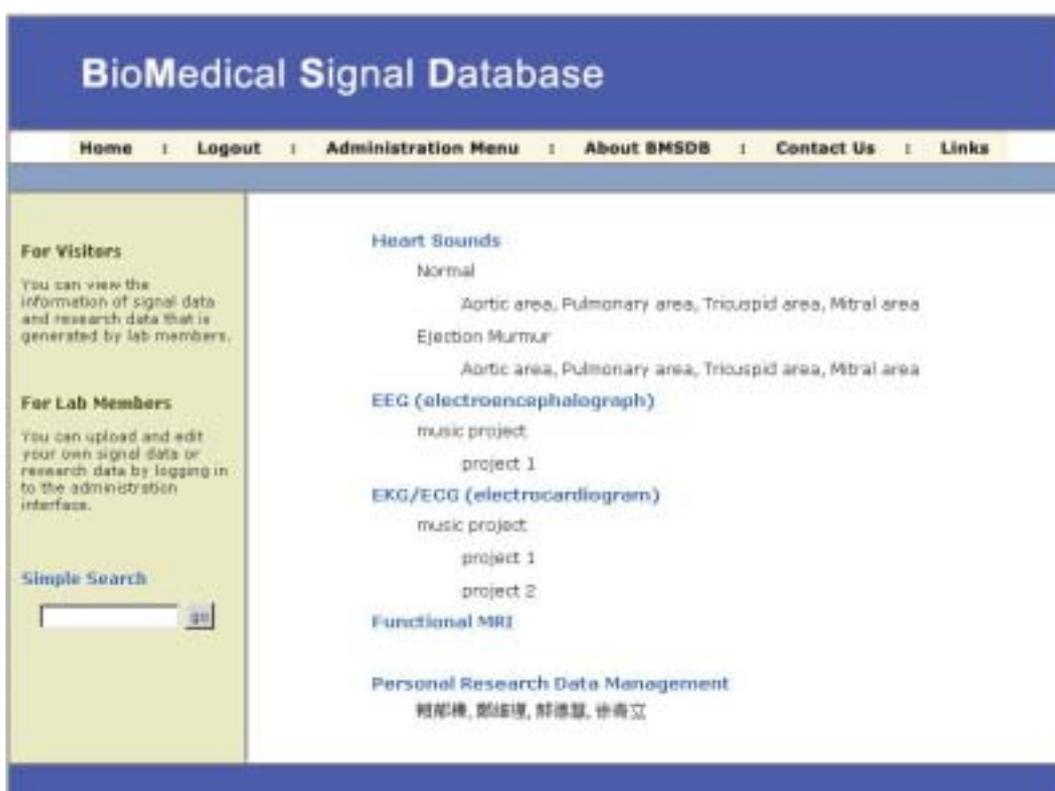
(i)、 首頁：



(ii)、 登入頁面：



(iii)、 登入成功之首頁：



(iv)、 心音資料列表：

BioMedical Signal Database

Heart Sounds - Ejection Murmur

Aortic area Pulmonary area Tricuspid area Mitral area
Data Owner: 楊坤瑾

#	Data ID	Sample ID	File Format	Interpretation	Sampling Date	Download
1	1	1	bit	original	2001-12-21	
2	2	1	bit	processed by 40Hz~100Hz bandpass filter		
3	9	2	bit	original	2001-12-21	
4	10	2	bit	processed by 40Hz~100Hz bandpass filter		
5	17	3	bit	original	2001-12-21	
6	18	3	bit	processed by 40Hz~100Hz bandpass filter		
7	25	4	bit	original	2001-12-21	
8	26	4	bit	processed by 40Hz~100Hz bandpass filter		
9	33	5	bit	original	2002-02-26	
10	34	5	bit	processed by 40Hz~100Hz bandpass filter		
11	41	6	bit	original	2002-02-26	
12	42	6	bit	processed by 40Hz~100Hz bandpass filter		

[Next](#) [Last](#)
[Back to Data Set / Back Home](#)

(v)、 心音訊號檔案資訊：

BioMedical Signal Database

Heart Sounds

Data ID:

Data ID: 1

Sample ID: 1

File Format: bit

Sampling Rate: 5000

Sample Length: 120

Sampling Date: 2001-12-21

Interpretation: original

[Back to Data List / Back Home](#)

(vi)、 心音訊號處理頁面：

BioMedical Signal Database

Signal Processing Applications

Data ID:

Data ID:	1	Sample ID:	1
File Format:	bit	Sampling Date:	2001-12-21
Sampling Rate:	5000 Hz	Sample Length:	120 sec
Interpretation:	original		

Time Domain Plot:

Time Domain ROI Plot:
 from: second to: second

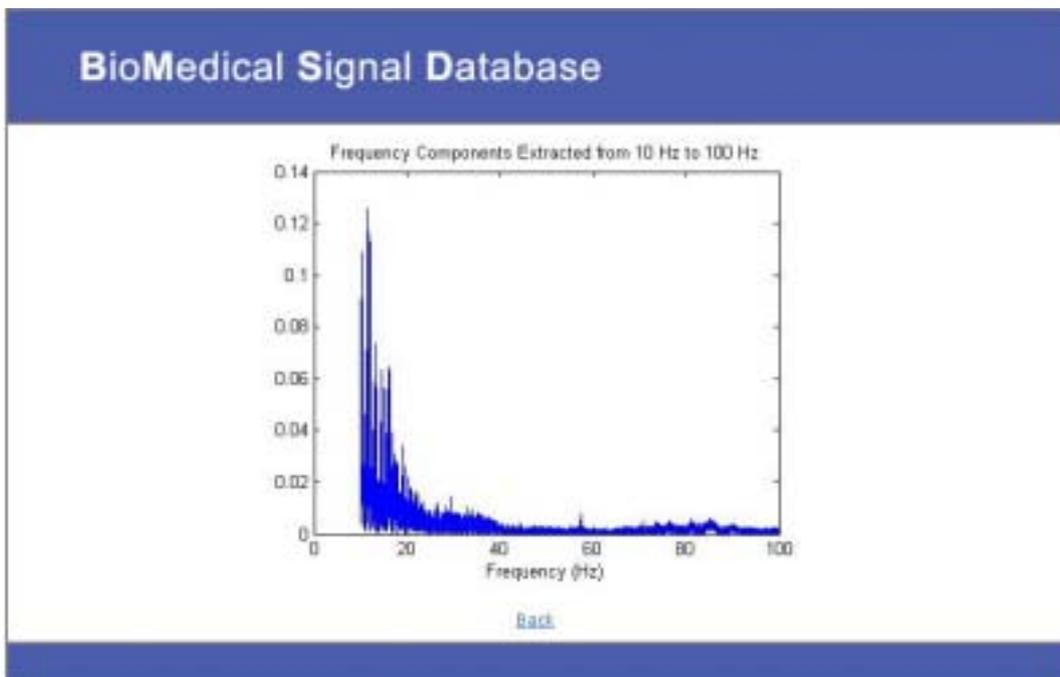
Frequency Domain Plot:

Frequency Domain ROI Plot:
 from: Hertz to: Hertz

Filter Analysis:

[Back to Data Information / Back Home](#)

(vii)、訊號處理結果呈現：(Frequency Domain ROI Plot)



(四) 病歷電子化資料收集

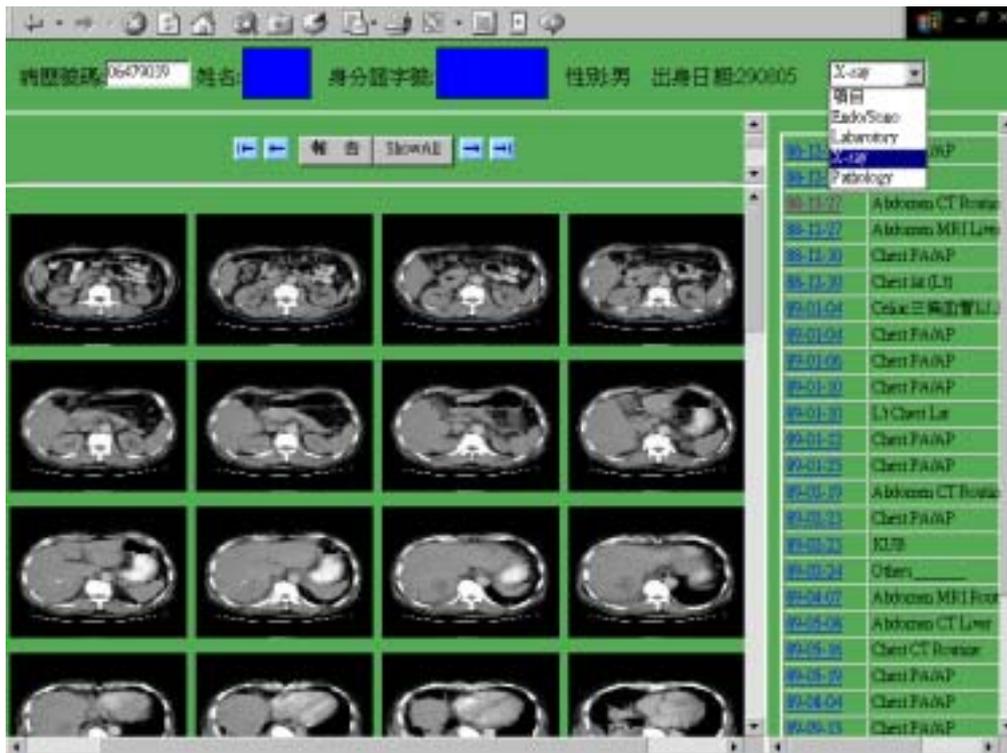
本研究之成果包括醫學超媒體之蒐集與衍生之研究，資料收集如下表所示：

	SOAP	內視鏡	放射	檢驗	病理
報告	2,228,972	201448	546,076	18,993,380	420,924

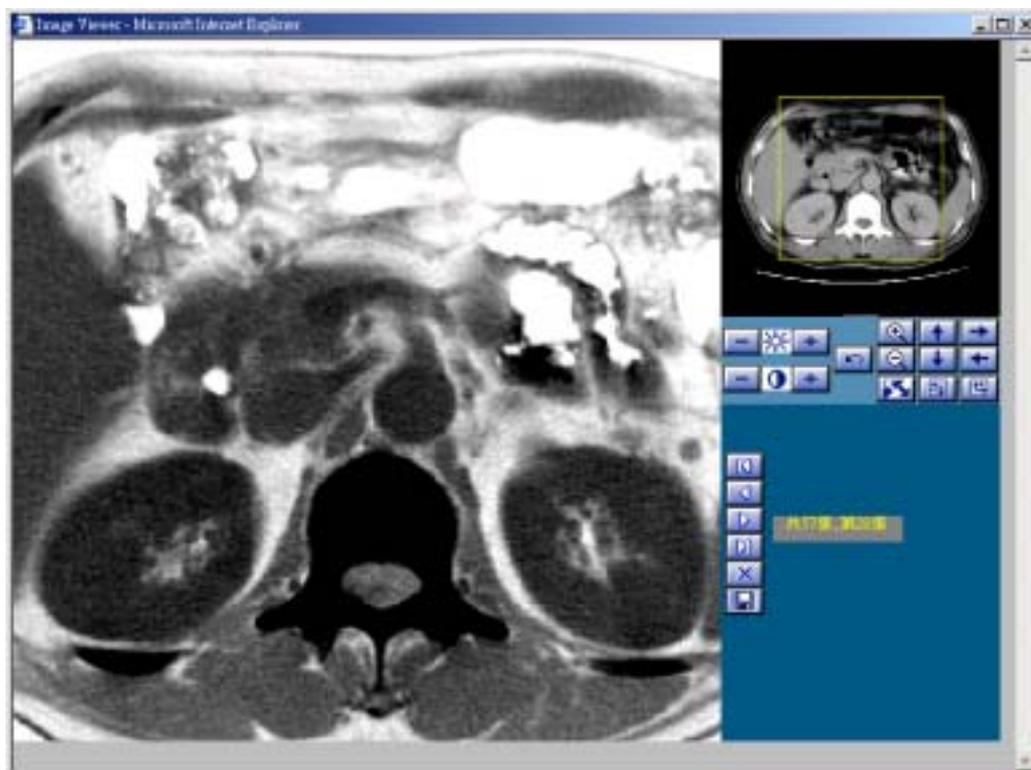
影像		2,469,904	2,354,352		
----	--	-----------	-----------	--	--

衍生之研究與系統如下：

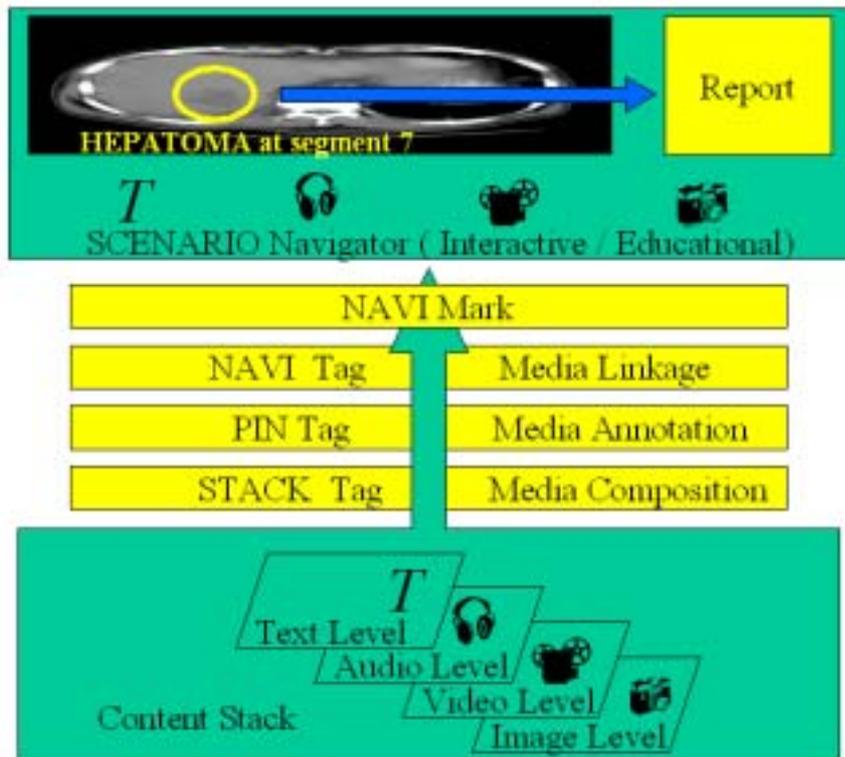
1. 臨床資訊站



2. DICOM Viewer



3. E-Learning NAVI: Navigation With Medical Text, Audio, Video and Image



4. 利用線上分析處理提供醫學資訊服務

甲乙兩院門診科別人次

科別	甲醫院人次	乙醫院人次								
28 神經內科	343	204	61	139	25	53	5	98	3	
29 乙醫院人次	513	416	325	96	247	36	102	3	5	
30 放射內科	359	147	85	31	35	40	8	32	2	
31 乙醫院人次	516	207	75	70	44	50	60	9	3	
32 骨科	94	58	47			57			1	
33 乙醫院人次	141	86		69			85		1	
34 婦產科	135	99	59	87	42	51	64		3	
35 乙醫院人次	209	185		116	170	84	100	125	5	
36 眼科	205	191	49	30	49	62	43	55	2	
37 乙醫院人次	320	286	61	76	48	75	119	61	4	
38 復健醫學科	373	271	300	23	175	301	40		4	
39 乙醫院人次	927	412		263	31	274	128	62	7	
40 腎臟科(專科醫學)	90	14	27							
41 乙醫院人次	134	22		42						
42 放射代造科	62	46					92			
43 乙醫院人次	101	76						149	1	
44 泌尿科(泌尿醫學)	77	89	31					49		
45 乙醫院人次	134	143	81	52					1	
46 甲醫院人次 的加總*	5544	3668	1294	767	1056	987	853	878	58	
47 乙醫院人次 的加總*	8330	5907	1064	1864	1214	1600	1756	1248	97	

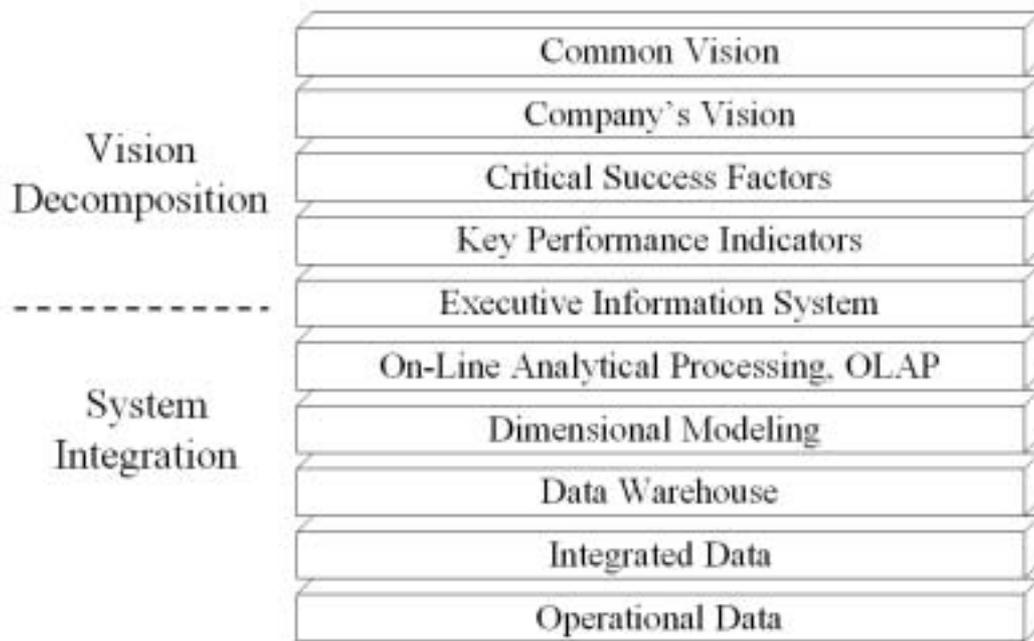
5. 動式臨床影像實驗室



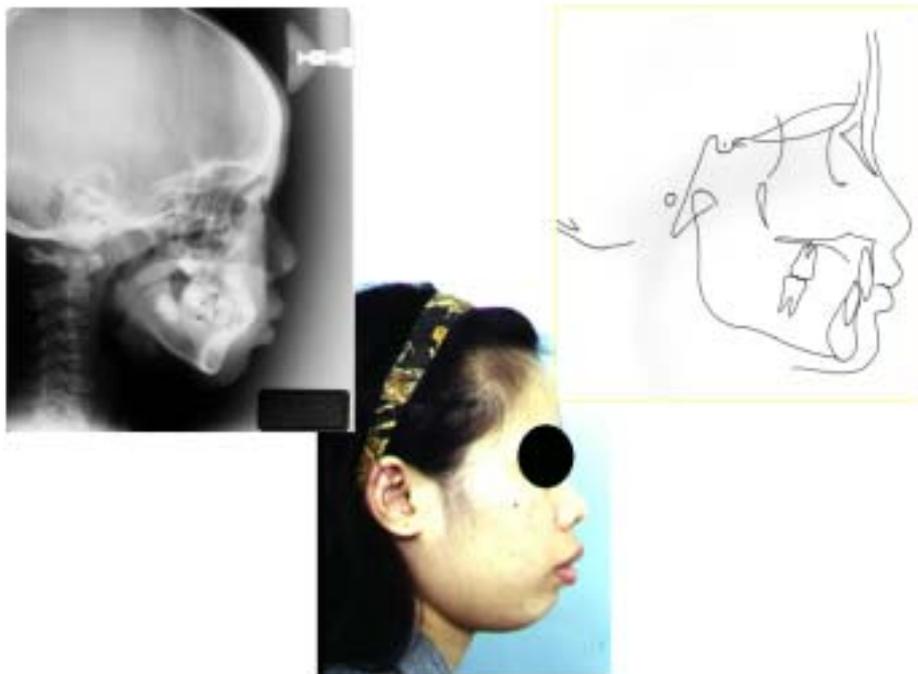
6. The Mobility and Consistency of PDA in Mobile Based e-Hospital



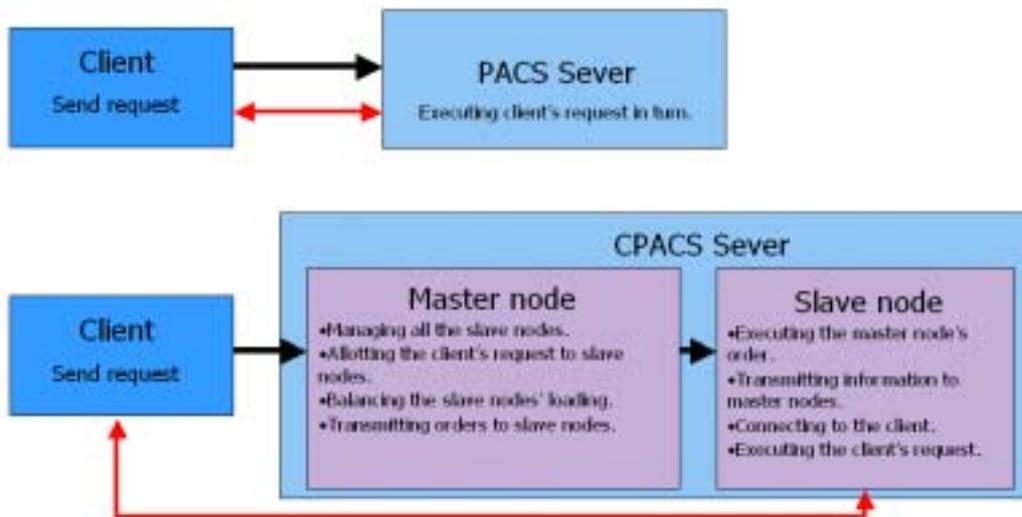
7. Integrating hospital strategy alliance executive information system using a Vision Decomposition method



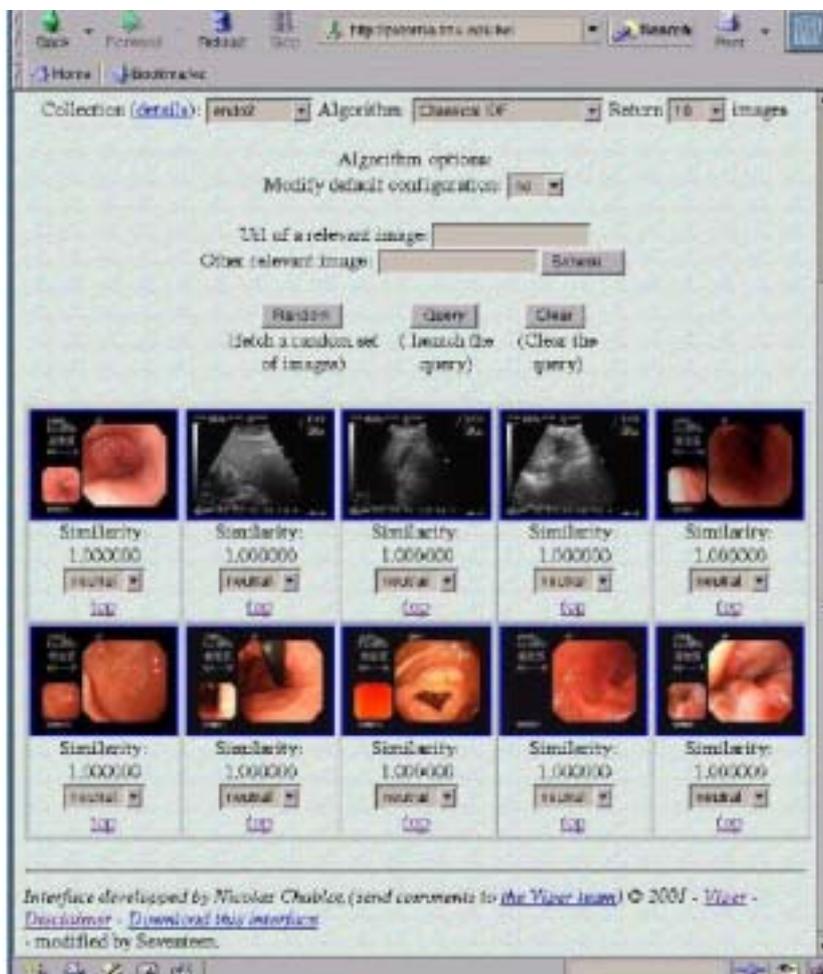
8. Evaluation and improvement of computer simulation of the results in plastic surgery



9. Cluster PACS



10. GIFT open source study



透過本研究改變傳統以物件(Object)為基礎之存取提昇為以醫療記錄物件集(Set of Object)為基礎之存取方式，並重新定義資料整合之方法及兩者之間資料存取之方式，並改善客戶端系統與伺服器端系統之程序資分別將其功效簡述如下

資料整合

1. 整合分散在各處不同檔案格式之醫學資料，使提供的醫療記錄物件具有醫學意義。
2. 醫療記錄物件成為一個獨立授權之單位，不在受限於傳統以資料表或檔案為授權單位之困擾。

資料存取

1. UMRL 所提供是將一個邏輯上的醫學資源名稱以名稱予以命名(Naming)與予以歸類，傳統每個 URL 所代表即是一真實醫學資料的位置，就使用者的操作觀點，UMRL 只要選取資料名稱即可，而不用熟記資料來源的真實位置。
2. UMRL 所提供是醫學資料記錄，傳統每個 URL 所代表即是一單一的醫學資料項目。而這集合可包含跨院所的多種資料格式、醫學資料的種類，傳統 URL 所提供僅是單一院所的單一種資料格式及醫學資料的種類。

客戶端

1. 以 Java Platform 實作，達成跨平臺之特性。
2. 滿足醫師要求且便利操作之使用者介面與影像處理之功能之功能。
3. 重新修改製程，能擺脫重新存取網路之現象，提高客戶端影像處理之效能、網路頻寬之需求與伺服器服務之負載。

伺服器

1. 資料移動性(Mobility)：醫學資料均以邏輯的名稱定義相對應，資料可被移動及更改儲存實體位置。
2. 資料延展性(Scalability):依後端伺服器的狀況，將醫學資料儲存於不同的後端伺服器上，以提高整個系統的效能。
3. 可信賴度(Reliability)：資料都是以邏輯的命名方式，不因伺服器更名或者移動，而造成問題。
4. 存取彈性化(Flexibility)：因應各種醫學資料之需求條件，動態產生相對應的醫學資料集合。
5. 任意存取(Accessibility)：以透過類別名稱方式來描述對應實際位置，改善傳統利用檔案架構的存取資料的方式。使用者不用了解其所儲存的資料的架構，僅需對所要的資料提供名稱即可取得所需的資料。
6. 容錯性(Fault-tolerance)：系統的每一筆資料複製到不同伺服器上。在下層的伺服器故障時，不會影響資料取得。

醫學研究超媒體資源儲存中心完成後，每個研究者或研究計畫暨是醫學超媒體資源的使用者又是資源的提供者，並且宛如擁有一群專業的資訊人員為其研究資訊的儲存、備份、安全管理與使用效能並控制媒體格式之複雜度，安全舒適地專心於研究的方向，運用資訊而非受制於資訊，提昇研究能量。

(五) 成立統計諮詢中心並完成下列各案：

- a. 完成本校「生物統計資源現況及需求問卷」(附件一)，統計需求調查結果如附件

- 二。
- b. 完成「統計諮詢線上預約系統」上線，提供統計諮詢預約服務。網頁範例如附件三。已提供 85 人次統計諮詢服務。
 - c. 完成「資料分析建議系統」上線，提供資料分析建議服務。網頁範例如附件四。已提供 203 人次資料分析建議服務。
 - d. 完成收集主觀性品質監測與評價指標。統計諮詢線上預約系統使用滿意度問卷如附件五。調查結果如附件六。資料分析建議系統使用滿意度問卷如附件七。調查結果如附件八。
 - e. 本校專任教師近三年獲得校內外研究計畫明顯提昇：由八十九學年度之 234 件增加至九十一學年度之 307 件，增加幅度達 31.2%。台北醫學大學專任教師近三年獲得校內外研究計畫之統計表詳細資料如下表。

學年度 機構	八十九 (89.08.01-90.07.31)		九十 (90.08.01-91.07.31)		九十一 (91.08.01-92.07.31)	
	件 數	金 額	件 數	金 額	件 數	金 額
校外						
國 科 會	130	89,085,980	118	99,298,900	142	129,007,103
衛 生 署	20	19,924,086	25	20,376,780	36	52,110,680
國家衛生研究 院	--	--	1	2,294,000	1	2,044,000
其他公私立機 構	31	20,678,383	47	35,617,340	58	69,138,961
小 計	181	129,688,449	191	157,587,020	237	252,300,744
校內						
本校種子基金 計 畫	53	7,974,000	67	9,734,600	70	12,616,693
合 計	234	137,662,449	258	167,321,620	307	264,917,437

- f. 本校專任教師近三年發表之原著性論文明顯提昇：由 2000 之 309 件增加至 2002 之 477 件，增加幅度達 54.4%。臺北醫學大學專任教師近三年發表之原著性論文統計表詳細資料如下表。

出版年份 論文種類	2000	2001	2002
SCI	159	191	230
SSCI	4	2	11
其他論文	146	171	236
總計	309	364	477

附件一：生物統計資源現況及需求問卷

本校獲國科會補助提昇私大研發能量專案-建立先進之醫學研究資訊基礎架構之子計劃：建立統計計算中心，擬收集學校及兩院同仁使用統計現況及需求，便於統計計算中心規劃適當服務，敬請撥冗填寫下列問題。感謝您的支持與合作。

1. 服務單位

校本部 附醫 萬芳醫院

2. 職稱

教授 副教授 助理教授 講師

醫師 藥師 醫檢師 護理師 營養師

其他專業人員

3. 您從事教學、服務或研究時，需使用統計方法嗎？

無 偶爾 常常

4. 您使用統計時，曾面臨困難嗎？

是 否

5. 您需要校方提供統計諮詢服務嗎？

非常需要 需要 不需要 非常不需要

6. 您期望統計計算中心提供哪些諮詢服務？（可複選）

假設的擬定

研究設計

統計方法的選擇

統計軟體的使用

統計結果的判讀

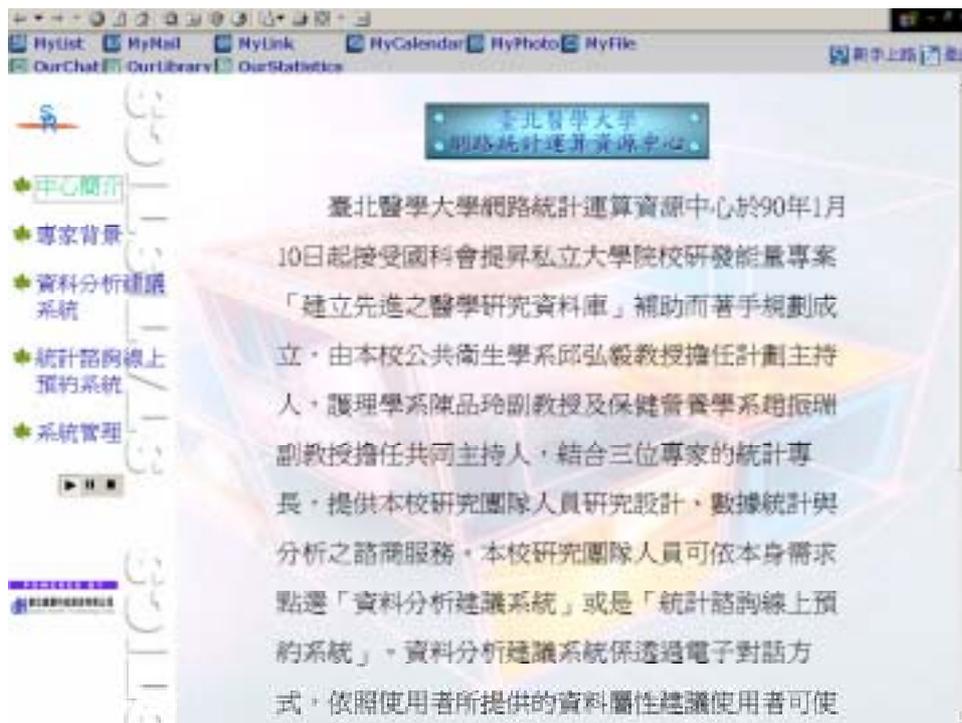
撰寫論文結果

統計需求調查結果

項目	人數	百分比
服務單位		
校本部	28	84.8
附醫	3	9.1
萬芳醫院	2	6.1
職稱		
教授	2	6.1
副教授	7	21.2
助理教授	2	6.1
講師	19	57.6
藥師	2	6.1
其他專業人員	1	1.5
統計使用需求		
無	0	0.0
偶爾	8	24.2

常常	25	75.8
統計使用困難		
是	32	97.0
否	1	3.0
統計諮詢服務需求		
非常需要	21	63.6
需要	12	36.4
不需要	0	0.0
非常不需要	0	0.0
諮詢服務種類		
假設的擬定	12	36.4
研究設計	19	57.6
統計方法的選擇		
統計軟體的使用	28	84.8
統計結果的判讀	27	81.8
撰寫論文結果	12	36.4

附件三：統計諮詢線上預約系統網頁範例



預約服務

歡迎使用臺北醫學大學網路統計運算資源中心統計諮詢線上預約系統，您可以利用本系統來預約本中心專家群所提供的諮詢服務。

在開始預約專家諮詢時間之前，請您先至**專家背景**了解本中心各位專家之研究領域，以選擇適當之專家為您服務；本系統提供二個月以內之預約服務，並請您先準備好研究計劃之相關背景資料，包括研究名稱、研究性質(臨床研究或是基礎研究)、研究資料來源/人體資料、動物實驗動

臺北醫學大學「網路統計運算資源中心」申請同意書

歡迎進入「網路統計運算資源中心」(以下簡稱本中心)，為保障使用權益，使用者請於使用本中心各項系統前，詳細閱讀本同意書所有內容。當您使用本中心各項系統時，表示您已閱讀、了解並同意接受本申請書所有注意事項及規範。本中心有權於任何時間修改或變更本同意書之內容，建議您隨時注意該書之修改或變更。

- 為能有效提昇本中心各項系統之便捷性，會員於登錄時應提供完整且符合真實之個人資料及研究計劃背景。您所登錄之資料事後有變更時，應隨時於線上更新之。如果您提供之個人資料違反或破壞服務宗旨，本中心保留隨時終止您所使用之各項服務之權利。您所提供之個人資料本中心均視為機密，決不將個人資料洩漏給本中心外的第三者。
- 本中心僅開放給臺北醫學大學教員及附屬，萬芳醫院醫療照護專業人員，由於資源有限，未免濫用，競爭損害您的權益，請勿將私人帳號借於他人，以免自身權益受損。
- 會員應遵守線上預約諮詢系統之流程，重視預約諮詢制度之精神，

請先填寫相關資料,有*為必填欄位

*姓名: Ping-Ling Chen

身份:

*研究名稱:

*聯絡電話:

手機電話:

*電子郵件信箱: pichen@tmu.edu.tw

補助機構:

研究性質調查: 臨床研究 基礎研究

研究資料來源: 人體資料(如疾病資料、環境暴露等) 動物實驗數據 體外實驗數據

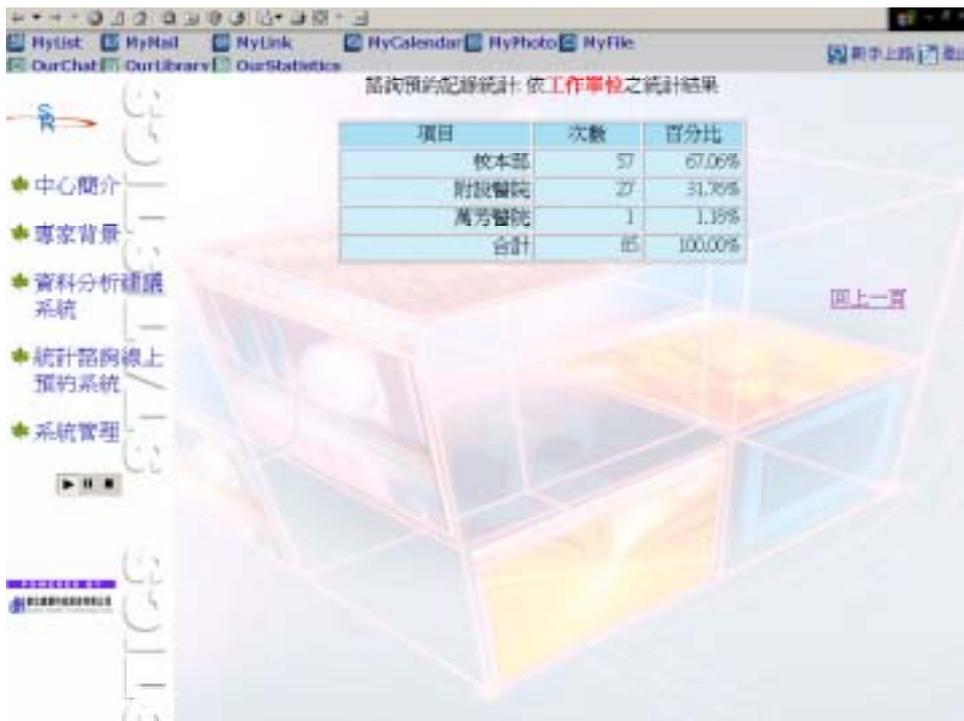
研究計劃進行析: 草擬研究計劃 收集資料 數據分析

狀態: 撰寫報告 回覆檢核審查意見

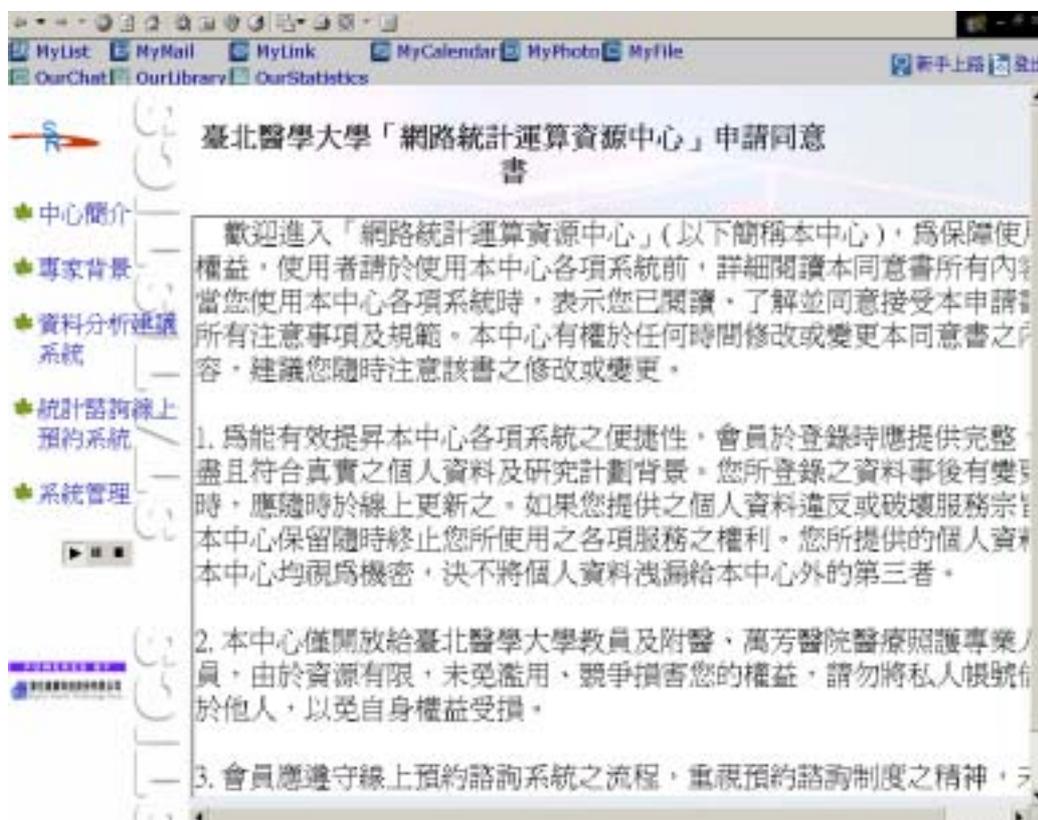
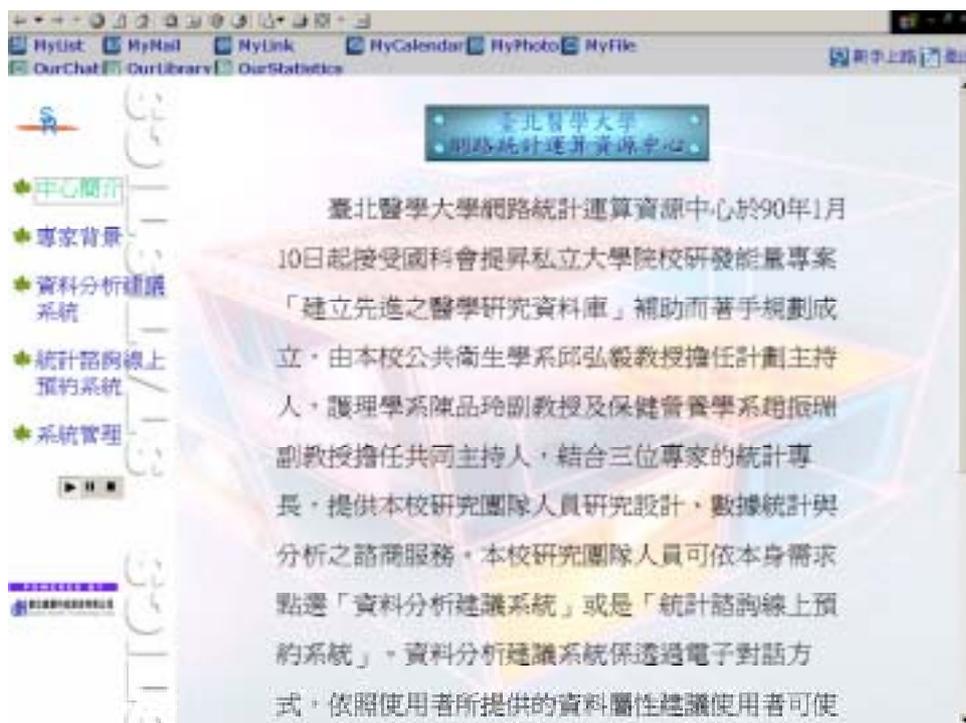
Pin-Ling Chen 您已預約成功

本系統將會寄送一封電子郵件到您的預約電子郵件信箱(pichen@tmu.edu.tw),並再次提醒您此一查詢的相關事宜。

預約專家:	陳品玲
預約時間:	2004-03-18 16:00
取消預約:	如果您取消此次的預約紀錄,請於預約時間前三天取消,並請輸入取消密碼。 取消密碼: 12345
注意事項:	請隨您攜帶相關文件資料,如:實驗研究設計、問卷、資料等等。以上請儘量帶齊資料,以便我們協助您。 請您於預約時段至"護理系"與"陳品玲"會面,若臨時變更,請電02771379與"陳品玲"聯繫



附件四：資料分析建議系統網頁範例



請先填寫相關資料,有*為必填欄位

*姓名: Ping-Ling Chen

身份:

*研究名稱:

*聯絡電話:

手機電話:

*電子郵件信箱: pichen@tmu.edu.tw

補助機構:

研究性質調查: 臨床研究 基礎研究

研究資料來源: 人體資料(如疾病資料、環境暴露等) 動物實驗數據 體外實驗數據

研究計劃進行析: 草擬研究計劃 收集資料 數據分析

狀態: 撰寫報告 回覆檢核審查意見

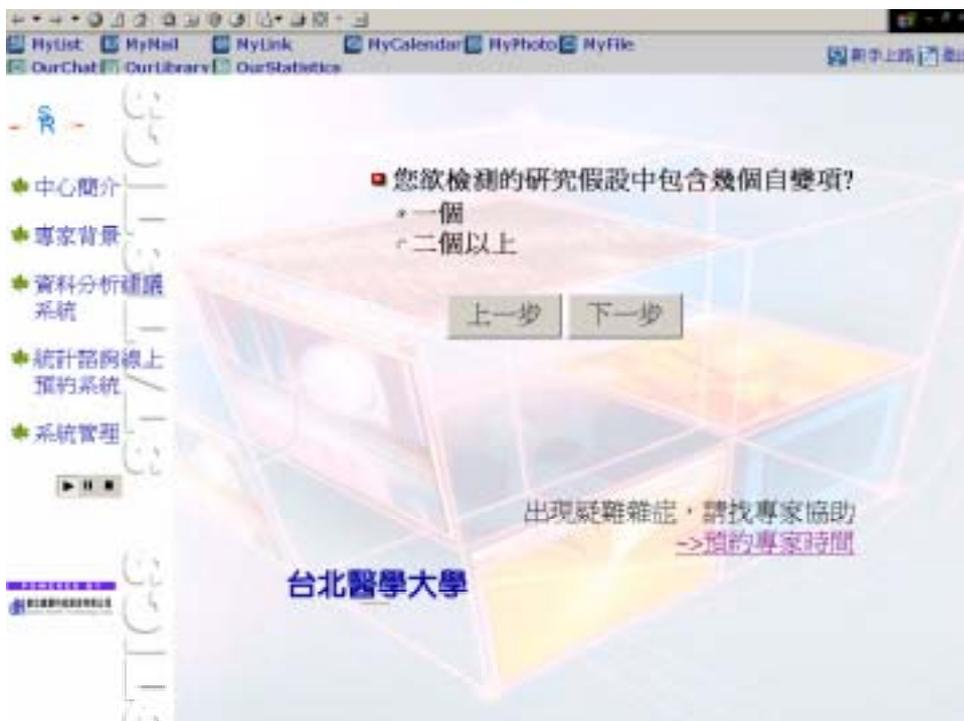
族群參數(population parameter)分佈不是常態分佈(normal distribution)或樣本數過少(如 $n < 30$)，而不呈現常態分佈

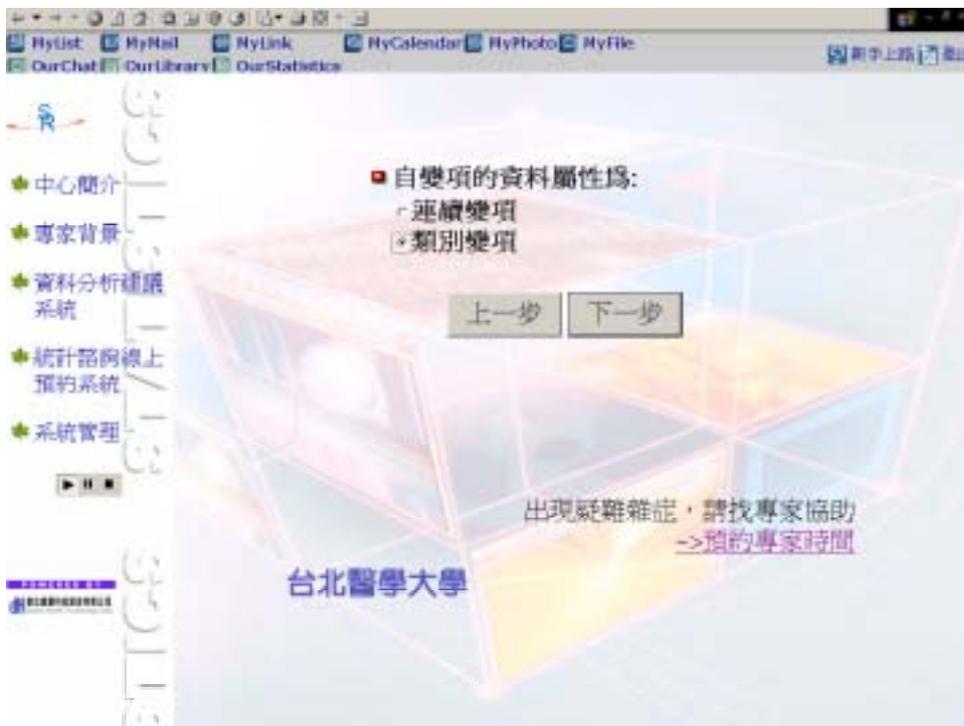
是,請進行無母數分析(nonparametric analysis)

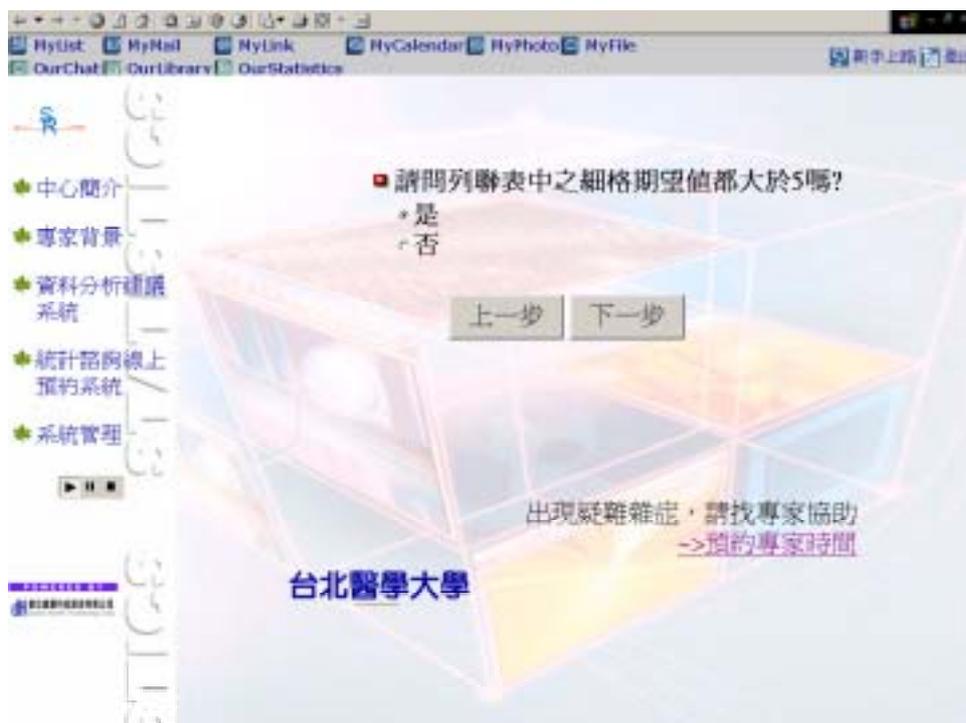
否,請進行母數分析(parametric analysis)

出現疑難雜症，請找專家協助
->預約專家時間

台北醫學大學









統計時間: 91年2月 - 92年2月

結果代碼之數值	檢本數	對照數	屬有數	統計次數	結果內容
1-4	27	19	2	48	請進行 spearman's rho 或 Spearman rank correlation coefficient
1-5	8	8	4	18	請進行 sgs test 或 Spearman rank correlation coefficient
1-8	3	0	8	3	請進行 Mann-Whitney test 或 wilcoxon test
1-9	8	0	8	0	請進行 Kolmogorov-Smirnov test
1-12	2	0	8	2	請進行 Kruskal-Wallis one-way ANOVA (一種自變數) 或 Friedman two-way ANOVA (多種自變數)
1-13	8	0	8	0	請進行 Kolmogorov-Smirnov test
3-3	8	9	3	28	請以第一檢本數定比與檢本率
3-7	7	0	3	18	請先定若干個因子存在上下兩種本檢定比較發生率,或若干個因子存在9,請將病人年齡資料分類并依進行比較
3-8	4	1	8	5	改變檢本數時請進一步檢定多種變數結果的趨勢,請以發生率趨勢分析進行檢定
3-9	8	2	2	12	請使用 log-rank 檢定
3-11	4	2	8	6	請用 Cox 存活分析模式進行檢定
3-3	9	9	3	28	請請您與統計專家商量其他檢定和列進行資料分析,歡迎使用統計師諮詢系統
3-7	7	1	8	8	您可以選擇使用 Logistic Regression
3-18	8	1	1	18	您可以選擇使用 Single Cases Regression
3-11	4	3	1	8	您可以選擇使用 Correlation
3-13	7	0	8	7	您可以選擇使用 ANOVA, ANCOVA檢定是否有顯著差異,請繼續使用 Multiple comparisons
3-15	5	1	8	6	您可以選擇使用 McNemar test
3-16	3	2	1	6	您可以選擇使用 Independent t-test
3-17	3	1	8	4	您可以選擇使用 paired t-test
3-18	5	2	8	7	您可以選擇使用 Chi-square
3-19	2	0	1	3	您可以選擇使用 Fisher's Exact test
合計	122	81	20	203	

附件五：統計諮詢線上預約系統使用滿意度問卷

1. 您覺得本網頁入口的連結圖示顯而易見？

非常同意 同意 不同意 非常不同意

2. 您是否滿意統計諮詢線上預約的方式？

非常滿意 滿意 沒意見 不滿意 非常不滿意

3. 您是否滿意線上預約申請所需的時間？

非常滿意 滿意 沒意見 不滿意 非常不滿意

4. 您是否滿意統計諮詢專家的服務態度？

非常滿意 滿意 沒意見 不滿意 非常不滿意

5. 在您使用過統計諮詢線上預約系統後，請您對本系統整體感覺給個分數

優異，85 分以上 良好，75 分~85 分 勉強，60 分 待改進，60
分以下 沒有意見

6. 若日後仍有提供統計諮詢服務，您是否會繼續使用？

會 不會 請簡述原因_____

附件六：統計諮詢線上預約系統使用滿意度調查結果

(N=60)

項 目	人數	百分比
您覺得本網頁入口的連結圖示顯而易見？		
非常同意	2	3.3
同意	52	86.7
不同意	6	10.0
非常不同意	0	0.0
您是否滿意統計諮詢線上預約的方式？		
非常滿意	2	3.3
滿意	58	96.7
沒意見	0	0.0
不滿意	0	0.0
非常不滿意	0	0.0
您是否滿意線上預約申請所需的時間？		
非常滿意	25	41.7
滿意	35	58.3
沒意見	0	0.0
不滿意	0	0.0
非常不滿意	0	0.0
您是否滿意統計諮詢專家的服務態度？		
非常滿意	57	95.0
滿意	3	5.0
沒意見	0	0.0
不滿意	0	0.0
非常不滿意	0	0.0
在您使用過統計諮詢線上預約系統後，請您對本系統整體感覺給個分數		
優異，85 分以上	47	78.3
良好，75 分~85 分	13	21.7
勉強，60 分	0	0.0
待改進，60 分以下	0	0.0
沒有意見	0	0.0
若日後仍有提供統計諮詢服務，您是否會繼續使用？		
會	60	100.0
不會	0	0.0

附件七：資料分析建議系統使用滿意度問卷

7. 您在本網頁中是否有獲得您想要的資訊？

有 有，但不完整 沒有

8. 您是否滿意本網頁的版面配置？

非常滿意 滿意 沒意見 不滿意 非常不滿意

9. 您是否滿意「資料分析建議系統」的使用方式？

非常滿意 滿意 沒意見 不滿意 非常不滿意

10. 您是否滿意「資料分析建議系統」的處理速度？

非常滿意 滿意 沒意見 不滿意 非常不滿意

11. 您是否滿意「資料分析建議系統」所提供的資訊？

非常滿意 滿意 沒意見 不滿意 非常不滿意

12. 在您使用過「資料分析建議系統」後，請您對本系統整體感覺給個分數

優異，85 分以上 良好，75 分~85 分 勉強，60 分 待改進，60
分以下 沒有意見

附件八：資料分析建議系統使用滿意度調查結果

(N=74)

項 目	人 數	百分比
您在本網頁中是否有獲得您想要的資訊？		
有	40	54.1
有，但不完整	34	45.9
沒有	0	0.0
您是否滿意本網頁的版面配置？		
非常滿意	21	28.4
滿意	50	67.6
沒意見	3	4.1
不滿意	0	0.0
非常不滿意	0	0.0
您是否滿意「資料分析建議系統」的使用方式？		
非常滿意	11	14.9
滿意	62	83.8
沒意見	1	1.4
不滿意	0	0.0
非常不滿意	0	0.0
您是否滿意「資料分析建議系統」的處理速度？		
非常滿意	9	12.2
滿意	65	87.8
沒意見	0	0.0
不滿意	0	0.0
非常不滿意	0	0.0
您是否滿意「資料分析建議系統」所提供的資訊？		
非常滿意	7	9.5
滿意	53	71.6
沒意見	13	17.6
不滿意	1	1.4
非常不滿意	0	0.0
在您使用過「資料分析建議系統」後，請您對本系統整體感覺給個分數		
優異，85 分以上	52	70.3
良好，75 分~85 分	19	25.7
勉強，60 分	3	4.1
待改進，60 分以下	0	0.0
沒有意見	0	0.0

附件九、發表論文清單

1. Lin CS, Li YC, Mok MS, Wu CC, Chiu HW, Lin YH. Neural Network Modeling to Predict the Hypnotic Effect of Propofol Bolus Induction. *Journal of the American Medical Informatics Association, Symposium Supplement*, 2002:450-3. (SCI)
2. Chang PL, Li YC, Lee SH. The differences in health outcomes between Web-based and paper-based implementation of a clinical pathway for radical nephrectomy. *BJU Int.* 2002 Oct;90(6):522-8. (SCI)
3. Lee JJ, Liu CH, Wu JH, Lee TK, Li YC, Chiang IJ, Chiu HY, Lu SY. Overview of Extremity Replantation in Taiwan by the National Insurance Database form 1996 to 2000. *Journal of Plastic Surgical Association.*2002(10):217-230.
4. Chen BH, Hsu CY, Lin MC, Li YC. Exchangeable Multimedia Framework for Digital Library in Medicine. In *Proceedings of the AMIA 2002 Annual Symposium*, 2002:994.
5. Liu CH, Chiang IJ, Li YC. Medical Data Mining - Experience of Knowledge Discovery in Two Clinical Databases. In *Proceedings of the AMIA 2002 Annual Symposium*, 2002:1085.
6. Hsu MH, Li YC, Chiu WT. Outcome Prediction after Moderate and Severe Head Injury Using an Artificial Neural Network. In *Proceedings of the AMIA 2002 Annual Symposium*, 2002:1049.
7. Lin CC, Li YC, Bai YM, Hsiao MC, Tsai SJ, Wu CH, Ou-Yang WC, Liu CY. Test-Retest Reliability of an Internet-based Self-assessment Program for Depression. In *Proceedings of the AMIA 2002 Annual Symposium*, 2002:1083
8. Chiang IJ, Hsu JY, Fuzzy Classification Trees for Data Analysis, *Fuzzy Sets and Systems*, Vol. 103, No. 1, 87-99, 2002.
9. Chiang IJ, Shieh MJ, Hsu JY, Wong JM, Apply Fuzzy Classification to Colon Polyp Screening, *Communications of IICM*, Vol. 5, No. 2, 7-12, 2002.
10. Chiang IJ, Shieh MJ, Hsu JY, Wong JM, A Medical Decision Support System for Polyp Screening by Using Fuzzy Classification Trees, *AMIA*, 2002.
11. Liu CH, Wu SJ, Li YC, Chiang IJ, Liu KS, Lu SY, Chiu HY, Prognostic Prediction of Digital Representation—Data Mining the CGH Database of 800 Cases, *Proceeding of the Inaugural Congress of the World Society for Reconstructive Microsurgery*, Bologna, Italy, 2001, 93-95.
12. Chiang IJ, Lin TY, Index Miner: a data mining system, In *Proceedings of the Annual International Computer Software and Applications Conference*

- (COMPSAC'2001), Chicago, IL, 2001, 613-614.
13. Liu CH, Li YC, Chiang IJ, Medical Data Mining—Experience of Knowledge Discovery in Two Clinical Databases, In Proceedings of the MIST2001, Taipei, Taiwan, 2001, 211-217.
 14. Lin F, Chiang IJ, Longitudinal Study of Atherosclerosis Risk Factors, In Discovery Challenge Workshop of the 6th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'02), Helsinki, Finland, 2002.
 15. Chiu HW, Kao T, A mathematical model for autonomic control of heart rate variation, IEEE Engineering in Medicine and Biology Magazine, 2001, 20: 69-76.
 16. Chiu HW, Wang TH, Huang LC, Tso HW, Kao T, The influence of mean heart rate on measures of heart rate variability as markers of autonomic activities after autonomic blockade and during exercise: a model study, accepted by Medical Engineering & Physics, 2002
 17. Chiu HW, Kao T, Hsiao HC, Kong CW, Cardiac Aliasing in Atrial Rate Variation Response to Fixed-Rate Ventricular Pacing. Computers in Cardiology 2001. 505-509
 18. Tsai AC, Chiu HW. Relationship between heart rate variability and electrolyte concentration in chronic renal failure patients under hemodialysis. International Journal of Bioelectromagnetism 2002, 4, 307-308.
 19. Yang KC, Hsu CY, Yang TF, Using Acoustical Energy Ratio in Multiple Frequency Band for Cardiac Murmur Analysis, in Proceeding, Conference on Biomedical Engineering and Technology, December, 2002.
 20. Chan JC, Chiu HW, Hsu CY, Integrating Information Technology in Life Science Course for Senior High School Students, in Proceeding, Conference on Biomedical Engineering and Technology, December, 2002.
 21. Chan JC, Hsu CY, Chen HS, Developing a Health Education Learning System on the WWW, in Proceeding, Conference on Health and Management, Taiwan, Oct. 2002.
 22. Lin MC, Li YC, Hsu CY, Chiang IR, Discovering the gene and disease relationship from biomedical literature database by applying the information retrieval technology, International Medical Informatics Symposium and Show in Taiwan 2002, Oct. 2002.
 23. Chen BH, Lin MC, Hsu CY, Li YC, Exchangeable Multimedia Framework for Digital Library in Medicine, AMIA Symposium, Nov. 2002.

24. Chan JC, Hsu CY, Chen HS, Integrating the Learning Pathway and Knowledge base Management into a Web-Based Learning System for Health Education, International Medical Informatics Symposium and Show in Taiwan 2002, Oct. 2002.
25. Hsu CY, Wang PY, Li YC, Kao MC, Hsieh D, Jian WS "A Study on the Effect of Physician's Prescriptions with a Drug Interactions Checking Integrated Computerized Order Entry System", The FASEB Journal. 16(4):A324, 2002.
26. Hsu CY, Pan LC, A Portable Device Using Speech Recognition and Synthesis Applied to Rehabilitation of Subjects with Speech Disorder, Proceeding, 2001 Annual Conference on Biomedical Engineering Technology and Education, December, 2001.
27. AC Tsai, HW Chiu. Relationship between heart rate variability and electrolyte concentration in chronic renal failure patients under hemodialysis. International Journal of Bioelectromagnetism 2002, 4, 307-308.
28. Bing-Hao Chen, Ming-Chin Lin, Chien-Yeh Hsu*, Yu-Chuan Li, Exchangeable Multimedia Framework for Digital Library in Medicine, AMIA Symposium, Nov. 2002.
29. C. C. H. Liu, S. J. H. Wu, Y.-C. Li, I-Jen Chiang, K.-S. Liu, S.-Y. Lu, and H.-Y. Chiu, Prognostic Prediction of Digital Representation-Data Mining the CGH Database of 800 Cases, Proceeding of the Inaugural Congress of the World Society for Reconstructive Microsurgery, Bologna, Italy, 2001, 93-95.
30. C.Y. Hsu, P.Y. Wang, Y.C. Li, M.C. Kao, D. Hsieh, W.S. Jian "A Study on the Effect of Physician's Prescriptions with a Drug Interactions Checking Integrated Computerized Order Entry System", The FASEB Journal. 16(4):A324, 2002.
31. Chan JC, Hsu CY, Chen HS. Developing a Health Education Learning System on the WWW, Journal of Health Management, Vol. 1, No. 1, 2003, pp. 19-29.
32. Charles C. H. Liu, Yu-Chuan Li, and I-Jen Chiang, Medical Data Mining-Experience of Knowledge Discovery in Two Clinical Databases, In Proceedings of the MIST2001, Taipei, Taiwan, 2001, 211-217.
33. Chien-Yeh Hsu and Li-Chern Pan "A Portable Device Using Speech Recognition and Synthesis Applied to Rehabilitation of Subjects with Speech Disorder", in Proceeding, 2001 Annual Conference on Biomedical Engineering Technology and Education, December, 2001.
34. Chih-Ying Huang and Chien-Yeh Hsu, Building a Web-Based Environment

- for Biomedical Signal Processing, International Medical Informatics Symposium and Show in Taiwan 2003, Oct. 2003, pp130-133.
35. Chiu HW, Lin LS, Kuo MC, Chiang HS, Hsu CY(2003). Using Heart Rate Variability Analysis to Assess the Effect of Music Therapy on Anxiety Reduction of Patients. *Computers in Cardiology* (30); 469-472 (EI)
 36. Chiu HW, Wang TH, Huang LC, Tso HW and Kao T(2003).The influence of mean heart rate on measures of heart rate variability as markers of autonomic function: a model study. *Medical Engineering & Physics*.(25):475-481. (SCI)
 37. Developing a NLP and IR-Based algorithm for Analyzing Gene-disease Relationship, International Joint Meeting EuroMISE 2004. Symposium on Biomedical Informatics and Biomedical Statistics Education, April, 2004.
 38. Forrest Lin and I-Jen Chiang, Longitudinal Study of Atherosclerosis Risk Factors, In Discovery Challenge Workshop of the 6th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'02), Helsinki, Finland, 2002.
 39. H.W. Chiu, and T., Kao, (2001) "A mathematical model for autonomic control of heart rate variation", *IEEE Engineering in Medicine and Biology Magazine*, 20: 69-76.
 40. Hung-Wen Chiu, Ti-Ho Wang, Lu-Chou Huang, Han-Wen Tso and Tsair Kao (2002) "The influence of mean heart rate on measures of heart rate variability as markers of autonomic activities after autonomic blockade and during exercise: a model study", accepted by *Medical Engineering & Physics*
 41. HW Chiu, T Kao, HC Hsiao, CW Kong Cardiac Aliasing in Atrial Rate Variation Response to Fixed-Rate Ventricular Pacing. *Computers in Cardiology* 2001. 505-509
 42. Jen Chiang and Jane Yung-jen Hsu, Fuzzy Classification Trees for Data Analysis, *Fuzzy Sets and Systems*, Vol. 103, No. 1, 87-99, 2002.
 43. Jen Chiang and Ming-Jium Shieh, Jane Yung-jen Hsu, and Jau-Min Wong, Apply Fuzzy Classification to Colon Polyp Screening, *Communications of IICM*, Vol. 5, No. 2, 7-12, 2002.
 44. Jen Chiang and T. Y. Lin, Index Miner: a data mining system, In *Proceedings of the Annual International Computer Software and Applications Conference (COMPSAC'2001)*, Chicago, IL, 2001, 613-614.
 45. Jen Chiang, Ming-Jium Shieh, Jane Yung-jen Hsu, and Jau-Min Wong, A Medical Decision Support System for Polyp Screening by Using Fuzzy Classification Trees, *AMIA*, 2002.
 46. Ju-ching Chan, Chien-Yeh Hsu, Hon-Sun Chen, Developing a Health Education Learning System on the WWW, in *Proceeding, Conference on Health and Management*, Taiwan, Oct. 2002.

47. Ju-ching Chan, Chien-Yeh Hsu, Hon-Sun Chen, Integrating the Learning Pathway and Knowledge base Management into a Web-Based Learning System for Health Education, International Medical Informatics Symposium and Show in Taiwan 2002, Oct. 2002.
48. Ju-ching Chan, Hung-Wen Chiu, Chien-Yeh Hsu, Integrating Information Technology in Life Science Course for Senior High School Students, in Proceeding, Conference on Biomedical Engineering and Technology, December, 2002.
49. Kun-Chang Yang, Chien-Yeh Hsu, Ten-Fang Yang, Using Acoustical Energy Ratio in Multiple Frequency Band for Cardiac Murmur Analysis, in Proceeding, Conference on Biomedical Engineering and Technology, December, 2002.
50. Lai YH, Hsu CY, Lee KL(2003). Development of an Interactive Multimedia E-learning System for Training Dental Clinic Assistants. Journal of Korean Society of Medical Informatics, Vol. 9, pp. s335-s337.
51. Lin CS, Li YC, Mok MS, Wu CC, Chiu HW, Lin YH.(2002).Neural Network Modeling to Predict the Hypnotic Effect of Propofol Bolus Induction. Journal of the American Medical Informatics Association, Symposium Supplement.:450-453. (SCI)
52. Ming-Chin Lin, Yu-Chuan Li, Chien-Yeh Hsu, I-Ren Chiang, Discovering the gene and disease relationship from biomedical literature database by applying the information retrieval technology, International Medical Informatics Symposium and Show in Taiwan 2002, Oct. 2002.
53. Shih CL, Chiu HW (Kyoto, Japan, 2003) Automatic spot detection of cDNA microarray images using mathematical morphology methods. IEEE EMBS Asian-Pacific Conference on Biomedical Engineering 2003
54. Tsai AC, Chiu HW. (2002) Relationship between heart rate variability and electrolyte concentration in chronic renal failure patients under hemodialysis. International Journal of Bioelectromagnetism, (4), 307-308.
55. Yen YT, Hsu CY(2003). Developing an IR and NLP-Based Algorithm for Biomedical Literature Searching. Journal of Korean Society of Medical Informatics, Vol. 9, pp. s184-s186.