

醫療資訊系統之整合機制：以異質性資料庫與無線網路為例

黃嘉政 吳光閔

南華大學資訊管理所

摘要

於早期資料庫興起時，各醫療機構紛紛導入資訊系統，管理本身的醫療資源，但是在今日這個網際網路與無線網路蓬勃的時代裡，若是醫療資訊沒有透過這個時代的技術，加以整合利用，將是一種資源的浪費，就好比 921 地震時，若是受傷的民眾資料在地震中損毀，那就會造成急救上的延誤，形成社會成本更大的損失；因此，在各醫療機構的分散式資料庫中，最近經由一些學者對資料庫中各種異質性分析之結果，本研究在不用改變原醫療機構的資料庫情形下，提出一種可行性的醫療資訊資料庫之架構，利用一個前置處理，配合知識資料庫，及加值處理機制，來整合醫療資訊系統，並創造其中的醫療附加價值，但此系統架構最大的優勢是使得分散的醫療資源，可以獲得整合，提供現有醫療機構、研究教學單位，甚至於利用無線網路，給予遠距醫療系統使用，發揮其整體最大之效用。

[關鍵詞]：異質性(Heterogeneous)、分散式資料庫(Distributed databases)、知識資料庫(Knowledge databases)、無線網路(Wireless network)、遠距醫療系統(Telemedicine System)

壹、前言

在過去資訊發展的幾年中，很多的醫療機構及醫院，不斷地建構屬於自己的資訊系統，而這些資訊系統都是內部或區域性的系統，所以在資料庫的設計上皆相異很大，因此，在今日這個資訊膨脹、且快速的網際網路世界裡，想要充份地利用這些醫療資訊的資訊，勢必要把這些分散式資料庫做整合，當然，醫療資訊系統在整合資料庫之前，一定會先遇到資料庫異質性的問題，這也是整合資料庫的最大困難點，也是本研究所要針對的議題。

早期，只有少部分的學者，對於分散式異質資料庫做基本的定義[1]，其中意味者要在異質的分散式資料庫中，分析其異質性的關係，事實上，也已經有學者開始在進行這些分析模

型，例如 MIPS 系統[2,3]，但是都是都局限於其它產業，直到後來才有人對醫療產業做異質分散式資料庫的異質性分析。

最近這幾年，開始有學者在分析資料庫異質性問題，探討各家醫療機構，相異之醫療資訊系統的資料庫，其中各資料庫之名稱、關連結構、資料值、語意、資料模式及時間環境上之異質性，並探討其異質性之比較[4]，其關係幫助本研究進行其醫療資訊系統的異質性的定義及分析及醫療資訊的整合。另外，若整合其醫療資訊系統之後，其醫療資訊亦可以透過無線網路，提供給遠距醫療使用；而目前在無線網路通訊系統方面，其系統目前大致上有 GSM System、GPRS System、CDMA System、Satellite Communication System 及 802.11 無線

區域網路 (WLAN, wireless Local Area Network) ...等, 其技術也發展到相當成熟的階段, 例如一個特殊的無線網路(ad hoc network)[5], 它是一種不需要收發器 (Access Point) 的高頻寬無線傳輸網路, 不須依靠現有的公眾網路, 便可以做通訊之機制[6,7,8,9], 如此一來便可以在遠距醫療時, 利用行動裝置, 可以方便且快速的存取或接收整合後之醫療系統上, 即時與非即時的醫療相關資源。

本研究是針對各個醫療機構, 已分析過的分散式資料庫, 其中各種相異性問題, 再做進一步的歸類分析, 以致於能做實際上之運作; 其方法是先分析其哪些醫療資料, 對診療的重要性或具有其重大影響, 再進行資訊和加值處理的轉換, 轉換後之醫療資訊再載入一個新的系統中, 利用這種前置處理的方法, 以整合這種具有異質性的分散式資料庫, 一方面讓醫療資訊能適用, 且快速地應用於這個國際網路與無線網路蓬勃的時代中, 另一方面也促使醫療資源的高度, 且充份地利用。

下一個段落, 本研究會先進行問題的描述及做一個基本的定義; 而在第三段, 會介紹醫療分散式資料庫的異質性與遠距醫療無線網路的分析結果; 第四段, 則是本研究針對醫療系統分散式資料庫的異質性, 所提出的一個前置處理方法; 第五段, 會做一個總結論與對未來之展望。

貳、問題描述

現在很多醫療機構都有建構自己的醫療資訊系統, 以往各個醫療機構皆是依自我之需求, 定義本身的資料

庫結構, 如表格 (Table) 的名稱、欄位 (Attributes) 之名稱及其所包括於該屬性下, 值的內容, 屬性的型態...等, 皆各有所異,

由於目前全民健康保險的實行, 因此在醫院的分級制度的落實顯得更加的重要, 家庭醫師在各地區所扮演的角色也就相當的重要, 加上電腦資訊化的趨勢, 使得病人病歷格式也會有所不同, 再加上最基層醫院並未與其他醫院作彼此的網路連結, 因此在重大災難發生例如 921 大地震, 某些醫院遭到損毀, 若這些醫療資訊已有做整合, 就不會有病歷資料的遺失, 因而發生就醫延誤或是作病人病症作最有效診斷, 如病人有無藥物過敏、近期就醫記錄。

目前許多研究也只針對各分散病歷資料庫提出各種目前發生的各種問題, 並沒有一個比較好的解決方式, 就連之前 Extranet 在轉診之應用的研究當中, 提出一個醫院資訊架構而已, 並未對資訊重複、資料格式定義提出更詳細的辦法, 現今我們將延續這項研究, 作一個延伸探討。目前我們在整個醫療資料庫發生幾個重要問題, 如下:

1. 無法提供雙向更新: 對於轉診的病人在完成就醫出院後, 其就醫記錄未能更新到原醫院, 因而產生資料的不一致。
2. 完整的就醫記錄: 對於病人而言就醫記錄最為詳細, 可能是在其最常就醫的醫院。但也不是最具有最完整病歷資料, 若病人有轉診或發生事故而到他處就醫, 這家原醫院也將不具有其轉診後的就醫記錄。
3. 資訊重複性高: 由於病患就醫是屬

於不可預測之行爲，因爲一般人若發生了緊急事故，可能就會在各地區之醫院擁有數個相同的就醫記錄，因此在其設計資料庫的過程將會有資料重複之疑慮。

4. 格式定義：各醫院對於本身的資料庫或病人的病歷格式並沒一個固定的格式與標準導致各種異質的產生，因此對於分散式資料庫的整合將有相當的困難。

參、醫療分散式資料庫的異質性與遠距醫療無線網路的分析

3.1 醫療分散式資料庫的異質性分析

在 2000 年時，H.T. El-Khatib, M.H. Williams, L.M. MacKinnon, D.H. Marwick 等多位學者曾提出對不同的醫療機構中，其資料庫設計的結構皆不大相同，若要整合這些資料庫之資訊，必定會有異質性的現象，如表一所示，由表中可以發現其異質資料的特性有哪些狀態，以下在進一步做詳細的說明。

表一 分散式資料庫會產生的異質性分類表

資料來源：Information and Software Technology 42(2000)

類 型	內 容
1. 名稱的異質	名稱同義字
	名稱同音異義字
2. 關連結構的異質	關連型式
3. 值的異質	數值對數值的關係
	字串對字串的關係
	數值對字串的關係
4. 語意的異質	結構
	不完整的關係
	哪些資料的呈現
	獲得哪些資料
	在概念層的差異
5. 資料的異質	範例的差異
	行為的差異
	相依性的差異
	在限制上的差異
	隱含值
	關連鍵
6. 時間上的差異	定義域的差異
	不同時的稅金導致非同步更新

3.1.1 名稱異質

一個相同的值，如某個 A 值可能會儲存在不同系統的不同資料庫中，或者是儲存在不同資料庫的屬性中，以致於不能做最簡單的查詢及轉換，這種就稱爲名稱的異質性。

3.1.2 關連結構異質

在資料庫的表格 (Table) 中，可能表格的名稱相同，但是其中表格的寬度 (Table's Degree)，也就是屬性 (Attributes) 的多寡不同，導致關連上的問題。

3.1.3 值的異質

就是在不同系統中，其數值所表示的意義不盡相同，如數值對數值的關係之差異，就病患體重而言，有的以公斤，另外一個可能會是磅來儲存資料，造成資料的異質性；當然，字串對字串的也有值異質性的問題，如一個可能是是“男”、“女”來儲存資料，另一個可能是以“Male”及“Female”或“M”及“F”，其它還有數值字串或結構，也會造會值的異質數。

3.1.4 語意異質

如同一個表格中，病患電話這個欄位屬性而言，其一個資料庫中可能是家裡的電話號碼，而在另外一個資料庫中，可能有的資料是空的 (NULL)，所呈現的資料不盡相同，有些可能在的位點或層級不同，造成資料庫中語意上的異質。

3.1.5 資料異質

資料的加入或刪除的時間上差異，造成的資料異質，如 A 醫療機構的資料庫可能保存 15 年，但是另一家醫療機構的資料庫卻保存 10 而已，不

同的保存時間，造成其資料異質性問題，當然也有其相依性問題，其資料庫本身的關連度高低，也會造成的異質性問題。

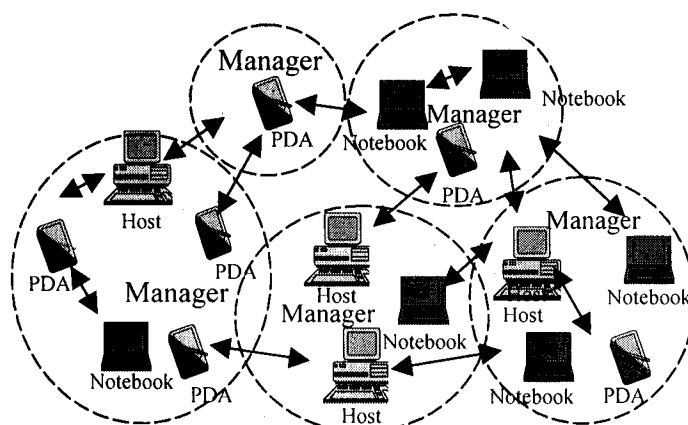
3.1.6 時間上的異質

這個問題是說在一個值的語意表示上逾時後，其定義域 (Domain) 的改變，所造成的資料異質性，如一個電話的區域號碼，在一月由二碼增加至三碼，但是在這個時間之後，有的資料庫在更新至三碼，但是，有些卻維持二碼而已，由於時間上的變化，所造成的資料異質性；那若是某些資料的改變，好比稅率的改變，所形成的非同步更新，也會造成異質性這種情形發生。

3.2 遠距醫療的無線網路

在遍遠的地區，其醫療的設備不足，而且在交通或者通訊上皆很不方便，所以醫療資訊也相當有限，若能利用無線網路結合整合後之醫療系統，使醫療資訊系統能有效運用無線通訊的傳輸方式，建立起無距離的遠距醫療所需要之相關醫療資訊。

現今無線通訊系統，大致上可分為下列幾種無線通訊系統：GSM(Global System for Mobile Communications) System、CDMA(Code Division Multiple Access, 分碼多重存取) System、GPRS(General Packet Radio Services) System、WLAN(Wireless Local Area, 無線區域網路)與 Satellite Communication System...等系統，如果醫療資訊系統在整合異質性之後，能有效利用行動電話，或者是通訊機具的可移動性，以克服在遍遠或交通不便之地理位置限制，而能快速的傳輸遠距醫療所需之資訊。



圖一 無線 Ad hoc 網路

資料來源：Broch, 1998, Hou, 1999, Sivakumar, 1997, Lin, 1997

此外，面對於無線行動通訊系統的範圍，會因為無線基地台的建構，與成本之間的效益之關係，但是在遠距醫療中，又要面對較頻寬的文字、圖片，甚至於影像...等，大量醫療資訊的傳輸問題。因此，近期所發展的藍芽 (Bluetooth) 科技通訊技術，與 IEEE 802.11a 無線區域網路傳輸技術，其這些技術能有效的結合一些電子裝置，例如：桌上型電腦、筆記型電腦、個人行動助理 (PDA)，與行動電話...等，其可組成一個無線之傳遞系統，而成為一個不用收發器 (Access point) 的高寬頻之無線 Ad Hoc 網路，其網路架構如圖一所示，其通訊方式不須以原有之公眾通訊網路，即可進行網路通訊輸傳。

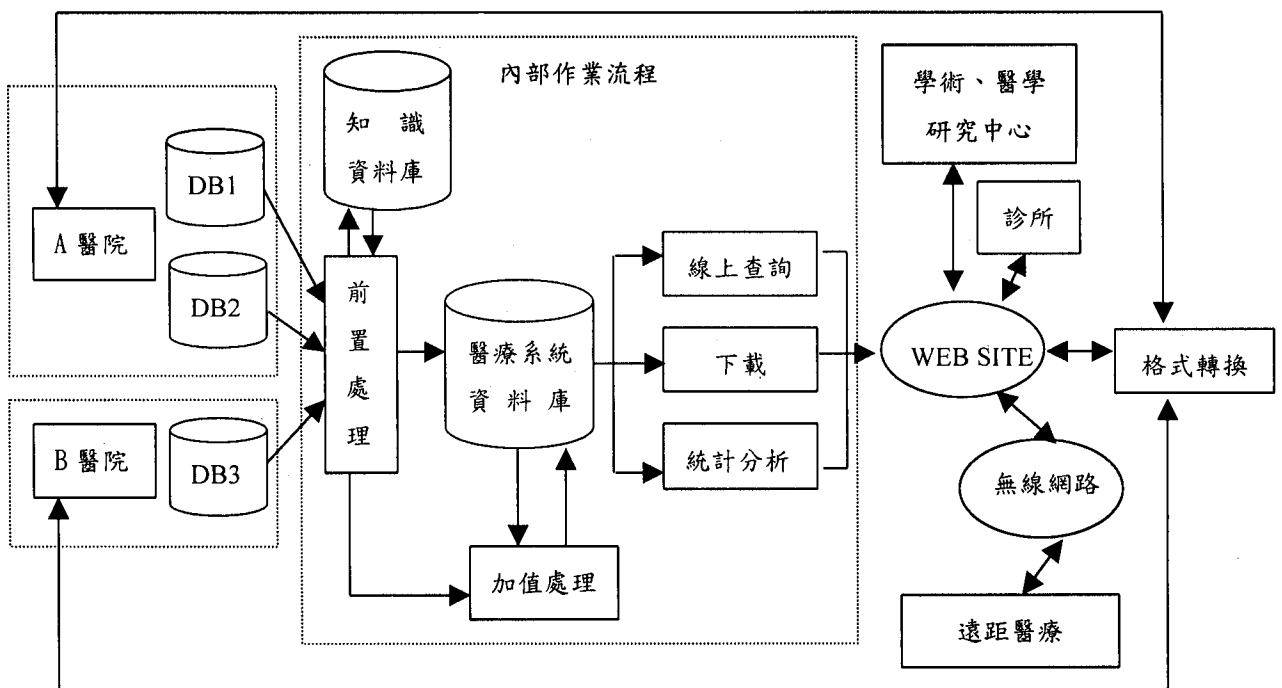
另外，還有現今太空科技的進步，所以在衛星的通訊上也非常的發達，而且衛星通訊的傳輸範圍廣大，而且不易受地形及地理位置影響，這種種的好處，可以帶給通訊無限的便利。雖然衛星傳輸有通訊費用高，通訊頻道少...等缺點，但無疑地它廣闊的傳輸範圍...等，各項便利卻能促使在遠距醫療時，能從整合後之醫療資

訊系統，即時而快速的接收到其醫療所需之相關資訊。

肆、對異質性的前置處理方法

本研究對於異質性分散式資料庫處理的一個系統架構圖，首先，是由外部的各醫療機構，把各自的資料庫資訊(如圖二中 A 醫院和 B 醫院的 DB1, DB2, DB3 資料庫)彙集進入本系統的內部作業流程，接著在配合知識資料庫的邏輯及法則，做資料的前置處理，完成後再進入加值處理的資料處理機制，並產生附加價值，最後，再將這將資訊傳到網際網路上，透過網

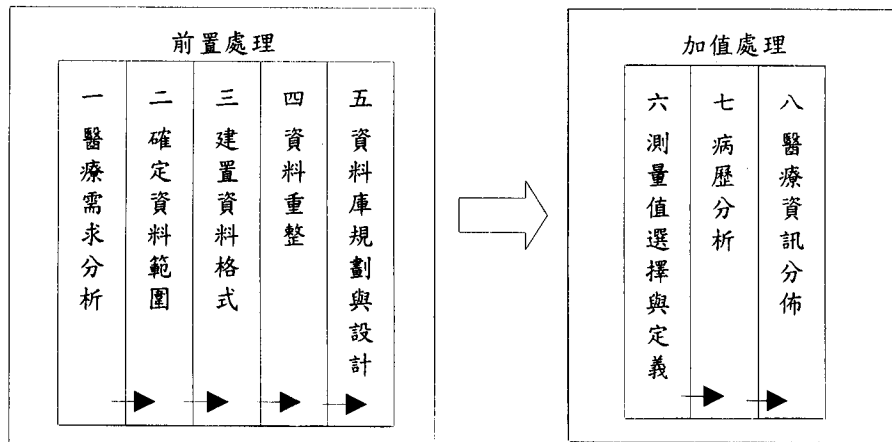
際網路資源分享的傳輸特性，可以供給各個醫療研究中心研究之醫療資訊，提供偏遠之診所相關之資訊，或者透過無線網路，提供遠距醫療所需之相關醫療資訊，甚至透過格式的轉換，使原先之醫院(如圖二中 A 醫院和 B 醫院)可以得到本身自己醫院之格式，而無需浪費醫療資源，更改醫療現有之系統，以助醫療服務品質之提昇，其整體的系統架構圖如圖二所示。



圖二 系統架構

在作業流程中，主要分成前置處理和加值處理兩二類，在前置處理中，會先做醫療需求分析、資料範圍的界定、資料格式的建置、資料重整及資料庫的規劃與設計，主要目的是做彙集資料之定義、轉換及重整；另

外加值處理，就是再將其所得之醫療資訊做分析，尋找出醫療資訊前後之關連性，以提供醫療上的附加價值，以下會在進一步說明前置處理和加值處理內部各項工作程序，其作業程序圖如圖三如示。



圖三 作業程序

4.1 醫療需求分析

- 1.時間效用：病歷對於病人的效用，在於診斷時間的掌握，在以往病歷對於不同醫師不具有任何價值但這往往會增長診斷時間，甚至重頭檢查起這樣將會浪費很多的醫療資源，有時這將是決定病人生死的重要關鍵。
- 2.研究價值：病歷對於醫院而言由其是已死的病人，可能不具有任何的價值，但是若能對這些病歷加以的整理，或許可以察看出那一種醫療方式對於病人是較好的，甚至可以從過去的記錄去發現一些疾病的治療方法。
- 3.考核價值：對於醫師而言，他們可以透過去醫療病歷仔細研究所有醫療的成功和失敗，再和他人比較然後定期檢討低於標準或高於標準的原因，這對於日後的醫療品質提昇有相當的幫助。
- 4.附加價值：對於一些衛生機構而言，龐大的病歷資料對他們也是可以創造附加價值，衛生人員可從過去的病歷資料去分析出一些傳染病的資

料，甚至他的流行區分佈狀況，這都大量病歷資料可以創造出來的附加價值，例如資料擷取（data mining）。

4.2 確定資料範圍

對於異質資料庫我們要整合這些資料庫，必須先對這些資料進行資料範圍界定，才能讓我們的病歷資料更容易去定出統一的格式，也是為我下面幾個步驟的前端處理，以下就幾個例子作簡略說明：

- 1.日期：在蒐集各醫療機構的資料時，其各單位的日期資料可能分為幾種，例如 88 年 10 月 1 日、88/9/1、1999/9/1、88.9.1 這些都是將蒐集到的資料範圍，因此亦是允許的範圍。
- 2.姓名：在姓名表示方面，其各單位的姓名資料可能有中、英文之情形產生，因此將資料定義的範圍，以中、英文名字之表示為主，例如王小明，Davie Wang。畢竟在台灣使用中文或英文名字是較常見的，因此對於往後資料轉置成本之降低，有相當大的幫助。
- 3.性別：在性別表示方面，其各單位的

性別資料可能分為幾種，例如可能收到的性別資料有”男”、”女”、”M”、”F”...等表示方式，這些都是可能蒐集到的資料範圍。

4.3 建置資料格式與轉換機制

將來在彙集來自各個不同醫療機構的資料庫資料時，可採取幾個建構資料格式的方式。

- 1.異質性較低：對於一些比較屬於同質的資料，直接採取該項資料類型為主要的資料格式，其優點是，在轉置的過程或時間上，就會減少很多的轉置時間，甚至可以減少往後的轉置成本。
- 2.較短資料型態：在重建資料過程中，可以找一個或定義一個較短且有意義的值，這樣未來在查詢及資料比對的效能上，將有莫大的幫助；而且對於整個資料庫容量而言，也就相對的縮小，例如疾病名稱可以用較短的縮寫表示。
- 3.時間擴大：對於各醫療機構資料庫資料的儲存年限之限制，影響到資料異值的情形，如先前所提 10 年、15 年的資料年限設置，在這裡可以採取不設限的方式，因為這些歷史醫療病歷，可以做為加值處理的資料量，並且藉由過去的記錄及資訊，使醫生或研究學者更加了解疾病的發展過程，並從中找出解決或預防之道。
- 4.編碼轉換：對於這種因時間上的異質，我們可以先建立起一個編碼轉換來重新定義表格名稱、欄位名稱或資料的編碼方式，以適合未來的改變，雖然這種方式可能較花費時間，但卻能在不改變原有醫療機構的資料庫下，做資料的轉置。

5.名稱更正：不同醫療機構中，可能對不同疾病，其所定義的名稱縮寫相同，例如說一樣是“PNE”這個縮寫，某家醫療機構可能是代表肺病“Pneumonia”之意，但另一家醫療機構，可能另一家醫療機構代表肺塵埃病“Pneumocniosis”，如此一來將會導至資料的同名異質，所以就必須先針對這些名稱作重新的更正，以確定未來轉置後資料能達到同名同義。

6.欄位擷取：對於關連結構異質方面，可以將各個資料庫中比較常用，也就是說其權重較重的屬性欄位取出，作為整合後資料庫的規劃基礎。

4.4 資料重整

在病歷內容的設計上，是由很多不同類型的資料所組成，然而病歷的用途不完全在本身的記錄內容，而是在如何快速且有效的利用記錄，因此在其記錄資料的前後順序上，就顯得特別重要，大概可以分為幾項分類：

- 1.不分類別只依時間先後排序
對於醫師而言整個醫療過程，醫師所要知道的是對於整個病人最新的醫療記錄或報告，但此種分類對於醫師而言雖然可以馬上獲得病人資料，但是對疾病比較不具有通盤性了解，以一般小型診所為主。
- 2.先按類別分在按同一類的記錄重新排序
反之先分類再排序，將使得醫師在對於病人的相關疾病能夠很清潔全盤了解，甚至是疾病之間的關係如併發症起因。

4.5 資料庫規劃設計

對於已經重整過的資料，可以將其轉成文件式資料庫，以文件方式呈

現，例如電子病歷，對於醫師而言可以直接在該文件表單上作複診記錄、開藥單據、補寫病歷記錄，對於轉診、呈遞可以直接以 e-mail 方式送到下一個醫師或醫院的手中。不同醫院之間的資料庫可作批次抄寫以求資料庫的完整性。

4.6 測量值選擇與定義

在這裡系統會選定一些比較常用到的數據，例如某一種病的罹患率為何，又其對男性、女性，或對老人、壯年人、青年人及小孩的罹患比率統計分析，再者如就診率也是，像是某個病人到那一個醫院的就診次數，必須清楚地對這些測量值做定義，以防止錯誤或資料異質的情況發生。

4.7 病歷分析

這是在本架構中的加值處理，也是在醫療上的一種附加價值，我們可以利用一些像 data mining 的技術或結合專家系統的推理機，藉由這些工作尋找出一些疾病發生的各項因素，像一些慢性病（糖尿病）或遺傳疾病發生的原因，可以提供給醫師一些臨床經驗作為參考，或者在醫療技術上之突破。

4.8 醫療資訊分佈

在一固定的期間，將會定期為整個醫療結果作評估與檢核，直接公布於網站上，讓各個醫療機構之間，對於彼此間醫療的情況或醫療資訊更加地了解，達成醫療知識的分享。也可以對於病人作線上共同會診，以期降低醫療資源的浪費。

伍、結論

在醫療機構未進入網際網路時代前，其醫療系統都局限於本身的醫療

資源，供自己本身的研究或醫療用途，若要轉診的話，也要透過醫療系統先把轉診資料轉換成轉診單，一方面如果要進行慢性病或遺傳疾病的研究時，其能提供的資訊亦非常有限，但這些都是奠定未來發展的基礎。

本研究已針對各個不同醫療機構所開發的資訊系統，根據其異質性的問題，提出了一個系統架構，其架構內容是配合所定義的知識資料庫，做前置處理、加值處理和資料轉換，此方式可以再不更改原醫院的醫療資訊系統下，解決醫療機構異質資料庫的整合問題，並且透過網際網路與無線網路的無遠弗界之特性，促使醫療資源能有效的整合和完善地利用，且達到醫療資訊共享。

展望未來，期望醫療團體、政府或較大型的醫療機構，能協商各醫療機構，訂定協定與標準，再透過數位時代的資訊組織工具，把醫療資源整合落之實際，提昇現有醫療服務與遠距醫療之品質，並且在醫療技術上有所重大之突破，相信這將是全人類之福祉。

參考文獻

- 1.A. K. Elmagarmid and C. Pu, "Introduction: special issue on heterogeneous databases (guest editors)," *ACM Computing Surveys*, Vol. 22, 1990, pp. 175-178.
- 2.L. M. Mackinnon, D. H. Marwick and M. H. Williams, "A model for query decomposition and answer construction in heterogeneous distributed database systems," *J. Intelligent Inf. Sys*, Vol. 11, 1998, pp.

- 69-87
- 3.W. J. Austin, E. K. Hutchinson, J. R. Kalmus, L. M. Mackinnon, K. G. Jeffery, D. H. Marwick, M. H. Williams and M. D. Wilson, "Processing travel queries in a multimedia information system," *Proceedings of information and Comms chnologies in Tourism*, Springer, Berlin, 1994, pp. 64-71
- 4.H. T. El-Khatib, M. H. Williams, L. M. MacKinnon and D. H. Marwick, "A framework and test-suite for assessing approaches to resolving heterogeneity in distributed databases," *Information and Software Technology*, Vol. 42, 2000, pp. 505-515
- 5.Huang, J. -F. and Chang W. -C., "Mobile Switch Management in Wireless ATM Networks," *Workshop on Internet and Distributed System*, Vol. 2, 2000, pp. 478-483
- 6.Broch, J., "A Performance Comparison of Multi-hoc Wireless Ad Hoc Network Protocols," *Proceeding of the MobiCom*, 1998, pp.85-97
- 7.L. M. Mackinnon, D. H. Marwick and M. H. Williams, "A model for query decomposition and answer construction in heterogeneous distributed database systems," *J. Intelligent Inf. Sys*, Vol. 11, 1998, pp. 69-87
- 8.Hou T. -C., "Adaptive Clustering for Mobile Ad Hoc Networks," *ISCOM*, 1999, pp.254-258
- 9.Sivakumar D. R., "Routing in Ad Hoc Networks Using Minimum Connected Dominating Sets," *IEEE ICC*, 1997, pp.376-380