

建立「疫疾」中醫藥典籍 文獻擷取索引典

陳逸光

慈濟大學 醫學資訊系

花蓮

(2003 年 11 月 21 日受理，2004 年 2 月 10 日接受刊載)

今年（2003），在亞洲及加拿大所爆發之流行病——嚴重急性呼吸道症候群（SARS），醫療、衛生及生物科技界精英皆投入抗 SARS 最前線，人人談 SARS 色變。中國古代當大疫流行，統治者承受相當大的壓力，所以史籍上很少詳細記錄嚴重大規模的疫情，經常以寥寥數語便略過，因為封建社會認為“疫”是天災不詳之兆。中國歷代醫家，對抗疫的貢獻相當大，如東漢末年張仲景之所以著《傷寒論》就是因為大疫流行之故。倉促中及電子典籍來源所限，本研究在短短三個月內整理出 22 本電子中醫藥典籍，約一千萬字文獻資料，並建立「疫」索引典內含 37 個關鍵詞，共篩選出二萬一千餘個與索引典相關之條文，並且以動態資料庫形式，放置在 TCMET 網際網路上供學者專家們參考、下載及研究。

關鍵詞：中醫藥典籍，疫，資訊擷取，索引典，網際網路。

前 言

公元 2003 年全球醫療衛生界最大事件；莫過於在本年初發生的嚴重急性呼吸道症候群 SARS (Severe Acute Respiratory Syndrome)，醫界發現 SARS 是由一種冠狀病毒所引起，這是一種過去只有在動物身上才能找到的病毒，如今卻找上人類作為宿主，由於人類的免疫系統缺乏對抗冠狀病毒 (Corona Virus) 的抗體¹，所以造成一場廣泛性的大流行。在這次事件中最不幸的是這種病毒一旦侵犯人體，有些患者會造成肺部嚴重浸潤 (Infiltration) 甚至發生纖維化 (Fibrosis)，最後導致呼吸衰竭 (Respiratory Failure) 而死亡。中國歷代史書上有關傳染病記載和歷代醫家創立的醫學理論及治療方法，是研究中國古代對抗傳染病的一個重要議題，也是本文一個主要研究動機；藉由電子計算機的幫忙擷取及整理中醫古籍文獻，已經是時代的趨勢，在電子文獻擷取議題中將資訊組織起來構成人類的知識，是資訊擷取科技的一項重要任務，由文獻資料演變成知識此一過程中，資訊「集結」(cluster) 是一項讓電腦認知人類知識的重要依據，「集結」需要有

索引典 (Thesaurus) 足夠代表一個專題或一概念而具邏輯意義」² (Logical Meaning)。過去筆者曾經發表過「黃帝內經」詞典庫³及「河間六書方劑擷取」⁴之資訊擷取 (Information Retrieval) 有很高的準確率 (Precision Rate) 和回收率 (Recall Rate)。本研究以「傳染病」為題，尤其是以「疫」症為資訊擷取對象，針對特定知識 (「疫」)，一方面可以深入探討「集結」之資訊認知，另一方面對中國歷代傳染病的醫學文獻資訊作整理，藉由網際網路之便利，希望對 SARS 研究工作者提供一些有關「傳染病」的電子中醫古籍文獻資訊，對人類醫療衛生作點微薄貢獻，更重要的是身為中醫人理應各展所長多方面發揚中醫。準確率和回收率是兩個有關文獻擷取效率最常用的指標³，有關「疫」疾的中醫古籍文獻篩選過程中，本研究盡可能從歷史學者著作或自臨症經驗中挑選較有代表性的「字彙」，如「疫癘」這個字彙對本文主題「高危險流行性傳染疾病」有較高的準確率，用單一個「疫」字來搜尋則準確率下降而回收率相對會提高，針對上述之議題本研究從傳染疾病的觀點整理出「疫」相關的索引典 (Thesaurus)。「傳染病」代表醫學上的一種概念，而利用上述具代表性之詞典擷取技術，將為學者專家們提供許多有關「疫」症之重要中國醫學典籍文獻信息。

當研究主題確定後，首要工作是要篩選能代表有關「傳染病」知識的中醫索引典，索引典的建造對資訊擷取的效率有決定性的影響。本研究中所收集的中醫古籍文獻共有 22 種，約一千萬字，文獻選擇類別包括：理論性典籍 (黃帝內經)、臨床醫籍 (傷寒、溫病)、方劑學 (千金方、外臺秘要)、藥物學 (經史證類備急本草、本草綱目) 等，運用電腦文獻資料處理技術，能將所有電子中醫古籍轉換成資料庫，再經過條件邏輯運算篩選出相關知識條文，本研究已成功建置「傳染病」相關的索引典及一千萬字的中醫古籍文獻資料庫。最後運用文獻擷取技術，整理出與「疫」索引典相關條文共二萬一千餘條，並已在線上加入至 TCMET 網頁上以動態資料庫形式⁵供全球網友使用。

材料與方法

一、材料

(一) 收集中醫古籍文獻庫

本研究共收集電子文獻 22 種 (見附件一): B1 (黃帝內經) B2 (《金元四大家著作》: 劉河間、張子和、李東垣、朱丹溪) B3 (明·張介賓·《景岳全書》) B4 (明·王肯堂·《女科證治準繩》) B5 (宋·陳自明·《婦人良方》) B6 (明·武之望·《濟陰綱目》) B7 (唐·王燾·《重訂外台秘要》) B8 (宋·御纂·太平惠民和劑局方) B9 (清·汪昂·《醫方集解》) B10 (唐·孫思邈·《備急千金要方》) B11 (清·柯琴·《傷寒來蘇集》) B12 (清·雷豐·《時病論》) B13 (清·葉天士·《溫病條辨》) B14 (清·戴天章·《廣瘟疫論》) B15 (清·柳寶詒·《溫熱逢源》) B16 (清·吳有性·《瘟疫論》) B17 (清·葉天士·《臨證指南醫案》) B18 (清·薛生白·《溫熱病篇》) B19 (宋·唐慎微·《經史證類備急本草》) B20 (清·張璐·《張氏醫通》) B21 (清·沈金鰲·《雜病源流犀燭》) B22 (明·李時珍·《本草綱目》); 上列之電子書籍可參考“電子中醫藥古籍文獻”網站上之資料。

(二)文獻資訊處理

資料庫使用微軟之 SQL 2000 伺服器，在網際網路上以動態伺服器頁 (Active Server Page)⁵ 形式供網友瀏覽「疫」疾資料庫，內容包括關鍵詞之書目、出處及上下文資料等資訊都有線上即時提供閱覽之功能。

(三)網際網路架構

網際網路架構基本與 2002 年 TCMET 相同 (參考筆者論文³⁻⁵ 材料與方法項)。

二、方法

(一)建立「疫」疾相關中醫索引典

本研究試從歷史學家的文獻資料中，尋找中國歷代有關「疫」疾的醫學文獻資料，以篩選出在“傳染病”此一專題上有比較高精確率 (Precision) 及回收率 (Recall) 的關鍵詞。

從歷史文獻中，有相當多的證據可以用來說明某些詞彙是具“流行”與“傳染”特性之記錄⁶。文獻記載如：

(春秋之時) 溫氣疫癘，千戶滅門；—《論衡》

建安二十二年 (217 AD)，癘氣流行。家家有僵尸之痛，室室有號泣之哀。

或闔門而殪，戶覆族而喪。—曹植〈說疫氣〉

寶應元年 (762 AD)，江東大疫，死者過半。—《新唐書》

從古代中國有關大規模流行之疫疾紀錄中，可篩選出“疫疾”、“疫癘”、“瘟疫”、“疫”...等詞彙 (見附件)。

從筆者臨床經驗中，外感六淫相關的文字如：“風寒”、“感冒”、“溫病”等關鍵詞，可以預期與「疫」相關之“精確率”雖然會較低，但仍然會有不錯的“回收率”。“精確率”低表示從文獻所收集的信息中有很多雜訊，“回收率”高代表被篩選出來之文獻與「疫」疾正向^{4,5}。由於電腦科技給文獻擷取帶來之便利³，本研究盡量提供不同程度“精確率”及“回收率”的關鍵詞，以建立“疫”索引典，使文獻資訊「集結」(cluster) 有更高之知識代表能力。

(二)文獻資料庫建置

上述之 22 種電子中醫古籍文獻資料皆以文本 (Text) 形式記載，為方便在網際網路上讀取資料，本研究所有電子文獻檔皆以資料庫形式存取。文本轉換成資料庫之方法，可參考筆者曾以“網際網路中醫藥典籍文獻動態資料庫研究”⁵ 為題發表之論文。

(三)“疫”文獻篩選及分類

在建立“疫”疾中醫索引典後，本研究以微軟之 Visual Basic 6.0 為工具，將文獻資料讀出並與索引典中之關鍵詞逐項比對，符合篩選條件的資料將轉成資料庫格式儲存。文字比對之方法可以參考筆者曾發表之“黃帝內經貫珠詞典庫之研發與應用”中有關 n-gram 文字擷取論著³。

(四)網頁與資料庫整合

網頁可以藉著動態伺服器頁 (Active Server Page) 形式供網友瀏覽「疫」疾資料庫中所有經整合後之

文獻資料，本研究用三層網頁來引導讀者閱覽資料，最下層網頁顯示出“關鍵詞”所在位置前後 20 行全文資料（見結果），網頁建置之方法已詳載於“網際網路中醫藥典籍文獻動態資料庫研究”⁵文獻中。

結 果

（一）索引典擷取「傳染病」文獻中相關句子所出現之頻率

本研究共整理出與“疫”相關之詞彙共 37 個及相關條文 21,972 處（見附件），每個詞彙在 22 種中醫文獻之出處及全文內容皆以關聯資料庫方式儲存。

（二）「疫」疾重要資訊

“疫”疾索引典（見附件），關鍵詞“疫”與“瘟疫”兩者中，很容易看出單一“疫”字所篩選出來之文獻數量會比“瘟疫”多，但眾多擷取出來之資料與“傳染病”之“疫”無關，所以單一“疫”字之“精確率”較差，但比起“瘟疫”詞彙相對會有較好的“回收率”；意謂單一“疫”字在篩選出來之文獻資料有很多是與“傳染病”之“疫”是相關的，相對於“瘟疫”詞彙而言，較不容易遺漏資料。索引典中所有關鍵詞與“疫病”皆有一定程度的“精確率”及“回收率”。有些關鍵詞是從臨床角度來看，例如“發熱”、“惡寒”等關鍵詞，因為“傳染病”經常會有這些症狀，雖然“精確率”較差，而“回收率”則需要慢慢細讀文獻內容才能決定，本研究仍然將這些“症狀”詞彙一併提供給讀者，是因為這些症狀對學者們研究疫病將會提供相當重要之信息。

（三）網際網路執行情況（見 www.tcmet.com.tw/infection/）

網頁第一層（Infect_Keyword.asp）提供所有索引典之進入點；網頁第二層（Infect_KeyCont.asp）提供關鍵詞整行文獻資料，還有書目、篇、章、節等文獻出處；網頁第三層（Infect_detail.asp）除提供網頁第二層之資料外，尚有上下文共 40 行。

討 論

〔嚴重急性呼吸道症候群〕（SARS）在 2003 年造成中國大陸、香港、台灣、越南、新加坡及加拿大傳染大恐慌，死亡人數約千餘人，而實際感染 SARS 病毒人數約一萬人疫情雖然不是很嚴重，但造成社會極度恐慌⁹，人人談[煞]色變，醫界關切 SARS 在本年冬天是否會再度爆發流行，現代所說的〔流行病〕，傳統中醫文獻常稱為疫、疫癘、瘟、染易、時行、溫病、傷寒⁶⁻⁹等尤其以“疫”一詞最為常見，醫聖張仲景所著《傷寒雜病論》其序記載“「余宗族素多，向餘二百、建安紀元以來，猶未十稔，其死亡者三分有二，傷寒十居其七。」”，可見當年（200 AD）的傷寒就是指疫症，東漢末年赤壁之戰，其時曹營中有疫病存在，有些學者認為是感染血吸蟲病（Schistosomiasis），但亦有學者認為另有原因¹⁰，可見從歷史文獻推斷病原（Pathogen）為何物實在是不容易，像今年大流的 SARS 為冠狀病毒（Corona Virus）由動物傳給人類，病毒可能只是短暫的流行，也許自此之後便消聲匿跡。中國醫學在古代抗疫的經驗，亦相當之豐富，除傷寒論外，還有很多瘟疫專著如吳有性於崇禎十五年（1642）撰成《瘟疫論》，清代醫家戴天章（1644~1722）

著《廣瘟疫論》都是相當著名之瘟疫專家，另外值得一提的中醫典籍《外臺秘要》，其之所以重輯是因為在宋朝皇祐年間（1049~1053）疫病流行，朝廷命御醫重新編輯醫籍唐王燾所著《外臺秘要》給全國醫家學習應用，所以在《宋·重訂唐王燾外台秘要方》序中有云：“臣昨南方州軍連年疾疫瘴癘，其尤甚處，一州有死十餘萬人...如有疾疫瘴癘之處於《聖惠方》內寫錄合用藥方出榜曉示，及遍下諸縣，許人抄筭，仍令秘閣簡《外臺秘要》三兩本，送國子監”，王燾《外臺秘要》自序：貶守房陵量移大寧郡途中，「自南徂北，既僻且陋」，「提攜江上，冒犯蒸暑」，造成「染瘴嬰痾，十有六七，死生契闊，不可問天」的慘劇。可看出當時王燾，亦是因為「痢疾」流行造成親人死傷慘重而編輯《外臺秘要》一書。近代歷史學家亦認同「疫」一詞最常用於描述〔流行病〕之發生，古之「疫」與近代〔流行病〕相似，皆有在一短時間內爆發大量病例，然後慢慢減少至消失為止；若死傷慘重之疫疾通常會以「疫癘」一詞形容其嚴重程度。近代歷史學家亦指出，中國史書記載疫疾經常未明確指出疫疾之名，對疫疾症狀、特徵與病程往往不足，有關疫疾的發生率、盛行率及死亡率，很少有精確的說明，甚至會有「瞞事實的情況發生」⁶，其中一個可能之原因是因為“疫”是天災，不祥之兆，對於古代君王而言，大都會造成民怨、人口傷亡、賦役不足、治安敗壞和戰爭。

我們不難發現所謂「疫」、「疾疫」、「疫癘」或「瘴癘」多有大規模流行的特色。從歷代中國醫學文獻中擷取與「疫」相關的資料是本研究首要目的，本文自歷史學家論著及筆者個人臨床經驗對疫疾的認知，共挑選出 37 個關鍵詞，這一組索引典可規為疫疾之〔集結〕，而資訊學上的〔集結〕觀念可擴展為知識之表現，而「索引典」(thesaurus)則進一步紀錄詞彙之間更多的關係，用以擴展或縮小檢索詞彙的主題範圍¹¹。從「疫」索引典(附件)中「瘟」、「疫」、「疫瘟」出現頻率總數分別為 535 筆，1728 筆及 298 筆；所以「疫瘟」一詞之精確率(Precision)最高，但回收率(Recall)相對較低，若讀者希望在最短時間內找出與「疫」相關之文獻，「疫瘟」一詞會較適合，但缺點是很多與「疫」相關的資料會遺漏掉。反之，若研究人員有很多時間慢慢閱覽篩選資料，較會優先選擇單「疫」一詞之文獻出處，因為有較好之回收率所以資料較不易流失，但精確率下降較易造成所擷取之資料與疫疾無關。

SARS 今年(2003)四月間在台灣引發流行恐慌，本研究在今年七月間完成，此期間共收集了約一千萬字電子資料，文獻共 22 種---「疫疾」專著諸如傷寒論、溫病學、《外臺秘要》和「疫疾」以專論方式散見在其他中醫經典名著中。本研究以資料庫型式開發資料擷取功能，並編寫電腦程序，以便將大量中醫藥典籍文本資料(TEXT)轉入資料庫中。日後，若能夠收集更多電子中醫藥典籍資料，便可藉此電腦程序更新資料，提供更完整的「疫疾」資料庫。最後，本研究在 22 種中醫藥古籍中，找出與疫疾索引典相關的條文 21,972 處，所有被擷取出來之文獻資料及其出處，皆在網際網路上分三層列出(見結果)，讀者可根據文獻擷取之結果，進一步研究中醫歷代醫家診治疫疾之知識。

結 論

自古至今疫疾對人民、社會、國家及帝王都會造成嚴重的威脅，中醫古籍對疫疾的流行病觀並無詳細記錄，但歷代醫家對疫疾的病因、病機、處方及用藥等經驗，都相當之豐富。漢·張仲景因為疫疾造成家族死傷慘重，史籍與醫籍記載「疫」發生之時間相吻合；中國醫籍經常間接提供歷史的證據，而且能深入

探討當時人民在醫療上抗「疫」的經驗是相當寶貴的。本研究針對中醫藥抗疫文獻作了一翻整理，並登錄在網際網路（www.tcmet.com.tw）供全世界學者專家研讀。

參考資料

1. 賴明詔，2003 春之煞：SARS 流行的科學和社會文化回顧，SARS 的病毒學，聯經出版社，台北，pp.1-23，2003。
2. 吳美美，網路資源組織的三個層次，大學圖書館，2：27-35，1998。
3. 陳逸光，黃帝內經貫珠詞典庫之研發與應用，中醫藥雜誌，13：49-57，2002。
4. 陳逸光，電腦輔助擷取河間六書方劑專論資料，中醫藥雜誌，12：1-9，2001。
5. 陳逸光，網際網路中醫藥典籍文獻動態資料庫研究，中醫藥雜誌，11：43-50，2000。
6. 張嘉鳳，「疾疫」與「相染」--以「諸病源候論」為中心試論魏晉至隋唐之間醫籍的疾病觀，臺大歷史學報，27：37-82，2001。
7. 林富士，東漢晚期的疾疫與宗教，中央研究院歷史語言研究所集刊，66：695-745，1995。
8. 李家庚、余新華，中醫傳染病學，中醫傳染病學的基本概念，中國醫藥科技出版社，北京，p.3，1997。
9. 林富士，2003 春之煞：SARS 流行的科學和社會文化回顧，中國中古時期的瘟疫與社會，聯經出版社，台北，p.86，2003。
10. 陳勝崑，中國疾病的歷史與地理 (24)：赤壁之戰與傳染病，當代醫學，10：100-102，1983。
11. 曾元顯，數位文件之資訊組織與主題分析自動化之技術與應用，台北市立圖書館館訊，20：23-35，2002。

附件：「疫」索引典在 22 種書目中出現之頻率統計表*

書目 詞典	書目																						總數
	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	B21	B22	
1. 瘟	0	15	121	2	2	0	1	13	12	0	0	33	1	25	1	98	4	0	12	67	71	57	535
2. 疫	0	41	129	6	6	2	40	36	27	43	0	64	23	394	69	271	15	0	67	220	121	154	1728
3. 瘟疫	0	10	84	1	1	0	1	12	8	0	0	17	0	22	1	54	4	0	4	20	24	35	298
4. 疫癘	0	10	13	0	0	0	4	7	1	3	0	2	0	2	3	0	3	0	4	40	15	13	120
5. 風溫	1	10	2	0	0	0	4	2	4	3	1	48	29	0	14	1	70	0	6	65	21	7	288
6. 溫病	4	30	21	1	0	1	49	2	10	41	8	51	197	0	112	49	1	4	22	139	61	27	830
7. 溫毒	0	3	2	0	0	1	16	0	2	4	0	25	22	0	1	2	0	0	6	25	8	11	128
8. 發斑	0	28	32	1	3	4	15	1	34	1	0	12	5	9	4	14	4	0	3	66	38	20	294
9. 發疹	0	1	0	0	0	0	1	0	0	1	0	6	2	7	1	0	9	0	1	1	1	1	32
10. 風寒	21	199	313	62	83	56	60	26	89	51	45	46	46	45	20	21	45	1	45	366	138	162	1940
11. 痢	0	346	407	154	61	90	639	128	96	230	0	201	97	12	1	27	151	20	670	526	198	1437	5491
12. 屍注	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	9
13. 陰陽易	2	0	2	0	0	0	15	0	0	2	2	0	0	0	0	0	0	0	1	20	2	13	59
14. 尸鬼	0	4	9	0	0	0	0	0	0	6	0	0	1	0	0	0	0	0	4	0	0	5	29
15. 鬼毒	0	0	1	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	13	6	0	17	40
16. 客邪	0	22	6	1	1	2	6	0	5	2	0	0	3	1	3	11	32	2	0	107	13	1	218
17. 戰汗	0	2	4	0	0	0	0	0	0	0	0	0	7	25	9	22	2	0	0	17	5	0	93
18. 伏邪	0	0	0	0	0	0	0	0	0	1	1	7	3	4	76	20	25	0	1	29	11	5	183
19. 時行	3	1	19	6	3	3	18	14	8	36	0	8	1	3	1	7	1	0	29	74	21	48	304
20. 感冒	0	5	26	3	4	7	0	7	15	0	0	8	1	0	5	17	1	0	0	94	29	2	224
21. 春溫	0	4	1	0	0	0	0	0	1	0	1	33	7	0	5	10	15	0	1	4	21	0	103
22. 夏溫	0	1	0	0	0	0	0	0	2	0	0	1	0	0	1	2	2	0	0	9	5	1	24
23. 秋溫	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	4
24. 冬溫	0	2	1	1	0	1	11	0	2	4	0	38	9	0	4	7	19	0	2	62	15	1	179
25. 尸厥	1	4	7	2	2	1	0	0	1	13	0	1	0	0	0	0	0	1	1	4	0	7	45
26. 腸痧	0	2	0	0	0	0	0	0	2	0	2	3	1	0	0	1	0	0	0	3	11	0	25
27. 痧	0	2	0	0	0	0	0	1	3	0	2	39	8	0	5	1	44	0	0	12	732	0	849
28. 發熱	9	257	710	211	263	187	126	41	163	95	246	52	22	72	44	66	53	3	52	1126	242	166	4206
29. 惡寒	13	264	237	89	102	68	83	14	157	66	230	40	39	48	56	31	17	10	8	537	173	52	2334
30. 項強	3	24	18	9	13	8	21	12	21	21	26	2	3	8	3	3	1	0	6	62	32	19	315
31. 背強	0	4	7	10	6	10	12	11	3	13	6	0	0	0	2	1	0	0	4	22	2	6	119
32. 相染	0	1	7	0	0	0	16	0	0	9	0	0	2	0	0	1	0	0	4	2	1	5	48
33. 染易	0	0	3	0	0	0	5	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	10
34. 疾疫	0	0	0	0	0	0	5	0	1	3	0	0	0	0	0	0	0	0	2	0	0	0	11
35. 疫疾	0	1	2	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	5	0	1	9	21
36. 注易	0	0	0	0	0	0	7	0	0	2	0	0	0	0	0	0	0	0	1	0	0	2	12
37. 疰	1	12	20	24	14	2	107	9	3	87	0	11	1	0	0	0	0	0	203	44	43	243	824

* “疫” 相關文獻擷取總數共 21,972 筆

B1 (黃帝內經) B2 (金元四大家) B3 (景岳全書) B4 (女科準繩) B5 (婦人良方) B6 (濟陰綱目) B7 (外台秘要)
 B8 (和劑局方) B9 (醫方集解) B10 (備急千金要方) B11 (傷寒來蘇集) B12 (時病論) B13 (溫病條辨)
 B14 (廣瘟疫論) B15 (溫熱逢源) B16 (瘟疫論) B17 (臨證指南醫案) B18 (薛生白溫熱病) B19 (經史證備)
 B20 (張氏醫通) B21 (雜病源流犀燭) B22 (本草綱目)

BUILDING AN EPIDEMIC DISEASE THESAURUS RETRIEVAL FROM TRADITIONAL CHINESE MEDICINAL LITERATURE

Yee-Guang Chen

Department of Medical Informatics, Tzu Chi University

Hualien, Taiwan

(Received 21th November 2003, accepted 10th February 2004)

Early 2003, an incidence of epidemic disease---SARS (Severe Acute Respiratory Syndrome) was occurred mainly in Asia district and Toronto, Canada. Health professional and biomedical specialist were dedicated themselves to the frontier of anti-SARS, and people endangered stressfully. Lethal epidemic contagious diseases occurred in ancient China, which used to be considered as a disaster from the heaven and considered as a sign of bad fortune, gave ancient emperor in China a lot of trouble in governing their country. In fact, most of the ancient official historical documentation of the severe epidemic regime was over-simplified or skipped and very few occurrences were recorded precisely. Historically, there were many ancient Chinese medical physicians dedicated themselves in treating for epidemic diseases and some scholars put their effort on collecting effective medical formula from their ancestor. Each modern traditional Chinese doctor knows that a famous medical doctor, *Juei-Tain Cheng*, in the Han Dynasty (200 AD) wrote a medical book “*Shang Han Lun*” because of a lethal outbreak of an epidemic disease---*Shang Han* which causing death to a lot of his relatives. In this study, a set of 22 volumes of electronic medical literatures from the ancient China was collected and stored into the database. Additionally, a thesaurus was defined with 37 keywords relating to “Epidemic Disease”. Around twenty-one-hundred-thousand of sentences relating to the Epidemic thesaurus were retrieved from the medical literature database. All the extracted information was stored in the data repository and showed on the TCMET (Traditional Chinese Medicinal Electronic Text) internet dynamically.

Key words: Classical Chinese Medicinal Literature, Epidemic, Information Retrieval, Thesaurus, WWW.