國立臺北護理健康大學資訊管理研究所碩士論文計畫書(草案)

National Taipei University of Nursing and Health Science

Graduate Institute of Information Management

指導教授：黃衍文 博士

*Advisor: Ean-Wen Huang Ph.D.*

中文標題

English Title

研究生：徐明睿 撰

*Name: Ming-Ray Hsu*

中華民國一百零一年十二月六日

December 6, 2012

國立臺北護理健康大學資訊管理所碩士論文計畫書摘要

研究所組別：資訊管理研究所

論文名稱：

指導教授：黃衍文 教授

研究生：徐明睿

時間：民國101年12月

# 論文摘要

隨著衛生署在電子病歷的推動下，將國際已經發展的健康資訊交換第七層協定(Health Level Seven, HL7) -臨床文件架構(Clinical Document Architecture, CDA)作為標準，讓醫院間透過此標準規範，彼此能夠進行資料交換，達到病歷資訊交換目的，以降低醫療成本，提升病人的醫療品質與安全性。然而在各醫院機構所自行發展的電子病歷系統，都有自己一套的系統規格、資料欄位及代碼與架構，因此在資訊傳輸交換上就會產生欄位對應的錯誤。

研究首先針對電子病歷交換標準單張與醫學詞彙代碼標準的資料進行了解，並探討電子病歷之架構與相關醫學詞彙代碼對應關係之研究，最後建立此系統以實作的方式來完成該研究計畫，預期將建置一套符合本土化的電子病歷交換標準之資料庫與CDA編碼產生之功能，以達到電子病歷交換標準之代碼與名稱的一致性。

關鍵字: 健康資訊交換第七層、電子病歷、醫學詞彙代碼

# 目　　錄

[論文摘要 I](#_Toc312972175)

[目　　錄 II](#_Toc312972176)

[圖目錄 III](#_Toc312972177)

[表目錄 IV](#_Toc312972178)

[第一章 緒論 5](#_Toc312972179)

[第一節 研究背景 5](#_Toc312972180)

[第二節 研究動機 5](#_Toc312972181)

[第三節 研究目的 6](#_Toc312972182)

[第二章 文獻探討 8](#_Toc312972183)

[第一節 電子病歷 8](#_Toc312972184)

[第二節 醫療資訊標準 9](#_Toc312972185)

[壹 臨床文件架構 9](#_Toc312972186)

[貳 醫療編碼系統 13](#_Toc312972187)

[第三節 代碼對應之問題 16](#_Toc312972188)

[第四節 XML 18](#_Toc312972189)

[第三章 研究方法 20](#_Toc312972190)

[第一節 研究步驟 20](#_Toc312972191)

[壹 研究流程 20](#_Toc312972192)

[貳 系統建置流程 21](#_Toc312972193)

[第二節 系統需求分析與設計 23](#_Toc312972194)

[壹 需求分析與系統架構 23](#_Toc312972195)

[貳 系統分析與設計 26](#_Toc312972196)

[第三節 相關技術及開發工具 33](#_Toc312972197)

[第四章 預期成果 34](#_Toc312972198)

[第五章 可能遭遇困難 35](#_Toc312972199)

[參考文獻 36](#_Toc312972200)

# 圖目錄

[圖 1 CDA文件結構示意圖 11](#_Toc332236801)

[圖 2 StructureBody架構分解圖 12](#_Toc332236802)

[圖 3 研究流程架構圖 21](#_Toc332236803)

[圖 4 文件分析子系統架構圖 24](#_Toc332236804)

[圖 5 資料介面對應子系統 24](#_Toc332236805)

[圖 6 資料管理子系統架構圖 25](#_Toc332236806)

[圖 7 系統架構圖 26](#_Toc332236807)

[圖 8 含有ICD-9-CM的區段範例 27](#_Toc332236808)

[圖 9 文件分析流程圖 28](#_Toc332236809)

[圖 10 資料庫擷取內容 29](#_Toc332236810)

[圖 11 代碼一對一情況 29](#_Toc332236811)

[圖 12 代碼對應近似情況 30](#_Toc332236812)

[圖 13 代碼無對應情況 30](#_Toc332236813)

[圖 14代碼對應組合情況 31](#_Toc332236814)

[圖 15 轉換代碼流程圖 32](#_Toc332236815)

# 表目錄

[表 1 ICD-9-CM對應到ICD-10-CM一對多情況範例 16](#_Toc332236816)

[表 2 ICD-9-CM對應到ICD-10-CM多對一情況範例 17](#_Toc332236817)

[表 3 ICD-9-CM對應到ICD-10-CM無對應情況範例 17](#_Toc332236818)

[表 4 對應情況範圍值列表 29](#_Toc332236819)

# 緒論

## 研究背景

醫療資訊在近年來以改善病患照護品質及醫院間病歷互通整合為重要的發展指標，國內在電子病歷的發展推動上亦投入相當龐大的資金，其目的是為了加速全國醫院實施電子病歷資料交換，行政院衛生署在近五年來推動了「國民健康資訊建設計畫NHIP」，透過醫護人員、資訊相關技術人員與政府的合作，應用國際間資料交換標準與規範來制定電子病歷標準單張所需要的通用格式及醫學詞彙代碼，加速醫療院所實施電子病歷系統以提升醫療照護品質及病人安全，是推動電子病歷交換的重要基礎。

在臨床術語編碼系統中，利用字母與數字等方式來描述病人疾病或健康照護問題，並依照各種不同的用途發展各式各樣的編碼系統，其目的是為了解決醫療人員的溝通與資料記錄儲存等問題。臨床編碼系統包含：診斷編碼系統、檢驗檢查編碼系統、藥劑編碼系統、處置編碼系統等。每種臨床術語編碼系統或是不同版次的編碼系統針對同一種疾病的敘述方式彼此會有些差異，因此這些差異會造成在不同系統或版本在對應上的困難。近年來學者也提出許多方法如；通用術語服務(Common Terminology Services, CTS)、統一醫學語言(Unified Medical Language System, UMLS)及各種本體論(Ontology)以處理臨床代碼轉換之問題。

然而在電子病歷交換中包括了許多醫學術語與詞彙代碼的使用，詞彙代碼可以明確表示疾病或進行流行病學上的統計研究分析。近年來，在行政院衛生署所公告的電子病歷交換標準規範書中，制訂了醫院進行資料交換所需使用到的詞彙代碼與欄位格式，透過統一的標準規範來進行電子病歷資料交換，然而在各醫院間都有各自的醫療資訊系統與代碼系統，因此需要有統一的標準代碼，才能達到資料交換時內容的一致性。

## 研究動機

衛生署所公告的電子病歷交換標準規範書中雖然定義了完整的交換互通格式與代碼，但各醫院所使用的編碼系統不完全相同的情況下，會造成內容交換時無法共通，即使產生出標準的電子病歷交換文件進行資訊交換，也無法完全保證接收者是否能了解這些資料內容，使醫院資訊無法達到真正的分享與交換。

目前存在的醫療編碼系統種類繁多，且每一種編碼系統都有不同的用途與目的，因此在進行評估時，會因為不同編碼系統之性質而有所差異，無法互相比較，且編碼系統的代碼增修會隨著時間不斷的更新，在這樣的情況下，收集完整的詞彙代碼來處理不同的電子病歷交換文件有其必要性。

為了讓醫院在進行電子病歷交換文件時的代碼與名稱可以一致，本研究欲參考通用術語服務(CTS)來設計一套電子病歷詞彙代碼系統平台，希望可以協助編輯人員進行電子病歷文件製作時，作為參考使用。

本研究將以行政院衛生署公告的電子病歷交換標準文件為例，收集電子病歷交換文件使用到的詞彙代碼與欄位規範的系統設計，實作並建置一套電子病歷之查詢與編碼系統平台。

## 研究目的

本研究目的在於建立一套以電子病歷為基礎的詞彙代碼編碼系統平台，透過統一的詞彙代碼資料庫與編碼系統，可以協助使用者在編輯CDA文件時候參考運用，系統平台提供查詢代碼與選取的方式，便可產生文件區段編碼。本研究之具體目的如下列所述：

1. 建置一個醫學詞彙代碼編碼系統平台，讓使用者在編輯CDA文件的過程中可以參考運用，進一步促進資料交換。
2. 系統收集電子病歷交換標準書中所需要使用到的詞彙代碼，以建置出一套符合國內電子病歷交換標準之詞彙代碼資料庫。
3. 系統自動分析電子病歷的欄位與代碼對應關係，讓使用者可直接在平台上透過查詢與選取代碼的方式產生CDA文件區段編碼，達到電子病歷交換時的代碼與名稱一致性。
4. 參考HL7 CTS及行政院衛生署所公告知電子病歷交換標準規範書為例來實作本研究之編碼產生模型，進行建置系統之雛型。

# 文獻探討

本章節依照研究目的首先探討電子病歷-臨床文件架構(Clinical Document Architecture, CDA)、各項電子病歷醫學詞彙代碼使用，由於本研究收集各項電子病歷所需使用的醫學詞彙代碼，因此會探討通用術語服務(Common Terminology Services, CTS)等相關代碼對應研究，來進行詞彙代碼系統開發時所面臨的問題。

## 電子病歷

近幾年政府大力的推動電子病歷，使得電子病歷在台灣更加普及化，而在電子病歷的問題焦點，也隨著發展階段而有所不同，逐漸由資料資訊化轉變為院際間資料的交換與共享。

電子病歷是一種包含過去、現在或未來病人的生理、心理及社會的紀錄，以電子化的方式來處理，主要用於協助醫療或其相關服務[3]。病歷的製作依照醫療法第67條「病歷應包含醫師依醫師法執行業務所製作之病歷、各項檢查、檢驗報告資料及其他各類醫事人員執行業務所製作之紀錄。」，而電子病歷則是將紙本的病歷電子化，並依照醫療機構電子病歷製作及管理辦法來建置及運作。簡單來說，電子病歷依據簡文山等人整理應包含下述四點：(1)有關收集儲存個人終生的健康資訊總集合；(2)存取或提調個人健康或公共衛生資料的管理機制；(3)提供醫學知識及支援臨床決策資訊管道，以促進病人安全及增進醫療照護效率和提昇醫療品質；(4)有效支援整體健康照護體 [4]。

在台灣電子病歷發展的現況，可以由行政院衛生署於2005年12月所發表的「醫療院所病歷電子化現況調查」中了解，根據資料顯示，有50%的醫院病歷已進行電腦化，約有30%的醫院病歷電子化已開始院內整合並漸漸邁向院際分享及交換之應用 [5]。院際交流的需求使得電子病歷的標準格式逐漸成為一個重要的議題，另一個使電子病歷標準格式成為關注的因素是移動病患的問題。所謂的移動病患指的是，某些族群因為工作的關係需要在世界各地來來往往，若健康出現狀況，語言能力又有所限制，導致無法確切交待自己的身體狀況，電子病歷的標準化就可以讓醫療提供者在病患的授權之下容易的取得其資訊，避免錯誤的處置及不必要的風險及成本 [6]。

## 醫療資訊標準

資訊的標準有許多不同的面向，在醫療資訊的領域中，學者Blair將標準分為：識別標準(Identifier)、訊息格式標準(Message Format)、內容結構標準(Content and Structure)、臨床資料編碼標準(Clinical Data Representations)、機密安全與認證(Confidentiality, Data Security, and Authentications)等等 [7]。資料交換標準可以用來與不同機構的醫師共享電子的出院文件及產生其他的臨床資料在電子病歷(EHR)中，也可以將臨床資料傳送到其中 [8]。發展標準的目的是，在不同的資訊系統提供一個可以互通且安全的環境進行醫療資訊的交換 [9]。

### 臨床文件架構

HL7臨床文件架構 (Clinical Document Architecture, CDA)是一個用於臨床文件上的標準可延伸標記式語言(eXtensible Markup Language, XML)格式，且是根據HL7第三版參考資訊模型(Reference Information Model, RIM)、資料型別及詞彙等標準。CDA標準也被國際標準組織(International Organization for Standardization, ISO)及美國國家標準協會(American National Standards Institute, ANSI)認可 [10]。

在早期，CDA稱為病人紀錄架構(Patient Record Architecture, PRA)，而CDA R1在1991年時被HL7會員所認同且在2000年被ANSI認可，它是根據初期HL7 RIM草案、資料型別及詞彙而來。CDA R2在2005年1月被HL7所認可，隔年成為ANSI的標準。HL7的文件結構化工作小組(Structured Documents Workgroup, SDWG)也開始發展CDA R3.0，於2011年年初已進行了計畫初步的投票 [10]。

臨床文件架構(CDA)是以交換為目的，並定義其結構及語意的文件標記規範，不對文件內容有所定義，且文件交換的方法、文件的產生及管理都不在此標準規範之內。其內容主要是使用XML編碼，此編碼是自HL7 v3的參考資訊模型(RIM)及資料型態(Data Type)所延伸 [11]。

CDA的組成分為兩個部分，Header及Body。其組成結構如圖 1。Header主要功能是用來識別文件的分類，且也同時提供認證、活動、病人與參與醫療服務提供者的資訊；Body則是屬於臨床報告的部分 [12]。

<ClinicalDocument>

<structuredBody>

</ClinicalDocument>

</structuredBody>

Header

Body

圖 1 CDA文件結構示意圖

* Header

CDA Header的主要目的是為了使得臨床文件可以在機構間交換，且能對管理臨床文件有幫助，並協助編譯病人的臨床文件，使其成為終身的電子健康紀錄。Header內容主要可以分為三類 [12]：

1. Header Attributes：定義文件識別之資訊、版本、時間等等。
2. Header Participant：對文件的參與者定義了不同的角色，共九種：產生文件(author)、管理(custodian)、描述對象(病患)(recordTarget)、文件確認(authenticator)、文件最後確認者(legalAuthenticator)、文件接收者(informationRecipient)、輸入資料者(dataEnterer)、可提供描述對象相關資訊者(informant)、文件其他參與者(participant) 。
3. Header Relationships：定義病患在醫療過程中的其他資料，共五個部分：相關文件(relatedDocument)、產生此文件之醫令(inFullmentOf)、所要執行之項目(documentationOf)、臨床資料(componentOf)、相關之同意書(Authorization)。

* Body

Body的部分含有各種醫療行為之紀錄，有兩種資料型態，一種是非結構化區塊(NonXMLBody)，另一種是結構化區塊(structuredBody) [11]：

1. NonXMLBody：非結構化的資料，不是XML格式的資料，會將資料直接存放至<text>標籤中，或其他非文字之檔案。
2. StructureBody：結構化標記的資料，由一個或多個組件(section component)所組成，且可接受巢狀之組合。透過section的方式來描述臨床內容，如檢驗結果、診斷內容等等，且Section間不會互相影響。在Section中對於text加以描述的為Entry-level(如服藥方式、劑量等等)，在此level中需要用到各種通用代碼，如健保碼、國際疾病編碼等等，詳細結構如圖 2所示。

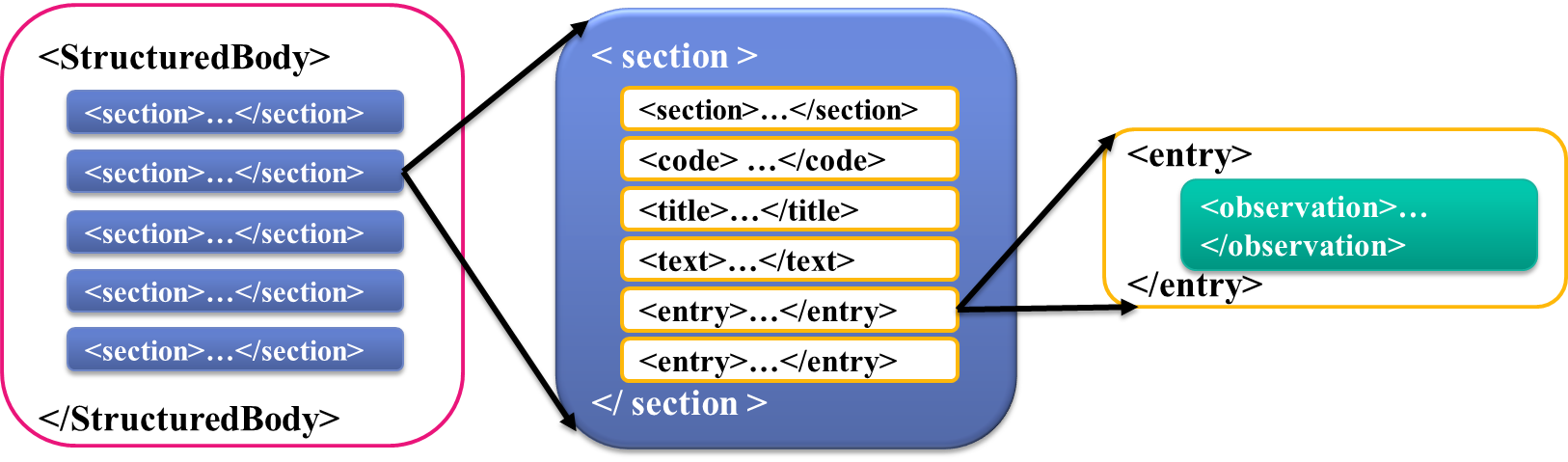


圖 2 StructureBody架構分解圖

以CDA為基礎所進行的研究像是：運用以CDA格式發展出來的連續性照護文件(Continuity of Care Document, CCD) 來對應不同來源系統的用藥史，目的是為了用藥的一致性，並運用CDA互通性(interoperable)的特性，來交換用藥資訊，解決未來用藥上及藥物過敏等等的問題，以做為決策支援[13]。基於CDA所規範的標準格式來改善醫療資訊交換系統，使病人資訊更能有效的被共享於院際之間，減少病人在醫療上的風險，提高管理疾病的能力並減輕醫療人員與病患之間的不對稱性 [14]。

### 醫療編碼系統

使用編碼的立意在於一個詞彙可以有很多不同的意義及概念，而相同的事件可以有很多不同的表示方法，雖然人可以經由知識及經驗來辨別，但電腦卻當成不同的事物來作業，因此利用編碼的方式來統合統一件事的表示方式。在醫療上的編碼有多用途如：發病率與死亡率是組成健康資訊系統(Health Information System, HIS)重要的一部份，他們的編碼可以統一資料的收集及分析，且可以在區域及國家間進行有意義的比較 [15]。醫療編碼(Medical Coding)是將疾病、傷害及醫療過程的敘事性描述轉換成字母或數字也就是代碼。代碼詳盡的用以精確地描述診斷，但醫學並不是一門精確的科學，代碼的發展是為了尋找醫療保健的所有原因。醫療編碼的使用通常包含 [16]：

1. 確認必須進行評估及提醒其他醫療照護人員有危及生命的過敏症狀
2. 向支付醫療費用支付者申報所執行之醫療服務
3. 協助行政工作像是人員編制與調度、及增加或減少醫療服務
4. 比較不同區域所執行的新計劃或設施

#### ICD-9-CM

1975年WHO於日內瓦召開修訂ICD-9的國際研討會，除了基本的分類更新，為了使國際疾病分類能更有彈性且可在不同情況下使用，ICD-9中加入了社會、健康、文化及環境影響健康的因子、精神疾病及周產期死亡等分類。以統計為目的的ICD-9共有兩冊，包括疾病字母索引、系統列表，其編碼擴充到了第四碼，為了更詳盡敘述疾病狀態。1977年到1980年，美國國家衛生統計中心(National Center for Health Statistic; NCHS)逐漸發展出ICD-9臨床應用的修訂版，就是ICD-9-CM(Clinical modification)，強調更貼切臨床的實際應用，使評估醫療作業更具比較性，且利於分析醫療費用，也較適合進行流行病與臨床之研究。其共有三冊，第一冊為疾病代碼列表，第二冊為疾病字母索引，第三冊為手術與處置代碼列表及字母索引 [17,18]。

ICD-9-CM在健保費用申報上佔有重要的地位，1981年引進後，於1992年勞保局規定以ICD-9-CM系統來申報住院診療費用，2000年1月起西醫全面採用ICD-9-CM來進行申報的動作，而中醫方面則是在2002年5月全面採用 [19]，目前台灣所施行的TW-DRGs也是以ICD-9-CM為基礎。

#### ICD-10-CM/PCS

ICD-10經WHO第43屆衛生大會認可，於1994年用42種語言公告採行。1992年WHO出版4位碼的ICD-10，其對死因統計是相當足夠的，但對於醫院找尋個案，特別是醫療保險申報或醫療資源使用分析，不夠詳細，因此許多國家對ICD-10進行修訂，對於某些診斷碼擴充到更細項的5位碼或6位碼 [20]，如：澳洲所修訂的ICD-10-AM、加拿大所修訂的ICD-10-CA、德國所修訂的ICD-10-GM及美國所修訂的ICD-10-CM等等 [21]。

ICD-10-CM是美國的臨床修訂版，於2007年7月公告，2009年1月美國聯邦政府通過法案，規定於2013年10月1日起，所有醫院住院申報皆得使用ICD-10-CM及ICD-10-PCS [22]。ICD-10-CM診斷碼是由美國國家衛生統計中心(National Center for Health Statistics; NCHS)來開發及維護，而ICD-10-PCS處置碼則由美國老人醫療保險與失業貧窮救助服務局(Centers for Medicare and Medicaid Services; CMS)類似台灣的中央健保局，來開發及維護，此兩單位每年都會對於所維護的編碼進行更新 [23]。 WHO所出版的ICD-10包括三冊，第一冊是註碼表清單(tabular list of code)，第二冊為選擇死因規則與統計列表指引，第三冊則是字母索引(alphabetic index)。美國所修訂的ICD-10-CM只有兩冊，一冊為註碼表清單，另一冊為字母索引 [20]。

ICD-10-PCS是由CMS的前身HCFA(Healthcare Financing Administration)於1995年委託3M公司來發展，針對ICD-9-CM第三冊處置碼的缺失進行改善，研發出更好的處置分類系統(procedure Coding System; PCS)。

發展ICD-10-PCS有四個主要的目標 [24]：

1. 完整性(completeness)：基本上來說，不同的編碼應該要有一個獨立的編碼
2. 可擴展性(expandability)：當新的處置產生時，ICD-10-PCS的結構要使其容易增加，變成一個獨立的編碼
3. 多軸向性(multiracial)：每一個處置碼的個別位碼都有其獨立的意義存在
4. 標準化術語(standardized terminology)：ICD-10-PCS須包括術語使用的定義，避免同一處置有多重術語或一種術語涉及多個處置

## 代碼對應之問題

猶如研究動機所述，每一種醫療編碼系統都有其發展的目的及其用途，因此對於同一件事的敘述方式就會不同，像是用於保險支付及作為統計分析其編碼的性質與立意就不相同，即便是同一性質之編碼系統也會因為版本更新而同一件事情產生不一樣的敘述。因此在代碼的對應上就會產生部分編碼無法精確對應的狀況發生。

以ICD-9-CM與ICD-10-CM/PCS的對應為例，ICD-10-CM/PCS較ICD-9-CM多增加了如：側性、細分類、代碼重新歸類及標準化醫學字詞等等之修訂，此些因素皆造成對應上之問題，如對應上產生一對多、多對一或是無對應的狀況如表 1、表 2及表 3所示 [25, 26]。

表 1 ICD-9-CM對應到ICD-10-CM一對多情況範例

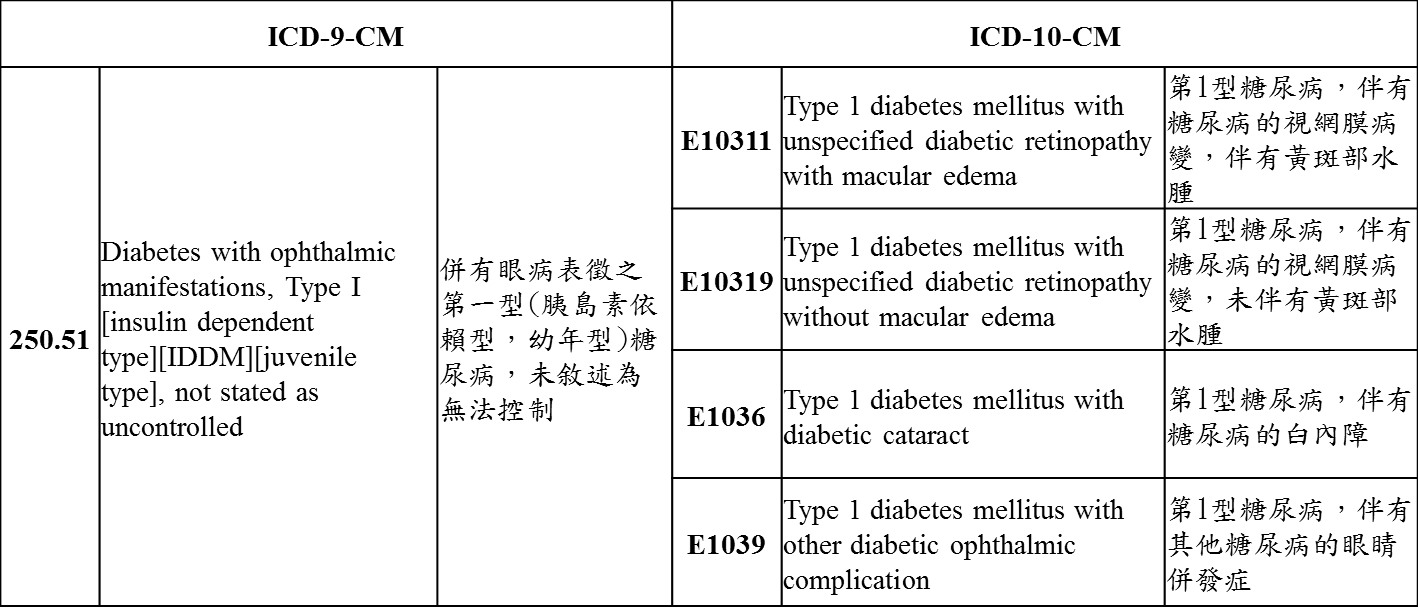


表 2 ICD-9-CM對應到ICD-10-CM多對一情況範例

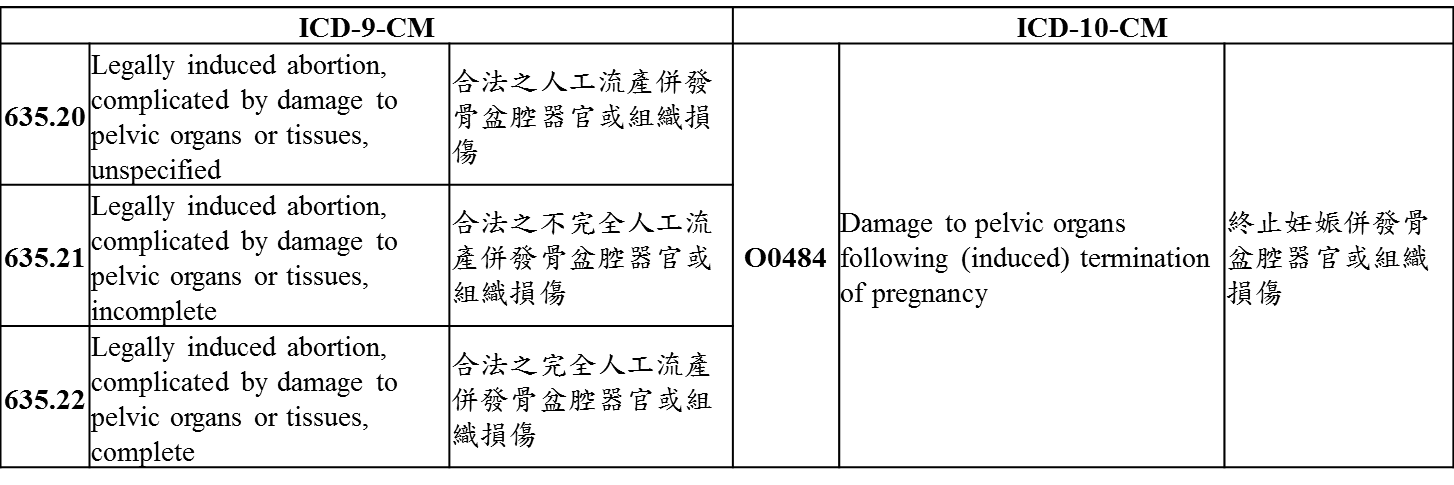
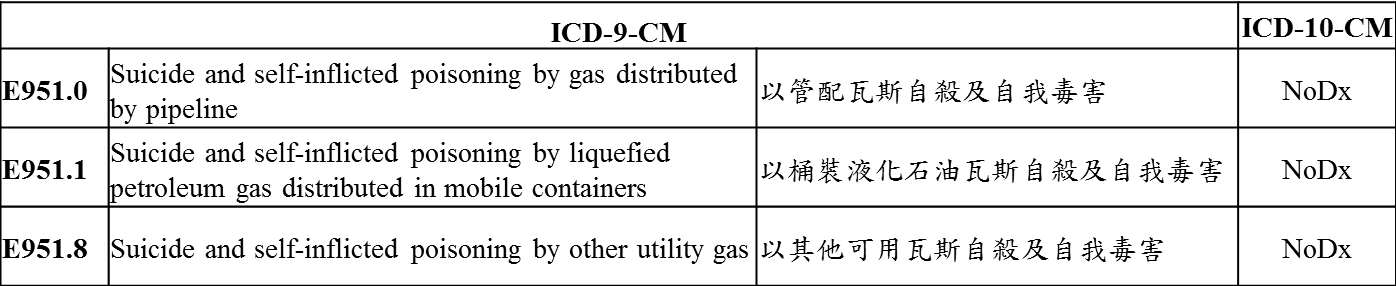


表 3 ICD-9-CM對應到ICD-10-CM無對應情況範例



對於建立對應關係在異質的編碼系統一直以來被視為是一項具有挑戰性且困難的任務，近年來也有相關的研究在尋找較好的解決方法。Fung, Bodenreider等學者利用統一醫學語言系統(Unified Medical Language System, UMLS)中的語義知識(semantic knowledge)為基礎發展對應演算法來對應兩個不同的分類及術語系統 [27]； Elkina, Brown等學者也是利用類似的方法來進行對應的研究，將ICD-9-CM代碼中隱含之語義製成對應表，方便ICD-9-CM代碼與其他代碼間進行對應 [28]；Dolin et al.學者應用了以語義為基礎的描述邏輯的LALR(lexically assign, logically refine)策略，合併兩個定義關係相同的重疊術語 [29]；Cardillo et al.等學者利用Web本體語言(Web Ontology Language, WOL)邏輯分析相對應的代碼，來完成兩代碼系統之間的對應 [30]。

## XML

可擴展標記語言(eXtensible Markup Language, XML)，描述資料物件的類型及部分描述了資料物件的程式設計行為。XML是一個應用於描繪文件輪廓或根據標準通用標記語言(Standard Generalized Markup Language, SGML)限制的形式。根據Bray et al. 研究整理，XML的設計目標是 [31]：

1. XML應該要直接可用於網際網路上
2. XML應該可支援使用於各種應用上
3. XML應能與SGML相容
4. 處理XML文件應該易於寫入程序
5. 在XML文件中可選擇的功能應減至最少，最好為無
6. XML文件應該要是人可閱讀的且明確理解的
7. XML設計應可迅速編寫
8. XML文件設計應該要正式且簡潔的
9. XML文件應該要易於建立
10. 在XML中簡化的標記是不重要的

XML的優勢是在設計上擁有極大的彈性及以可延伸階層式結構化型態來描述資料。XML的標記主要是針對資料的描述，且可根據使用者之需求來自行定義標記之名稱及屬性，這也被廣泛使用在電子商務中，許多資訊廠商如Adobe、IBM、Microsoft、Sun、Oracle、Netscape等等也採用了此項技術，影響各行各業都有XML的相關應用 [32]，如用於醫療資訊交換電子病歷的設計所使用的CDA就是採用XML的技術，並符合HL7標準之需求，來處理臨床上的資料，使其經過標記後利於資訊的轉換。

# 

# 研究方法

本章節依序說明研究步驟、系統需求分析與設計、相關技術應用及開發環境與工具。

## 研究步驟

本研究主要是設計一個以電子病歷為基礎架構的詞彙代碼查詢與CDA編碼製作為主，將參考衛生署電子病歷交換中心所公告的電子病歷交換標準書，並實作此平台。研究首先對於標準書所需要的欄位代碼進行資料收集及詞彙代碼對應的關係進行了解，並探討電子病歷之架構及相關電子病歷之研究，最後以實作的方式建立此系統平台完成本研究之計畫。研究流程從研究背景、動機及目的、文獻探討、研究方法與預期成果如圖 3所示。

### 研究流程



圖 3 研究流程架構圖

### 系統建置流程

#### 研究電子病歷架構

電子病歷是以HL7 CDA R2為基礎，並使用XML之技術來描述整份文件，需了解CDA文件各項欄位中的詞彙代碼與欄位資料的對應關係。

#### 研究電子病歷交換標準書欄位及格式標準規範

本研究依據行政院衛生署電子病歷推動專區(<http://emr.doh.gov.tw/allowance100.aspx>)公告之電子病歷單張交換標準書作為參考範本，並分析文件架構中所制定的欄位與格式規範需存放的資料內容。

#### 研究醫學詞彙代碼

本研究以電子病歷交換標準書所使用到的詞彙代碼為例，探討每個標準書中的欄位與代碼對應關係，發現標準書的欄位中所使用的代碼都有對應的詞彙關係，因此在系統設計上需要有一套標準的詞彙代碼，並分析詞彙代碼所屬欄位之情況。

#### 建置詞彙代碼資料庫

收集電子病歷交換中需要使用到的詞彙代碼，並分析代碼與欄位之對應關係，將需要對應的代碼分類建立於資料庫中，以方便使用，但目前國內並未針對詞彙進行統一的定義，造成在語義上混淆與誤解，因此在部分詞彙代碼參考HL7所定義的詞彙代碼標準，建置一套符合國內使用的詞彙標準資料庫。

#### 系統需求及系統設計

本研究之系統依使用者可區分為標準書欄位分析、資料介面對應、文件編碼產生及資料庫管理等四項需求，再依照前述所提之電子病歷單張標準格式參考範例中，分析研究每張標準書所需要的欄位，進而完成本研究系統之設計。

#### 子系統開發

本系統包含三個子系統，資料介面對應、文件編碼產生及資料庫管理等，將逐步完成各部份子系統功能之研究設計。

#### 系統整合與測試

完成各項子系統建置後，進行整合與測試，進行代碼與標準書的欄位對應情況是否符合文件中所制定的格式規範。

#### 研究成果撰寫

本研究將撰寫系統平台建置的操作說明與結果，及討論實作過程中所遇到的挑戰及其研究之限制，提供研究者後續參考。

## 系統需求分析與設計

本章節先進行系統需求分析及系統架構，再依其規劃系統之功能，以建置出符合研究之系統。

### 需求分析與系統架構

本研究在建置一套電子病歷詞彙代碼系統平台，系統將電子病歷交換標準規範所需的欄位與醫學詞彙代碼進行整理與對應關係建立，讓使用者可以了解各欄位需要哪些詞彙代碼並提供查詢功能，主要是為了促進醫療資訊的一致性。本系統另外含有詞彙代碼的區段編碼產生，使用者可以透過系統詞彙代碼查詢與選取的方式，系統會把使用者選取的代碼產生區段編碼後顯示於畫面上，讓使用者了解電子病歷文件內容的關係。

本研究系統包括單張欄位對應、醫學詞彙代碼查詢、區段代碼產生及資料庫管理等四個子系統，將依序分析之。

#### 單張欄位對應子系統

單張對應子系統主要功能為，將電子病歷文件的資料欄位與詞彙資料庫進行對應關係建立，由於各表單內容因為文件的不同有所差異，且每個欄位對應的詞彙代碼也不盡相同，透過子系統可以得到單張文件對應的欄位與詞彙代碼關係，其架構圖如?所示。



圖 4 單張欄位對應子系統

#### 醫學詞彙代碼查詢子系統

醫學詞彙代碼查詢子系統主要是提供使用者查詢電子病歷文件使用到的詞彙名稱與代碼。由於行政院衛生署所公告的四類電子病歷交換標準書、中醫、西醫、牙醫門診單及電子健康照護等單張都有相對應的醫學詞彙代碼，因此本研究需要將收集到的詞彙代碼與電子病歷單張進行對應關係，讓使用者可以快速選取到符合電子病歷單張所需要的辭彙代碼類別，其架構圖如圖 5所示。



圖 5 醫學詞彙代碼查詢子系統

#### 區段代碼產生子系統

區段代碼產生子系統主要是將從電子病歷欄位與詞彙代碼對應關係產生該區段編碼，使用者透過電子病歷單張的欄位與詞彙代碼的選取可以產生編碼提供使用者參考，因應各電子病歷的欄位與代碼不同的對應關係，在進入本子系統前會先將上述兩子系統進行處理，以達到研究之目標。

#### 資料庫管理子系統

資料庫管理子系統主要是產生一個介面使管理者可因應不同單張欄位與詞彙代碼系統的更動而進行管理的動作。其子系統架構圖如圖 6所示，在該子系統中可進行查詢、新增、修改及刪除資料庫等動作。



圖 6 資料管理子系統架構圖

圖 7為根據上述需求所建置詞彙代碼系統之架構圖。使用者透過系統平台進入詞彙代碼查詢子系統後，進行標準單張及詞彙代碼的選取，因此在平台中會先到單張欄位對應子系統進行欄位與詞彙代碼對應處理後，回傳對應的詞彙代碼類別提供使用者選取，接著在區段代碼產生子系統會顯示該欄位需要輸入的資訊與代碼確認以進行編碼產生，將區段編碼資料呈現結果給使用者了解，提供使用者編輯電子病歷文件時參考使用。



**使用者**

**管理者**



**代碼資料庫**



**醫學詞彙代碼查詢子系統**

**單張欄位對應子系統**

分析電子病歷單張欄位

選取符合詞彙代碼類別

**區段代碼產生子系統**

標準規範單張選取

欄位、詞彙代碼選取

**比對**

**對應**

**瀏覽結果**

輸入單張欄位資料

詞彙代碼確認

**維護**

圖 7 系統架構圖

### 系統分析與設計

由圖 7可知，本系統主要的功能為詞彙代碼查詢與編碼產生的平台，其工作分為電子病歷文件分析，分析後的資訊可以提供使用者找到電子病歷欄位與詞彙代碼的對應關係及代碼資料庫的管理等等。此章節將探討單張欄位對應、醫學詞彙代碼查詢、區段代碼產生及資料庫管理等子系統的分析與設計。

#### 單張欄位對應子系統

本系統首先將行政院衛生署公告的四類電子病歷交換標準單張包括醫療影像報告、血液檢驗報告、門診用藥及出院病歷摘要，門診層級單張包括西醫、中醫及牙醫門診單，電子健康記錄摘要單張包括生理量測資料、收案評估、居家訪視、服務轉介、會員轉案及藥事安全等單張進行文件欄位分析，判斷該欄位對應的詞彙代碼類別有哪些，在將對應的關係建置於系統平台中，讓使用者選取單張時就可以取得該欄位的詞彙代碼類別的名稱以進行資料查詢與應用，圖 8為電子病歷區段欄位使用到LOINC及ICD9CM辭彙代碼的範例。

|  |
| --- |
| <!--診斷-->  <component contextConductionInd="true" typeCode="COMP">  <section classCode="DOCSECT" moodCode="EVN">  <code code="29308-4" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="Diagnosis"/>  <title>診斷</title>  <entry>  <observation classCode="CNOD" moodCode="EVN">  <code code="003.0" codeSystem="2.16.840.1.113883.6.2" codeSystemName="ICD9CM" displayName="Salmonella gastroenteritis"/>  </observation>  </entry>  </section>  </component> |

圖 8 電子病歷區段欄位使用到LOINC及ICD9CM辭彙代碼的範例

由於在每張電子病歷交換單張的欄位所使用的詞彙代碼不盡相同，因此本研究需要分析欄位與詞彙代碼對應關係建置於系統平台中如圖 9所示。

對應關係的研究主要是將單張欄位具有CodeSystem的屬性名稱時，表示此欄位需要具有該詞彙代碼類別的資訊，因此本研究將欄位名稱對應到的詞彙代碼類別建立關聯資料庫，來完成單張欄位對應子系統的功能，在對應的過程中欄位與詞彙代碼類別的關係可能是一對一、一對多或沒有對應的狀況，在資料庫中都需要建置完整的位應關係，協助使用者在編輯欄位時可以快速了解詞彙代碼類別的使用。



圖 9 單張欄位對應內容

#### 詞彙代碼查詢子系統

由文件分析子系分析可知，所匯入之電子病歷為何單張，此子系統會依據符合文件分析子系統之單張，將其格式匯入至所對應之介面欄位。從技術的角度來說明，就是讓CDA格式的電子病歷單張經由判斷的方式來對應到網頁介面欄位上，讓使用者可清楚了解電子病歷內所含之內容資訊，使電子病歷的內容在呈現上較具人性化。

#### 區段代碼產生子系統

以ICD-9-CM轉換至ICD-10-MC/PCS為例，將含有ICD-9-CM屬性值的標籤中的代碼(code)屬性取出，如圖9綠色線條圈選之部分，帶入對應資料庫進行比對，以完成轉換工作。在進行代碼轉換之作業時，會遇到數種無法精確對應的狀況，因此在對應資料表的設計便加上了對應情況的標記，方便轉換作業進行。

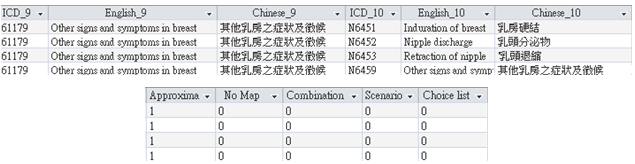


圖 10 單張欄位對應資料庫內容

在圖11中可以看到在兩代碼系統對應的資料欄位後方，有5個對應情況的資料欄位，其範圍值如表 4。

表 4 對應情況範圍值列表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 近似  Approximate | 無對應  No Map | 組合  Combination | 細項情況  Scenario | 代碼選項  Choice List |
| 0/1 | 0/1 | 0/1 | 1-9 | 1-9 |

1. 近似(Approximate)：標記為0時，代表兩代碼系統可以精確的對應，也就是說可以達到一對一的情況，如圖 11所示。當標記為1時表示ICD-9-CM對應到ICD-10-CM呈現一對多的狀態，如圖 12所示。



圖 11 代碼一對一情況



圖 12 代碼對應近似情況

1. 無對應(No Map)：標記為0時，表示有ICD-9-CM編碼系統可以有效的對應到ICD-10-CM編碼系統，而當標記為1時，表示ICD-9-CM編碼系統無法有效的對應到ICD-10-CM編碼系統，如圖 13所示。



圖 13 代碼無對應情況

1. 組合(Combination)：標記為0時，表示所對應到的代碼非並用之關係，若標記為1時，需要多個代碼才可充分表示來源代碼系統的對應。在組合的情況下會有細項情況(Scenario)及代碼選項(Choice List)兩種情況的標記。
2. 細項情況(Scenario)：當標記數值相同時代表代碼為同組的情況，可根據代碼選項(Choice List)來做組合，不同組的對應碼不可做交叉組合。
3. 代碼選項(Choice List)：標記數值為同組對應碼之順序，通常標記數值1為主診斷，數值2為次診斷，如圖 14所示。對應結果為”Q65.01為主診斷, Q65.32為次診斷”及”Q65.02為主診斷及Q65.31為次診斷”兩組。



圖 14代碼對應組合情況

圖 15為代碼轉換之流程是以ICD-9-CM對應到ICD-10-CM/PCS，分析過後會遇到的對應情況的發生來繪製。由文件分析子系統分析所匯入之文件後，便可得到ICD-9-CM的代碼值，本子系統會將其值帶入資料庫來做比對，並會經過五個判斷過程。首先會先判定所匯入的代碼值是否為無對應(No Map)，若為無對應的狀況，則取出ICD-9-CM代碼之對應字串，帶入至完整的ICD-10-CM/PCS來做比對，而為不是無對應的狀況則往下判斷該代碼是否為近似值(Approximate)。如不為近似值，則該代碼轉換到ICD-10-CM/PCS為精確的轉換，如為近似值則繼續判斷該碼是否有組合(Combination)之狀況。若不為組合碼之情形，則將所對應的到ICD-10-CM/PCS列表，讓使用者可選取與病歷資料狀況最為符合之狀況，若為組合碼之狀況則需進一步分析細項情況(Scenario)，若細項情況的值只出現1時，則判斷代碼選項(Choice List)中數值1及數值2之個數，若各為1則可直接列出數值1為主要診斷而數值2為次要診斷，若其各數多於1時，則會請使用者先選擇代碼選項為1的選項，再依選出之情況列出數值為2之選項供使用者選取。若細項情況的值出現的範圍超過1，則需請使用者先行選擇細項情況值，選擇出值後再重複執行判斷代碼選項之步驟。

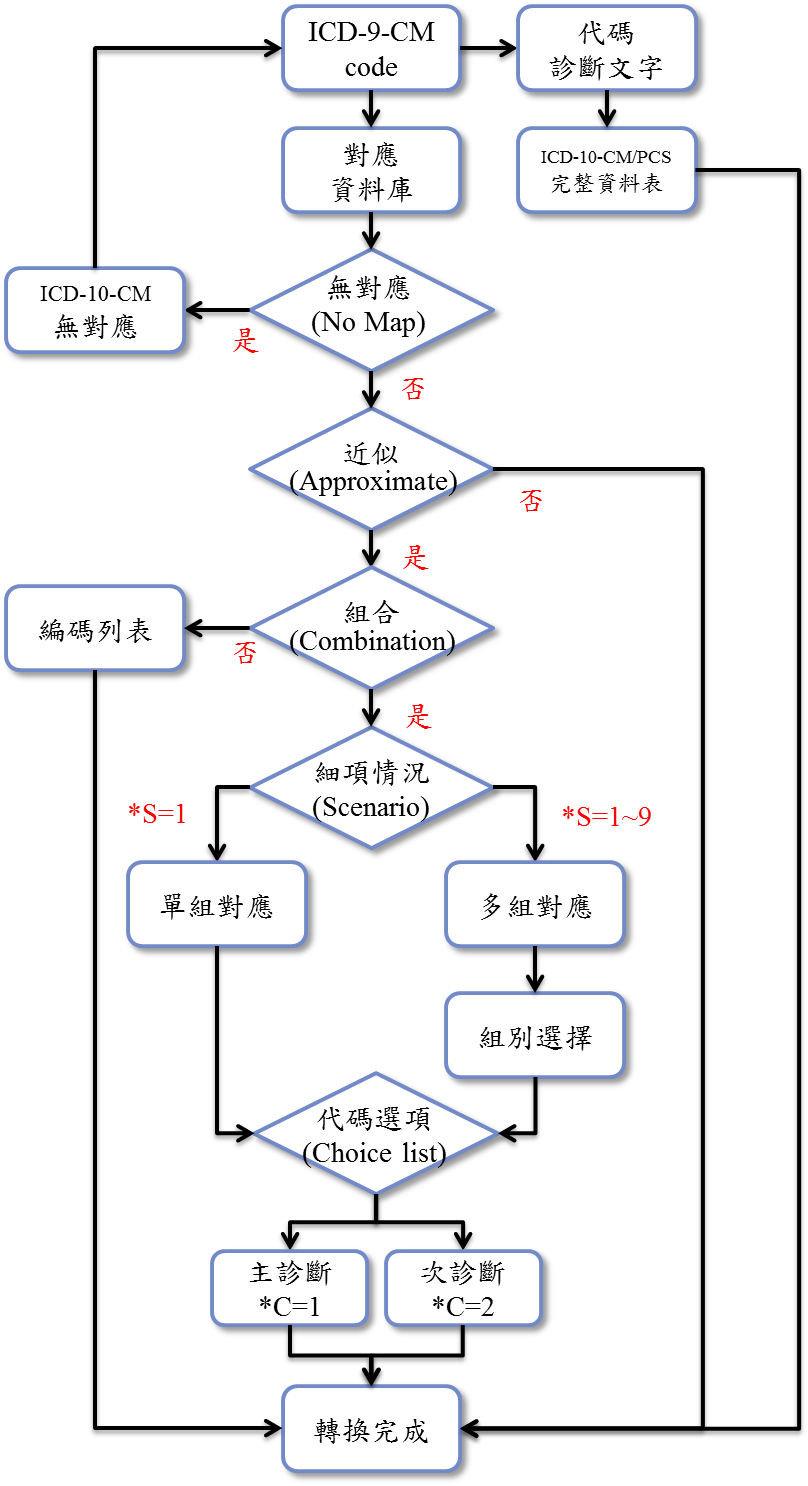


圖 15 轉換代碼流程圖

註：\*S為細項情況的係數，\*C為代碼選項係數

#### 資料庫管理子系統

本系統資料庫含有來源代碼系統及目標代碼系統資料表，資料表中會含有該代碼的分類、架構及其他細項狀況。為因應未來代碼可能因科技的進步而有所修編，因此設計此一管理子系統，使代碼資料庫可便於管理。系統中含有新增、修改及刪除等功能，同時含有查詢之功能，用以確認資料庫是否依功能變動成功。

## 相關技術及開發工具

本研究準備使用的開發環境、開發工具、開發程式語言及資料庫如下：

* 開發環境：Microsoft Window 7
* 開發工具：Microsoft Visual Studio 2008 Ultimate
* 開發語言：Visual Basic.NET
* 資料庫：Microsoft Access 2010

# 

# 預期成果

以目前行政院衛生署所公告的電子病歷交換標準規範書的角度來看，透過標準規範產逐步產生文件來進行醫院間的資訊交換，可預期在未來進行電子病歷交換及編輯CDA文件的需求會越來越廣泛。本研究將先收集到電子病歷標準規範所使用到的詞彙代碼與名稱進行資料庫建置，並以Visual Studio軟體進行系統平台開發，其中平台會針對電子病歷標準規範書中所制定了欄位進行資訊分析，並針對CDA文件所需要使用到的詞彙代碼進行查詢與選取的動作。

本研究預期將完成設計出一個電子病歷文件編碼平台，提供使用者可以針對電子病歷標準規範書中的欄位進行區段編碼產生。本研究以行政院衛生署公告的四類電子病歷單張交換標準、中醫、西醫、牙醫門診及電子病歷健康紀錄摘要為例，此系統預期可以將公告的四類電子病歷標準書進行CDA文件區段編碼產生，讓使用者只需要將標準書中的某一欄位透過詞彙代碼查詢與選取的方式，即可產生該CDA文件中的XML區段編碼，協助使用者在編輯CDA文件時使用，以達到電子病歷進行資料交換時的一致性，期望在建立此系統平台後，作為電子病歷文件進行編輯時的參考，對未來在電子病歷交換有所助益。

# 可能遭遇困難

* 醫學詞彙本地化：在各醫院使用的詞彙不盡相同，而國內並未統一標準，本研究僅對電子病歷交換標準所需使用到的詞彙進行收集，以及參考行政院衛生署計畫『HL7 V3 中文化』中詞彙翻譯的部分進行詞彙資料庫建置，因此系統平台針對未定義的詞彙以HL7 V3 Vocabulary 所使用的詞彙標準作為參考依據。
* 電子病歷交換標準單張：本研究參考行政院衛生署公告的四類單張格式、中醫、西醫、牙醫門診及電子病歷健康照護為例，將上述單張所使用到的欄位與代碼進行定義，在單張中的欄位有使用到詞彙部分都可以讓使用者進行編碼產生，為了讓使用者可以透過簡單的詞彙代碼查詢與選取的方式產生區段編碼，在每個單張的欄位資料編碼內容方式不盡相同，因此在系統平台上需要針對每一個單張欄位設計標準模組，才能提高使用者在編輯文件時的效率。

# 

# 參考文獻