

行政院國家科學委員會專題研究計畫 期末報告

運用關聯規則探勘技術建構台灣新移民女性醫療健康需求模型(GM03)

計畫類別：個別型
計畫編號：NSC 101-2629-E-040-001-
執行期間：101年08月01日至102年07月31日
執行單位：中山醫學大學醫學資訊學系

計畫主持人：曾明性
共同主持人：林靜儀
計畫參與人員：碩士班研究生-兼任助理人員：廖芳彬
碩士班研究生-兼任助理人員：沈瑜豐
博士班研究生-兼任助理人員：吳慧卿

公開資訊：本計畫涉及專利或其他智慧財產權，1年後可公開查詢

中華民國 102年10月31日

中文摘要： 台灣進入多元文化社會，政府及有關單位對於新移民權益(尤其新移民女性)重視，已看到許多政策努力推動。但是，醫療服務及輔導是否滿足新移民女性的處境與權益，需要更深入的剖析，以避免產生不合適的移民政策或是過度強調「同化」與「適應」的移民輔導服務。本研究希望透過資訊科技、臨床醫學及社會人文的跨領域合作，探討大陸及外籍配偶的健康權，並提出促進新移民女性健康促進方案的可能思考方向。

本論文為首次以內政部 92 年及 97 年大陸及外籍配偶生活調查之大型資料庫進行跨年度資料庫分析，融入「客戶導向」的觀念，藉由剖析新移民女性人口學及社經條件特性，建構滿足新移民女性差異化的醫療健康需求模型。有別於以往的研究文獻均僅呈現出新移民女性各項醫療健康需求的整體百分比，並無法確切指出各項醫療健康需求下不同族群特性為何？為深入剖析各項醫療健康需求下真正有需求的新移民女性族群之人口特性及社經條件，本研究運用關聯規則探勘 (Association Rule Mining) 技術分析全國性調查的大型資料庫，找出更精確的人口屬性、社經條件與醫療健康需求之關聯性；最後，結合臨床醫學與社會學的觀點，提出「分眾化」的性別化的健康照護系統，及加強社會支持網絡團體與醫療機構整合合作，從協助新移民女性的「賦能」出發，建構促進新移民女性健康的整合社會支持系統，作為政府及民間團體推動外籍與大陸配偶各項醫療保健服務措施、及政府制定相關服務政策之參考。

中文關鍵詞： 新移民女性、醫療健康需求、關聯規則探勘、客戶導向

英文摘要：

英文關鍵詞：

目錄

中文摘要	II
英文摘要	III
壹、前言	1
1.1 研究背景	1
1.2 研究目的	2
1.3 計畫重要性	3
貳、文獻探討	3
2.1 新移民女性的個人健康	4
2.2 新移民女性的家庭健康	4
2.3 新移民女性的健康照護資源	4
參、研究方法	5
3.1 Apriori 演算法	6
3.2 指標的計算公式與物理意義	7
3.2.1 支持度(Support)	7
3.2.2 信賴度(Confidence)	7
3.2.3 增益值(Lift)	7
3.3 改良型 Apriori 演算法(I)	7
3.4 改良型 Apriori 演算法(II)	10
肆、結果與討論	14
4.1 新移民女性醫療衛生需求-主要需求人次統計	14
4.2 改良型關聯規則演算法之效能探討	15
4.3 新移民女性醫療衛生需求-關聯規則模型探討	17
4.3.1 關聯規則模型一	17
4.3.2 關聯規則模型二	18
4.3.3 關聯規則模型三	19
4.3.4 關聯規則模型四	20
伍、性別影響評估	21
5.1 性別影響評估檢視清單	21
5.2 簡易性別相關參考依據	23
陸、結論	24
參考文獻	25
國科會補助專題研究計畫成果報告自評表	27

中文摘要

台灣進入多元文化社會，政府及有關單位對於新移民權益(尤其新移民女性)重視，已看到許多政策努力推動。但是，醫療服務及輔導是否滿足新移民女性的處境與權益，需要更深入的剖析，以避免產生不合適的移民政策或是過度強調「同化」與「適應」的移民輔導服務。本研究希望透過資訊科技、臨床醫學及社會人文的跨領域合作，探討大陸及外籍配偶的健康權，並提出促進新移民女性健康促進方案的可能思考方向。

本論文為首次以內政部 92 年及 97 年大陸及外籍配偶生活調查之大型資料庫進行跨年度資料庫分析，融入「客戶導向」的觀念，藉由剖析新移民女性人口學及社經條件特性，建構滿足新移民女性差異化的醫療健康需求模型。有別於以往的研究文獻均僅呈現出新移民女性各項醫療健康需求的整體百分比例，並無法確切指出各項醫療健康需求下不同族群特性為何？為深入剖析各項醫療健康需求下真正有需求的新移民女性族群之人口特性及社經條件，本研究運用關聯規則探勘(Association Rule Mining)技術分析全國性調查的大型資料庫，找出更精確的人口屬性、社經條件與醫療健康需求之關聯性；最後，結合臨床醫學與社會學的觀點，提出「分眾化」的性別化的健康照護系統，及加強社會支持網絡團體與醫療機構整合合作，從協助新移民女性的「賦能」出發，建構促進新移民女性健康的整合社會支持系統，作為政府及民間團體推動外籍與大陸配偶各項醫療保健服務措施、及政府制定相關服務政策之參考。

關鍵詞：新移民女性、醫療健康需求、關聯規則探勘、客戶導向

英文摘要

Taiwan has already entered the plural cultural society, government and the relevant units have concerned the new immigrant's rights and demands (especially new immigrant women) for a long time. Also, they have already planned a lot of policies and made great efforts to promote the new immigrants women living quality. However, what are the really medical care rights and interests for new immigrant women? It needs deeper analysis to discover immigrants demand truly, avoid excessively emphasizing 'assimilation' and 'accommodation' to produce inappropriate immigration health care policies.

The paper is a comparative analysis between two survey databases of Ministry of Interior, R.O.C. Proposing the idea incorporated 'the customer need' in the design of the health care system, the related analysis between demographic and social-economic characteristics of new immigrant women hopes to construct a differentiated medical treatment and health care service of every sector of new immigrant women. The improved 'association rule' is applied to prospect the nationwide large-scale database, which can find out more accurate relationships between specific population attribute and health care needs.

Finally, this study combined medical specialty and sociology to carry on the evaluation and analysis of relevant discussions. This paper hopes to provide suggestions of an integrated health care system of social networks and medical institutions for new immigrant women health care service policies.

Keywords: new immigrant women, health right, association rule

壹、前言

1.1 研究背景

性別平等是一種價值，而不是屬於特定人口社群才享有的福利，性別平等也不等於婦女福利的狹隘意義。性別主流化的政策目的要求政府全盤檢討勞動、福利、教育、環保、警政、醫療等政策，隱藏的性別不平等，性別主流化的分析就是企圖透過目標設定的反省、行動方案政策規劃評估，及性別統計等研究方法，打造一個符合性別正義的社會。「性別統計」對於婦女健康的意義為何？1975年第一次世界婦女大會於墨西哥市舉行，是首次提出婦女統計重要性的國際論壇。1985年於奈洛比(Nairobi)舉行第三次世界婦女大會，首次提出性別主流化(Gender Mainstreaming)的概念，會後發展策略共識是「發展中的婦女」(WID; Women in Development)，統計的焦點則從婦女統計(Statistics on Women)發展為兩性統計(Statistics on Women and Men)，正式將性別主流化議題納入整個統計體系。

1995年在北京舉行的第四次世界婦女大會對性別統計的發展方面邁出重要一步，《北京宣言行動綱領》戰略目標 H3 宣布各國政府同意採取一系列行動，來「產生及發布性別基礎資料和資訊，用以規劃及評核」。許多國家相繼成立性別統計單位，制定具體方案，將性別主流化議題納入整體官方資料的編製中。我國的性別統計工作也自 1999 年婦權會第五次決議開始進入由國家的層級進行性別分類統計資料的時代。

統計的性別主流化(Gender Mainstreaming of Statistics)是將所有資料蒐集、整理、分析，並將不同性別的不同影響因素納入考量。而統計的每一步驟皆須考量性別的影響，並考量社會與文化因素，以避免資料蒐集、分析與發布的性別偏見(Gender-based Biases)。所以性別主流化的統計原則就是將焦點放在「人」上面，進行跨項目、跨複分類與跨時空，透過此原則進一步婦女健康政策的規畫與推動。

全球資本主義市場化不僅表現在資本勞動行為，也造成國際移民行動，其中婚姻跨界遷移呈現的多元化社會樣貌，對於移入的接收國產生許多文化、社會的衝擊，所以跨國婚姻移民代表國家對「國民」的範疇重新定義，國家人口及社會政策也必須審慎思考新移民的定位並保障其基本權益，建立尊重包融的多元移民文化社會。

台灣自 1970 年代開始出現婚姻移民，1980 年代政府推動南向政策，促進東南亞外籍新娘的大量移入。從內政部戶政司「各縣市外籍配偶人數按國籍分與大陸（含港澳）配偶人數」統計資料可知，1987 年 1 月至 2011 年 12 月止，台灣婚姻移民移入人口已達 43 萬人。(內政部, 2011)新移民移入改變台灣社會人口結構，同時也突顯健康需求、生活適應、政治參與、文化衝擊、勞動就業、子女教養與學習等層面的特殊社群結構問題。

以結婚登記之外籍與大陸港澳配偶人數比例而言，2010 年結婚登記對數中雙方均為我國籍者為 14 萬 3,811 對(占當期總結婚對數 86.99%)。中外聯姻者 2 萬 1,516 對(占 13.01%，較 92 年之高峰減少 18.85 個百分點)，其中配偶為大陸港澳地區人民者計 1 萬 3,463 對(占 62.57%)，為東南亞國籍者 4,887 對(占 22.71%)，其他國籍者 3,166 對(占 14.71%)。依性別統計，外籍新娘 1 萬 7,426 人占 80.99%，遠高於外籍新郎之 4,090 人占 19.01%。(內政部, 2012)

以生母原屬國籍(地區)別區分，2010 年出生嬰兒之生母國籍(地區)為本國籍者有 18 萬 1,230 人(占 92.17%)，大陸港澳地區者 8,937 人(占 4.55%)，外國籍者 6,460 人(占 3.29%)；近 8 年來，每百名出生嬰兒中約維持有 5 名生母為大陸港澳地區配偶者，生母為外國籍者則由 93 年之 8 名降至本期之 3 名(內政部, 2012)。

近年來政府及民間團體已投入相當多時間及經費關心大陸及外籍配偶之生活適應問題，相關政策方案包括，外籍配偶生活適應、子女教養、優生保健、就業服務、人身安全等跨部會研議和規劃整合措施。但到了地方政府，由於財務狀況差距，提供的協助差異大，呈現資源供給不均的現象。現行戶籍制度規定大陸及外籍配偶若要得到社會福利給付，需具有中華民國國民身分證，東南亞外籍配偶至少需要四年才可取得，大陸配偶則需要更長的時間，約八年至十一年。除了少數例外情況，政府規定她們申請公民身份時，先生需要擔任保證人，使得大陸及外籍配偶缺乏自主權，處於社會中劣勢地位。面對外籍配偶仍將持續成為台灣婚姻的重要角色，跨國婚姻孕育出「新台灣之子」的比例逐年提高的趨勢，對外籍及大陸配偶為主體的福利政策規劃的審思是有重要意義的。

從族群融合及社會公平正義的觀點來看，新移民女性已成為台灣社會人口組成的重要生力軍，但在健康照護體系中，因「社群隔離」的因素，使其陷入社會弱勢，不論在醫療利用、公部門提供之生活照顧輔導、接受職業訓練課程者，皆呈現較本國女性為低的參與率。

台灣社會對大陸及外籍配偶「生殖化」與「疾病化」的刻板印象，雖然一方面鼓勵其努力生產，一方面卻也擔心大陸及外籍配偶身體健康不好，影響台灣下一代健康，政府單位在矛盾心理下，只好制定與外籍配偶生育及節育有關的相關衛生規定。當政府的健康照護制度設計是由上而下、強調生育為目的之「優生」觀點，對將外籍配偶的身體當成「生產工具」，忽略女性新移民真正的身心需求，雖然內政部已推動「外籍配偶生活適應輔導實施計畫」多年，但不論提供衛教手冊、單張或透過研修課程、網路傳播健康照護相關資訊，對新移民女性來說，體會到的是政府一再強調生育及幼兒保健重要性，對於婦女自我健康自主照護的知能促進卻十分薄弱。當原本以衛教為出發點的體貼，卻演變為制度的推動，相關的公共衛生管理制度儼然將她們置於受監控的環境，「生育」功能的發揮與否使大陸及外籍配偶陷入被接納或被排斥的矛盾情境中。因為經濟弱勢、文化邊緣化、社會階層化等因素使其被限制於特定地區生活，落入「非自願性的社群隔離」，除將使新移民女性無法就近有效利用健康資訊，也因社群隔離的結果造成自信心退縮，若再加上本地民眾對新移民不瞭解而降低互動支持的態度，將導致部份新移民女性，寧可捨近求遠請娘家寄藥到台灣，或者與其同屬移民原生國的親友交流自行療癒的作法，也不願意前往醫院看病。

本計畫的問題意識是：新移民女性的真正醫療健康需求是什麼？因此，本研究將透過內政部對新移民女性的跨 5 年資料庫的調查，運用改良型關聯規則探勘技術，建構台灣新移民女性醫療健康需求模型，嘗試提出規劃符合大陸及外籍配偶婦女健康權的醫療健康福利政策建議。

1.2 研究目的

新移民女性融入台灣社會，擔負生育及家庭照護任務，從人權及性別平等的角度來看，協助新移民女性提昇自我健康促進能力，免除婚姻移民背負的污名化包袱，擺脫社會經濟地位及在移入家庭中之性別階級弱勢地位，是值得探討的重要社會及文化問題。政府對新移民女性提供一系列「外籍與大陸配偶照顧輔導措施」，不只跨部會分工，也要求各縣市政府積極配合辦理各項活動，投入相當多的人力與經費，是十分值得肯定的施政方向，然而就成效統計表來看，十分可惜的是除活動參與率低外，活動雷同與制式化，似乎淪為政令宣導與無感的團體活動。需求面(新移民女性)與供給面(社會福利政策及服務提供者)為什麼會出現認知落差？原因是值得研究探討。

本計畫首先擬改良資料探勘的「關聯規則」(Association Rule)演算法以提升其搜尋效率，並用此改良型關聯分析技術探勘全國性調查的大型資料庫，期能找出更精確的人口屬性與健康需求之關聯性，繼之，結合醫學專業與社會學的觀點進行相關探討評析，期建構出台灣新移民女性健康促進的醫療需求模型，作為政府及民間團體推動外籍與大陸配偶各項醫療保健服務措施，及政府制定相關服務政策之參考。

1.3 計畫重要性

台灣已進入多元文化社會，政府及有關單位對於新移民權益(尤其女性)重視，已看到政策努力的腳步。但是，這些醫療服務及輔導是否能真正滿足新移民女性的處境與權益問題，是需要更深入的剖析，以真實的滿足及回應她們的需求，避免產生不合適的移民政策或是過度強調「同化」與「適應」的移民輔導服務。本研究希望透過資訊工程科學、醫學及社會學的跨領域合作與對話，讓大陸及外籍配偶的健康權及公民權不再被視為邊緣化，以期逐步改善此一弱勢社群的權益。

內政部 2003 年之全面普查外籍與大陸配偶生活狀況調查及 2008 年移民署辦理外籍與大陸配偶生活需求調查，都在藉以瞭解新移民女性的生活需求。然而從文獻回顧中，雖然已有各領域專家學者投入關懷與研究(內政部, 2003, 2008; 王秀紅 & 楊詠梅, 2002; 王素美, 林育秀, 李卓倫, 陳文意, & 梁亞文, 2011)，但由於分析方法多侷限於單向變數的描述性統計，僅就各項輔導需求加以排出優先性或比例，未能更精確就其人口學屬性與健康醫療需求進行關聯分析，建構台灣新移民女性與健康促進輔導措施的醫療需求模型，殊為可惜。施政措施若能針對新移民個別族群的差異性，導入客戶導向的服務概念，相信能更有效運用經費及提高政策效益，最重要的是，真正感動新移民、深化對台灣土地及社會的認同，促進族群融合及社會公平正義。

資料探勘中的關聯分析技術，Apriori 演算法最常被應用，但是傳統 Apriori 演算法在大型資料庫應用上常因重複計算而導致搜尋效率不佳。本研究整合 Dunkel & Soparkar (1999)與 Zhang et al. (2009)的方法，提出 Column-wise and Vector-based 計算方式來減少整體計算量、加速搜尋常出現的項目集合，提出改良型 Apriori 演算法；並將此演算法運用在發掘台灣新移民女性與醫療健康需求的關聯性探討。

新移民女性社群隔離及健康照護弱勢現象，不僅是新移民自身健康問題，更包括家庭、社會體制、健康照護專業及社會支持網絡問題。新移民提供了許多照護人力的支援與協助，新移民婦女的健康權還有更廣泛的健康照護問題需要討論，所以除了生育、育兒，還有身體自然老化所帶來的疾病(如：更年期、糖尿病、高血壓、婦女癌症...等)。新移民女性的社群隔離狀態若持續加深，將對整個台灣社會的生活及人口品質提出挑戰，關注女性新移民健康照護問題，必須從社群隔離現象所帶來的困境，並加入人口學特質的需求分析，相信本研究的結果將對整個健康照護體系及社會穩定產生正向助力。

貳、文獻探討

內政部 2003 年全面普查外籍與大陸配偶生活狀況調查及 2008 年移民署辦理外籍與大陸配偶生活需求調查，目的都在瞭解新移民女性的醫療衛生輔導及生活照顧輔導之需求。從文獻回顧中，已有各領域專家學者投入關懷與研究(內政部, 2003a, 2008a; 王秀紅 & 楊詠梅, 2002; 王素美, 林育秀, 李卓倫, 陳文意 & 梁亞文, 2011)，但由於分析方法多以單變數描述性統計為主，僅就各項輔導需求加以排出優先性或比例，未能更精確就人口學屬性與健康醫療需求進行關聯分析，並建構台灣新移民女性與健康促進輔導措施的醫療需求模型，較為可惜。

新移民女性社群隔離及健康照護弱勢現象，不僅是新移民自身健康問題，更包括家庭、社會體制、健康照護專業及社會支持網絡問題。新移民婦女的健康還包括更全面的生命週期的健康照護問題，所以除了生殖、育兒，還有身體老化的疾病(如：更年期、糖尿病、高血壓、婦女癌症...等)。

女性因婚姻移民進入台灣社會及家庭，在文化適應過程中，可能產生個人、社會網絡、社區環境邊緣化，心理上經歷認同、連結及逐漸邊緣化歷程(Hall, Stevens, & Meleis, 1994)。新移民對台灣文化生活認同與學習並非一朝一夕達成，若調適不順利，將加深其與社會隔離的情況。例如語言溝通

能力不佳，便無法融入文化與生活，始終會認定自己為「邊緣人」。而移民者的文化融合能力，「經濟獨立」亦為關鍵的適應因素，女性因為性別差異更加凸顯勞動條件的落差，移民者語言與生活技能不足，造成經濟依賴，更加孤立與邊緣化(Diwan, Jonnalagadda, & Balaswamy, 2004)。

移民及性別的雙重性造成大陸及外籍女性配偶的社群隔離現象，使得新移民婦女健康更易陷入弱勢，相對於本國女性，在獲得健康資訊、醫療及健康照護都較困難。移民適應過程中，因為生活調適所帶來的壓力、無力感、生活習慣改變，都是影響移民者健康的危險因子(Millar & Stephens, 1993)。雖然，就實際的大陸及東南亞籍配偶之跨國婚姻來看，許多新移民女性都是屬於身體條件較佳者，但面對生活環境改變，自我健康照護能力也較本國婦女弱(廖建彰 & 詹婉瑜, 2005)，若再加上社群隔離因素，無法充分獲得健康照護知識與資源，非常容易成為身心疾病高危險的婦女。新移民女性的健康弱勢可以從三方面進行討論：

2.1 新移民女性的個人健康

壓力影響健康已是不爭的事實，而移民壓力使個人難以適應新文化和新環境，特別是人際關係、角色、語言、價值觀與態度的調適困難，身心壓力極易轉化為嚴重的身心健康問題，造成憂鬱、精神分裂及焦慮等(Hovey & King, 1996)。移民者生活調適的衝擊，包括工作、經濟、文化價值、生活調適、語言障礙、喪失社會地位等(Hattar-Pollara & Meleis, 1995)，可能導致語言問題、背痛、焦慮、憂鬱等症狀(D'Avanzo, Frye, & Froman, 1994)，更常見包括頭痛、無胃口、體重減輕、失眠、月經不順，甚至孤寂、焦慮、擔心、悲傷、失落、憂鬱等心理問題(Lien, Margaret Conway, & Wong, 2003)。

特別是語言溝通問題是新移民女性就醫最大障礙，因為語言障礙使得就診時無法適當表達自身的感受，也無法完全理解醫師指示，讓求醫過程充滿挫折感。

2.2 新移民女性的家庭健康

跨國婚姻中，夫妻雙方成長背景相異，文化及價值觀差異影響婚姻期望，若生活適應及語言溝通障礙，將更容易造成婚姻維繫的壓力。新移民女性承受生育、照顧子女、奉養長輩等責任，生育尤其被視為首要任務。新移民女性經常在尚未完全調適狀況下，無預期懷孕生育，結果將對婦女及其子女健康造成很大影響。若男方（配偶）原本即為弱勢家庭或社會的高危險族群（社經地位及健康條件較差），子女較不健康的機率更高。(吳慧卿、曾明性、林宗弘, 2012)。因文化差異，導致坐月子觀念的不適應及經驗不足，幼兒照顧困境增加。跨國婚姻中的家庭暴力、兒童虐待及藥物濫用亦是常見的現象(李瑞金 & 張美智, 2004)。

2.3 新移民女性的健康照護資源

新移民女性對生活環境陌生，多會選擇離家近的診所就醫，學者研究指出新移民女性對於診所醫師的處置滿意度相當高，但因就診需家人陪同，形成就醫過程的不便；語言表達的限制，造成新移民女性對就醫方便性及等候的滿意度較低(周天給, 許怡欣, 李丞華, & 湯澡薰, 2006)。此外，也讓新移民女性感受到不舒服的就醫過程，包括醫事人員的歧視與不友善，使其在利用健康資源時，經常是被邊緣化的族群(戴君倚 & 鍾聿琳, 2002)。當又需擔負照顧家庭慢性病患或老人的照護者角色時，也同樣面臨資源取得不足的困境。

新移民女性來台後多是居住於社會資源相對較貧乏少的鄉村、山地或離島，因為自身生活適應能力不足，較難尋求有效的社會支持網絡系統，影響新移民女性的自我健康促進及醫病互動的能力(林淑玲 & 蕭仔伶, 2007)。

近年來，台灣醫療體系已在衛生署的政策導引下，加強對新移民女性健康照護服務提供，例如對無健保身分者提供醫療補助，減少經濟問題的健康排擠；衛生所護理人員進行訪視了解需要，但又因衛生人力有限，尤其大都會縣市，可能反而會越來越慢接受到訪視服務(周天給 et al., 2006)。

新移民女性的本國籍配偶，有許多是社經地位較差、年齡偏高或身心障礙者，而新移民女性在原生國家庭也多來自社會經濟地位較為弱勢階層，兩者結合，十分容易成為社會弱勢族群(廖建彰 & 詹婉瑜, 2005)。所以，針對跨國婚姻移民的婦女健康研究，應將討論議題分為「移民適應與壓力」、「經濟困境與健康」、「語言困難與健康」及「特殊健康問題」等面向，且必須理解，這些問題都不是單獨存在的構面，而是互相交纏的社會及文化因素的結果。學者指出，以印尼外籍配偶為例，由於語言障礙，無法瞭解醫師指示的用藥方法，陪同的家人也因教育程度較低，無法理解醫病對話，醫護人員若再因語言及刻板印象，較缺乏耐心傾聽，就容易造成解釋病情的時間較短，進而形成就醫心理障礙及後續醫療的困難(王秀紅 & 楊詠梅, 2002)。

參、研究方法

在數位資訊時代，持續性的紀錄過程會產生的大量資料，這些大量資料中所潛藏的、未知有用資訊或知識，如果能有效挖掘出來，將可帶來更多的效益。資料探勘 (Data Mining) 可從大量的資料中找出未知、正確且有用的隱藏知識，是知識發現的重要技術。Chen et al. (1996) 認為資料探勘可從資料庫中萃取出重要、事先未知、潛在有用的資訊；Berry & Linoff (1997)認為資料探勘可針對大量的資料，應用自動或半自動的方式進行分析，以找出有意義的關係或法則。Fayyad et al. (1996)與Feyyad (1996)將知識發現流程分為資料選擇、資料前置處理、資料轉換、資料探勘、解釋與評估等步驟，如下圖所示。Han et al. (2011)將資料探勘方法分為四大類，分別為關聯分析(Association Analysis)、分類與預測(Classification and Prediction)、群集分析(Clustering Analysis)、推估與偏差分析(Evolution and Deviation Analysis)。資料探勘相關技術以往常用於挖掘商業知識，近年來也被運用於挖掘醫學知識，例如：(吳瑞堯 & 周駿賢, 2011)運用資料探勘技術結合全民健康保險研究資料庫進行六大死因慢性疾病之關聯分析。關聯分析又稱為購物籃分析(Basket Analysis)，因擅於分析購物籃中購物品項間的關係而得名，可從眾多交易中(多個購物籃)分析出哪些商品同時被顧客購買的頻率較高，分析結果可協助業者訂定相關商品促銷策略，可應用於市場行銷及客戶社群分析等許多應用上 (Khan et al., 2008)

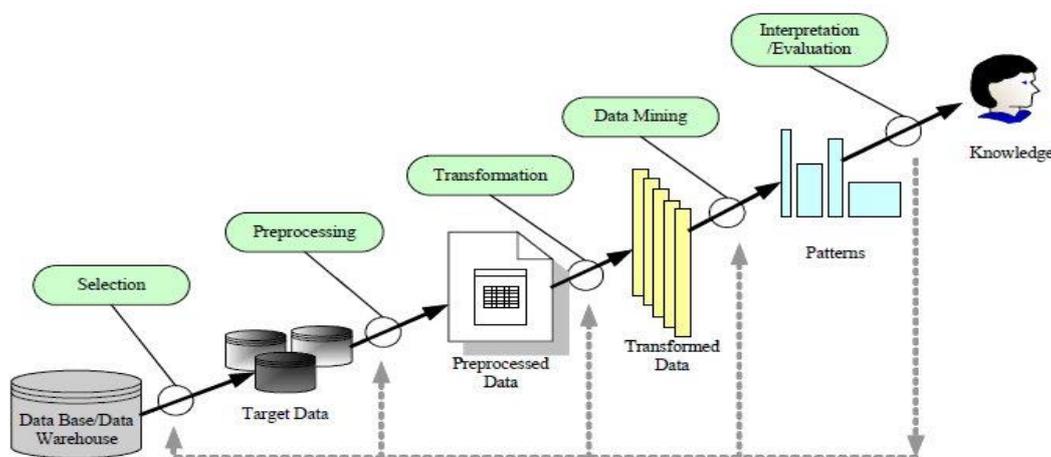


圖 1 知識發現流程 (Fayyad et al., 1996)

關聯分析由 Agrawal 等學者於 1993 年提出(Agrawal et al., 1993; Agrawal & Srikant, 1994)，主要觀念建立在條件機率上，並且運用支持度(Support)與信賴度(Confidence)篩選關聯規則，以關聯規則 $X \rightarrow Y$ 而言，其支持度 $s =$ 包含項目 X 的交易數量，信賴度 $c =$ (同時包含項目 X 、 Y 的交易數量) / 包

含 X 的交易數量。舉例而言，若在某賣場的交易中，50 筆有購買尿布，50 筆有購買啤酒，同時購買尿布及啤酒的交易有 30 筆，則對於買尿布就會買啤酒的關聯規則(以 $X \rightarrow Y$ 表示)來講，其支持度 $s = 50$ ，信賴度 $c = 30 / 50 = 0.6$ 。關聯分析的目的是找出具備足夠支持度與信賴度的關聯規則，所以一個關聯規則要成立必須同時滿足預先設定的最小支持度與最小信賴度。

關聯規則探勘(Association Rule Mining)可分為兩大主要步驟：(1)找出常出現的項目集合(Frequent Itemsets)，即找出常出現且符合最小支持度門檻值的項目集合；(2)產生關聯規則，即根據所訂定之最小信賴度來篩選其項目集合。關聯規則眾多演算法中，以 Agrawal et al. (1993)所提出的 Apriori 演算法為最具代表性且經常被應用在資料探勘的關聯分析上。

3.1 Apriori 演算法

Step 1: 演算法先掃瞄所有交易紀錄，產生單一項目的候選項目集合(Candidate Item Sets)，如下圖步驟 1 所示之 C1。

Step 2: 將數量較少的項目集合排除，只保留滿足最小支持度的項目集合，稱為大型項目集合(Large Item Sets)，如下圖步驟 2 所示之 L1。

Step 3: 對 L1 中的各集合進行組合，成為具有兩個項目的集合，產生兩個項目的候選項目集合，如下圖步驟 3 的 C2。

Step 4: 排除 C2 中支持度小於 2 的項目集合，得到下圖步驟 4 的大型項目集合 L2。

Step 5: 對 L2 中的各集合進行組合，成為具有三個項目的集合，產生三個項目的候選項目集合，如下圖步驟 5 的 C3。

Step 6: 因為 C3 的兩個項目集合支持度都大於 2，所以 L3 中的大型項目集合與 C3 相同，如下圖步驟 6 的 L3。

Step 7: 因為無法再由 L3 產生四個項目的候選項目集合 C4，結束項目集合產生程序。

Step 8: 從 L2 及 L3 中的項目集合產生符合最小信賴度的關聯規則。對每個項目集合 l 的子項目集合(非空集合) a 檢查 $a \rightarrow l-a$ 之信賴度是否高於最小信賴度，若是，則產生該關聯規則。例如對項目集合 ABC，須檢視 $A \rightarrow BC$, $B \rightarrow AC$, $C \rightarrow AB$, $AB \rightarrow C$, $AC \rightarrow B$, $BC \rightarrow A$ 等關聯規則是否高於最小信賴度。

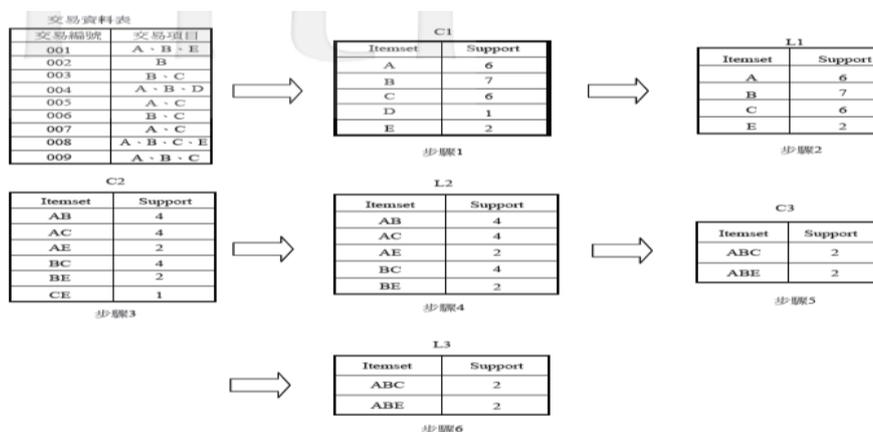


圖 2 Apriori 演算法範例(吳瑞堯 & 周駿賢, 2011)

上述即為 Apriori 演算法的演算過程，利用此方法可以獲得相當多之關聯規則。針對每一個關聯規則($X \rightarrow Y$)可以用支持度(Support)、信賴度(Confidence)及增益值(Lift)等三個指標來評估其效用和重要性。

3.2 指標的計算公式與物理意義

3.2.1 支持度(Support)

一個項目集 X 的支持個數(Support Count)定義為支持項目集 X 的交易總數，而項目集 X 的支持度則為支持項目集 X 的交易個數佔交易總數的比例。關聯規則 $X \rightarrow Y$ 的支持度為項目集 X 與 Y 同時出現的機率，即：

$$\text{Support}(X \rightarrow Y) = P(X \cap Y) \quad (1)$$

關聯規則要有意義，必須此關聯規則相對於全部的資料是顯著的。因此以最小支持度篩選掉不顯著的關聯規則。

3.2.2 信賴度(Confidence)

發生項目集 X 時，將發生結果項目集 Y 的機率，即：

$$\text{Confidence} = P(Y|X) = P(X \cap Y)/P(X) \quad (2)$$

此為關聯規則的預測強度。此指標之物理意義為當前提項目 X 發生時，可推得結論為結果項目 Y 的信賴度。關聯規則要有意義，信賴度要達到一定顯著水準，也就是有信心當前提項目 X 發生時，可推得結果項目 Y 也發生。因此以最小信賴度篩選掉不顯著的關聯規則。

3.2.3 增益值(Lift)

當支持度與信賴度都大於或等於最小限制時，接著判斷其增益值是否大於 1，即計算信賴度與項目 Y 發生的機率是否大於 1，若大於 1 代表為顯著的關聯規則。

$$\text{Lift}(X \rightarrow Y) = P(Y|X)/P(Y) \quad (3)$$

其代表的物理意義為關聯規則的條件機率比母體中原本發生的機率大才具有意義。若某關聯規則其支持度與信賴度都達到顯著水準，有可能是原本結果項目 Y 發生機率就大，因此必須比較此關聯規則信賴度與原本結果項目 Y 發生之機率，若比值大於 1，代表此關聯規則的信賴度大於原本結果項目 Y 發生之機率，滿足此條件才為顯著的關聯規則。

3.3 改良型 Apriori 演算法(I)

Dunkel & Soparkar(1999)提出 Column-wise Apriori Counting 方式修正原始 Row-wise Apriori Counting 方式來搜尋常出現的項目集合，可提高其搜尋演算效能，其演算法如下圖 3 所示：

```
1)  $L_1 = \{\text{frequent 1-itemsets}\}; // \text{Column-wise scan}$ 
2) for ( $k = 2$  until  $L_{k-1} \neq \emptyset$  step 1) do begin
3)  $C_k = \text{apriori-gen}(L_{k-1}); // \text{New candidates}$ 
3a') forall candidates  $c \in C_k$  do begin
4) forall ( $k-1$ )-subsets of  $c$  do
5) Verify the subset is frequent;
6) if  $c$  has no infrequent subset
7) Count the support of  $c$  by conjunction;
8) end
9)  $L_k = \{c \in C_k \mid c.\text{count} \geq \text{minsup}\};$ 
10) end
11)  $\text{Answer} = \cup_k L_k;$ 
```

圖 3：Column-wise Apriori 演算法 (Dunkel & Soparkar, 1999)

Zhang et al.(2009)基於向量內積運算提出 Vector-based 的演算法來降低候選頻率項目集合的數量，可有效改善修剪(Pruning)運算量及支持度的計算量，其演算法如下圖 4 所示：

```

Input: a transaction database  $D$ , the minimum threshold
of support  $minSupNum$ 
Output: the set of frequent itemsets  $L$ 
1) for all transaction  $t \in D$  do
2)   generate  $TV$ ;
3)  $L_1 =$  (frequent 1-itemsets);
4)  $C_2 = L_1 \times L_1$ ;
5)  $L_2 = \{c \in C_2 \mid sup(c) \geq MinSupNum\}$ ; //  $sup(c)$  is the
   result of the formula (4)
6) for ( $k=3; L_{k-1} \neq \emptyset; k++$ ) do begin
7)   for( $j=k; j \leq m; j++$ ) do
8)     generate  $CIV_{i_j}^{k-1}$ ;
        $C_k =$  candidate_gen( $L_{k-1}$ );
        $L_k = \{c \in C_k \mid sup(c) \geq MinSupNum\}$ ;
9)   end
10) end
11) Return  $L = \cup L_k$ ;
candidate_gen( frequent itemsets  $L_{k-1}$  )
1) for all ( $k-1$ )-itemset  $l \in L_{k-1}$  do
2)   for all  $i_j \in L_{k-1}$  do
3)     //  $S$  is the result of the formula (2)
       if for every  $r(1 \leq r \leq k)$  such that  $S[r] \geq k-1$  then
4)       add  $l \cup \{i_j\}$  to  $C_k$ ;

```

圖 4：Vector-based Apriori 演算法 (Zhang et al., 2009)

由於 Zhang et al. (2009) 的 Vector-based Apriori 演算法，其向量內積運算係採 Row-wise 方式計算較為繁雜，故本研究擬整合 Dunkel & Soparkar (1999) 與 Zhang et al. (2009) 的方法，藉由 Column-wise and Vector-based 計算方式來減少整體計算量、加速搜尋常出現的項目集合，提出改良型 Apriori 演算法，以下利用九筆交易編號(T1, T2, T3, T4, T5, T6, T7, T8, T9)之五種交易項目(A, B, C, D, E)為例，如下表 1 所示，並設定最小支持個數 $S_{min}=2$ 進行本改良型演算法的演算步驟說明：

表 1：交易明細範例

交易編號	交易產品項目
001	A、B、E
002	B
003	B、C
004	A、B、D
005	A、C
006	B、C
007	A、C
008	A、B、C、E
009	A、B、C

Step 1: 依照上述九筆交易明細範例，建立初始對應矩陣。

表 2：初始對應矩陣

	A	B	C	D	E	S_{min}
T1	1	1	0	0	1	3
T2	0	1	0	0	0	1
T3	0	1	1	0	0	2
T4	1	1	0	1	0	3
T5	1	0	1	0	0	2
T6	0	1	1	0	0	2
T7	1	0	1	0	0	2
T8	1	1	1	0	1	4
T9	1	1	1	0	0	3
Support	6	7	6	1	2	

上表 2 即為 1-候選項目集合 $C1=\{A, B, C, D, E\}$ 。

Step 2: 篩選 1-候選項目集合 $C1$ ，求取 1-大型項目集合 $L1$ 。

由於設定最小支持個數 $S_{min}=2$ ，因此刪除 $C1$ 集合中之項目 D ，刪除後對應表如下表 3 所示：

表 3：大型項目集 1 之對應矩陣

	A	B	C	E	S_{um}
T1	1	1	0	1	3
T2	0	1	0	0	1
T3	0	1	1	0	2
T4	1	1	0	0	2
T5	1	0	1	0	2
T6	0	1	1	0	2
T7	1	0	1	0	2
T8	1	1	1	1	4
T9	1	1	1	0	3
$S_{support}$	6	7	6	2	

上表 3 即為 1-大型項目集合 $L1=\{A, B, C, E\}$ 。

Step 3: 進入 $Count_2$ ，重新計算各交易的項目總和 S_{um} 與支持個數 $S_{support}$ 。

進行各交易總和篩選，若小於配對標準 $Count_i$ 者，代表無法組成 i 個項目的配對，故將此項目刪除，其中 i 代表配對項目數。例如目前正進行 C2 的配對，配對項目數為 2，則設定 $Count_2=2$ 。因此，刪除交易總和 S_{um} 小於 2 的交易編號 T2，並重算第二階段各項目的支持個數如下表 4 所示：

表 4：第二階段各項目的支持個數

	A	B	C	E	S_{um}
T1	1	1	0	1	3
T3	0	1	1	0	2
T4	1	1	0	0	2
T5	1	0	1	0	2
T6	0	1	1	0	2
T7	1	0	1	0	2
T8	1	1	1	1	4
T9	1	1	1	0	3
$S_{support}$	6	6	6	2	

總共有 $\{A, B, C, E\}$ 4 個項目的支持個數大於或等於最小支持個數 $S_{min}=2$ ，可形成 $C^4_2=6$ 種配對結果，得到 2-候選項目集合 $C2=\{AB, AC, AE, BC, BE, CE\}$ 。

Step 4: 篩選第二候選項目集合 C2，求取 2-大型項目集合 L2。

各項目的向量分別為： $A=(1,0,1,1,0,1,1,1)$ ； $B=(1,1,1,0,1,0,1,1)$ ； $C=(0,1,0,1,1,1,1,1)$ ； $E=(1,0,0,0,0,0,1,0)$ 。計算各向量內積和： $\langle A,B \rangle=4$ ； $\langle A,C \rangle=4$ ； $\langle A,E \rangle=2$ ； $\langle B,C \rangle=4$ ； $\langle B,E \rangle=2$ ； $\langle C,E \rangle=1$ 。選取內積大於或等於最小支持個數 $S_{min}=2$ 的配對數目，刪除 $\langle C,E \rangle$ 項目，形成 2-大型項目集合 L2，求得 $L2=\{(A,B), (A,C), (A,E), (B,C), (B,E)\}$

Step 5: 進入 $Count_3$ ，重新計算各交易的項目總和 S_{um} 與支持個數 $S_{support}$ 。

設定 $Count_3=3$ ，因此刪除交易總和小於 3 的交易編號，刪除 T3, T4, T5, T6, T7，並重算第三階段各項目的支持個數如下表 5 所示：

表 5：第三階段各項目的支持個數

	A	B	C	E	S_{um}
T1	1	1	0	1	3
T8	1	1	1	1	4
T9	1	1	1	0	3
$S_{support}$	3	3	2	2	

總共有{A, B, C, E}4個項目的支持個數大於或等於最小支持個數 $S_{min}=2$ ，可形成 $C_3^4=4$ 種配對結果，得到第三候選項目集合 $C3=\{ABC, ABE, ACE, BCE\}$ 。

Step 6: 篩選第三候選項目集 $C3$ ，求取 3-大型項目集合 $L3$ 。

各項目的向量分別為: $A=(1,1,1)$ ； $B=(1,1,1)$ ； $C=(0,1,1)$ ； $E=(1,1,0)$ 。計算各向量內積和: $\langle A,B,C \rangle=2$ ； $\langle A,B,E \rangle=2$ ； $\langle A,C,E \rangle=1$ ； $\langle B,C,E \rangle=1$ 。選取內積大於或等於最小支持個數 $S_{min}=2$ 的配對數目，刪除 $\langle A,C,E \rangle$ 、 $\langle B,C,E \rangle$ 項目，形成 3-大型項目集合 $L3$ ，求得 $L3=\{(A,B,C), (A,B,E)\}$ 。

3.4 改良型 Apriori 演算法(II)

Huang, Liao & Lin (2011)提出 A data cutting and sorting method (DCSM)方式來搜尋常出現的項目集合，可提高其搜尋演算效能，其演算法如下圖 5 所示：

Step1 Transfer the original database into the original Boolean matrix, BM_0

```

/* DB → BM0
1: for i = 1 to NR
2:   for j = 1 to Ndb
3:     if Oj ∈ Ri
4:       E0ij = 1
5:     else
6:       E0ij = 0
7:     endif
8:   endfor
9: endfor

```

Step2 Establish the first large itemsets, L_1

```

10: NT1 = NT0
11: for x = 1 to NT0
12:   if SP0,x < Smin
13:     delete I0,x
14:     NT1 = NT1 - 1
15:   endif
16: endfor
/* the remained items become the elements of the first
   large itemsets, H1,x,
/* where x = 1, 2, ..., NT1
17: NI1 = NI0
18: for x = 1 to NI0
19:   if SI0,x = 1
20:     delete T0,x
21:     NI1 = NI1 - 1
22:   endif
23: endfor
24: I1,j = sort( H1,x) /* recalculate the support of each item,
/* the items of BM1 are arranged in the order of their
   supports

```

Step3 Establish the second large itemsets, L_2

```

25: for x = 1 to NT1 - 1
/* extract the item which has the minimum support
26:   NT2 = NT1
27:   delete T1,i where E1ix = 0
28:   for y = 1 to NT1 - x
29:     if SP1,y ≥ Smin
/* (I1,x, I1,y) become the element of the second large
   itemsets, H2,x

```

```

Step4 Establish the third large itemsets,  $L_3$ 
30:   for  $z=1$  to  $NT_2 - x - y$ 
31:     if  $SP_{2z} \geq S_{min}$ 
        /*( $I_{1,x}, I_{1,y}, I_{1,z}$ ) become the element of the third large
        itemsets,  $H_{3,x}$ 
32:     endfor
33:   endif
34: endfor

```

圖 5 DCSM 演算流程 (Huang, Liao & Lin, 2011)

本研究修正 DCSM 原始 Row-wise Apriori Counting 方式，以下利用九筆交易編號(T1, T2, T3, T4, T5, T6, T7, T8, T9)之五種交易項目(A, B, C, D, E)為例，並設定最小支持個數 $S_{min}=2$ 進行本改良型演算法的演算步驟說明：

Step 1: 依照上述九筆交易明細範例，建立初始對應矩陣，如表 6 所示。

表 6 初始對應矩陣

	A	B	C	D	E	S_{um}
T1	1	1	0	0	1	3
T2	0	1	0	0	0	1
T3	0	1	1	0	0	2
T4	1	1	0	1	0	3
T5	1	0	1	0	0	2
T6	0	1	1	0	0	2
T7	1	0	1	0	0	2
T8	1	1	1	0	1	4
T9	1	1	1	0	0	3
$S_{support}$	6	7	6	1	2	

表 6 即為 1-候選項目集合 $C1=\{A, B, C, D, E\}$ 。

Step 2: 篩選 1-候選項目集合 $C1$ ，求取 1-大型項目集合 $L1$ 。

由於設定最小支持個數 $S_{min}=2$ ，因此刪除 $C1$ 集合中之項目 D，刪除後對應表並依 $S_{support}$ 由大到小排序，如表 7 所示：

表 7 大型項目集 1 之對應矩陣(排序後)

	B	A	C	E	S_{um}
T1	1	1	0	1	3
T2	1	0	0	0	1
T3	1	0	1	0	2
T4	1	1	0	0	2
T5	0	1	1	0	2
T6	1	0	1	0	2
T7	0	1	1	0	2
T8	1	1	1	1	4
T9	1	1	1	0	3
$S_{support}$	7	6	6	2	

表 7 即為 1-大型項目集合 $L1=\{A, B, C, E\}$ 。

Step 3: 建立項目 E 的 2-削減矩陣 R_{E2} 。

因要進行配對所以先刪除 $S_{um}<2$ 的交易 T2，繼之，挑選 $S_{support}$ 最低的项目 E 進行 E 的 2-削減矩陣建構，刪除項目 E 為 0 的交易 T3, T4, T5, T6, T7, T9，建立 R_{E2} ，如表 8 所示：

表 8 項目 E 的 2-削減矩陣 R_{E2}

E/	B	A	C	S_{um}
T1	1	1	0	2
T8	1	1	1	3
$S_{support}$	2	2	1	

Step 4: 建立項目 E 的 2-大型項目集合 L_{E2} 。

(1) 刪除 $S_{support}<2$ 的項目 C，建立項目 E 的 2-大型項目集合 L_{E2} ，如表 9。

表 9 項目 E 的 2-大型項目集合 L_{E2}

E/	B	A	S_{um}
T1	1	1	2
T8	1	1	2
$S_{support}$	2	2	

求得 $L_{E2}=\{(E,B), (E,A)\}$

(2) Step 5: 建立項目 EA 的 3-削減矩陣 R_{EA3} ，如表 10。

表 10 項目 EA 的 3-削減矩陣 R_{EA3}

EA/	B	S_{um}
T1	1	1
T8	1	1
$S_{support}$	2	

Step 6: 建立項目 EA 的 3-大型項目集合 L_{EA3} 。

刪除 $S_{support}<2$ 的項目，建立項目 EA 的 3-大型項目集合 L_{EA3} 如表 10 求得 $L_{EA3}=\{(E,A,B)\}$ 。

Step 7: 遞迴 Step4~6，建立項目 EA 以外的 3-大型項目集合 $L_{E?3}$ 。

回到 Step 4 的表，尋求另一配對 EB，但已無其他項目，故停止遞迴。

Step 8: 遞迴 Step3~7 建立項目 E 以外的 2-削減矩陣 $R_{?2}$ 。

(3) 回到 Step 2 的表 7。

(4) 挑選 $S_{support}$ 最第二低的項目 C 進行 C 的 2-削減矩陣建構，刪除項目 C 為 0 的交易 T1, T4，建立 R_{C2} ，如表 11 所示：

表 11 項目 C 的 2-削減矩陣 R_{C2}

C/	B	A	S_{um}
T3	1	0	1
T5	0	1	1
T6	1	0	1
T7	0	1	1
T8	1	1	2
T9	1	1	2
$S_{support}$	4	4	

(3)刪除 $S_{support}<2$ 的項目，建立項目 C 的 2-大型項目集合 L_{C2} ，如表 12。

表 12 項目 C 的 2-大型項目集合 L_{C2}

C/	B	A	S_{um}
T8	1	1	2
T9	1	1	2
$S_{support}$	2	2	

求得 $L_{E2}=\{(C,B), (C,A)\}$

(5) 建立項目 CA 的 3-削減矩陣 R_{CA3} ，如表 13。

表 13 項目 CA 的 3-削減矩陣 R_{CA3}

CA/	B	S_{um}
T8	1	1
T9	1	1
$S_{support}$	2	

(5) 刪除 $S_{support} < 2$ 的項目，建立項目 CA 的 3-大型項目集合 L_{CA3} ，同上表，求得 $L_{EA3} = \{(C,A,B)\}$ 。

Step 8-?: 遞迴 Step 3~7 建立項目 E, C 以外的 2-削減矩陣 $R_{?2}$ 。

(1) 回到 Step 2 的表 7。

(2) 挑選 $S_{support}$ 最第三低的項目 A 進行 A 的 2-削減矩陣建構，刪除項目 A 為 0 的交易 T1, T4，建立 R_{A2} ，如表 14 所示：

表 14 項目 A 的 2-削減矩陣 R_{A2}

	B	A	S_{um}
T1	1	1	2
T3	1	0	1
T4	1	1	2
T5	0	1	1
T6	1	0	1
T7	0	1	1
T8	1	1	2
T9	1	1	2
$S_{support}$	7	6	

(6) (3) 刪除 $S_{support} < 2$ 的項目，建立項目 A 的 2-大型項目集合 L_{A2} ，如表 15。

求得 $L_{A2} = \{(A,B)\}$ 。

表 15 項目 A 的 2-大型項目集合 L_{A2}

A/	B	S_{um}
T1	1	2
T4	1	2
T5	0	1
T7	0	1
T8	1	2
T9	1	2
$S_{support}$	4	

(4) 由於表 15 只剩項目 B，所以結束建立項目 A 的 3-大型項目集合搜尋。

(5) 回到 Step 2 的表 7，只剩項目 B，所以停止所有計算。

結果: $L2 = \{(A,B), (A,C), (A,E), (B,C), (B,E)\}$ 。

$L3 = \{(A,B,C), (A,B,E)\}$ 。

由上述案例說明，可知改良型 Apriori (I) 和改良型演算法 (II) 求得的大型項目集合 $L2$ 及 $L3$ 與傳統 Apriori 演算法求得結果是完全相同。

改良型 Apriori 演算法 (I) 演算過程只需搜尋資料庫一次、而且每一階段候選項目集合與大型項目集合的搜尋量有效減少、支持度的計算量亦明顯簡化，並可同步計算各可能關聯規則的信賴度與增益值，而改良型 Apriori 演算法 (II) 演算過程減少冗餘資料，並透過拆解資料表，且交易項目的支持度進行排序，經由刪除重複的搜尋時間，可縮減所需處理的項目，還可計算各可能關聯規則的信

賴度與增益值。

本計畫將運用改良型 Apriori 演算法(I)和改良型 Apriori 演算法(II)，就 240,837 筆紀錄的九十二年外籍與大陸配偶生活狀況調查資料庫與 13,345 筆紀錄的九十七年外籍與大陸配偶生活需求調查資料庫中，經由設定不同的最小支持度與增益值大於 1 的門檻值、並檢驗其最高信賴度條件下，發掘台灣新移民女性與醫療健康需求之間的關聯規則。

肆、結果與討論

4.1 新移民女性醫療衛生需求-主要需求人次統計

針對「醫療衛生需求-列為主要者」答題，有效填答樣本數為 2003 年 240,837 人次，2008 年為 13,345 人次。2003 年調查結果之主要醫療衛生需求前三項為「提供幼兒健康檢查」（4.69%）、「協助加入全民健康保險」（4.13%）、「提供育嬰育兒知識產前產後指導」（3.81%）。2008 年的調查結果之主要醫療衛生需求前三項為「提供醫療補助」（23.02%）、「提供幼兒健康檢查」（7.34%）、「提供育嬰育兒知識產前產後指導」（3.64%）（表 16，圖 6）。

「主要需求」的填答代表新移民女性內心最渴切得到協助的選項，比較特殊的現象是，2003 年的選項中並無「醫療補助」，結果「提供幼兒健康檢查」及「協助加入全民健康保險」是重要的，到了 2008 年的調查，政策上已經提供新移民女性享有全民健保的補助，所以此項需求已經降到僅 1.62%，而是「提供醫療補助」及「提供幼兒健康檢查」。「提供醫療補助」高居 23.02% 代表的意涵是什麼？推測原因可能是新移民女性的家庭經濟條件面對當前的就醫環境的自負額（掛號費、部分負擔、病房差額）仍力有未逮，醫療制度及醫療院所的競爭，導致許多自費項目的出現，對於經濟能力較差的家庭構成相當程度的就醫壓力，因此醫療衛生需求的最主要需求是「錢」，這對號稱全民保險的台灣健保制度來說，是值得深思的問卷結果。

表 16 新移民女性醫療為生輔導需求（主要需求）人次統計百分比

新移民女性醫療衛生輔導需求-主要需求	2003年 (N=240837人次)		2008年(N=13345人次)	
	人次	%	人次	%
協助加入全民健康保險	9,950	4.13	216	1.62
協助就醫時的溝通	5,372	2.23	382	2.86
提供傳染病疾病的知識	7,160	2.97	389	2.91
提供育嬰育兒知識產前產後指導	9,168	3.81	486	3.64
避孕方法介紹	944	0.39		**
提供產前檢查醫療補助訊息		**	148	1.11
提供幼兒健康檢查	11,288	4.69	979	7.34
提供生育保健知識		**	176	1.32
提供醫療補助		**	3,072	23.02
其他	130	0.05	40	0.30
未答 (missing value)	196,825	81.73	7,457	55.88

註：「**」為當年度調查問卷題目未設計該題

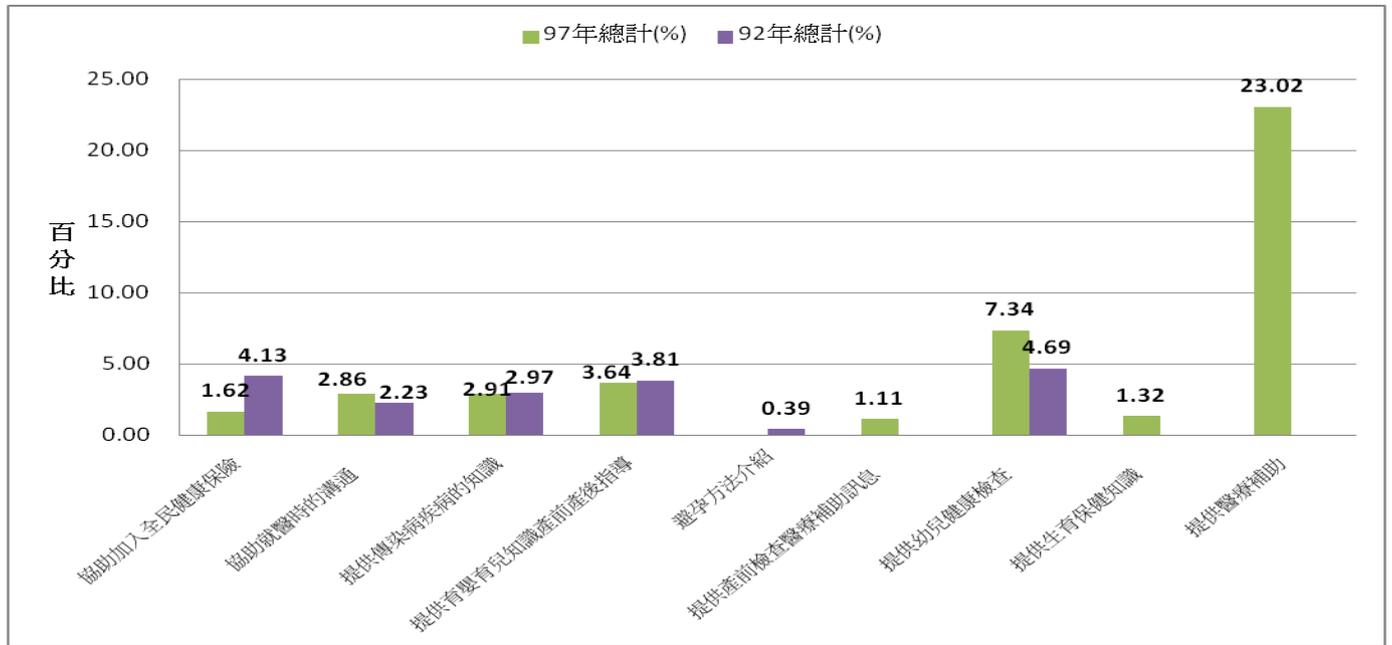


圖 6 新移民女性醫療為生輔導需求 (主要需求) 人次統計百分比

就實際原始資料庫檢視，「醫療衛生需求」題組中，排除答「其他」及未回答者，2003 年的有效樣本為 18.22%，2008 年為 43.82%。特別 2003 年為普查型的調查，有效份數相當低，僅 18.22%。2008 年的有效填答份數雖然提高到 43.82%，但卻 23.02% 集中於「提供醫療補助」。

由於探勘出的關聯規則數量繁多，本計畫以「支持度>0.03」、「增益值>1」為基本門檻，再經由「信賴係數值最高的前三項」為本研究探討的關聯規則篩選的標準，以新移民女性之「個人社經資料」(國籍、學歷、在台資格、居住時間、婚姻方式)和新移民女性配偶(學歷、在台資格、居住時間、婚姻方式)為自變項，「醫療衛生需求」之各項需求為依變項，建構自變項為 1 至 4 項的新移民女性醫療衛生需求關聯規則模型。

4.2 改良型關聯規則演算法之效能探討

本計畫 92 年的資料庫有 63812 筆資料和 32 個屬性，97 年的有 8424 筆資料和 45 個屬性，用來測試的參數為 6 種不同的支持度，時間計算由程式開始執行至有關「醫療衛生需求」的關聯規則搜尋出來為止，進而比較改良型 Apriori(I)、改良型 Apriori(II)、Weka Apriori 與 Weka FP-growth 四種方法的時間差異，其結果如(圖 7，圖 8)所示。

研究工具：Weka 3.7.10 與 Microsoft Visual Studio 2008 程式 C# 語言來撰寫演算法，筆記型電腦，CPU：Intel(R) Core(TM) i7-2637M 1.70GHz，RAM：4.00GB，win7 64 位元作業系統。

表 17 不同支持個數效能比較表(92 年資料集)

支持度	0.2	0.15	0.1	0.05	0.03	0.02
改良型 Apriori (I)	14.02(s)	23.864(s)	44.099(s)	58.131(s)	67.076(s)	74.97(s)
改良型 Apriori (II)	13.57(s)	22.84(s)	34.72(s)	47.83(s)	59.79(s)	66.38(s)
Weka FP-growth	2(s)	3(s)	4(s)	30(s)	96(s)	168(s)
Weka Apriori	13(s)	32(s)	73(s)	195(s)	504(s)	NA

註：「NA」代表軟體無法跑出結果

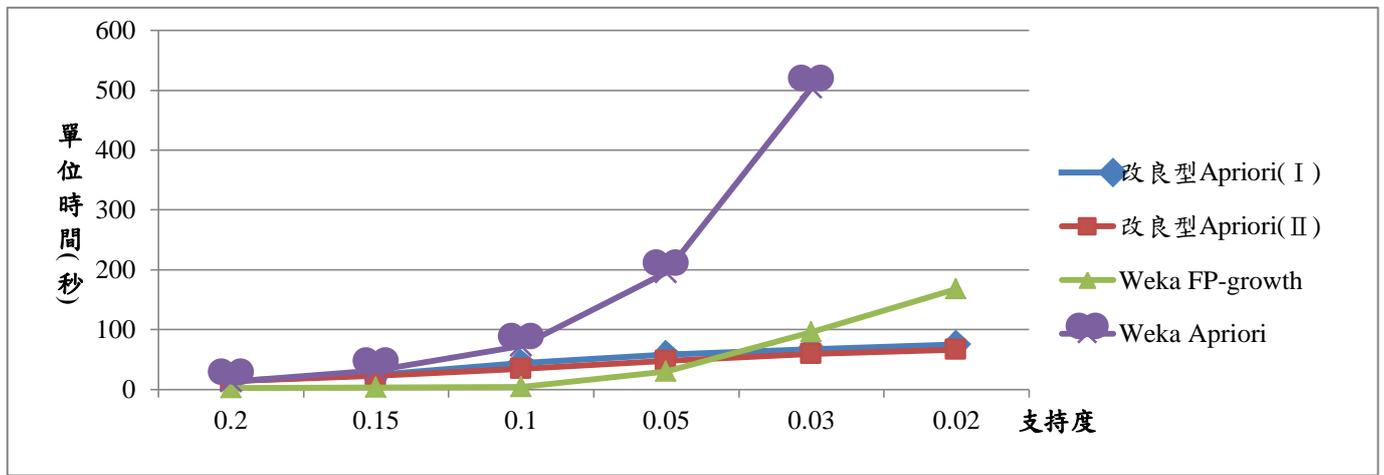


圖 7 演算法效能比較圖(92 年資料集)

從(表 17, 圖 7)得知在最小支持度 0.03 和增益值>1 的情況下:

1. 改良型 Apriori(I)的運算時間較 Weka Apriori 的運算時間快了 7.51 倍。
 2. 改良型 Apriori(II)的運算時間較 Weka Apriori 的運算時間快了 8.43 倍。
 3. Weka FP-growth 的運算時間較 Weka Apriori 的運算時間快了 5.25 倍。
- 效能結果:改良型 Apriori(II) > 改良型 Apriori(I) > Weka FP-growth。

表 18 不同支持個數效能比較表(97 年資料集)

支持度	0.2	0.15	0.1	0.05	0.03	0.02
改良型 Apriori (I)	1.638(s)	29.021(s)	46.847(s)	61.468(s)	62.664(s)	66.414(s)
改良型 Apriori(II)	2.43(s)	15.38(s)	36.54(s)	67.56(s)	78.82(s)	81.62(s)
Weka FP-growth	1(s)	1(s)	2(s)	16(s)	NA	NA
Weka Apriori	6(s)	18(s)	58(s)	NA	NA	NA

註:「NA」代表軟體無法跑出結果

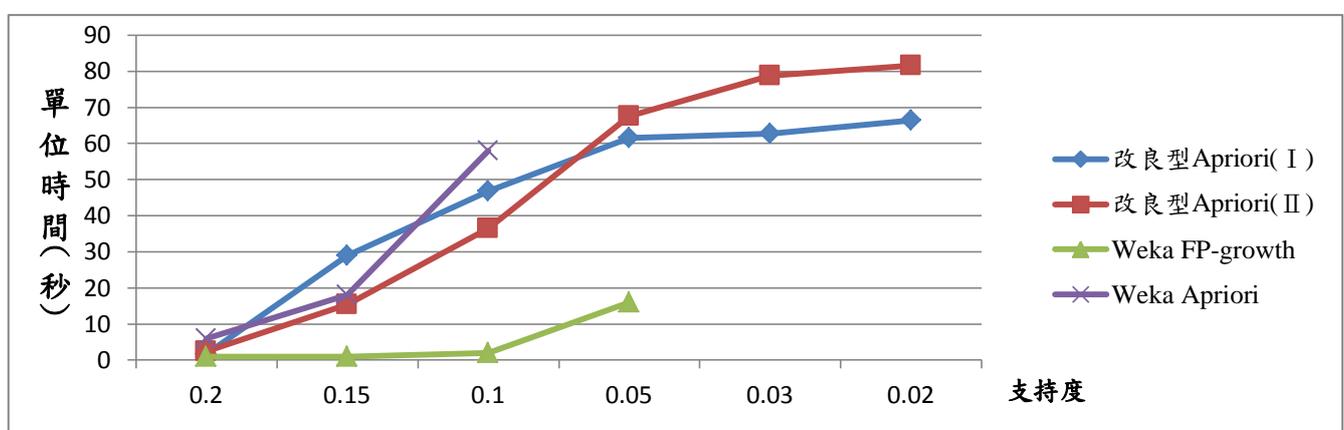


圖 8 演算法效能比較圖(97 年資料集)

從(表 18, 圖 8)得知在最小支持度 0.03 和增益值>1 的情況下, 改良型 Apriori(I)的運算時間較 改良型 Apriori(II) 的運算時間好, 而 Weka FP-growth 和 Weka Apriori 在這情況下, 程式無法回應且無法跑出結果。

4.3 新移民女性醫療衛生需求-關聯規則模型探討

4.3.1 關聯規則模型一

就單一變項而論，2003 年的移民政策尚對外籍配偶的政策有所保留，所以在台資格為「停留」（也就是以探病及探親事由申請在台旅行證）以居住台灣的外籍配偶，對於「協助加入全民健保」的醫療需求的相關程度最高。2004 年 11 月 30 日行政院公布「外籍配偶照顧輔導基金收支保管及運用辦法」，藉由政府預算法合法整合政府及民間資源，設置「外籍配偶照顧輔導基金」，由內政部為主管機關，提供外籍配偶剛入境時期前四個月等待期，無法納入健保的空窗期的醫療費用補助，此一補助範圍包括無健保身份前可申請 5 次產檢費用補助（每次 600 元）、避孕（女性避孕器及男性結紮）的醫療費用等。所以到了 2008 年的調查，協助加入全民健保的需求已降到 1.62%。

醫療費用由於「外籍配偶照顧輔導基金」的協助，降低了新移民女性的就醫門檻，但是值得留意的現象是，居住北部的的新移民女性對於育嬰、育兒及產前產後的衛教指導的需求關聯高，而已居住台灣 4 年至 8 年的新移民女性表達需要政府提供幼兒健康檢查的協助。代表受訪的新移民大多已生育子女，但是因為文化適應及語言溝通的問題，使得她們對於養育下一代出現焦慮，是值得進一步討論的議題。

此外，受訪時來台之婚姻情況為其第二次結婚者，呈現出「提供醫療補助」的高關聯規則需求。是因為第二次結婚來台者，其配偶家庭經濟條件可能臨界於低收入戶及中低收入戶的界線，導致面對醫療環境中諸多自費項目（掛號費、非健保房之病房差額、部分負擔乃至自費醫材），出現對家庭經濟的壓力，所以表達需要醫療補助的需求？也是未來需要再釐清的問題。

表 19 單項社經條件變項之醫療衛生需求關聯規則

年度:2003 年

自變項(X)	依變項(Y)	支持度	信賴度	相關程度
在臺居住時間 (滿四年但未滿六年)	協助加入全民健康保險	0.0432	0.4974	2.1936
與台灣配偶生育子女人數 (1 人)	提供育嬰、育兒知識	0.0883	0.2309	1.4236
與台灣配偶生育子女人數 (2 人以上)	提供幼兒健康檢查	0.0904	0.382	1.4853

年度:2008 年

自變項(X)	依變項(Y)	支持度	信賴度	相關程度
本次婚姻(二次結婚以上)	提供醫療補助	0.0988	0.7382	1.4052
與我國配偶生育子女數(1 人)	提供幼兒健康檢查	0.0768	0.2221	1.3264
在臺居住時間(未滿四年)	提供育嬰育兒知識及產前產後指導	0.0421	0.1845	2.2206

4.3.2 關聯規則模型二

選擇 2 個社經條件為變項建構模型時，在 2003 年時，因為無證件居台及第二次結婚的背景之下來台者，對於加入全民健保的需求相關最高。與模型一的結果相似的是，來自港澳地區及育有 2 子女者，需要政府提供幼兒檢查。

而 2008 年的調查，發現 3 組雙社經條件變項的模型，在臺居住時間(未滿四年)且同意週遭的人對其很友善者對於育兒相關知識的需求高；配偶教育程度為高中、高職畢業含五專 4、5 年級者及與我國配偶生育子女數為 1 人者，需要提供幼兒檢查的協助，顯示本國配偶的教育程度較高及少子女數者，對於兒童健康的意識相對較高；最後，在模型一之二次結婚需要提供醫療補助者，在本模型出現了「配偶教育程度(不識字、自修、小學畢業)」的背景條件，部分應證了第二次結婚而來台，教育程度影響勞動參與能力，所以呈現對醫療費用補助的實質需求。

表 20 雙項社經條件變項之醫療衛生需求關聯規則

年度:2003 年

自變項(X)	自變項(X)	依變項(Y)	支持度	信賴度	相關程度
配偶婚姻 (二次以上)	在臺居住時間 (未滿一年)	協助加入全民健康保險	0.0326	0.6637	2.9272
與台灣配偶生育子女人 數(1 人)	在臺居住時間 (未滿一年)	提供育嬰、育兒知識	0.0354	0.2652	1.6351
國籍別 (大陸地區、港澳地區)	與台灣配偶生育子女 人數(2 人以上)	提供幼兒健康檢查	0.0394	0.4092	1.5911

年度:2008 年

自變項(X)	自變項(X)	依變項(Y)	支持度	信賴度	相關程度
本次婚姻 (二次結婚以上)	配偶教育程度(不識字、自 修、小學畢業)	提供醫療補助	0.0479	0.8358	1.5911
配偶教育程度(高中、高職 畢業含五專 4、5 年級)	與我國配偶生育子女數(1 人)	提供幼兒健康檢查	0.0381	0.2503	1.495
在臺居住時間(未滿四年)	週遭的人對我很友善(同 意)	提供育嬰育兒知識產前 產後指導	0.032	0.2053	2.47

4.3.3 關聯規則模型三

以 3 個社經條件變項為模型時，2003 年時出現了並不意外的結果，即無長期居留資格的二次婚姻來台者，最需要協助加入全民健保。但隨著「外籍配偶照顧輔導基金」政策的推動，2008 年，已不復出現此關聯需求。

而延續模型二的現象，個人生活零用錢主要來源(配偶、配偶父母、子女提供)及與台灣配偶生育子女人數 2 人以上者，不論在 2003 年及 2008 年都是對「提供幼兒檢查」的高需求相關，只是在 2003 年是以大陸港澳地區之外籍配偶為主，到了 2008 年就呈現為對台灣生活滿意者亦是表達需要加強兒童健康檢查的協助。這是值得討論的問題，為什麼新移民女性對於幼兒健康檢查的焦慮如此高？對新移民女性來說，是個人的社會適應問題讓其難以帶小孩接受檢查，還是衛生政策出現盲點？

2008 年，在臺居住時間八年以上、本次婚姻為二次以上婚姻及配偶目前沒有工作者，呈現需要提供醫療補助的高相關性。顯示經濟條件主要受到配偶的工作能力影響。

表 21 三項社經條件變項之醫療衛生需求關聯規則

年度:2003 年

自變項(X)	自變項(X)	自變項(X)	依變項(Y)	支持度	信賴度	相關程度
配偶婚姻 (二次以上)	與台灣配偶生育 子女人數(0 人)	在臺居住時間(未 滿一年)	協助加入全民健 康保險	0.0318	0.6885	3.0367
與台灣配偶生育 子女人數 (1 人)	在臺居住時間(未 滿一年)	移民女性工作狀 況(無工作)	提供育嬰、育兒知 識	0.0312	0.2712	1.6719
國籍別(大陸地 區、港澳地區)	與台灣配偶生育 子女人數(2 人以 上)	個人生活零用錢 來源(配偶或子女 提供)	提供幼兒健康檢 查	0.0313	0.4155	1.6153

年度:2008 年

自變項(X)	自變項(X)	自變項(X)	依變項(Y)	支持度	信賴度	相關程度
在臺居住時間(滿 八年以上)	本次婚姻(二次結 婚以上)	配偶目前是否有 從事工作(沒有)	提供醫療補助	0.0323	0.8591	1.6354
個人生活零用錢 主要來源(配偶、 配偶父母、子女提 供)	與我國配偶生育 子女數(2 人)	整體而言,我在台 灣的生活很快樂 (同意)	提供幼兒健康檢 查	0.0412	0.26	1.553
在臺居住時間(未 滿四年)	個人生活零用錢 主要來源(配偶、 配偶父母、子女提 供)	配偶本次婚姻(首 度結婚)	提供育嬰育兒知 識產前產後指導	0.0306	0.2263	2.723

4.3.4 關聯規則模型四

以 4 個社經條件變項為模型時，2003 年支持度、信賴度及相關程度最高的關聯是「協助加入全民健康保險」需求，而綜合來看，本次婚姻是第二次結婚、來自大陸港澳、尚未生育者及在台時間未滿 1 年者，對此的關聯程度最高。

2008 年的模型則是「提供醫療補助」為最高支持度、信賴度及相關程度的醫療衛生需求。延續關聯規則模型三的結果，除了居住時間八年及配偶沒工作的因素繼續延續之外，還加上是來自大陸港澳地區及配偶教育程度低（不識字）者，對醫療衛生輔導的主要需求是「提供醫療補助」。

2008 年的關聯規則模型提供許多醫療衛生政策的改進想像空間，首先，來自大陸及港澳地區的國籍因素提供對新移民女性的「分眾化」的政策設計思考，而居住時間達八年以上的新移民女性卻在調查「醫療衛生輔導需求」時，仍是表達經濟援助的回應。這樣的結果或許不應單純視為這群新移民女性對醫療費用負擔焦慮，有可能是經濟生活的壓力心理投射所致，把「提供醫療補助」的調查認為是另一種金錢資助的可能來源？

表 22 四項社經條件變項之醫療衛生需求關聯規則

年度:2003 年

自變項(X)	自變項(X)	自變項(X)	自變項(X)	依變項(Y)	支持度	信賴度	相關程度
國籍別(大陸地區、港澳地區)	配偶婚姻(二次以上)	與台灣配偶生育子女人數(0人)	在臺居住時間(未滿一年)	協助加入全民健康保險	0.0304	0.7037	3.1037
與台灣配偶生育子女人數(1人)	個人生活零用錢來源(配偶或子女提供)	家庭生活費用主要來源(本人)	在臺居住時間(未滿一年)	提供育嬰、育兒知識	0.0306	0.2727	1.6811
國籍別(大陸地區、港澳地區)	與台灣配偶生育子女人數(2人以上)	個人生活零用錢來源(配偶或子女提供)	家庭生活費用主要來源(本人)	提供幼兒健康檢查	0.0301	0.4168	1.6205

年度:2008 年

自變項(X)	自變項(X)	自變項(X)	自變項(X)	依變項(Y)	支持度	信賴度	相關程度
國籍別(大陸地區)	在臺居住時間(滿八年以上)	配偶教育程度(不識字、自修、小學畢業)	配偶目前是否有從事工作(沒有)	提供醫療補助	0.0318	0.8774	1.6702
在臺居住時間(未滿四年)	本次婚姻(首度結婚)	個人生活零用錢主要來源(配偶、配偶父母、子女提供)	配偶目前是否有從事工作(有)	提供育嬰育兒知識產前產後指導	0.0325	0.2303	2.7712
個人生活零用錢主要來源(配偶、配偶父母、子女提供)	與我國配偶生育子女數(2人)	我居住的區域少有災害與意外的地方(同意)	整體而言,我在台灣的生活很快樂(同意)	提供幼兒健康檢查	0.0356	0.2726	1.6284

伍、性別影響評估

本計畫參照財團法婦女權益促進發展基金會編印之「性別影響評估操作手冊」的『性別影響評估』7大類目10項原則，並使用『性別影響評估檢視清單』所列29個問題，及『簡易性別相關參考依據』來評估本研究計畫，得到以下結果：

5.1 性別影響評估檢視清單

性別參考依據 ^(註1)	檢視清單	說明
基本條件		
2,3	1.方案規劃發展時，是否徵詢女性的意見？	有。本計畫共同主持人為具醫學與性別研究專長且曾擔任行政院婦權會委員之婦產科醫師，協同主持人為具性別、科技與社會研究專長之醫社系講師，皆為女性。
8	2.方案中是否有納入女性的觀點？	是。本計畫從新移民女性的健康醫療角度切入，瞭解其困境與需求。
2,3	3.方案執行過程中，是否有徵詢女性的意見？	是。本計畫共同主持人及協同主持人皆為女性，共同提供專業領域意見。
1,8	4.方案是否涉及對女性有助益的行業或領域？	是，本計畫主要討論新移民女性之健康權及醫療需求，很明確是對女性有助益之領域。
8	5.方案中是否有任何可能對女性不利的潛藏因素或影響？ 6.若有，如何修改方案以降低這些負面影響？	無。
1,8	7.方案是否包含女性相關的議題與權益？	是，本計畫探討新移民女性之健康權及醫療需求，與女性權益相關。
1,4	8.方案是否蒐集與使用性別分類統計資料或資訊？	是。本計畫蒐集與分析性別分類統計資料，使用內政部「九十二年外籍與大陸配偶生活狀況調查資料庫」與「九十七年外籍與大陸配偶生活需求調查資料庫」，並參考政府相關單位提供之性別統計資料，綜合多方面性別分類統計以進一步提出深度的性別資料庫探勘分析。
2	9.方案是否與女性相關網站有連結？	有。
7	10.方案的預算項目是否有分配給女性相關領域？	有。
7	11.預算規劃是否有含納女性需求的項目？	有。包含收集、會議、訪談、統計資料處理與性別相關之資料，如國際會議、論文、調查資料庫...等。
6	12.審核方案的預算時，是否徵詢女性的意見？	有。本計畫共同主持人及協同主持人皆為女性，提供專業領域意見。

性別參考依據(註1)	檢視清單	說明
1	13.方案是否考慮到對兩性的影響?	有。本計畫探討大陸與外籍配偶婦女在台灣受到的健康照護弱勢現象，這不僅是新移民婦女自身健康問題，更包括兩性平等的家庭、社會體制、健康照護專業及社會支持網絡問題，對兩性影響深遠。
5	14.如何針對女性(婦女團體)宣傳方案的成果?	透過研討會、論文發表及後續研究計畫深化議題，提供政府及婦女團體促進政策推動參考。
9	15.是否會將女性納入結論小組?	是。本計畫共同主持人及協同主持人皆為女性，將在計畫執行會議共同參與及提出專業領域意見。
研究 Studies		
1	24.該研究是否考量到性別?	是。本計畫將研究資料庫的人口社群扣緊在新移民女性，透過資料探勘關聯分析技術指出提高新移民婦女健康的政策設計建議。
1	25.該研究中的女性相關議題是否受到檢視/重視?	是。本計畫透過資料探勘方法找出更適合各分群新移民女性的醫療健康需求，降低社會健康照護政策中的不平等及不利因素。
1	26.此研究是否探討如何提升女性的技能?	是。本計畫分析將建議針對不同社群之新移民女性提出分群的健康照護服務措施及自我健康照護技能，也間接提出新移民女性參與健康照護社會支持網絡的勞動參與機會。
1, 2	27.該研究是否有助於打破女性職業性別隔離障礙?	是。本計畫為跨性別及跨領域(資訊工程、醫學、社會學)進行整合研究，打破職業性別隔離障礙。
資料庫 Databases		
2	29.方案是否考慮到資料庫中的性別分析統計資料和資訊	是。本計畫蒐集與分析性別分類統計資料，使用內政部「九十二年外籍與大陸配偶生活狀況調查資料庫」與「九十七年外籍與大陸配偶生活需求調查資料庫」，並參考政府相關單位提供之性別統計資料，以進一步提出深度的性別資料庫探勘分析。

註 1：

本欄位之數字指符合『性別影響評估』7大類目10項原則。

【參考依據1】指評估項目中的『目標』項目。【參考依據2】指評估項目中的『結合方式』項目。【參考依據3,4】指評估項目中的『方法』項目。【參考依據6,7】指評估項目中的『預算』項目。【參考依據8】指評估項目中的『評估』項目。【參考依據9】指評估項目中的『受益人』項目。

5.2 簡易性別相關參考依據

檢視清單	性別參考依據	說明
在方案設計及執行過程中，有徵詢女性的意見	相關依據 (2) 方法 (3 和 4)	有。本計畫共同主持人為具醫學與性別研究專長且曾擔任行政院婦權會委員之婦產科醫師，協同主持人為具性別、科技與社會研究專長之醫社系講師，皆為女性。
此方案有蒐集並使用性別分類統計資料	目標 (1) 方法 (4)	是。本計畫蒐集與分析性別分類統計資料，使用內政部「九十二年外籍與大陸配偶生活狀況調查資料庫」與「九十七年外籍與大陸配偶生活需求調查資料庫」，並參考政府相關單位提供之性別統計資料，綜合多方面性別分類統計以進一步提出深度的性別資料庫探勘分析。
協助女性循序漸近地參與方案及規劃執行	評估 (8)	是。本計畫共同主持人及協同主持人皆為女性，將在計畫規劃、執行、共同參與會議及研擬專業領域意見。
此方案有助於女性發展	目標 (1) 評估 (8)	是。本計畫分析將建議針對不同新移民女性提出分群的健康照護服務措施及自我健康照護技能，也間接提出新移民女性參與健康照護社會支持網絡的勞動參與機會。
努力找出對女性不利的潛在因素	評估 (8)	是。本計畫探討新移民女性在台灣的健康照護弱勢現象，這不僅是新移民女性自身健康問題，更包括兩性平等的家庭、社會體制、健康照護專業及社會支持網絡問題，對兩性影響深遠。
此方案重視女性	目標 (1)	是。本計畫為以女性為主題的研究，且著重影響新移民女性健康的議題。
女性的參與應納入方案評估項目中	評估 (8)	是。本計畫共同主持人及協同主持人皆為女性，將在計畫規劃、執行、共同參與會議及研擬專業領域意見。
女性能充分獲得方案中的相關資訊	散播、宣傳 (5)	是。主動透過研討會、論文發表及後續研究深化議題，提供政府及婦女團體促進政策推動參考。

從以上兩種檢驗清冊瞭解，本計畫為具有性別意識、對女性友善及促進性別平等的研究計畫。

陸、結論

性別主義的研究指出，各個年齡層都存在性別不平等，而女性特別是因為生命歷程中經歷教育、勞動市場、家庭分工及公共政治參與的不平等，進一步造成惡化老年女性生活陷入不利情境的地位。Choudhury and Leonesio (1997)指出，越來越多老年女性被迫處於貧病交迫，遭受嚴重的社會隔離及資源剝奪，與女性生命歷程中社會、文化及經濟交互不平等的累積有關。

而女性身體醫療化的醫療政策，導致對女性的健康關注著重於女性生殖器官的疾病預防為主，如：子宮相關病變、更年期的身體。這個現象在新移民女性醫療政策更是明顯，由此兩年度的調查問卷設計中看到，對於新移民女性醫療衛生輔導需求仍是圍繞「孕期前後指導」、「避孕」、「育兒」為主，可見對女性老化及生命週期的問題意識較無關懷。

女性經歷長期累積的社會經濟條件基礎的不平等之後（勞動地位、工作條件），老年之後經濟自主能力遠低於男性，女性的生命歷程肩負母職、照顧者、職場工作者的多重角色，若再加上家庭關係改變（成為女性單親戶長或老年喪偶獨居），沈重的身心壓力加上經濟弱勢，這個現象在相對弱勢的新移民女性身上，相信將會更加嚴重。

根據內政部戶政司的統計，女性外籍配偶的比率逐年下降中，站在公民平等健康權的角度，現階段對於新移民女性的醫療健康照護需求應該從「生殖」功能的維繫轉向照護女性的老化生命週期的政策方向。全世界都正處於人口老化的定時炸彈威脅中，尤其已開發國家的人們都越來越長壽，出生率急速下降、性比例失衡，老年女性普遍佔老年人口的多數時，將嚴重衝擊社會組織的運作模式，特別是健康照護體系需求如何滿足老年女性的議題。人口老化產生的社會議題，包括：失能的發生率將繼續攀高，導致對健康及社會經濟支持的服務需求提高；重新思考退休的意義（新的兼納性別意識的勞動及工作概念，老年收入來源的開拓，老年休閒型態）；貧窮的普遍化（弱勢群體因為慢性病，更凸顯生命延長卻經濟困頓的問題）；老化的社會心理議題（孤單，喪親，照顧支持社群存在的迫切性）。

本研究藉由逐步推演的關聯規則模型中發現，老化與性別兩個面向原因結合，再加上新移民女性在社會適應過程的特殊與障礙；新移民女性可能因為年輕來台時，進入的家庭就屬於社會中下經濟階層，加上身份的特殊性，在勞動市場中落入低薪及低工作地位的工作條件，新移民女性年老時更無擁有獲得充足的養老退休金的可能。當女性較男性長壽，加上當初婚姻狀態可能具有男大女小的年齡差距的婚姻型態時，新移民女性將比本國女性更容易陷入獨居老化的情境，經濟條件加速惡化，變成貧病交迫族群的機會很大。本研究建議，從新移民女性健康需求的關聯規則模型、經濟不平等及生命週期(老化)等角度加以思考，依據新移民女性不同特徵族群設計醫療健康政策。

例如：以 2008 年的關連規則模型四來看，針對在臺居住時間未滿四年、本次婚姻為首度結婚、配偶目前有工作、個人生活零用錢主要來源為家人(配偶、配偶父母、子女)提供的族群，所表達的「提供育嬰育兒知識產前產後指導」的醫療需求滿足，及與我國配偶生育子女數 2 人、及同意在台灣的生活很快樂且自認生活住居安全者表達「提供幼兒健康檢查」的結果來看，透露新移民女性未能充分善用全民健保提供新生兒健康照護的服務。這個現象並非全民健保政策的問題，而是如何對新移民女性運用社會資源能力的「賦能」(empowerment)，造成新移民女性能力不足的因素包括語言、經濟能力、社會階層化的問題，因此，建議加強社區支持團體（新移民家庭/女性成長團體、協會）及社區醫療機構（衛生所、診所、地區醫院）的合作，鼓勵新移民女性勇敢走出家庭，進入成長團體交流分享育兒知識及善用全民健保的醫療服務，有效的社會網絡支持團體與醫療機構的合作，對於促進新移民女性健康與育兒能力的自信，將會有相當大的幫助。

資料來源

本論文使用資料來源為內政部「九十二年外籍與大陸配偶生活狀況調查」及「九十七年外籍與大陸配偶生活需求調查」資料庫。該資料由中央研究院調查研究專題中心學術調查研究資料庫釋出，作者感謝上述機構及林宗弘研究員提供資料協助，然本文內容概由作者自行負責。

參考文獻

- Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases.
- Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. paper presented at the Proc 20th Int Conf Very Large Data Bases VLDB
- Berry, M., & Linoff, G. (1997). Data mining techniques: for marketing, sales, and customer support 1997. John Willey & Sons.
- Chen, M. S., Han, J., & Yu, P. S. (1996). Data mining: an overview from a database perspective. *Knowledge and Data Engineering, IEEE Transactions on*, 8(6), 866-883.
- Chu, C. J., Tseng, V. S., & Liang, T. (2008). An efficient algorithm for mining temporal high utility itemsets from data streams. *Journal of Systems and Software*, 81(7), 1105-1117.
- D'Avanzo, C. E., Frye, B., & Froman, R. (1994). Stress in Cambodian refugee families. *Journal of Nursing Scholarship*, 26(2), 101-106.
- Diwan, S., Jonnalagadda, S. S., & Balaswamy, S. (2004). Resources predicting positive and negative affect during the experience of stress: A study of older Asian Indian immigrants in the United States. *The Gerontologist*, 44(5), 605.
- Dunkel, B., & Soparkar, N. (1999). Data organization and access for efficient data mining.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37-54.
- Feyyad, U. (1996). Data mining and knowledge discovery: Making sense out of data. *IEEE expert*, 11(5), 20-25.
- Hall, J. M., Stevens, P. E., & Meleis, A. I. (1994). Marginalization: a guiding concept for valuing diversity nursing knowledge development. *Advances in Nursing Science*.
- Han, J., Kamber, M., & Pei, J. (2011). *Data mining: concepts and techniques (2nd ed.)*: Morgan Kaufmann.
- Han, J., Pei, J., Yin, Y., & Mao, R. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data mining and knowledge discovery*, 8(1), 53-87.
- Hattar-Pollara, M., & Meleis, A. I. (1995). The stress of immigration and the daily lived experiences of Jordanian immigrant women in the United States. *Western Journal of Nursing Research*, 17(5), 521-539.
- Hovey, J. D., & King, C. A. (1996). Acculturative stress, depression, and suicidal ideation among immigrant and second-generation Latino adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(9), 1183-1192.
- Huang, Y. Chao, & C. S. Lin. "An Improved Data Mining Method Applied to the Search of Relationship between Metabolic Syndrome and Lifestyles." *Proceedings of World Academy of Science, Engineering and Technology*. Vol. 60. 2011.
- Khan, M. S., Muyebe, M., & Coenen, F. (2008). A weighted utility framework for mining association rules.

- Li, Y. C., Yeh, J. S., & Chang, C. C. (2008). Isolated items discarding strategy for discovering high utility itemsets. *Data & Knowledge Engineering*, 64(1), 198-217.
- Lien, P., Margaret Conway, M., & Wong, J. (2003). The Contours and Sources of Ethnic Identity Choices Among Asian Americans*. *Social Science Quarterly*, 84(2), 461-481.
- Millar, W., & Stephens, T. (1993). Social status and health risks in Canadian adults: 1985 and 1991. *Health reports/Statistics Canada, Canadian Centre for Health Information= Rapports sur la sante/Statistique Canada, Centre canadien d'information sur la sante*, 5(2), 143.
- Ordonez, C. (2006). Association rule discovery with the train and test approach for heart disease prediction. *Information Technology in Biomedicine, IEEE Transactions on*, 10(2), 334-343.
- Park, J. S., Chen, M.-S., & Yu, P. S. (1995). An effective hash-based algorithm for mining association rules. Paper presented at the Proceedings of the 1995 ACM SIGMOD international conference on Management of data.
- Song, M., & Rajasekaran, S. (2006). A transaction mapping algorithm for frequent itemsets mining. *Knowledge and Data Engineering, IEEE Transactions on*, 18(4), 472-481.
- Zaki, M. J. (2000). Scalable algorithms for association mining. *Knowledge and Data Engineering, IEEE Transactions on*, 12(3), 372-390.
- Zhang, X., Liao, P., & Wang, H. (2009). A New Association Rules Mining Algorithm Based on Vector.
- 內政部. (2003a). 九十二年外籍與大陸配偶生活狀況調查. 台北: 內政部.
- 內政部. (2003b). 九十二年外籍與大陸配偶生活狀況調查【原始數據】。取自中央研究院調查研究專題中心學術調查研究資料庫 <http://srda.sinica.edu.tw>
- 內政部. (2008a). 97 年外籍與大陸配偶生活需求調查結果摘要分析.
- 內政部. (2008b). 九十七年外籍與大陸配偶生活需求調查【原始數據】。取自中央研究院調查研究專題中心學術調查研究資料庫 <http://srda.sinica.edu.tw>
- 內政部. (2011). 近期各月人口統計資料, from <http://www.ris.gov.tw/37>
- 內政部. (2012, 2012/01/07). 一〇一年第一週內政統計通報(100 年結婚及出生登記統計), from http://www.moi.gov.tw/stat/news_content.aspx?sn=5870
- 王秀紅 & 楊詠梅. (2002). 東南亞跨國婚姻婦女的健康. [The Health of Southeast Asian Women in Transnational Marriages in Taiwan]. *護理雜誌*, 49(2), 35-41.
- 王素美, 林育秀, 李卓倫, 陳文意 & 梁亞文. (2011). 台灣新移民女性與醫療衛生需要. *澄清醫護管理雜誌*, 7(1), 19-31.
- 吳瑞堯 & 周駿賢. (2011). 運用資料探勘技術於六大死因慢性疾病之研究. *資訊管理學報*, 18(1), 187-211.
- 吳慧卿、曾明性、林宗弘. (2012). 新移民家庭父母社經地位對子女健康狀況影響因素變遷分析：跨五年資料庫比較研究初探. Paper presented at the 2012 社會變遷與健康發展研討會, 台灣:台北.
- 李瑞金, & 張美智. (2004). 從文化觀點探討東南亞外籍配偶在臺灣之生活適應: 社區發展季刊.
- 周天給, 許怡欣, 李丞華 & 湯澡薰. (2006). 外籍新娘對基層婦產科診所之評價. *臺灣家庭醫學雜誌*, 16(1), 51-63.
- 林淑玲 & 蕭仔伶. (2007). 台灣之女性新移民健康照護的社群隔離現象. [Social Segregation and the Health Care of Female Immigrants in Taiwan]. *護理雜誌*, 54(4), 67-72.
- 廖建彰 & 詹婉瑜. (2005). 東南亞籍與大陸女性配偶生活健康關注. *慈濟護理雜誌*, 4(1), 12-16.
- 戴君倚 & 鍾聿琳. (2002). 影響週產期女性新移民適應之因素及護理省思. *護理雜誌*, 49(1), 34-38.

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：研究內容與原計畫相符程度完全相符、並已達成預期目標。

2. 研究成果在學術期刊發表或申請專利等情形：

論文：已發表 未發表之文稿 撰寫中 無

專利：已獲得 申請中 無

技轉：已技轉 洽談中 無

其他：（以 100 字為限）

本計畫經費無國外差旅費補助，故無國外學術研討會論文發表。本計畫執行過程已發表三篇國內學術研討會論文。學術期刊論文目前撰寫中。

(1) 吳慧卿、曾明性、林宗弘 (2012)，「新移民家庭父母社經地位對子女健康狀況影響因素變遷分析：跨五年資料庫比較研究初探」，社會變遷與健康發展學術研討會，台灣大學。

(2) 吳慧卿、林靜儀、廖芳彬、曾明性 (2013)，「台灣新移民女性醫療健康需求探討：運用關聯規則探勘」，2013 性別與健康研討會，成功大學。

(3) 廖芳彬、吳慧卿、林靜儀、曾明性 (2013)，「改良關聯規則演算法加速台灣新移民女性醫療健康需求之探勘」，第十九屆資訊管理暨實務研討會，台中科技大學。

3. 依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

學術成就方面：本計畫以內政部 92 年及 97 年大陸及外籍配偶生活調查之大型資料庫進行跨年度資料探勘研究，改良關聯規則演算法以提升其執行效能，可快速找出口屬性、社經條件與醫療健康需求之關聯性，嘗試建構出台灣新移民女性醫療健康需求模型。

技術創新方面：本計畫藉由整合 Column-wise、Vector-based、Data Cutting and Sorting Method 等計算方式來減少整體計算量、加速搜尋高頻項目的集合，提出兩種改良型關聯規則演算法，同時探討與資料探勘廣用軟體 Weka 中的 Apriori 和 FP-growth 演算法的效能比較。

社會影響方面：本計畫經由改良型關聯規則演算法建構出台灣新移民女性醫療健康需求模型，並嘗試提出符合大陸及外籍配偶婦女健康權的醫療健康福利政策建議。

國科會補助計畫衍生研發成果推廣資料表

日期:2013/10/31

國科會補助計畫	計畫名稱: 運用關聯規則探勘技術建構台灣新移民女性醫療健康需求模型(GM03)
	計畫主持人: 曾明性
	計畫編號: 101-2629-E-040-001- 學門領域: 性別主流科技計畫
無研發成果推廣資料	

101 年度專題研究計畫研究成果彙整表

計畫主持人：曾明性		計畫編號：101-2629-E-040-001-					
計畫名稱：運用關聯規則探勘技術建構台灣新移民女性醫療健康需求模型(GM03)							
成果項目		量化			單位	備註(質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等)	
		實際已達成數(被接受或已發表)	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比			
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	3	2	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力 (本國籍)	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	1	60%	篇	撰寫中
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		章/本
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力 (外國籍)	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	<p>本計畫經費無國外差旅費補助，故無國外學術研討會論文發表。本計畫執行過程已發表三篇國內學術研討會論文。學術期刊論文目前撰寫中。</p> <p>(1)吳慧卿、曾明性、林宗弘 (2012)，「新移民家庭父母社經地位對子女健康狀況影響因素變遷分析：跨五年資料庫比較研究初探」，社會變遷與健康發展學術研討會，台灣大學。</p> <p>(2)吳慧卿、林靜儀、廖芳彬、曾明性 (2013)，「台灣新移民女性醫療健康需求探討：運用關聯規則探勘」，2013 性別與健康研討會，成功大學。</p> <p>(3)廖芳彬、吳慧卿、林靜儀、曾明性 (2013)，「改良關聯規則演算法加速台灣新移民女性醫療健康需求之探勘」，第十九屆資訊管理暨實務研討會，台中科技大學。</p>
--	---

	成果項目	量化	名稱或內容性質簡述
科 教 處 計 畫 加 填 項 目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與(閱聽)人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文： 已發表 未發表之文稿 撰寫中 無

專利： 已獲得 申請中 無

技轉： 已技轉 洽談中 無

其他：（以 100 字為限）

本計畫經費無國外差旅費補助，故無國外學術研討會論文發表。本計畫執行過程已發表三篇國內學術研討會論文。學術期刊論文目前撰寫中。

(1) 吳慧卿、曾明性、林宗弘 (2012)，「新移民家庭父母社經地位對子女健康狀況影響因素變遷分析：跨五年資料庫比較研究初探」，社會變遷與健康發展學術研討會，台灣大學。

(2) 吳慧卿、林靜儀、廖芳彬、曾明性 (2013)，「台灣新移民女性醫療健康需求探討：運用關聯規則探勘」，2013 性別與健康研討會，成功大學。

(3) 廖芳彬、吳慧卿、林靜儀、曾明性 (2013)，「改良關聯規則演算法加速台灣新移民女性醫療健康需求之探勘」，第十九屆資訊管理暨實務研討會，台中科技大學。

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

學術成就方面：本計畫以內政部 92 年及 97 年大陸及外籍配偶生活調查之大型資料庫進行跨年度資料探勘研究，改良關聯規則演算法以提升其執行效能，可快速找出人口屬性、社經條件與醫療健康需求之關聯性，嘗試建構出台灣新移民女性醫療健康需求模型。

技術創新方面：本計畫藉由整合 Column-wise、Vector-based、Data Cutting and Sorting Method 等計算方式來減少整體計算量、加速搜尋高頻項目的集合，提出兩種改良型關聯規則演算法，同時探討與資料探勘廣用軟體 Weka 中的 Apriori 和 FP-growth 演算法的效能比較。

社會影響方面：本計畫經由改良型關聯規則演算法建構出台灣新移民女性醫療健康需求模型，並嘗試提出符合大陸及外籍配偶婦女健康權的醫療健康福利政策建議。