

# 以群集分析技術探討同縣市醫學中心婦產科病患就醫特性 —以台中市為例

研究生：馮秀卿

指導教授：黃欽印 博士

東海大學工業工程與經營資訊研究所

## 摘要

健保政策丕變，同縣市之醫院競爭愈加激烈，而瞭解醫療市場、病患族群，擬訂正確的經營策略，才能鞏固醫療市場佔有率。台中市醫院及診所充斥，屬醫療市場一級戰區，因此本研究期望能利用健保資料庫，以資料探勘群集分析技術挖掘病患就醫特性等訊息，將資料探勘結果提供醫院擬訂經營策略之參考。

本研究將 93 年 1 月 1 日至 12 月 31 日台中市二家醫學中心婦產科住院病患健保申報資料投入分析，結果發現：一、利用描述性統計分析方法：1.發現婦產科病患年齡層集中於 26-45 歲。 2.住院主診斷以生產最多、其次為癌病化學治療。 二、利用群集分析方法：1.確可依病患特性分為數個族群，包括論病例計酬案件群、婦科疾病群、婦癌重大傷病群等 3 個群集。 2.婦產科重大傷病患者：於婦科疾病群內之主診斷大多為子宮頸惡性腫瘤，主處置以徹底腹式子宮頸根除手術為主；於婦癌重大傷病群內之主診斷及主處置為癌病化學治療或放射線治療。 3.台中市醫學中心婦產科住院病患就醫特性確實有差異：其中 (1)論病例計酬案件群：西屯區醫學中心病患年齡較高，產科病患較多有合併症或併發症；北區醫學中心產科病患較婦科多。(2)婦科疾病群：西屯區醫學中心病患年齡較高，主診斷以子宮頸惡性腫瘤最多，北區醫學中心以子宮頸原位癌最多。(3)婦癌重大傷病群：西屯區醫學中心主診斷以癌病化學治療為主，北區醫學中心以癌病化學治療外，其次為放射線治療。

**關鍵字：**群集分析、婦產科、就醫特性、卡方檢定、k-means

# **A Cluster Analysis on the Patient Characteristics and Treatments of Obstetrics and Gynecology in Medical Centers of the Same Districts - A Study on Two Medical Centers in Taichung City**

Student : Hsiu-Ching Feng

Advisor : Dr. Chin-Yin Huang

Department of Industrial Engineering and Enterprise Information  
Tunghai University

## **ABSTRACT**

As healthy insurance policy changed, the competition between hospitals of the same district becomes more intense. In order to consolidate health care market share, it is important to know the medical market and patient characteristics for developing suitable business strategies. With hundreds of hospitals and clinics, Taichung city has an intense medical market competition. Via data mining and cluster analysis with National Health Insurance Research Database (NHIRD), a study on patients' characteristics and treatments of Obstetrics and Gynecology is presented as a reference for Taichung city's hospitals to develop business strategy.

By analyzing the NHIRD in 2004 for two medical centers in Taichung city, the results are as below. First, by using descriptive analysis, we can conclude that (1) the patients of Obstetrics and Gynecology are concentrated in the 26 to 45 age bracket; (2) the main diagnoses are "normal delivery" and "cancer chemotherapy".

Second, with cluster analysis, we found that: (1) Cluster analysis method is useful to divide patients into three groups, Case Payment Group, Gynecological Disease Group, and Gynecological Cancer Group. (2) The characteristics of major diseases' patients are different between groups. In Gynecological Disease Group, the main diagnoses are "malignant neoplasm of cervix uteri" and the main treatment is "radical abdominal hysterectomy". In Gynecological Cancer Group, The most main diagnoses and treatments are "cancer chemotherapy" and "radiotherapeutic procedure." (3) The characteristics of Obstetrics and Gynecology inpatients appear differences between the two medical centers. There are at least three differences in the study: (a) In Case Payment Group, Situn District medical center has older patients and more complications or merge sickness at its Obstetrics patients. In North District medical center, the amount of Obstetrics patients is more than Gynecology's. (b) In Gynecological Disease Group, Situn District medical center has older patients with "malignant neoplasm of cervix uteri" as its popular main diagnoses. The number of patients with "carcinoma in situ of cervix uteri" is the largest in North District medical center. (c) In Gynecological Cancer Group, more patients of Situn District medical center were treated with "cancer chemotherapy", and patients of the North District medical center were commonly treated with "cancer chemotherapy" and "radiotherapeutic procedure."

**Keywords : Cluster Analysis, Obstetrics and Gynecology, Patient Characteristics and Treatments,  $\chi^2$  test, k-means**

# 目錄

中文摘要.....	
英文摘要.....	
誌謝.....	
目錄.....	
圖目錄.....	
表目錄.....	
第一章 緒論.....	01
1.1 研究背景與動機.....	01
1.2 研究目的與預期效益.....	02
第二章 文獻探討.....	03
2.1 病患就醫特性之相關研究.....	03
2.2 婦產科之醫療特性.....	04
2.3 婦產科病患就醫特性之相關研究.....	05
2.4 資料探勘技術.....	05
2.5 群集分析技術.....	09
2.6 群集分析的步驟.....	10
第三章 研究方法與材料.....	11
3.1 研究流程.....	11
3.2 研究測量變項.....	11
3.3 研究工具與方法.....	14
第四章 研究結果.....	16
4.1 樣本資料之描述性統計分析.....	16
4.2 群集分析結果.....	22
4.2.1 群集特徵分析.....	22
4.2.2 醫院別群集分析.....	31
第五章 研究討論與結論.....	45
5.1 討論.....	45
5.1.1 台中市醫學中心病患就醫特性.....	45
5.1.2 群集分析.....	47
5.2 結論與建議.....	51
5.2.1 群集前.....	51
5.2.2 群集後.....	51
5.3 研究限制.....	54
參考文獻.....	55

# 圖目錄

圖 3.1 研究流程圖.....	11
圖 3.2 資料探勘及分析流程圖.....	15
圖 4.1 群集分佈圖.....	22
圖 4.2.1 群集特徵長條圖-醫院別 .....	26
圖 4.2.2 群集特徵長條圖-案件分類.....	26
圖 4.2.3 群集特徵長條圖-年齡層 .....	26
圖 4.2.4 群集特徵長條圖-給付類別 .....	26
圖 4.2.5 群集特徵長條圖-住急性病床天數 .....	26
圖 4.2.6 群集特徵長條圖-轉歸情形.....	26
圖 4.2.7 群集特徵長條圖-部份負擔.....	27
圖 4.2.8 群集特徵長條圖-DRGs.....	28
圖 4.2.9 群集特徵長條圖-主診斷 .....	29
圖 4.2.10 群集特徵長條圖-主診斷補充分類 .....	29
圖 4.2.11 群集特徵長條圖-主處置 .....	30
圖 4.3.1 群集 1 內醫院別與年齡比較圖.....	31
圖 4.3.2 群集 1 內醫院別與住急性病床天數比較圖.....	32
圖 4.3.3 群集 1 內醫院別與案件分類比較圖.....	33
圖 4.3.4 群集 1 內醫院別與給付類別比較圖.....	33
圖 4.3.5 群集 1 內醫院別與轉歸情形比較圖.....	33
圖 4.3.6 群集 1 內醫院別與部份負擔類別比較圖.....	33
圖 4.3.7 群集 1 內醫院別與 DRGs 比較圖.....	35
圖 4.3.8 群集 1 內醫院別與主診斷比較圖.....	36
圖 4.3.9 群集 1 內醫院別與主處置比較圖.....	36
圖 4.4.1 群集 2 內醫院別與年齡比較圖.....	37
圖 4.4.2 群集 2 內醫院別與住急性病床天數比較圖.....	37
圖 4.4.3 群集 2 內醫院別與案件分類比較圖.....	38
圖 4.4.4 群集 2 內醫院別與給付類別比較圖.....	38
圖 4.4.5 群集 2 內醫院別與轉歸情形比較圖.....	39
圖 4.4.6 群集 2 內醫院別與部份負擔類別比較圖.....	39
圖 4.4.7 群集 2 內醫院別與主診斷比較圖.....	40
圖 4.4.8 群集 2 內醫院別與主處置比較圖.....	40
圖 4.5.1 群集 3 內醫院別與年齡比較圖.....	41
圖 4.5.2 群集 3 內醫院別與住急性病床天數比較圖.....	42
圖 4.5.3 群集 3 內醫院別與轉歸情形比較圖.....	43
圖 4.5.4 群集 3 內醫院別與部份負擔類別比較圖.....	43
圖 4.5.5 群集 3 內醫院別與主診斷 vcode 比較圖.....	44
圖 4.5.6 群集 3 內醫院別與主處置比較圖.....	44

# 表目錄

表 2.1 國內醫療產業資料探勘相關研究.....	07
表 2.2 國外醫療產業資料探勘相關研究.....	08
表 2.3 群集分析方法比較.....	10
表 3.1 研究變項之操作型定義及變項轉換說明.....	12
表 4.1 病患就醫屬性之描述性分析.....	17
表 4.2.1 DRGs ( 診斷關聯群項目 ) .....	19
表 4.2.2 主診斷 ( ICD-9-CM ) .....	20
表 4.2.3 主診斷補充分類 ( ICD-9-CM Vcode ) .....	21
表 4.2.4 主處置 ( ICD-9-CM ) .....	21
表 4.3.1 群集特徵及卡方檢定表-病患就醫屬性 .....	24
表 4.3.2 群集特徵及卡方檢定表-DRGs ( 診斷關聯群項目 ) 群集 .....	27
表 4.3.3 群集特徵及卡方檢定表-主診斷 ( ICD-9-CM ) 群集 .....	28
表 4.3.4 群集特徵及卡方檢定表-主診斷補充分類 ( ICD-9-CM Vcode ) 群集 .....	29
表 4.3.5 群集特徵及卡方檢定表-主處置 ( ICD-9-CM ) 群集 .....	30
表 4.4.1 群集 1-t 檢定-年齡及住急性病床天數.....	31
表 4.4.2 群集 1-卡方檢定-病患就醫屬性.....	32
表 4.4.3 群集 1-卡方檢定-DRGs ( 診斷關聯群項目 ) .....	34
表 4.4.4 群集 1-卡方檢定-DRGs ( 診斷關聯群項目 ) -產科項目 .....	34
表 4.4.5 群集 1-卡方檢定-主診斷 ( ICD-9-CM ) .....	35
表 4.4.6 群集 1-卡方檢定-主處置 ( ICD-9-CM ) .....	35
表 4.5.1 群集 2-t 檢定-年齡及住急性病床天數.....	36
表 4.5.2 群集 2-卡方檢定-病患就醫屬性.....	37
表 4.5.3 群集 2-卡方檢定-主診斷 ( ICD-9-CM ) .....	39
表 4.5.4 群集 2-卡方檢定-主處置 ( ICD-9-CM ) .....	40
表 4.6.1 群集 3-t 檢定-年齡及住急性病床天數.....	41
表 4.6.2 群集 3-卡方檢定-病患就醫屬性.....	42
表 4.6.3 群集 3-卡方檢定-主診斷補充分類 ( ICD-9-CM Vcode ) .....	43
表 4.6.4 群集 3-卡方檢定-主處置 ( ICD-9-CM ) .....	43
表 5.1 論病例計酬案件之 DRGs 類別分佈表.....	46
表 5.2 群集 1 西屯區醫院給付類別與 DRGs 產科項目交叉分析表.....	49