

以資料圖形化的方式呈現台灣家庭醫學 雜誌近五年的文章

江瑞坤¹ 高以信²

科技的精進使得資訊工具的發達與便利，讓人們很容易獲取資訊。然而從大量的資訊中如何有效率地獲得對自己有幫助的資訊，則是大家所關心的議題。資料圖形化方法，不僅是一種技能的提升更是一種態度的改變，對於哪些主題或關鍵字常被提到，常是研究的熱點。

本研究以2011-2015年間台灣家庭醫學雜誌所刊登的所有文章為研究對象。藉由文本資料探勘圖形化的方法，期望能幫忙學者及讀者在短時間內能「看到」台灣家庭醫學雜誌的風貌是本文的目的。本文呈現對標題、摘要、作者所屬的機構科部間、所屬單位、及跨院單位合作網絡圖的分析。

(台灣家醫誌2017; 27: 27-31) DOI: 10.3966/168232812017032701004

關鍵詞：資料圖形化、台灣家庭醫學雜誌

前 言

科技的精進使得資訊工具的發達與便利，讓人們很容易獲取資訊。然而從大量的資訊中如何有效率地獲得對自己有幫助的資訊，則是大家所關心的議題。教學醫院有三大功能：服務、教學、及研究。之前的學者在資料收集所需耗費大量的時間，經費與人力來作研究，如今可以有效率的解決此問題。

台灣家庭醫學雜誌是1991年3月由台灣家庭醫學會創辦，以季刊形式出版，平均每年出版約20篇文章，其文章類別包含原著論文、綜論、專論、個案報告

等。台灣家庭醫學雜誌近五年出版的文章數量如圖1（2015年只算到第3季）。隨著時間增長導致文章的累積，要想一窺其奧秘實在需要有快速方法。

資料圖形化方法，不僅是一種技能的提升更是一種態度的改變，對於那些主題或關鍵字常被提到，常是研究的熱點。進一步可將這些關鍵字出現的頻率轉變成圖案呈現「文字雲」，可讓讀者一目了然。藉由文本資料探勘圖形化的方法，期望能幫忙學者及讀者在短時間內能「看到」台灣家庭醫學雜誌的風貌是本文的目的。

¹佛教大林慈濟醫院家庭醫學科、²台南市立醫院家庭醫學科

受理日期：105年3月20日 修改日期：105年6月20日 同意刊登：105年8月7日

通訊作者：高以信

通訊地址：台南市東區崇德路670號

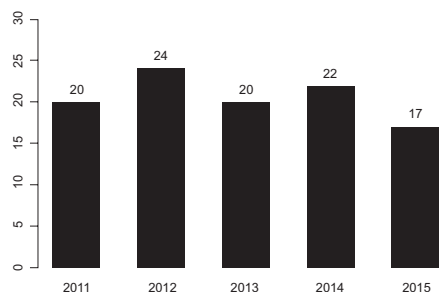


圖1 台灣家庭醫學雜誌出版的文章數量，2011-2015年（2015年只算到第三季）

材料與方法

從家庭醫學醫學會網站，以2011-2015年間台灣家庭醫學雜誌所刊登的所有文章為研究對象，共收集103篇文章。本文使用統計軟體R，文字雲分析時使用的套件包括jiebaR、wordcloud，繪製網絡圖時要外加使用套件plyr、stringr、gtools、及igraph，分析文字資料。套件jiebaR內，可以建立此次分析的「字典」及stop words，而且最終的結果可以再次以人工審視，將不要的詞放入stop words內，重新分析，並可依照關鍵詞的頻率來調整字體大小及顏色標示文字雲及繪製網絡圖^[1,2]。

結 果

對標題的分析

在台灣家庭醫學雜誌標題的分析可見其多樣性，出現「報告」19次、「相關」18次、「病例」17次、「因子」12次、「探討」12次、「研究」10次、「分析」9次，其餘為7次以下。文字雲呈現如圖2。由於台灣家庭醫學雜誌每一期都有病例報告類別的文章，因此「報

告」、「相關」及「病例」這三個詞出現最多次。其次關於原著論文類別的標題則常見相關因子探討這些字，導致「因子」及「探討」這兩個詞也是很常見。

對摘要的分析

將2011-2015年台灣家庭醫學雜誌內容摘要作詞之數目分析，得到以下的次數分布表，病人（包含患者、病患）共出現162次最多，其次為相關（因子），治療，男性及醫師。其中肥胖的議題出現55次，糖尿病（45次）、腦梗塞（38次）也是熱門的疾病主題。我們將有關病人的48篇文章做進一步探討，其中文章類別以個案報告20篇(41.7%)最多，其次為原著論文18篇(37.5%)。文字雲呈現如圖3。



圖2 台灣家庭醫學雜誌標題的分析，2011-2015年



圖3 台灣家庭醫學雜誌摘要的分析，2011-2015年

對作者所屬的機構及科部之間的分析

將2011-2015年台灣家庭醫學雜誌投稿作者之機構的次數，以成大（包括成功大學，成大醫學院，成大醫院，成大斗六分院）的39次最多，其次為馬偕（包括馬偕紀念醫院，馬偕紀念醫院新竹分院）23次，台大（包括台灣大學，台大醫學院，臺大醫院，及台大醫院金山分院）19次，中國（包括中國醫藥大學，中國醫藥大學附設醫院，及中國醫藥大學附設醫院台北分院）17次等。值得一提的是，基層醫療院所吉豐診所所有7次投稿。將2011-2015年台灣家庭醫學雜誌將3次及以上次數的機構以文字雲方式呈現如圖4。

台灣家庭醫學雜誌作者單位最多為家庭醫學科（部）118次，其次為醫學院11次、耳鼻喉科9次、醫學系9次、小兒科7次、社區醫學7次、內科6次、神經科6次，其文字雲如圖5。有95篇(92.2%)文章作者的單位有家庭醫學科（部）。本文將標題、摘要、及作者所屬機構的關鍵字中頻率排名前十名者以表格呈現（表1）。

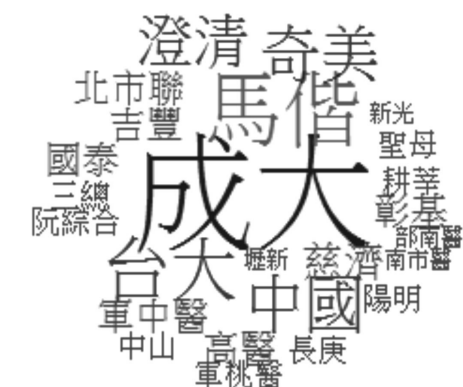


圖4 台灣家庭醫學雜誌作者所屬機構的分析，2011-2015年

對作者跨院及科部連結網絡圖

由於一篇文章的完成常需要多專科的合作，也可能作者身兼數職（例如兼職，在職進修等），以及文章常需要指導老師的教導等，造成一篇文章有多位作者，及多個機構。藉由資料圖形化方法，以「網絡圖」的方式可以清楚地呈現，那些機構科部之間常相互一起研究，但是同單位內則未呈現。本研究僅將機構科部之間的網絡呈現，但對於作者則限於隱私的考量並未呈現。本文發現地域性佔有相當大的重要性（例如台南市立醫院較易與奇美，成大合作），



圖5 台灣家庭醫學雜誌作者所屬單位的分析，2011-2015年

表1 標題、摘要、及作者所屬機構的關鍵字中頻率排名前十名者

標題		摘要		作者所屬機構	
關鍵字	頻率	關鍵字	頻率	關鍵字	頻率
病人	162	報告	19	成大	39
相關	100	病例	17	馬偕	23
治療	82	因子	12	台大	19
男性	75	探討	12	中國	17
醫師	72	相關	18	奇美	16
疾病	69	研究	10	澄清	15
肥胖	55	分析	9	慈濟	9
檢查	55	表現	8	北市聯	7
女性	50	危險	6	吉豐	7
糖尿病	45	家庭醫學	6	軍中醫	7

未來如何突破地域性做更好的跨院科系連結與整合則是另一項挑戰。其跨院及科部連結網絡圖如圖6。

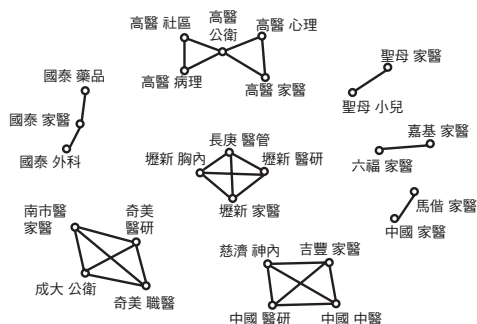


圖6 台灣家庭醫學雜誌作者跨院及科部連結網絡圖的分析，2011-2015年

標題內的關鍵詞是一篇文章中最精髓的詞彙，藉由圖書或網路上關鍵詞的搜尋，讀者很容易找到所需要的文獻，以及不同期刊所發表文章的重點。

摘要則是將一篇文章的內容以精簡的方式呈現出，讓讀者能在最短的時間得知文章的大意。本文藉由對摘要的分析，可以粗略知道台灣家庭醫學雜誌在2011-2015年發表的文章，以病人接受治療之相關因子較多，病人的疾病以肥胖、糖尿病及腦梗塞為最多。

關於台灣家庭醫學雜誌投稿作者的屬性，本文發現作者所屬的機構以醫學中心最多，基層醫療院所也不少。對於作者跨院及跨科部連結網絡圖可以知悉作者群之間常有地域性的特質，以及縱向的聯繫。由此可知台灣家庭醫學雜誌是一本兼具學術與實務的優良雜誌。

結 論

將2011-2015年間台灣家庭醫學雜誌所刊登的所有文章以標題，摘要，及作者所屬的機構及科部之間的關聯性分別以文字雲及網絡圖的方式呈現，使讀者能一窺本雜誌這五年來的全貌。

參考文獻

1. Wenfeng Qin. Package 'jiebaR'. <https://cran.r-project.org/web/packages/jiebaR/jiebaR.pdf>. Accessed June 17, 2016.
2. Oesper L, Merico D, Isserlin R, Bader GD. WordCloud: a Cytoscape plugin to create a visual semantic summary of networks. *Source Code Biol Med* 2011; 6: 7.

Visualization of Papers Published in Taiwan Journal of Family Medicine from 2011 to 2015

Jui-Kun Chiang¹ and Yee-Hsin Kao²

Established in March 1991 by the Taiwan Association of Family Medicine, the *Taiwan Journal of Family Medicine* (TJFM) publishes original papers, review articles, monographs, and case reports on a quarterly basis, leading to an annual publication of approximately 20 articles. It takes a modern approach to obtain in an effective and efficient manner needed information and helpful data from all the papers published by the journal during the past decades.

Based on the concept of data visualization, the article aims to utilize text mining to present TJFM papers in the form of visual pictorials. The frequencies of key words were calculated and showed as "word clouds" and network map links that allowed readers to understand the status of TJFM-published papers at a glance.

One hundred and three articles published in TJFM from 2011 to 2015 were collected and analyzed in terms of titles, abstracts, network map links between institutions and departments. The top 3 most frequently seen words in article titles were "report", "association", and "case", and those in abstracts were "patient", "association", and "treatment". National National Cheng Kung University appeared to be the institute hosting the greatest number of article authors. The network map links showed that authors tended to collaborate with colleagues in institutes located near their own. With the results of the data visualization, readers can easily know the most used words in the papers published in *Taiwan Journal of Family Medicine* during the last five years.

(*Taiwan J Fam Med* 2017; 27: 27-31) DOI: 10.3966/168232812017032701004

Key words: data visualization, Taiwan Journal of Family Medicine

¹Department of Family Medicine, Buddhist Dalin Tzu Chi Hospital, Chiayi; ²Department of Family Medicine, Tainan Municipal Hospital, Tainan, Taiwan.

Received: March 20, 2016; Revised: June 20, 2016; Accepted: August 7, 2016.