

中醫症狀標準化過程的研究

林劉育¹、劉建宏¹、陳仁義²、葉家舟^{3,4}、林迺衛¹、葉明憲³

¹國立中正大學資訊工程學系 ²南華大學資訊管理學系

³佛教慈濟綜合醫院大林分院中醫科 ⁴慈濟大學醫學系

摘要

中醫診斷分為望、聞、問、切(通稱四診)，透過四診得到病患所有的資訊，醫生藉由這些資訊去分析、診斷及處方用藥，此過程就是辯證。因此，症狀是疾病診斷的重要依據^[1]，然中醫發展悠久，諸多因素導致症狀描述不盡相同，如能將中醫症狀描述做標準化處理，對中醫辯證以及科學化分析有很大的助益^[2]。而且經由對於標準化症狀反復的探討及修正，可以發現症狀與症狀的組合規律，形成中醫的證候及病因病機的主要內容。

本研究使用《中醫證候學》^[3]的症狀描述做為標準化對象。書中對於證候、證象以及症狀的關係有完整的架構呈現。本實驗設計一套症狀標準化流程，且提出驗證標準化症狀的方式，透過症狀分群去分析證象的關聯性，觀察分群結果，修正標準化錯誤的症狀，重新分群直到結果符合書中的證象分群。以便為往後的專家辨證系統奠下最重要

的基礎。

關鍵詞：中醫、症狀標準化、證象、證候

前言

辯證就是分析、辨識疾病本質的過程。實際上是以臟腑、經絡、病因、病機等中醫基本理論為指導，通過對四診所獲取之臨床資料進行分析、歸納、推理，從而做出證候診斷結論的過程^[7]。在《臨證指南醫案》提到：「醫道在乎識證，立法，用方，此為三大關鍵。」「然三者之中，識證為要緊。」。

透過四診所得之症狀是構成證候的主要元素，而症狀又是診斷證候的主要根據，因此在辯證候之前應該先了解這些症狀本質才能認識證候病機。辨識症狀本質及證候病機，才是中醫臨床辯證最核心最關鍵的問題^[8]。

中醫症狀標準化的過程中，許多症狀有一詞多義的特性，主要分兩種情況：「一個症狀有兩個或多個內涵」及「症狀的內涵模糊，具體所指不明確」。前者如「眩暈」，能指頭

暈、頭昏或者是目眩，也是並見的症狀；後者如「面白」，可表示面色淡白、面色白、面色蒼白等。再來是多詞一義的問題，例如「便溏」、「大便溏薄」、「大便稀溏」、「大便溏」、「大便溏爛」、「大便稀薄」、「大便稀」、「大便多不成形」等；又如「精神疲乏」、「精神倦怠」、「精神困憊」、「精神困倦」、「精神困頓」、「神倦」、「神疲」等。而這些表述均出自於教科書、專家專著、政府制訂的有關標準等較為權威的書籍中。

本研究依照《中醫證候學》書中所描述的結構，將疾病門、證類、證、證候、證象、症狀的關係建立資料庫，透過標準化這些資料，產生出問診表來記錄病歷以及做統計分析。但在完成這些資料庫結構時，我們發現“症狀標準化”是非常重要的環。若症狀未有標準化則無法對資訊來源及其本質進行分析。

在 2008 年張啓明等人提出了症狀單元的概念，將每一個症狀單元由一個症狀部位和一個症狀性質所組成^[9]。將中醫症狀名稱科學化，定了一些規則^[5]：

- (一)單義性,一個術語只反映一種或一類具體事物的特有屬性。
- (二)準確性,術語的內涵(特有屬性)界定明晰而充分,使交流各方不會誤解。
- (三)統一性,不同學科不同民族語言之間的相關術語的內涵是統一的。

本研究依循此原則在單義，準確性和統一性上與張啓明等人一致，但在實行的程序上採用條件機率分析，以逐步修正標準化的症狀。避免陷於標準化症狀過於龐雜的缺點。科學化的標準不統一是目前症狀標準化研究中存在的主要問題^[6]。在標準化症狀的

過程中，須參考許多症狀的定義，標準化後的症狀才比較具有規範。^[10-17]

材料與方法

(一)資料來源:依據“中醫證候學”的 75 個證象之 1294 筆資料的 3516 個片段症狀描述。

(二)標準化原則

1. 標準化原則

本研究的症狀標準化過程，針對上述會依以下的原則進行標準化。

- (1)每個症狀描述將會分解成一個症狀單元且只具有一種內涵^[4]。每一個症狀單元只能有一種內涵，將只具有一種內涵的症狀單元做標準化，便可確認沒有其它標準化症狀跟此標準化症狀內涵相同，此標準化症狀具有獨立性，跟單一性且有意義。
- (2)每一個症狀單元以階層式結構來呈現分類跟屬性以及其屬性值^[5]。將症狀單元依內涵分成數個類別，每個類別都有屬性以及其屬性值，透過階層式架構去呈現症狀單元，可清楚知道症狀之間的關聯性，並可做簡單的分群。
- (3)將屬性做量化處理^[6]。在某些症狀單元具有連續性或者是程度上的不同，如疾病發生時間:下午到傍晚，或者是尿的量多，這些症狀單元的屬性，都可以定義之後量化，例如將時間切割成清晨，中午，傍晚，夜間等，使得時間有自己量化後的屬性值。
- (4)內涵相同的屬性描述合併成相同的屬性描述^[6]。在標準化的過程中每個

症狀單元所使用的屬性描述，可能會不同，如小便顏色和面色，都是指顏色這個屬性，在整合的過程中，會將相同含意的屬性合併在一起，並重新量化其屬性。

- (5)內涵相同的症狀描述句必須被標準化成相同的標準化症狀^[5]。此原則是本研究做標準化最根本的條件，因為中醫對同一症狀的用詞，會因許多原因而不一致。除了將這些描述確認是否為單一內涵的症狀單元之外，還要將有相同內涵的症狀描述，標準化成相同的標準化症狀。

在標準化過程中會參考相關研究，將標準化後的結構定為「症狀的狀態：描述句」。如表 1 中的「小便淋秘」，實際情況是指無法小便。在標準化的過程，會將其定義成「小便的情形:不通暢」。在階層化的過程，會加入症狀的分類以及屬性跟其屬性值，並合併討論。依照內涵決定其屬性值，產生階層式標準化如「排泄物:小便:流利情形:無尿出」。排泄物以及小便是症狀的分類，流利情形是屬性，無尿出是流利情形的屬性值，大致架構為「大分類:小分類:屬性:屬性值」，在這樣的架構底下去標準化中醫症狀，以便於將症狀定義釐清。

表 1 原始症狀標準化

原始症狀	標準化症狀	階層式標準化症狀
小便淋秘	小便的情形:不通暢	排泄物:小便:流利情形: 無尿出
小便不行	小便的感覺:無法小便	排泄物:小便:流利情形: 無尿出

2. 標準化後分群的原則：

將症狀標準化後，利用條件機率將症狀分群，分群的目的是在自動化辨出證象。方式有合併及分群兩步驟。症狀可以出現在許多證象裡面，一個證象內都有一群症狀描述，於是合併的過程就分成以下三點，將有關聯的症狀連結在一起。

- (1)每一個症狀自成一群。
- (2)同一個證象下的症狀視為同一群症狀群。
- (3)若群與群的症狀有交集則合併，若群和群之間再也無共同症狀，則中止合併。

依出現機率由低至高將症狀刪除，若刪除症狀可使症狀群分群，觀察症狀群可否再分，不行則分出此群；可以分群則繼續刪除症狀，直到所有症狀都刪除或完成症狀分群為止。

完成上述合併及分群的方式，產生原始分群結果，觀察原始分群結果，列出沒有類似病機的證象群，證象群個數超過兩個的會繼續分群，證象群個數剛好兩個的則依下列步驟進行修正，修正後再重新產生結果並觀察修正後的改變。

3. 症狀分群後的檢驗原則：檢查有否辨證錯誤。

(1)證象的特徵

列出分群錯誤的證象內出現次數較高的標準化症狀，驗證此群的共同症狀是否為每個證象內出現次數較高的標準化症狀，驗證符合表示此症狀為此群的主要共同症狀，反之表示此

症狀出現次數低不是此群的主要共同症狀，進行下一步驟檢驗。

(2)原始症狀的檢驗

將此群的共同症狀的原始症狀列出來，觀察原始症狀的內涵是否和共同症狀的內涵一致，觀察結果一致表示此分群結果正確，此分群證象繼續分群，反之則修正標準化症狀。

(3)修正標準化症狀

修改或新增每個階段的對照表內容，修改對照關係錯誤的表格，新增無此對照關係的表格，產生新的問診式標準化症狀。

(三)條件機率分析

條件機率所要傳遞的訊息是：在縮小了樣本空間(嘗試以一個較為熟悉的事件)的情況之下來重新檢視我們所關心事件的發生機率。例如，我們主要關心的事件為A！如果在相同樣本空間的事件B，我們也較為瞭解其相關性質！則會嘗試來探索事件B所擁有的新資訊是否可帶給我們進一步瞭解主要關心事件A的發生機率。可表為條件機率的型式：

$$P(A|B) \quad (1)$$

其意義是「在事件B已經發生的情況之下，事件A會發生的機率」。運算型式可轉換成爲

$$P(A | B)=P(A \cap B)/P(B) \quad (2)$$

相當於樣本空間已縮小爲在事件B的範圍中(分母部份)、主要關心的事件A則縮小爲兩者重疊的交集部份(A ∩ B；分子部份)。

在本研究中使用條件機率做爲疾病發生頻

率的計算，在一個資料庫D內，令資料庫內證象集合 $X: \{X_1, \dots, X_n\}$ 以及症狀集合 $Y: \{Y_1, \dots, Y_m\}$ ，則資料庫一筆記錄 D_a 表示爲：

$D_a: (X_i, \{\dots, Y_j, \dots\})$ 。在資料庫D中，證象 X_i 出現的機率由(1)知 $P(X_i|D)$ ，由(2)知機率可爲：

$$P(X_i|D)=P(X_i \cap D)/P(D)$$

分母部分因爲D是爲樣本空間， $P(D)$ 出現機率爲1，分子的部份在集合D內的元素，有出現 X_i 元素的機率，因此證象 X_i 出現的機率可以爲：

$P(X_i)$ 同理，症狀 Y_j 出現機率爲： $P(Y_j)$ 考慮在某個證象發生的條件下，症狀出現的可能透過(2)爲：

$$P(Y_j | X_i)=P(Y_j \cap X_i)/P(X_i)$$

意思是在所有包含 X_i 元素的集合 D' ，元素 Y_j 出現的機率，表示著在出現此證象的資料庫內，症狀出現的機率，本研究稱此條件機率高者爲此證象的主要症狀，利用這個機率值去做分群的分析。

(四)症狀標準化的程序

症狀標準化過程會分成五個階段，分割階段、標準化階段、階層化階段、問診化階段以及問診表。症狀做標準化的過程會按下列的流程。

1. 將原始症狀描述拆成有意義的片段，一個原始症狀的描述，是一段症狀描述，中間有句逗號分隔，我們先以句逗號做區分，將原始症狀描述拆成數段可表達的完整的片段症狀描述。
2. 接著將片段症狀描述依照本身的內涵，標準化成具有此內涵的標準化症狀，此

時的標準化症狀已具備是單一內涵的症狀單元。

3. 接著建立每個標準化症狀的分類及屬性，用階層式症狀去呈現，階層式症狀表示症狀的分類、屬性以及其屬性值，此階層式症狀已經符合所有標準化原則。
4. 再進行問診化的步驟，將階層式症狀的分類以及屬性做合併以利於中醫問診的過程，產生問診式症狀。
5. 將問診式症狀以問診表形式呈現出來。每一標準化階段中，都有一份一對一的對照表記錄著不同形式的症狀間的轉換，以利於新增症狀資料或者是調整標準化的分類跟屬性。

結果與討論

一、標準化過程及結果

(一)分割

將中醫證候學中的 75 證象之 1294 筆資料分別輸入並開始做切割，一筆證象的描述先依句點為分割點對每一段症狀描述做切割，再拆解成適合標準化的片段如下表 2。

表 2 原始症狀切割

原始症狀描述	片段症狀描述
小便短澀不利，尿頻	小便短澀不利
量少，色黃渾濁，色	尿頻量少，色黃渾濁，色如濃茶
如濃茶。小便澀痛，	小便澀痛
小便淋秘，尿道灼	小便淋秘
熱，點滴若難下，尿	尿道灼熱，點滴若難下
則號痛不已，少腹脹	尿則號痛不已，少腹脹滿難堪
滿難堪。腹脹，腹痛，	腹脹
二便虛急。小便癱	腹痛
閉。	二便虛急
	小便癱閉

在這個階段，原本在 75 個證象，1294 筆證象資料的症狀被切割成 3516 個片段症狀。

(二)標準化

標準化的目的就是將同義的不同中醫術語，轉換成相同的描述，以便記錄在資料庫中進行分析。例如「小便短澀」和「小便不利」的內涵就是指小便不通暢。但對電腦而言，這兩者是無法歸類在一起的，而轉換成「小便的感覺:不通暢」這樣的描述就可以。所以在參考專業及文獻的描述後，當標準化一筆資料後，再遇到相同的症狀描述，電腦即可透過資料庫自動產生標準化的症狀。

在標準化的過程中，定義了一個標準化症狀的架構，將症狀描述標準化成「症狀的狀態:描述句」。如下表 3 在此階段症狀已被標準化為單一內涵的症狀。但症狀之間的關聯性仍不清楚，無法對這症狀做分析。

表 3 標準化症狀

片段症狀	標準化症狀
小便短澀不利	小便的流利度:不流利，排尿困難
	小便的量:少
	小便的感覺:澀痛
尿頻量少，色黃渾濁，色如濃茶	小便的色:如濃茶
	小便的色:黃混濁
	小便的量:少
小便澀痛	小便的頻率:數
	小便的感覺:澀痛
小便淋秘	小便的流利度:不流利，排尿困難
	小便的量:少
	小便的感覺:澀痛
尿道灼熱，點滴若難下	小便的感覺:不流利，排尿困難
	小便的感覺:澀痛

片段症狀有 3516 個，在這個階段結束

後，產生了標準化症狀共有 3087 筆，表示有部分不同的症狀因為同義而被合併。符合症狀標準化的原則。

(三)階層化

在階層化階段會將所有標準化症狀分成數大類，每個大類之下再細分數小類，每一類都有其屬性及其相對應的屬性值，透過階層式架構可將一個證象內的症狀分類。透過此方法可以將兩者的關係，做進一步的分析，此階段會分為「歸類」以及「屬性分層」兩步驟去將標準化症狀轉換成階層式標準化症狀，簡稱為階層式症狀如表 4。

表 4 階層式症狀

標準化症狀	階層式標準化症狀
小便的情形:不通暢	二便:小便:情形:不通暢:不通暢
小便的量:少	二便:小便:量:少:少
小便的感覺:澀痛	二便:小便:感覺:澀痛
小便的色:如濃茶	二便:小便:顏色:紅赤:如濃茶
小便的色:黃混濁	二便:小便:顏色:黃:黃混濁
小便的頻率:數	二便:小便:頻率:很多次:數
灼熱的部位:尿道	寒熱:灼熱:部位:腰以下半身:尿道

在進行階層化後會使原本症狀之間的關聯性大幅提高，可知道一個證象內的症狀分布情形，原本標準化症狀共有 3087 筆資料，經由階層化後產生 2638 筆階層式標準化症狀，意謂著有 449 筆資料在階層化的合併過程成為同一症狀。階層式症狀因此具有分析能力，證象可藉由症狀的分類或屬性知道症狀的分布情形，進而去分析該證象間的關聯性。

(四)問診化

階層式症狀具有能讓病人將自身的症狀完整描述，或是透過系統輔助醫生詢問

病人症狀的能力。但階層式症狀描述太精細，問診時症狀過多以致無法問到重點。因此要將階層式標準化症狀減少以利於問診，依照標準化的五大原則進行合併，合併後症狀的內含不變，產生具有問診表性質的階層式標準化症狀，簡稱問診式症狀。

在問診化階段會依問診時的需求，將階層式症狀進行以下三步驟修正，

1. 合併類似症狀，將症狀本質呈現出來。

「二便:小便:顏色:紅赤:如濃茶」，症狀的本質即是「小便是紅赤色的」，無需再確認是如濃茶色般的紅赤還是如咖啡色般的紅赤，了解問題核心以利於辯證。

2. 將病患無法分辨的症狀，透過屬性合併修正為較模糊粗略的屬性。

當患者描述痛時，並無法知道是「痛」、「痠痛」或是「腫痛」，腫痛又有許多形式如「腫痛、脹痛、腫脹」等，如一一問病人，病人無法分辨其箇中差異，因此，須將其屬性做合併。

3. 將症狀依部位和分類拆成兩種症狀。

在問診階段時，如傷寒論者會先問「汗的有無」，來將病人可能患有的症狀作分別，將症狀拆成部位及症狀分類，有助於醫生釐清病人的病位及病機。

完成上述三點後產生如下表 5，上個階段的階層式症狀有 2638 筆資料，經過了問診化的過程後，問診式標準化症狀有 696 筆資料，成功的將症狀精簡成原來的 1/4，此階段是症狀做標準化的最後一階段，利用這些標準化症狀，可產生問診表協助醫生問診以及記錄病人病歷，也可透

過問診式症狀的階層去分析證象的組成元素及證象之間的關聯性。

表 5 問診式症狀

階層式症狀	問診式標準化症狀
二便:小便:情形:不通暢:不通暢	排泄物:小便:流利情形:無尿出
二便:小便:量:少:少	排泄物:小便:量:少
二便:小便:感覺:澀痛	排泄物:小便:感覺:灼熱疼痛
二便:小便:顏色:紅赤:如濃茶	排泄物:小便:顏色:紅赤
二便:小便:顏色:黃:黃混濁	排泄物:小便:顏色:黃
二便:小便:頻率:很多次:數	排泄物:小便:次數:一天很多次
寒熱:灼熱:部位:腰以下半身:尿道	溫度:寒熱:部位:下腹部

(五)問診表

在中醫問診時透過問診表有系統的去記錄病人症狀是很有用的輔助。設計問診

表有兩種形式，選擇型和是非型的問診表。如問顏色這屬性，問的症狀是面色，會有蒼白、紫色、紅潤、青、黑等。用是非型的問診表只問兩個問題就能將症狀記錄。

在問診化階段，問診式症狀已將症狀的屬性合併，而問診表的產生會將同性質的問題放在一起問，例如原本問診式症狀「頭面部:唇色:顏色:紅赤」及「頭面部:眼睛:顏色:紅」，在問診時會先問是否有頭面部問題，用問診式症狀有助於分清楚問診表的症狀，但在問診時，這兩問題都會是頭面部問題，意謂著病人若無頭面部問題，就不問這兩症狀，此為問診表方便之處。表 6 為前例所轉換出的問診表。

表 6 問診表部份

症狀	問診問題	屬性		
		會	不會	
小便的顏色:紅赤、淡黃、清澈。	1. 小便的顏色是深黃或紅赤色嗎?	會	不會	
小便的感覺:灼熱疼痛。	2. 小便会灼熱疼痛嗎?	會	不會	
小便的流利情形:點滴不盡、小便自行流出	3. 小便時會不會有一直尿不完的感覺?	會	不會	
小便的流利情形:無尿出、小便不流利	4. 小便時會不會一直想尿尿不出來?	會	不會	
小便的頻率:小便頻繁	5. 會不會頻繁地想小便?	會	不會	
小便的時間:夜間	6. 晚上會不會起來小便?	會	不會	
小便的內容物:白濁物質、血液、砂石	7. 有沒有看到小便中有下列物質?	白濁泡感	血液	結石
小便的味道:合併為臭味	8. 小便有沒有很重的尿味?	有	沒有	
小便的量:少、多、	9. 每次小便的量多還是少?	多	少	
小便的頻率:少	10. 會不會想小便?	會	不會	
小便的感覺:急促感	11. 會不會忍不住尿的感覺?	會	不會	

在這部份將中醫證候學的 66 個證象內所包含的症狀群，經標準化的過程產生每階段的標準化症狀，最後產生出問診表。標準化的過程中，不斷處理著症狀內涵的拆解分層與合併歸類，表示著症狀的標準化是可以不斷修正的。本研究標準化所用的資料為中醫證候學，使用書中證象的症狀做為標準化的資料，目的是希望能透過標準化症狀後，能了解證象與症狀間的關係，將證象的病機找出來。我們發現在階層式症狀中，症狀若被分成過度的小單位，會無法確認其真正代表性的階層；在問診式症狀中，若症狀被過度合併，使原有內涵變模糊，原本非此證象的代表性分類，因合併分類而使之成為代表性分類。因此過度的合併或過度的分層都不利於分析證象與症狀的關係。所以最好用分析的方式達到一種平衡。

二、證象關聯性的分析

在中醫證候學書中將 66 個證象分成 13 類如表 7，每一個證象分類都由類似病機的證象所組成，例如脾理有證象「脾理不宣」、「脾理不調」以及「脾理不實」等。

表 7 中醫證候學的證象分類表

脾理	經脈	絡脈	清空	清竅	血	陰
氣機	神志	津	液	營	陽	

透過將標準化症狀以合併及分群方式來產生證象分群，利用分群的症狀建立證象關聯性，對照分群結果是否符合書中的證象分群。本過程所用的資料型態為 66 個證象之 696 筆資料。

(一)合併

為了達到辨證的需求，所有收集到的標準化症狀，我們都會先用關聯症狀的方式將之合併成一群。合併流程如下：

1. 每個症狀自成一群。

總共有 696 筆標準化症狀，有 696 群。

2. 在同一證象底下的症狀，視為同一群。

總共有 66 個證象，將 696 筆標準化症狀分成 66 群，證象也分成 66 群。

3. 群和群之間有共同症狀則合併，沒則終止合併。

證象群和證象群之間有共同的症狀則合併成同一證象群，症狀也合成一群，最終 66 個證象合成一群以一群有關聯的症狀使之連接，稱有關聯的症狀為「關聯症狀」。

將有關聯的症狀連結在一起，意謂著有關聯的證象也連結在一起，才能透過將關聯症狀分群去分析證象關聯性。

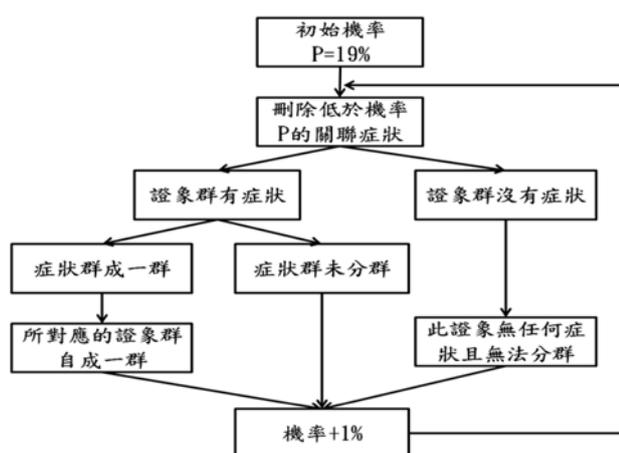
(二)分群

為了瞭解哪些症狀才是能否分群的關鍵，那就要觀察某一證象出現機率最高的「主要症狀」。一個證象可有很多「主要症狀」，但某個「主要症狀」也可能同時出現在不同的證象中。若刪除這個「主要症狀」後，一個證象群就因此失去彼此的關聯性而分成二群。則稱此症狀是使這個證象群分開的決策點。我們不斷重複這個過程，其目的在使證象群獨立分開。

在分群的過程中，從出現頻率低的關聯症狀開始刪除，不會失去證象的代表性。透過機率由低到高的刪除關聯症狀將證象分群，可以知下兩點：

1. 每證象分群內的症狀一定比被刪除的關聯症狀出現頻率高。
2. 分出來的證象群因此刪除掉的關聯症狀而和其他證象群分離。

隨著機率提升每個群的關聯症狀會越來越少，當刪除機率越大，存在於證象群的關聯症狀可能為其證象群的「最主要症狀」，圖 1 為分群流程圖。



每個證象內的症狀出現機率最低為 19%，因此初始刪除機率從 19% 開始。刪除所有證象內出現低於此機率的關聯症狀，刪除症狀後分兩種情況：(1) 因此關聯症狀合併在一起的證象群就會被分離出來，分離的證象群自成一群。以及(2) 有其它關聯症狀使此證象群合併在一起，表示著刪除此關聯症狀，不能將此證象群分離。若刪除的關聯症狀為此證象群的最後一個症狀，表示著刪除此關聯症狀後，此組證象也沒有任何症狀存在了。則此證象群無法再被分群，於是單獨一個證象被區分出來。

當判斷完所有情況後，機率增加 1% 並重複上述判斷，直到所有症狀都被刪除完或是所有證象群都被分群。如圖 2。

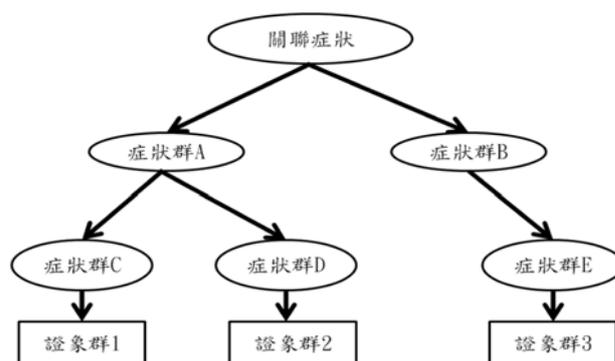


圖 2 證象分群流程圖

關聯症狀群透過刪除某一症狀後就分成症狀群 A 以及症狀群 B，每個症狀群只要有症狀可以使症狀群分群，就會一直往下分群，直到無法分群為止。症狀群可以與書上的證象群相對應，便可辨別出是何證象群了。當所有證象都各自被分成一群。便可以清楚地知道由症狀到證象已建構完成了。這是原始的證象分群。

三、標準化症狀的改進

由標準化症狀分群至產生出原始證象分群。我們發現原始證象分群結果可做為標準化症狀的改進依據。改進的方法如下：

(一) 原始症狀的檢驗

將此群的共同的標準化症狀的原始症狀列出來，觀察原始症狀是否和共同的標準化症狀之內涵一致。如果是一致則表示此分群結果正確，有相同內涵的證象群聚在一起。若發現內涵是不一致的，則檢查原始症狀，並重新標準化之。然後再對此原始證象分群進行分類，看是否可以繼續分群。

(二) 修正標準化症狀

若有重新標準化的症狀，則還要修改或新增每階段的對照表內容，修改對照關係錯誤的表格，新增無此對照關係的表

格，產生新的標準化症狀。

表 8 將 1294 筆證象以問診式標準化症狀作為分群資料進行分群

1	氣機沖逆	氣機不降	12	津氣蘊熾	津氣蘊蒸
2	清竅不宣		13	清空失養	清空不寧
3	陰精不固			清空不宣	
4	血滯不行	血熱蘊蒸	14	陽氣不行	陽氣不振
5	神志昏蒙		15	絡血妄行	絡瘀血溢
6	經脈不和	絡脈不宣	16	血虛失養	
7	氣化不行		17	腠理不實	
8	腠理不宣		18	津不化氣	經氣不宣
9	陰熱蘊蒸	陰熱蘊熾	19	經脈失養	經脈失榮
	陰熱蘊灼	腠理不調	20	氣虛失養	氣虛不充
10	神志不寧		21	熱迫津泄	氣機下陷
11	陰液消涸	津氣蘊灼		水谷不分	
	津氣蒸灼	津液枯涸	22	氣機不利	氣機鬱結
	津液消灼				

在第 4、9、11、18、21 群的分群的證象群沒有相同的病機。第 9、11、21 群則有三個以上的證象聚為一群，則會檢查原始症狀及標準化症狀後看可不可以繼續分群。而第 4、18 群的分群的證象群進行上述的步驟以及探討修正後的結果，表 9 中列出第 4、18 群的分群證象的主要共同症狀以及所對應的標準化症狀以及原始症狀以方便下一階段的分析。

表 9 分群證象的主要共同症狀

群	主要共同症狀	標準化症狀	原始症狀
4	皮膚:顏色:紅紫赤色	皮膚的色:紫紅	皮膚紅斑 皮膚瘀斑紫紅而暗
18	痛:部位:全身	腫脹的部位:全身	浮腫
	痛:合併型式的痛:腫脹:脹痛、腫痛、腫脹	脹滿的部位:背部	
		痛的部位:全身	

在第四群中，發現疹色跟斑色都被標準化成皮膚顏色的問題，且顏色被標準化後和原意有出入。在第十八群中，主要共同症狀「痛:合併型式的痛:腫脹:脹痛、腫痛、腫脹」的原始症狀描述，觀察在「津不化氣」中原始症狀描述內涵和「痛:合併型式的痛:腫脹:脹痛、腫痛、腫脹」完全不同，浮腫的問題全部被標準化成腫痛的問題，意謂著標準化過程有錯誤，這些原始症狀描述要重新被標準化。重新標準化再分群結果如修正後分群之表 10。此項結果已接近證象的辨別，對辨證又接近了一步。

表 10 修正後分群

1	氣機沖逆	氣機不降	15	津氣蘊熾	津氣蘊蒸
2	清竅不宣		16	清空失養	清空不寧
3	陰精不固			清空不宣	
4	血熱蘊蒸		17	陽氣不行	陽氣不振
5	神志昏蒙		18	絡血妄行	絡瘀血溢
6	經脈不和	絡脈不宣	19	血虛失養	
7	氣化不行		20	腠理不實	
8	腠理不宣		21	津不化氣	
9	陰熱蘊蒸		22	經氣不宣	
10	陰熱蘊熾		23	經脈失養	經脈失榮
11	腠理不調		24	氣虛失養	氣虛不充
12	血熱蘊熾	神志不寧	25	熱迫津泄	氣機下陷
13	津氣蘊灼			水谷不分	
14	陰液消涸		26	氣機不利	氣機鬱結

結論

本論文主要描述中醫標準化的過程，過程中將中醫症狀描述模糊的部份重新定義，將屬性做量化以清楚所描述的程度。透過標準化的流程，將症狀逐步的標準化成適合問診及研究分析的格式，每種格式的標準化症狀都有相關的對應資料，利於修改以及產生問診表，希望藉由標準化症狀，將中醫症狀

做數位化的記錄。

我們進行了將症狀做分群以達到證象分群的目的，透過分群標準化後的症狀，找出每個證象，進而由證象找出可能的證候，為辨證系統做了準備的工作。

另一點是我們依中醫證候學的分析發現類似病機的證象會同分在同一群，因此我們用可不可以被分群來做為我們標準化各種症狀的依據，並依此確認標準化症狀是否完善，且利用分群的共同症狀，來修正被標準化錯誤的症狀。所以更豐富了張啓明^[9]等人的看法。並提供標準化症狀的驗證條件。

期許未來能夠利用各個形式的標準化症狀，針對中醫症狀作更多進一步的研究以及分析，以及分析症狀-證候之間的關係以利於辯證系統的開發，並進行臨床上的診斷與應用。

誌謝

本論文由財團法人佛教慈濟醫院大林分院院內計畫 DTCRD-100-E-15「中醫診斷資料庫研發與應用系列三：中醫典籍症狀標準化的嘗試及自動化辨證系統的探討」支持。

(投稿日期：2012年3月8日)

參考文獻

1. 陳世波、劉保延、王永炎、周雪忠、倪青、何麗云：“症”及其分類在中醫臨床評價中的作用分析，中醫雜誌，2009，50(3): 5-7。
2. 王志國、王永炎：制定《中醫臨床診療術

語· 症狀體征部分》國家標準的重要性和迫切性，中國中醫科學院，2007，30(11): 729-739。

3. 李洪成：中醫證候學，中國醫藥科技出版社，2008。
4. 張啓明、王義國、張磊、于東林、王永炎：內涵最小的獨立症狀，北京中醫大學學報，2010,33(1):5-10。
5. 張啓明、田欣、李毓秋、李檬、李正旺、趙傳亮：做為科技術語的中醫症狀名的定義，北京中醫藥大學學報，2008.31[1]:5-6。
6. 趙暉、陳家旭：試論證候準段規範化的研究思路和方法，天津中醫藥，2008,6(4): 1672—1519。
7. 嚴石林、陳為、于宏波、沈宏春、王浩中、鄧瑞鎮：中醫辯證與症候 75C5 機辨識. 成都中醫藥大學學報 2010,(4/2)1004-0668。
8. 李慶生：試論中醫辨證與病機分析[J]，湖南中醫藥導報，1996,2[1]：3-5。
9. 張志斌、王永炎：證候名稱及分類研究的回顧與假設的提出，北京中醫藥大學學報，2003,26[2]:1-5。
10. 張啓明、王義國、王永炎：小便類症狀單元的確認方法，中醫雜誌，2008,49[12]: 1109-1111。
11. 于明珠、田欣、王義國、馬洪剛、武嘉興、張啓明：中醫咳痰的 24 個臨床特徵，環球中醫藥，2009,2[4]:273-275。
12. 白舒霞、田欣、王義國、張明鑫、閔麗芳、張啓明：中醫寒熱症狀 12 種臨床特徵，山東中醫藥大學醫報，2009,33[3]:3-4。
13. 張啓明、王義國、劉保延、王永炎：中醫症狀中的感覺症狀，中國中醫科學院中醫

12 台灣中醫臨床醫學雜誌 2012:18(1)

臨床基礎醫學研究所，2009,11[4]:
512-515。

14. 李檬、田欣、王義國、李毓秋、閔麗芳、
張明鑫、白舒霞、張啓明：帶嚇類症狀單
元的提取，中國中醫基礎醫學雜誌，
2009,15[6]:473。

15. 徐麗、田欣、師偉、王義國、張啓明：
汗的 14 個症狀單元，山東中醫雜誌，

2010,29[4]:219-220。

16. 閔麗芳、張啓明、王義國、張明鑫、白舒
霞：中醫胃脘痛的 25 個臨床特徵，中國中
醫基礎醫學雜誌，2009,15[5]:387-388。

17. 鄭紅、張啓明、王義國：中醫口部症狀單
元的提取及其臨床特徵分析，山東中醫雜
誌，2009,28[10]:675-676。

The Research on the Standardization Proccess of Traditional Chinese Medicine Symptoms

*Liu-Yu Lin¹, Chen-Hung Liu¹, Zen-Yi Chen², Chia-Chou Yeh^{3,4},
Nai-Wei Lin¹, Ming-Hsien Yeh³*

¹ National Chung Cheng University, Department of CSIE; ² NanHua University, Department of IM

³Buddhist Tzu Chi General Hospital, DaLin Branch, Department of TCM

⁴TzuChi University, Department of Medicine

Abstract

Syndrome pattern is one of core theories of Traditional Chinese Medicine. Symptom represent the most important sources of the information required for pattern identification. Symptoms are often due to geographical, climate, customs and physicians have different background narrative. These annoy doctors and researchers.

Therefore, we proposed the concept of “Symptom Standardization” . There exists 3516 sentences in Theory of Traditional Chinese Medicine Syndrome-complexes which describes symptoms and we would like to present them in the following manners:

1. One symptom unit only express one connotation.
2. One symptom has its own classification , properties and its property values.
3. Quantify a symptom description of the attribute
4. Unify a symptom description of the attribute.
5. Sentences with same meaning must standardize to the same result.

Symptom Standardization is used to produce traditional Chinese medicine symptoms Questionnaire to facilitate the collection of disease symptoms, and also used to analysis with algorithm to verify the theory of traditional Chinese medicine. This study will describe how symptoms standardization and its procedures, clustering the standardization symptom using statistical probability, through the clustering results verify standardization symptom.

Key word: Traditional Chinese medicine, symptom standardization, Syndrome pattern, pattern identification.