

## 論 述 二

# 實證臨床診療指引評估工具與 AGREE

## CPG Evaluation Tool and AGREE Instrument

陳杰峰<sup>1</sup> 王慈蜂<sup>2</sup> 羅德毓<sup>1</sup> 邱文達<sup>3</sup>

台北醫學大學市立萬芳醫院實證醫學中心<sup>1</sup> 家庭醫學科<sup>2</sup> 台北醫學大學<sup>3</sup>

### 前 言

以實證為基礎之臨床診療指引，可以適時反應客觀的文獻證據、提升照護品質。我國醫療照護品質，必須要有具公信力的審核標準，而臨床診療指引是其中重要工具。國內醫療界正推廣以實證為基礎之診療指引，在數年之內，應會產生相當數量的指引。惟目前國內缺乏實證診療指引之品質評核指標，亦沒有可供參考的指引評核範例。本研究期望以實證醫學為基礎，尋找適合在國內評核診療指引品質之工具，以評估國內的診療指引，進而選取高品質之指引，形成治療規範，甚或可以讓爭議審議之解決有所本。

臨床診療指引的評估工具中，美國醫學研究院（Institute of Medicine, IOM）於 1992 年出版的評估工具是第一個出現的。IOM 將「臨床診療指引」定義為：有系統發展出來的描述，用來輔助醫療人員與病患，針對特定的臨床狀況，作出適切的醫療照護，由此可知十幾年來，所有的臨床診療指引差異性相當大[1]。

目前國內各醫院對各疾病編製之臨床診療指引或流程，可找到的有：高血壓防治手冊、高血脂防治手冊、全民健保降血脂藥物給付規定（抗高血脂）、癌症共識手冊（子宮頸癌、口腔癌、肺癌、大腸直腸癌、胃癌、攝護腺（前列腺）癌、鼻咽癌等）、泌尿道感染、骨質疏鬆症、過敏性鼻炎等。關於這些臨床指引的品質、遵從度、可行性、是否有爭議、是否符合上述品質指標等，並沒有客觀的研究分析，國內尚無一套系統性的臨床指引制定流

程可供參考，亦沒有任何以實證為基礎的臨床診療指引。

實證醫學對於醫療品質的改善，主要透過臨床診療指引的建立。臨床診療指引是針對特定臨床病症，以系統性方式發展出來的診療陳述，可協助醫師及病人決定何為最適當的健康照護。臨床診療指引對醫療品質改善之重要性，文獻中多有描述，美國學者 Diamond 認為，實證醫學是以臨床研究中的最佳證據，應用於目前病人同時融入病人期望與醫師經驗的判斷，目的在提供各種醫療資訊及證據[3]。因此，實證醫學的主要目的在於，以嚴謹的方式整合目前最佳的技術文獻、臨床經驗、以及病人的價值觀，使得診斷、治療及照護獲得最佳結果，提升病人的生活品質。藉由臨床診療指引，可形成明確易使用的標準，較容易作客觀的決定；另外，臨床診療指引可教育病人或醫護人員目前的最佳治療方式，增進醫療的成本效益，或當作管控的工具。

隨著臨床診療指引的發展逐漸受到重視，有相當多的指引評估工具嘗試用來辨識臨床診療指引的品質，2000 年加拿大學者 Graham 等人，曾針對各種臨床診療指引評估工具，共比較 13 種評估工具[4]；另外，比利時的一份研究報告[5]由 98 篇文獻內，辨識 24 種指引評估工具，進行系統性文獻回顧的相似性與差異性之比較評估，為至今最完整的系統性文獻回顧。

在 24 種評估工具當中，共包含的 469 個問題，依據性質可將其歸類為 50 個項目及十個領域。這 10 個領域分別為：臨床指引的正確性（validity）、指引的可信度（reliability/reproducibility）

lity)、臨床的應用性 (clinical applicability)、臨床使用彈性 (clinical flexibility)、透明度 (clarity)、定期回顧 (scheduled review)、發展團隊 (development team)、指引的執行 (implementation)、指引的傳播 (dissemination)、指引的評估 (evaluation)。完整包含這 10 個領域的評估工具，僅有三個，分別是 SIGN、Helou and Ollenschlager 以及 Cluzeau，其中唯有 Cluzeau 評估工具通過效度研究。

根據 Cluzeau 所發展的評估工具，共有 Helou and Ollenschlager、Hutchinson、AGREE，這三種評估工具中唯有 AGREE 經過效度研究，且為量化分析，相較於 Cluzeau 工具顯得更加精簡與濃縮，讓等級的比較變得較為容易，是國際間最具公信力的臨床診療指引評估工具。

## AGREE 簡介

AGREE 是來自 13 個國家的專家，在歐盟資助下 (Biomedicine and Health Research, BIOMED 2) (BMH4-98-3669)，耗費 4 年 (1994-1998) 所完成。AGREE 第一次草稿由 195 位專家評讀 11 個國家，共 100 個指引；第二次草稿則從 100 個指引當中，每個國家隨機挑選三個指引，交由 70 位新遴選的專家來評讀。AGREE 評估工具共包含六個領域以及 23 個項目，每個領域可以用來評估臨床診療指引各獨立面向的品質[6]。

AGREE 主要用於臨床指引研究和評估的鑑定，始於指引評估者 (appraisers) 和政策制定者間，為國際合作及促進臨床診療指引的品質與有效性，成立並建構一個關於指引發展、發表及評估的共享組織架構。AGREE 之宗旨為發展評估指引的工具，並協調歐盟間指引的發展，讓重複的情況降至最低。

AGREE 工具評估內容包括評斷臨床診療指引的方法學、最後建議的內容和相關影響因素。AGREE 可以用來評估指引的預期有效性，達成預期效果的可能性，但並不評估臨床診療指引對病人所產生的健康結果衝擊。根據 AGREE 報告顯示，95% 的指引評估員認為，AGREE 工具對於指引的品質評估很有幫助，可廣泛應用。再者，AGREE

的信度令人滿意，國外研究 Cronbach's alpha 範圍約在 0.64 至 0.88 間，國內研究 Cronbach's alpha 約為 0.875，不同評估者之信度 (inter-rater reliability)，同一指引四位評估者信度約落在 0.57 至 0.91 範圍間。

AGREE 可評估的範圍相當廣泛，包含新的指引、現行的指引、現行指引的更新版，均為 AGREE 可評估之範圍。由地區、區域、國家、國際組織，或政府機構所發展的指引，AGREE 亦可評估。同時，AGREE 為通用性高的工具，可以評估各種疾病範圍的指引，包括疾病的診斷、健康促進、治療或醫療介入。

## 研究方法

### 一、中文化

為進一步瞭解評估臨床診療指引的工具，我們將 AGREE 評估工具進行中文翻譯，同時透過與 AGREE Research Trust 之間的國際合作，正式將我們所翻譯之成果，納入 AGREE 網站連結[7]，全球中文使用者均可自由連結下載，為 AGREE 臨床指引工具繁體中文版。

欲翻譯 AGREE 工具，目前所有語言翻譯版本均依照國外專家 Guillemin 的研究 "Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines." [2]，作為翻譯流程之根據。該研究之主題為評估工具的翻譯，包括健康取向的生活品質 (health-related quality of life)、病患滿意度問卷 (patient satisfaction questionnaire) 之翻譯經驗。同樣地，我們將 AGREE 中文化亦須以此為根據，翻譯流程共包括以下四點：

1. 分別請兩位專家，將 AGREE 英文版翻譯為其母語，兩位專家中至少有一位必須熟知臨床診療指引的發展、評估或相關方法學 (methodologies)。AGREE 工具中文化由國內實證醫學領域之兩位專家分別獨立地進行翻譯，一位為台北市立陽明醫院賴榮年主任 (台北市立陽明醫院實證醫學小組，2004)，一位為市立萬芳醫院陳杰峰主任，集兩位專家之力完成中文翻譯。兩種分別

完成的中文翻譯版本，再比較中文翻譯之間是否有不一致之處，完成翻譯初稿。

2. AGREE 中文翻譯初稿必須再交由以英文為母語，且在翻譯工作開始前對 AGREE 工具不熟悉之譯者，根據中文初稿，倒譯 (back translation) 成英文。
3. 經過兩位外國專家倒譯後，仔細比較該譯文與 AGREE 原始英文版本之間的差異。這項步驟是為了確認中文初稿的正確性與差異性，並可修正中文初稿中差異或含糊不清的地方。

最後經過修正完成後，找出我國現有之臨床診療指引，本研究以全國醫師聯合會高血壓診療指引為試作評估的對象，具體以 AGREE 中文工具評估，AGREE 臨床診療指引工具之內容共包含六大領域、23 項評估項目 (如表一)，逐項針對高血壓診療指引進行試評。

## 二、評估本國高血壓診療指引

為能評估特定疾病之醫療品質監測指標，本研究選擇本國高血壓臨床診療指引作為評估對象，以繁體中文文化 AGREE 工具試作。評估高血壓診療指引必須經由專業的評估人員，並經過完整專業訓練，本研究共訓練數名試評員參與文獻蒐集、評讀、AGREE 工具之翻譯及編修，經過訓練手冊的熟讀後，實際使用 AGREE 工具評估指引，訓練熟稔評核臨床診療指引之人員包含 4 步驟：

1. 二小時訓練課程。
2. 熟讀訓練手冊。
3. 完成實際評讀臨床指引，並以 AGREE 評估。
4. 計算分數之後，綜合討論。

本研究之試評員涵蓋了內、外科及家庭醫學科的醫師及醫管相關行政人員，對 AGREE 工具進行測試。AGREE 共包含六大領域及 23 個項目 (如表一)，對於指引的形成方法與流程進行評估分析，將指引分成三個大部分，第一是「強烈建議」，第二是「建議」，第三是「不建議」。計分方式視領域之總分決定，該總分由領域內各獨立項目分數總和而得，進而對該領域的得分百分比進行標準化。表二為試算之範例。

表一：AGREE 工具之中文

### 領域一：適用範圍與目的 (scope and purpose)

1. 清楚描述指引的目的。
2. 清楚描述指引所涵蓋的臨床問題。
3. 清楚定義適用的病患族群。

### 領域二：權益相關人的參與情形 (stakeholder involvement)

4. 指引發展團隊成員來自所有相關專業領域。
5. 納入病患的意見及喜好。
6. 清楚界定指引使用者。
7. 指引公告前已有使用者完成試作。

### 領域三：指引發展的嚴謹度 (rigour of development)

8. 運用系統性的方法搜尋證據。
9. 清楚描述選擇證據的標準。
10. 清楚描述形成指引建議的方法。
11. 指引的建議內容有考慮到健康效益、副作用及風險。
12. 指引建議與其支持證據間有明確的關聯性。
13. 指引公告前已經由其他外部專家審閱。
14. 提供指引定期更新的步驟與準則。

### 領域四：指引的明確性和代表性 (clarify and presentation)

15. 指引的建議明確不含混。
16. 清楚呈現不同的治療選擇。
17. 主要建議清楚易辨。
18. 指引在實際應用上有完整的配套工具。

### 領域五：應用性 (applicability)

19. 有討論到在推行指引時所遭遇到的組織障礙。
20. 有考慮到推行指引對成本費用的衝擊。
21. 說明該指引的主要評估監測標準。

### 領域六：編製的公正客觀及獨立性 (editorial independence)

22. 指引的編製不受贊助者的影響。
23. 指引中有記錄發展團隊成員的利益衝突。

表二：針對第一領域得分（適用範圍&目標），由 4 位評估者評分

	項目 1	項目 2	項目 3	總分
評估者 1	2	3	3	8
評估者 2	3	3	4	10
評估者 3	2	4	3	9
評估者 4	2	3	4	9
總分	9	13	14	36

最高分 = 4 分(完全同意)\*3 項目)\*4(評估者) = 48

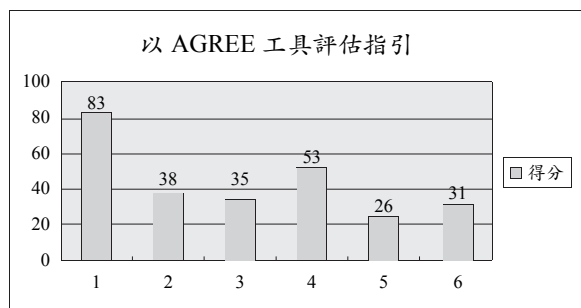
最低分 = 1 分(完全同意)\*3(項目)\*4(評估者) = 12

標準化領域分數：總分 - 最低分 / 最高分 - 最低分 = (36-12)/(48-12)\*100 = 67%

### 結果與討論

AGREE 評估項目中，「強烈建議」的定義為，被評核的分數多為 3 分或 4 分，或整個領域總分達到高於 60 分的結果；「建議」的定義為，被評核的分數得到 4 分與得到 1 分的次數相近，或整個領域的總分達 30 至 60 分。反之，「不建議」的定義為被評核的分數大部份皆為 1 分或 2 分，或整個領域的總分低於 30 分。按照此定義，整理高血壓指引評估評分與專家意見，高血壓指引評估的分數多介於 30 至 60 分之間，因此，經過中文文化的 AGREE 工具評估，發現此高血壓指引是可被建議使用的。

針對 6 個領域進行詳細分析（如圖一）領域分數分別為 83、38、35、53、26、31。經由圖一，可知六個領域中，僅第一領域達到 83 分，顯著高



圖一：以 AGREE 工具評估高血壓指引各領域綜合得分

於強烈建議的標準 60 分，代表高血壓指引在「描述指引適用範圍與目的」部份是相當詳細的，特別在「描述指引的目的」之項目，幾乎是完全同意。第 2 項「清楚描述指引所涵蓋的臨床問題」，評估者均給予高分。整體而言，本國的高血壓指引在第一領域「描述指引適用範圍與目的」，可說非常完善。

第二領域「權益相關人參與的情形」，領域得分為 43 分。由於本指引的發展團隊以醫師為主 [8]，因此在第 4 項「指引發展團隊成員來自所有相關領域」之得分較低，尤以第 7 項「指引公告前已有使用者完成試做」，這部份並沒有描述，事實上全世界幾乎少有指引於公告前，即有使用者報告，因此本項目的分數較低在預期之中。第 5 項「納入病患的意見及喜好」，本指引無法看出有納入病患的意見及喜好的過程，這是指引在製作的過程中較受爭議的部份，也許由於本土對於納入病患意見及喜好的文獻較少，因此這部份可以再加強。第 6 項：「清楚界定指引的使用者」，該指引中僅提到使用對象為全國醫師，至於是否為內科醫師、外科醫師、一般科醫師並沒有清楚的描述。

第三個領域「指引發展的嚴謹度」共有七個項目，本指引在該領域部份獲得 44 分，得分較高的是第 11 項目「指引建議的內容有考慮到健康效益、副作用及風險」，以及第 12 項目「指引與其支持證據間有明確的關連性」，這兩項皆獲得不錯的分數。至於第 14 項，提供指引定期更新的步驟與準則，本高血壓指引中並無清楚的描述或相關的紀錄。第 8、9、10、13 項目，也許在製作的過程中有操作，但是並沒有描述在指引中。本指引主要引用其他相關的臨床指引及文獻，並沒有描述第 8 項的運用系統性方法搜尋證據，這方面是可以再加強的。對於清楚描述選擇證據的方法並沒有相關的紀錄。第 10 項「清楚描述形成指引的方法」，本高血壓指引中並未描述指引的形成過程，只在首頁提及「醫師全國聯合會邀請學者專家根據醫療經驗及世界最新相關文獻」形成指引。第 13 項目「指引公告前已經由其他外部專家審閱」，則沒有紀錄。

第四領域部份，指引的明確性及代表性，在六個領域中為第二高分，領域得分為 53 分。其中在



第 16 項「清楚呈現不同的治療選擇」達到不錯的分數，這是國人寫作指引時的強項。第 15 項「指引的建議明確不含混」，及 17 項「主要建議清楚易辨」，都獲得了評估者的肯定。第 17 項中，由於本指引運用了許多粗體、底線及圖表等方式，因而呈現出最適當的建議，如高血壓指引第 7 頁高血壓新病人處置及選擇降壓劑的準則，有用粗體表示，可讓閱讀者清楚易辨。但由於本指引並無摘要版，以及參數的參考指引，所以第 18 項「指引在實際應用有完整的配套工具」分數較低，此項可列為將來發展指引時需加強的部份。

第五個領域「適用性」，這是所有領域中分數最低的，國外亦然，本指引在該領域只有 26 分。經過了解，本指引較少著墨於第 19 項「有討論到在推行指引前所遭遇的組織障礙」，而第 20 項「有考慮到推行指引對成本費用的衝擊」也較少描述。第 21 項「指引的主要評估監測指標」也沒有清楚的寫作。基本上，在本指引中，這三個項目都描述較少。

第六領域「編製的公正客觀及獨立性」的得分 31 分，主要因為高血壓指引中，並沒有詳細紀錄指引發展團隊成員的利益衝突。

我們搜尋相關文獻，找到一篇"Quality of Evidence-Based Pediatric Guidelines" [9]，針對小兒科指引進行評估，以問卷方式詢問荷蘭的 51 位專家，有關指引發展的 5 個最迫切主題，就 AGREE 的六個領域進行相關證據的評估。在本文獻中，總共比較了 215 份實證診療指引中，較具指標性的 17 份指引，以 AGREE 工具評估其品質，而獲得 6 大領域的分數，分別為 84、42、54、78、19、

40。我們以國內評估高血壓指引的分數比較，發現整體分數是類似的（如圖二）。這份荷蘭的報告，同意 17 個指引中，有 14 個值得推薦。

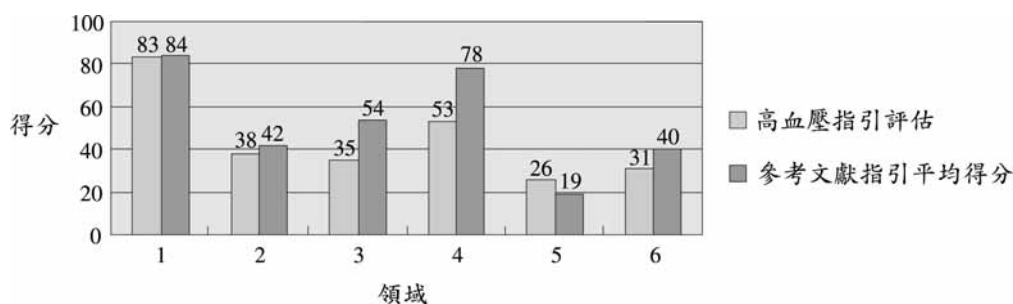
綜觀國內高血壓指引，是可被推薦使用的。參考 AGREE 訓練手冊[10]，以標準化領域分數作為是否可被推薦的依據，若該分數多半在 30 至 60 分之間，即可認定為可被推薦。高血壓指引在六大領域中，有四個領域均在 30-60 分之間，各評估員的細項評分資料中，獲得高分與低分的數量相似（62:76）。因此，經過 AGREE 完整地評估後，高血壓指引應可被推薦。

## 結論與建議

自 1990 年，美國醫學研究院將「臨床診療指引」定義為：針對特定的臨床狀況，以有系統發展出來的描述，來輔助醫療人員與病患，作出適切的醫療照護[1]後，全世界相關研究，已如火如荼地展開。國內在這方面的發展，目前尚在起步階段。

本研究經文獻搜尋，專家意見整合，實際評估本國高血壓指引後，認為中文化 AGREE 工具，是一部適合評估本國臨床指引的工具。

我們同意下列結論，首先，AGREE 評估工具是經由 AGREE Research Trust，此國際性組織所研發之工具，可用於歐洲及歐洲以外的國家，經由專家獲致共識，亦認為可以用於國內指引之評估。其次，這個工具提供了通用的標準，以促進品質流程之改善，可增進本國指引的發展。第三，這些評估標準不僅可用於計畫、執行及監督各指引發展計畫，更可用於國際上各指引的比較，因此，使用本



圖二：AGREE 工具評估國內外指引之比較

臨床指引評估工具於國內，將有助於國內臨床指引品質之提升。

本中文化 AGREE 工具是世界的公共財，可於 AGREE Research Trust 網頁自由下載使用[9]，提供國內未來評估臨床指引使用，以期台灣在臨床指引發展的起步階段，即能穩定開展，成長茁壯。

## 誌 謝

感謝中央健保局編號 DOH93-NH-1016 計畫，與國家衛生研究院之協助與指導，以及各界專家的參與，使本研究得以順利進行。

## 推薦讀物

1. Marilyn JF, Kathleen NL, Committee on Clinical Practice Guidelines, Institute of Medicine: Guidelines for Clinical Practice: From Development to Use. Washington DC: National Academies Press, 1993:27-8.
2. Guillemin F, Bombardier C, Beaton D: Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology* 1993; 46: 1417-32.
3. Diamond LH: Why evidence-based medicine and why now? – industry trend or event. *Health Management Technology* 2000; August: 37-8.
4. Graham ID, Calder LA, Hebert PC, et al: A comparison of clinical practice guideline appraisal instruments. *International Journal of Technology Assess in Health Care* 2000; 16:1024-103.
5. Vlayen J, Aertgeert B, Hannes K, et al: A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *International Journal for Quality in Health Care* 2003;17: 235-42.
6. 台北市立陽明醫院實證醫學小組：實證醫學的運用：臨床診療指引（2004 版），台北，合記圖書，2004:1-8。
7. The AGREE Collaboration: The AGREE Instru-

ment。網 址：<http://www.agreetrust.org/instrument.htm>。

8. 全國醫師聯合會：西醫基層總額提報之高血壓臨床治療指引。網址：[http://www.nhi.gov.tw/web-data/AttachFiles/Attach\\_656\\_1\\_高血壓\\_CPG\\_新版by\\_李德福\\_西基\\_GB\\_940302.pdf](http://www.nhi.gov.tw/web-data/AttachFiles/Attach_656_1_高血壓_CPG_新版by_李德福_西基_GB_940302.pdf)。
9. Boluyt N, Carsten R, Lincke C R, Offringa M: Quality of evidence-based pediatric guidelines. *Pediatrics* 2005; 115:1378-91.
10. The AGREE Collaboration: AGREE Instrument Training Manual。網址：[http://www.agreetrust.org/docs/AGREE\\_INSTR\\_MANUAL.pdf](http://www.agreetrust.org/docs/AGREE_INSTR_MANUAL.pdf)。

## 摘 要

臨床診療指引（clinical practice guidelines, CPGs）是針對特定的臨床狀況，有系統地發展出來的描述，用以輔助醫療人員與病患，作出適切的醫療照護[1]。隨著醫療照護品質日益受到重視，臨床診療指引在國外蓬勃發展，尤以歐洲等國家為盛。國內近年來雖有許多指引陸續發表，然而項目的選取、方法的運用尚在起步階段，缺少評核臨床診療指引的指標與範例。

本研究旨在探討臨床診療指引之評估工具－AGREE（Appraisal of Guidelines for Research & Evaluation）之中文化及國內適用性。AGREE 工具中文化過程遵照國外專家的研究[2]，以四大步驟進行：(1)請兩位以中文為母語的專家，分別翻譯 AGREE 英文版為中文(2)比較兩種中文翻譯版本不一致之處，完成第一次初稿(3)由 2 位以英文為母語之外國專家，將中文翻譯初稿倒譯（back-translation）為英文(4)比較 AGREE 原始英文版本與倒譯版的差異，完成標準版本。之後並以中文化 AGREE 工具評估特定疾病，本研究以國內一份高血壓指引作為評估對象，以數名評估員評讀，試作結果肯定中文化 AGREE 工具之功能，是適合評估本國臨床指引的工具。同時，本中文化 AGREE 工具順利獲得 AGREE Research Trust 正式繁體中文版之認證，已納入 AGREE Research Trust 官方網頁中，可供全球繁體中文使用者自由下載使用。

關鍵字：實證醫學，臨床診療指引，評估工具，AGREE。

## Abstract

Clinical practice guidelines (CPGs) is a promising tool to improve medical quality. In the next few years, there should be many CPGs published in Taiwan. With the increasing concerns about the guidelines, there is a need for ensuring the quality of CPGs.

The AGREE instrument is one of the most widely used appraisal instrument to assess clinical guidelines. We followed the methods for translating AGREE instrument [2], which including four steps : (1) Translation from English into Chinese independently by two translators. (2) The resultant two translations in Chin-

ese should be examined to work out differences and discrepancies between the two versions. (3) This draft should be back-translated from Chinese into English by two translators. (4) Discrepancies between the original English and the back-translated English version of the AGREE instrument should then be examined to identify and clarify ambiguities in the translation. To prove the AGREE instrument as a stable and adaptable tool for CPGs in our country, we chose a domestic hypertension treatment guideline to evaluate AGREE. After the trials, we concluded that the AGREE instrument is a stable tool to evaluate CPGs in Taiwan.

**Keywords:** Evidence-Based Medicine, AGREE, Clinical Practice Guideline, Evaluation