

CIPP 評鑑模式在醫學教育領域的適用性

楊聰財¹ 李明濱² 吳英璋³

醫學教育旨在培養(1)醫療(臨床醫學)、(2)公共衛生(預防醫學)及(3)生物醫學科學(基礎醫學)之專業人才；國際間較流行之評估工具主要針對個別的學習成效，包括標準病人、客觀結構式臨床測驗、臨床表現測驗模式、社區導向醫學教育品質模式等，多為醫學生臨床技能的測驗方法，缺乏足以涵蓋基礎醫學及公共衛生領域之醫學教育評鑑方式。評鑑制度在提昇醫學教育品質方面扮演著重要的角色。一九九八年美國「外國醫學教育與評鑑」國家委員會曾指出我國缺乏嚴謹的醫學教育評鑑制度，而認為我國醫學教育評鑑制度與美國的制度不能相比較。有鑑於此，教育部委託國家衛生研究院成立醫學院評鑑委員會(TMAC)，為我國醫學院評鑑開創了新的展望。然而我們仍缺乏一個具宏觀架構性的評鑑模式，是以在此提出CIPP模式。CIPP模式發展於1960年代晚期，當時主要用於教育機構之評鑑，係以明確的目標、測試和實驗設計為取向。CIPP模式涵蓋背景評鑑、輸入評鑑、過程評鑑及成果評鑑；本文在於介紹此一模式，期能整合各項醫學教育評鑑之元素及方法，建構完善的評鑑模式。CIPP評鑑模式可能適用於醫學教育領域的原因如下：(1)評鑑過程系統化；(2)兼重形成性與總結性的功能；(3)適用於各層級的醫學教育；(4)適用自評與外部評鑑。

關鍵詞：評鑑、醫學教育、CIPP 模式
(醫學教育 2003；7：218~33)

前 言

隨著世界醫學教育的潮流與發展趨向，台灣地區自1981年起，即開始醫學教育改革，針對課程設計、教學策略等作逐步的改革，任何改革都應附加設計評鑑方法，方能使品質更趨

完善而有關。台灣的「醫學教育評鑑」曾經過美國「外國醫學教育及評鑑」國家委員會(National Committee on Foreign Medical Education & Accreditation, NCFMEA)的批評，嗣後國家衛生研究院接受教育部委託成立了醫學院評鑑委員會(Taiwan Medical Accreditation Council, TMAC)，訂定了清楚的評鑑要點。本文試

輔仁大學醫學院醫學系、財團法人天主教耕莘醫院精神科¹ 臺大醫學院精神科、台北市立療養院² 台灣大學心理學系³
受理日期：2003年7月1日；接受日期：2003年8月28日
通信作者地址：李明濱 台北市信義區松德路309號台北市立療養院



圖回顧文獻，說明醫學教育內涵及其評鑑目的、國際間現行醫學教育評鑑狀況、我國醫學教育評鑑之發展，進而介紹一套發展於1960年代晚期，至今仍廣泛運用於其他教育界以及企業界，具有內部自我評估以及外部監控功能的CIPP評估模式，冀望對於整體醫學教育，提出一套評鑑功能更強且可行的方法。

醫學教育及其評鑑

醫學教育之內涵與目的

醫學教育是要培養(1)醫療（臨床醫學），(2)公共衛生（預防醫學）及(3)生物醫學科學（基礎醫學）之專業人才；其中每個領域又再分為數個專科及次專科。醫學教育可分為：(1)醫學生教育（醫學士教育），(2)專科醫師教育（住院醫師訓練），(3)細專科醫師教育（fellowship），(4)終生繼續教育^[1-3]。

以醫學生養成教育為例，醫學生學習的範圍相當廣泛，包括醫學知識、臨床技術、態度及價值觀等。醫學教育之目的在於培養醫學生下列基本能力^[4]：

1. 有效的溝通；
2. 基本的臨床技術；
3. 運用科學來進行診斷、處理、治療及預防；
4. 終生學習；
5. 自我反省、自我保護及個人成長；
6. 對於醫療相關之社會的及社區的議題的了解及反應；
7. 醫學倫理的判斷；
8. 解決問題；
9. 專業特質及角色的認定。

另外，美國醫學教育評鑑委員會（ACGME）認為，醫學生畢業時必須學會並接受評

鑑的技能（competencies）包括^[5]：

1. 病患之照顧；
2. 醫療知識；
3. 以實務為基礎的學習及促進；
4. 人際溝通技巧；
5. 專業素養；
6. 有系統的實習。

然而為測驗以上各項技能亦發展了許多評鑑工具，以下將介紹目前醫學教育評鑑概況，以及我國醫學教育評鑑的進展。

國際間現行醫學教育評鑑狀況

目前全球醫學院品質最佳的學校，美國仍是執世界牛耳，而澳洲、紐西蘭嚴格的評鑑制度，已經讓兩個國家的醫學院品質和美國不相上下，甚至澳、紐兩國醫生執照，是經由醫學評鑑來認定。

美國醫學院評鑑起源於1904年，起初由美國醫學會（American Medical Association, AMA）設立CME（Council on Medical Education）負責推動。CME於1942年改由美國醫學會與美國醫學院學會（Association of American Medical Colleges, AAMC）合辦，名稱亦隨之改為醫學教育聯繫委員會（Liaison Committee on Medical Education, LCME），該會於1976年改訂過去之十項評鑑標準，頒佈 "Accreditation of Schools of Medicine: Policy Documents and Guidelines" 以及附加的 "Supplemental Guidelines for Medical Schools with Branches of Multiple Campuses"，以做為評鑑之標準。美國自1994開始國會立法成立外國醫學教育評鑑委員會（Foreign Medical Education and Accreditation），針對世界各地醫學院的課程，醫學生畢業後的實習，住院醫師的訓練等基礎及臨床醫學教育展開評鑑的工作；先由各該被評鑑學校準備好書面文件依據評量的準則審查，在第二階段另有

駐地訪問審核。此後每七年複查一次，是否通過決定於上述審查結果，以及評鑑委員會的決策。審查通過率較高的國家多是歐美地區，亞洲部分僅印度、巴基斯坦、菲律賓、泰國等地區的醫學院通過審查；日本與台灣、韓國，及歐洲的義大利、奧地利、比利時、挪威等國之醫學教育尚未通過審查^[6]。

至於國際間較流行且為大部分學界所熟悉的評估工具主要是針對個人的學習能力表現評估，主要包括：標準病人 (standardized patient, SP)、客觀結構式臨床測驗 (objective structured clinical examination, OSCE)、臨床表現測驗模式 (clinical practice examination, CPX)、社區導向醫學教育品質模式 (community-oriented medical education, COME) 等，茲將其分述如下：

(一)標準病人^[7,8]

SP 是一位神經科醫師 Dr. Barrow 在 30 多年前所創，是利用正常人，經由教導，能夠正確扮演一位已知病史、理學檢查的人，此人不但要具有模擬病人症狀和病程的能力，能將臨床個案作完整、正確而且重複的演出，也要具有互動、評估及給予醫學生回饋的能力，SP 可以運用在教學上或評量上。它的優點包括：(1)在任何時間、任何地點均可使用。(2)可以持續扮演同樣的病人，供不同的學習者練習。(3)可以提供相同的病人經驗給不同學習者，而避免隨機學習。(4)可以避免錯誤處理真實病人。(5)作為處理真實病人前之橋樑。(6)作為練習技巧之對象。(7)可提供某些困難個案或情境之練習。(8)可配合教學需要，做暫停、或繼續之配合。(9)可控制學習過程。(10)教學者與學習者可一起討論下一步準備要做什麼。(11)學習者可對同一個 SP 反覆的練習。(12)學習者可得到立即的回饋。(13)可以事先設計評量標準，有一致的

評估方式。(14)在基本的病史詢問或理學檢查上，可以減少教師參與的時間。其缺點是：(1)需要時間訓練標準病人。(2)有些個案無法模仿，有些真實理學檢查無法假造。(3)一次不能涵蓋太多個案。(4)一個個案的處理能力不代表對另一個個案的處理能力。(5)一對一的教學或評估，耗費人力及時間。

近年來關於 SP 的相關研究越來越多，使用場合也越來越廣，從教學研究、醫學生評量、醫師執照考試、專科醫師再教育到門診看診品質評估，都有許多成果報導及前瞻設計的發表。Anderson 曾於 1993 對美國及加拿大 142 所醫學院做統計，在回覆的 138 所醫學院中，將 SP 使用在教學或評估上的共有 111 所，占 80.43 % 之高。

(二)客觀結構式臨床測驗^[7,9]

OSCE 是 1970 年代中期，由 Harden 所創，其目的是在測試一項單一的技巧，指在模擬臨床的場景下，使用模型標本、標準病人或真實病人、以及臨床檢驗數據或影像判讀之材料來測試醫學生的臨床能力。如在模型上打點滴，腹部理學檢查、X 光判讀、解讀檢驗數據等。它的目標在於正確有效的評量醫學生臨床見習課程的知識以及臨床診療的技術。好處是可以使用於大批的學生，並且多用於早期臨床教學。其主要構成包括：(1)測驗協調委員會：監督指導單位；(2)測驗負責人：測驗之規劃執行者；(3)測驗的技術、行為及態度的項目：由各測驗單位提出；(4)評分的定義（評估表）：依測驗之項目準備；(5)測驗場所：臨床技能中心；(6)測驗站：安排各站測驗時間長短、可重複測驗的解剖等標本模型、關聯性的測驗站、測驗問題呈現方式、標準病人測驗站、測驗動線的安排等；(7)標準病人或模擬病人；(8)時間管控；(9) OSCE 表現的評估：由標準病人評分

與立即回饋，或由指導教師依據錄影帶進行回饋及討論。

(三) 臨床表現測驗模式^[7,8]

CPX 是 1980 年代，南伊利諾州州立大學醫學院所發展出的一種評估模式，藉以了解醫學生臨床知識運用成效，作為來年與往後努力的借鏡。每次測驗包括 17-18 個病人個案，每個個案測驗時間約為 20 分鐘，醫學生先和病患問診及做身體檢查，然後在電腦輸入病人資料、初步診斷結果及治療計劃；然後取得病人實際的病情資料，據以修改診斷內容及治療計劃。目的在評量並協助醫學生熟練其臨床技能，可用以區別知識上或推論上的學習問題，作為採用適當的教學方案的參考。CPX 常於見習結束時使用，測量學生是否有能力綜合使用適當的臨床技能，完成完整的病患處理工作。

(四) 社區導向醫學教育品質模式^[10,11]

顧名思義，此項評鑑是針對社區醫學教育；由於學習的地點是在社區的基層醫療單位，其病人和醫療型態皆與醫學中心不同，各社區醫療單位間亦存在相當大的差異，而教師也以基層醫療單位的工作人員為主，故美國紐約州立大學水牛城分校家庭醫學科，遂發展出這套社區導向醫學教育教學品質的模式，希望藉此剖析影響社區基層醫療單位的教學因素，成為一套評估工具。

其理論依據為經驗學習理論（experiential learning theory），視學習為一種發展的過程，並且有四種形式：經歷實際經驗（如看過某一病例）、經驗的思考與投射、歸納概括經驗法則、反覆測試所得經驗法則。社區醫學教育的學習過程即是同時含有其中部分或全部之狀況，因此適用經驗學習理論為基礎。社區導向之醫學教育是為一個整體系統（如圖 1），由一些相互依賴的構成要素組成（input），透過人與人、人與環境的互動過程（process），而產生結果（outcome），對社區及民衆而言，結果即是服務，對學習者或學生而言則是醫療執業內容的了解。

既然將社區導向的醫學教育視為一個整體的過程，而非單獨的事件，而且學習地點在社區基層醫療單位，故教學品質的評估就應該考量 COME 的特性，評估內容包含下列各項：教學方式、課程內容、教學活動與時間的安排情形、學習者心態與參與程度、基層醫療團隊的服務能力、學習者在知識與技能的增進程度、學習者的滿意度等。紐約州立大學水牛城分校也依據此模式發展出一個創新的評量工具，稱為 Med IQ（Medical Instructional Quality），用以測量 COME 的教學品質。Med IQ 主要用於醫學生在基層醫療實習時（primary care clerkship）的教學評量，主要測量下列四部分：(1) 基層醫師作為老師時激發與協助學習的角色；(2) 場所作為有效學習環境的角色；(3) 學習者可

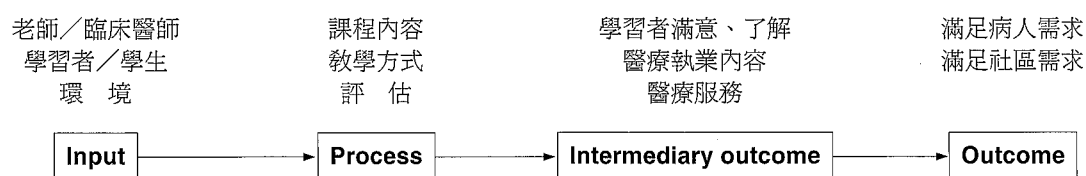


圖 1. 社區導向之醫學教育的組成要素

利用的學習機會；(4)學習者主動學習的態度。初步的信度與效度評估顯示，Med IQ是一個很不錯的評估工具，其評估結果可回饋給基層醫師、場所作為改善的參考，不過仍待更多驗證研究。

在2001年歐洲教育聯盟年會上，亦有許多學者提出新的成效評估方法，如加拿大的拉貝爾（M. Labelle）介紹了他們使用的書寫符合性測試（script concordance test），即由受試學生、全組同學及專家對於實際醫療問題給予和題意自符合（+2）到不符合（-2）之評定，再比較其吻合度，以決定學生學習之成效，此種方法頗有參考價值。此外，美國的哈登（R.M. Harden）提到，可重複使用之學習物件（reusable learning objects, RLO），就像拼湊玩具般，將每個小單位加以整合；以惡性黑色素癌（malignant melanoma）為例，這種RLO包括不同類似病變的照片、病患存活率之圖表、診斷流程等，哈登認為往後課程將由RLO透過學習指引（study guide），藉由課程譜圖（curriculum mapping）而成，而整個教學目標是以結果為基礎之教育（outcome-based education）。關於實習醫師及住院醫師的教學評估，醫學院學士後教育評鑑委員會（Accreditation Council for Graduate Medical Education, ACGME）認為，目前由於臨床醫師服務時間延長，影響教學時間的降低、醫院院際的整合、住院時間縮短、疾病急性度增加、照料破碎化、以及醫療失誤增加等，顯示教學環境之惡劣與成果不彰；因此他們建議用結果測量（outcome measurement）來改變教學，把良好之醫學畢業後教育和優質的病患照顧整合在一起，建立臨床角色典範（role model）的機制與方法，同時將人文（humanism）整合於訓練計劃中，相關的評估系統可以在JSA@ACGME.org的網站查到^[12]。

由上可之，除COME為比較全面的評鑑方

式外，目前發展出的評量模式多為醫學生臨床技能的測驗方法，至於基礎醫學及公共衛生領域之課程或教學評量，多是自製之量表，其評量內容不外是教學內容、教材內容、教學方法、教學整體評量、課程品質、學生回饋；進行的評估方式包括經由教師自評、學生的課程評估、學習滿意度問卷調查、教師互評、學生獲得之知能評量等^[13,14]。

我國醫學教育評鑑之發展^[15,16]

一九九八年美國教育部外國醫學教育及評鑑委員會（NCFMEA）曾批評我國缺乏嚴謹的醫學教育評鑑制度，認為我國醫學教育制度與美國制度不能相評比（non-comparable）。其中部份理由如下：(1)台灣無獨立而中立之醫學教育評鑑委員會；(2)我國未提出實據，說明對下列各項有關醫學院管理之質疑：(a)醫學院負責人（即院長等人）的經歷與資格標準；(b)醫學院執行教育有無符合大眾利益及專長理想；(c)醫學院教師之經歷與資格；(d)醫學院行政人員的能力、效率以及資格問題。(3)我國未提出實據，說明下列有關課程之質疑：(a)以筆試評量知識之外，有無評估學生看病的技巧；(b)除了基礎與臨床課程之外，是否有足夠的倫理、行為及社會人文等等課程。（對我國醫學院課程之此項缺點，美國教育者詬病已久）；(c)見習生（Clerkship students）的床邊教學如何進行。(4)我國未能對下列質疑，提出滿意的答覆：(a)如何確定入學學生性向，如何確定學生適合從醫；(b)如何考核學生。(5)教學設施與資源之利用方面之質疑：(a)教學設施與資源與學生人數比例是否適當；(b)師生比是否適當；(c)圖書館資源是否適當。(6)美方認為我國無明文規定評鑑周期。(7)美方認為我國未能明確說明醫學院評鑑工作，如何確保中立而避免偏袒與利益輸送，並如何有效地督責醫學院改進缺點。(8)美

方認為我國未能充分說明，評鑑是否以訪視方式進行。(9)美方質疑評鑑人員之資格與經歷，是否稱職。(10)認為我國沒有明文規定評鑑標準，也無規定評鑑不合格時之處置。

有鑑於此，教育部決定委託國家衛生研究院成立一個客觀獨立的醫學院評鑑委員會(TMAC)，民國九十年開始逐年展開週期性而遠較過去深入的評鑑訪視。民國九十一年三月，國內醫學評鑑制度獲美國 NCFMEA 評估通過。台灣醫學院評鑑制度與美國醫學院相容，除了與世界接軌有密切關係之外，也與美、台兩地醫學院是否能夠互相承認實習成績有關。

TMAC訪視之目標不是要強諸受評學校採用任何制式而統一之辦學方式，希望秉持公正、客觀及建設性運作，主要是依據美國模式注重下列幾個要點：(1)各校之辦學目標，亦即學校之使命，是否訂得明確易懂；(2)各醫學院是否將這目標對全校師生以及行政人員皆有清楚說明，醫學院所做的所有努力與所有措施是否與這個目標完全相配合。瞭解各校之指揮系統是否健全，是否合乎醫學教育機構之結構常理，使學校表述的目標易於達成。所謂指揮系統乃指包括董事會成員、校長、醫學院長、醫學系主任以及各科主任之任命機制、互動關係與配合度。訪視小組尤其注意醫學院與其附設醫院之隸屬關係與配合度；(3)醫學系之教育課程是否注意學生之人格成長、人文素養、社會問題之認知、倫理精神的紮根工作，而不只是知識與技術的傳授。使醫學生在進入臨床領域之學習之前，即培養了堅強不拔的濟人精神。訪視小組亦特別注意校方如何建立學生終身學習的習性；(4)醫學系是否對學生施以適量而平衡的生物醫學以及人類行為的基本知識，而這些教材是否考慮到如何與臨床醫學接軌；(5)醫學生在病房實習階段是否被賦予適當的臨床責

任，而其職責是否釐定清楚。同時，臨床教學課程的設計是否把學生納入醫療團隊，以激發學生學習動機；(6)校園文化如何經營，以薰陶學生之氣質；(7)校方如何評估學生學習之成效，學習成效也就是教學成效。校方除了由學生所填寫的評估表、或在國考的表現之外，是否訂有更客觀而可供自我檢討之評估方法。校方有沒有客觀可靠的方法評估學生看病的技能，而不是只測驗學生的臨床知識；(8)其他有關提升學生學習成效之計劃是否落實。

醫學教育評鑑制度在提昇醫學教育品質方面扮演著重要的角色，而醫學教育的好壞更直接影響國民健康與社會經濟。面對知識經濟的到來，教育的改革與整頓成為各方殷切的期待。醫學院評鑑制度的成功，對國內各項教育評鑑制度的設立指引了新的方向，對與國際教育的銜接更踏出了成功的一步。

CIPP 評鑑概述^[17,18]

發展源起

CIPP 模式發展於 1960 年代晚期，是當時最普遍的評鑑方法之一；以目標、測試和實驗設計為取向。1965 年美國國會通過中小學教育法案 (Elementary and Secondary Education Act, ESEA)，需要經由評鑑來決定教育補助款的發放，但當時的教育人員缺乏評鑑的訓練跟經驗，為了幫助教育機構改善他們的評鑑計畫，便成立了俄亥俄州立大學評鑑中心，並由 D. L. Stufflebeam 領導。

Stufflebeam 提出了幾項目標，作為發展評鑑的遵循法則：

1. 為教育機構提供評鑑服務。
2. 研究評鑑服務所得的經驗。
3. 建構更理想的評鑑方法。

4. 設計新的評鑑工具和策略，以實現新的評鑑觀念。
5. 訓練教育人員使用新的評鑑工具與策略。
6. 傳播與中心有關的工作及成果。

Stufflebeam認為，必須先決定關於方案實行部分的設計，例如：如何分配資源設備、如何溝通協調等，這些就稱為「實施性決定」(implementation decision)。爲了找出及解決運作上的問題，方案設計的工作人員和執行的工作人員必須接觸，持續觀察方案進行的活動，至少兩週一次向方案工作人員報告活動過程產生的結果。

另一個主要的決定種類是「裁決性決定」(recycling decision)，由於美國聯邦政府的補助以年度爲週期，這類決定關係到方案的終止或繼續、補助該增加或減少、方案的合併，以及方案是否制度化等問題。至此爲止已包含了過程評鑑 (process evaluation) 及成果評鑑 (product evaluation)；其中過程評鑑在當時是新的嘗試，結果評鑑則類似 Tyler 的目標模式。

後來有人質疑這樣的新評鑑方式忽略了對目標的評估，目標顯示方案的期望及企圖，是關鍵性的決定，於是 Stufflebeam 建議評鑑人員要先行評估和報告學生的需求及體系的問題，來協助選擇理想的目標，遂產生了「計畫性決定」(planning decision)，即背景評鑑 (context evaluation)。

但是另外有一個缺陷，就是沒有考慮到如

何決定要用哪些方法來達成既定目標和需求。這類決定包括對程序、期程、人力規畫、和預算等，必須在給董事會或基金會的計畫書中說明，而且一般會因爲選擇一種計畫而捨棄其它可能計畫，這就稱為「結構性決定」(structuring decision)，它建議採行可確立和評估各種方案設計相對優點的「輸入評鑑」(input evaluation)，來提供所需資訊。

至此，CIPP 架構經過幾次修訂、實驗及擴展後，基本模式已經完成，即以背景評鑑形成計畫性決定，幫助確立目標；以輸入評鑑做成結構性決定，協助方案設計的確定；以過程評鑑達成實施性決定，來指導實行過程，控制方案運作；以成果評鑑協助決策性的決定，以判斷及反應方案的結果（如表 1）；集合以上四種評鑑的第一個字母以方便記憶，CIPP 就成爲此一評鑑架構的簡稱^[17-19]。

類型定義

(一)背景評鑑 (context evaluation)

背景評鑑的基本原則是找出機構、方案、目標人口群或是個人所處的地位及環境，並提供執行的方向。主要目的是評鑑其對象整體的狀態，找出缺點與優勢，診斷問題以求解決改善。背景評鑑也需注意現存目標及優先順序是否滿足服務對象之需求。評鑑結果是調整現存目標和優先順序，以及改變需求滿足的參考基礎。背景評鑑應考慮下列問題：

1. 受評鑑機構有哪些需求需要達成？

表 1. CIPP 之決定類型與評鑑類型對照表

	意圖面	實際面
目的	計畫性決定 (由背景評鑑支持)	裁決性決定 (由成果評鑑支持)
方法	結構性決定 (由輸入評鑑支持)	實施性決定 (由過程評鑑支持)

2. 為達成這些需求應該訂定哪些目標？
3. 這些目標是否能得到各項資源的支持與配合？
4. 這些目標之中哪些最有可能達成？

背景評鑑有很多建設性的運用。它公開的分享對行政優缺點、需求、機會和問題優先性的看法，可促進學校行政溝通；同時針對地區緊急需求的建設性計劃。背景評鑑亦可使目標系統化，以便員工發展或修正程序；它也可以爭取到特定學校的支持。當然它也能夠時常幫助學生及家長將注意力放在更須進步的發展性區域。同時它可使用於如何刪減或促進某一方案。這些例子都是背景評鑑可用於支持個人或團體建立改進努力方向的考量。判斷成效的一個原則就是評估它們是否適當反應背景評鑑所找出的需求。最後，背景評鑑的紀錄也是用來驗證目標及優先性的極佳工具。

(二)輸入評鑑 (input evaluation)

輸入評鑑的主要在確定如何運用資源來達成方案目標，估算及找出合適的方法，以及說明和掌握方案的配置與延續性。它也可以找出推動方案的過程中，所需考量的當事人環境中的界線、限制和潛在可利用的資源。大體上，輸入評鑑的目的是協助當事人認知他們的各種需求及環境情勢，並且發展工作計劃；另一個重要功能是協助當事人避免浪費經驗在無用的改革計劃上，或使資源浪費減到最小。

輸入評鑑包括了一連串的步驟，但並沒有任何的順序。一開始可能是文獻收集與參考，以反映出特定需求的實際狀態；然後運用這些資訊來判斷是否還有潛在可行的解決策略，估算可行方法的效率及彈性，以便建議當事人是否要尋求新的解決辦法。然後，評鑑者要分析如何結合具有優勢的潛在計劃或意見；此外，實地參訪也可估計實際的資源和限制。

輸入評鑑主要的應用有：(1)協助評鑑委託者在其環境及需求的前提下，考慮各種可能的方案策略，發展適用的計畫；(2)提供判斷方案可行性的依據，避免資源的浪費；(3)提供專案補助申請計畫的參考；(4)評鑑現存方案與其他方案之間可能出現的競合情形；(5)說明何以選擇某種方案行動，而放棄其他方案行動；(6)檢視改革成本是否大於方案之利益；(7)提供對立衝突的團體一個架構或討論機會，以求達成共識；(8)此外，輸入評鑑研究的紀錄也能幫助掌權者對他們的選擇更負責。

(三)過程評鑑 (process evaluation)

基本上，過程評鑑是持續的檢視計劃的實施。目的是提供管理者及工作人員關於方案活動回饋，並且更有效管理可利用資源。另一個目的是提供修正或說明計劃需求的方針，因為不是計劃的所有層面都能獲得建議，且做出最初的決定後常會發現有瑕疵。還有一個目的是定期評估哪些參與者會接受方案並找出他們的規則。最後，過程評鑑提供廣泛的紀錄，包括方案實際執行與趨勢、以及各種成本的計算，還有觀察者和參與者對總體努力品質的評價。在實施過程評鑑時，要考慮下列問題：

1. 方案是否按計畫如期進行？
2. 在現行方案結束進入下個循環週期前，人員是否需要再接受訓練或輔導？
3. 現有的資源與設備是否適度而有效的加以利用？
4. 現行的方案在實施上最大的障礙為何？

一個方案的過程評鑑包含許多工作。首先要檢視方案計劃及監督先前方案的重要面向，像是員工發展會議、中心的發展及建立、學生的輔導、家長的聯繫、家教服務、員工訓練、課堂講學、實地考察、回家作業以及診斷測驗的運用等。另外還有一些找出重點的方法，評

鑑者可藉由顧問團及方案參與代表，定期徵詢他們所關心的議題和需求。

評鑑者對方案活動的觀察可以安排例行的資料收集活動；盡量不要太唐突，以免威脅或打擾到方案工作者，發展和諧的關係後，評鑑者便可使用一些結構性的方法。參訪中心的活動、檢視方案文件、參與員工會議和訪談重要參與者都可以大概了解方案的運作；亦可做一些資料收集的簡短報告，檢視他學到的，指出他認為重要的議題、計劃的誤差、以及不同人或次團體執行方案時不同的注意事項、敘述方案持續運作的特徵、並追蹤方案基本計劃的進展等。

過程評鑑方當事人要達成的目的是回饋員工，以期使他們完成所計劃的方案，或對持續進行的方案做必要的修正。一些管理者也會利用規律的過程評鑑的回饋機制作為維持員工盡責的工具。過程評鑑的紀錄對於責信也很有用，因為像是基金會、政策委員會或委託人基本上都想知道對方是否採納他們的提議。如果有人想做類似方案的話，過程評鑑也可以幫助其他人學習方案在做什麼；更甚者，過程評鑑是成果評鑑所必須的資料來源和說明，因為要了解方案的成果必須先知道他們在作什麼，那就要知道方案實際完成了什麼，就此方面而言，過程評鑑是為促進進步及支持責信，同時增進對研究狀況的了解。

四、成果評鑑 (product evaluation)

成果評鑑的用意在於測量、解釋和判斷方案的達成度，比較結果與目標之間的差異情形。成果評鑑主要目的是確定方案滿足服務族群需求的程度。成果評鑑應包含下列數項：(1)要廣泛檢視方案的效果，包括達成的和未達成的、正面的和負面的；(2)確認達成的預期和非預期結果；(3)提供決策者資訊，用以決定方案

是否繼續、修正或終止；(4)應該大範圍地從方案相關人口群取得意見並加以分析，有時也要跟研究中其他的方案結果做比較；(5)所達成的目標與之前狀態如何比較，以及達成目標所費成本有多值得；(6)對於方案計劃失敗，未達成目標或未滿足需求的解釋也相當重要；(7)最後，成果評鑑通常以一些既得利益者的角度來視察結果：從有不同需求或接受不同服務的方案領受者（個人或團體）收集資訊。

成果評鑑並無公式，但有許多可應用的方法。可以從先前的需求評估的結果開始，比較之前需求評估的概況來檢測成效；成果評鑑也可以從觀察者、雇主、或是方案領受者來評估成果。為使成果評鑑能符合期望的成果，評鑑者必須廣泛蒐尋非預期的結果，包括正面的和負面的。可以舉辦意見聽取會或團體會談以設定對於整個結果假設，並用臨床研究來追蹤假設是否成立；也可以使用立意抽樣以對方案效果有更深入的研究；或是讓觀察者視察方案與對照組的成果，以發展或證實兩者的區別，藉此發現方案的獨特貢獻。

成果評鑑結果的報告可以在不同階段提出。在方案進行期間可以不時提出期中報告，以顯示目標需求被達到和被滿足的程度。期末報告可總結已達成結果，同時對先前評鑑的需求、耗費的成本及計劃完成程度提出說明。追蹤報告可指出有哪些長期的影響。

成果評鑑的功能包含下列幾項：(1)決定方案是否繼續、重複或另案實施；(2)提供方案修正的方向，使需求得到更好的服務，並提高其成本效益；(3)協助有意採取此類方案者，衡量其利害得失；(4)成果評鑑具有心理上的暗示作用，亦即方案結果的優劣消長，會影響方案相關工作人員及實施對象的士氣；(5)評鑑結果是績效責信的報告重點，這些紀錄關係到方案是否繼續受到支持，或成為類似方案的參考依

據。

表 2 簡要的顯示出背景、輸入、過程、和成果的主要意義。這四類型的研究分別以相關的目的、方法和運用來定義。

評鑑流程

CIPP 在評鑑醫學教育領域上極具系統的觀點。它不只是管理的方針，也提供機構決策者持續評鑑服務。圖 2 說明 CIPP 模式的流程如何能導向維持和改善機構品質。

從圖 2 的左上角看，一個學校或機構的運作可能包括各種未經協調的評鑑活動，故他們需要定期接受特別的背景評鑑。背景評鑑的目的，有檢驗該機構服務對象的需求；以及掌握可利用的機會，如方案資金、改進教育技術、或尋求企業幫助等。找出機構內部有待解決的問題。評鑑機構目標和優先順序的績效。背景評鑑形成原因可能是機構內部例行性的評鑑，

或是對機構表現不滿意的反應；也可能是來自機構外部，像是機構被要求自我評鑑，或基金會要求的需求評估。總之背景評鑑有助於系統的革新、品質的提升，和提供更有效率的服務，以及診斷特殊的問題以謀求改進，並使機構成員了解該機構的優缺點。

基本上，背景評鑑的結果會決定此一系統是否需要進行改變。如果不需要改變，則機構成員維持往常的方案運作；如果考慮要有些改變，則方案的成員便要釐清待解決的問題和擬定其目標，並考慮有沒有明顯且適當的解決之道。

假如有明顯且適當的解決之道，則將其納入原機構所進行方案中使用，並加以評鑑。如果方案沒有顯而易見的解決之道，那麼就要進行輸入成本評鑑。輸入成本評鑑所進行的工作包括相關文獻的探討，可以請教有成功處理相關問題經驗的人員，或發揮機構成員的智力及

表 2. CIPP 所涵蓋的四種評鑑類型

	目 的	方 法	變遷過程中的決策
背景評鑑	定義機構的背景，找出目標人口群並評鑑其需求，找出發現需求的機會，診斷需求中的問題，並且判斷有效回應需求的方案目標。	使用系統分析、考察、參考資料、聽證會、面談、診斷測驗以及德菲爾技巧。	根據所提供的設置、與滿足需求或利用機會有關的目標，以及與解決問題相關的目的，也就是計劃需要的改變和提供判斷成果的基準。
輸入評鑑	找出和評估體系的能耐、各種方案策略，以及重要策略、預算、行程的程序設計。	藉著列細目及分析可利用的人力物力資源，分析策略及程序設計的適切性、彈性和經濟狀況；使用收集文獻、參訪典範之方案、提倡團隊和指導檢驗。	尋求支持的來源、分析策略和程序設計；也就是因應結構變遷的方案，並提供判斷執行程度的基準。
過程評鑑	找出過程中，程序設計或執行的缺點，改善前置方案的決策，並紀錄和評判各步驟和活動。	監督方案的潛在障礙並對意外保持警覺，獲取方案決策的所需資訊，詳述實際的程序並與方案執行工作人員保持互動及觀察。	執行和精緻化方案的設計與程序，也就是有效的過程管控，以及提供很多實際的過程以供日後說明成果。
成果評鑑	收集成果的敘述與意見，以及相關的目的和內容、輸入、過程的資訊，並且說明他們的價值和優點。	找出操作化的與可測量的結果指標，向利害關係人收集關於成果的意見，以及呈現品質與數量的分析。	決定是否持續、期限、如何修正或改變方案焦點，以及呈現清楚的成效紀錄（預期、非預期；正面、負面的）。

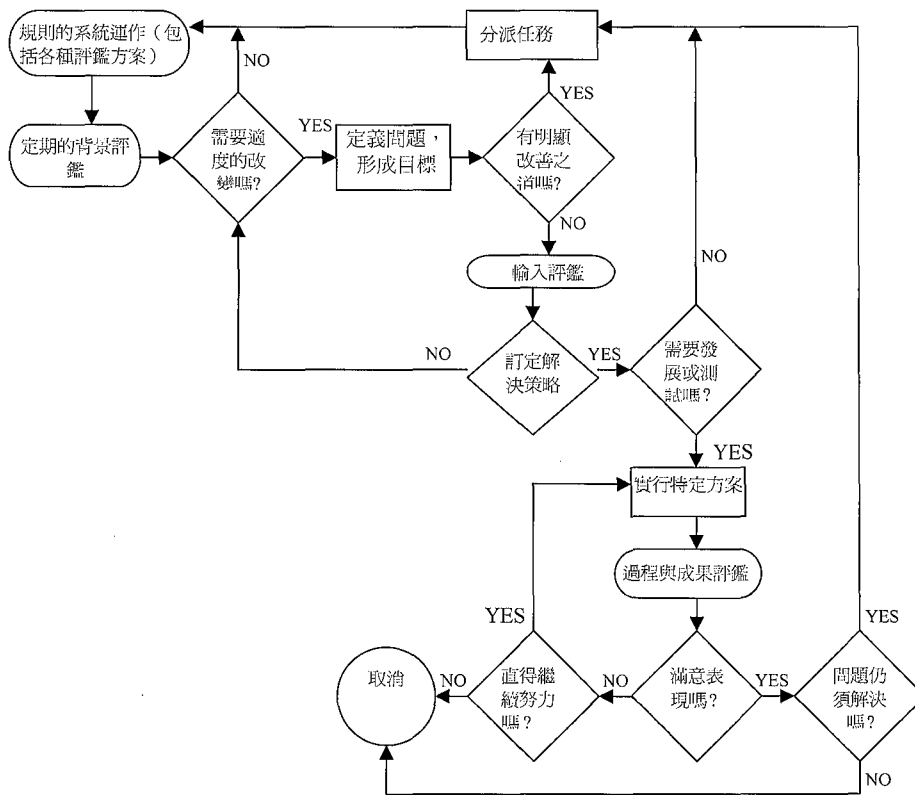


圖 2. CIPP 評鑑模式的流程與導向

創造力，也可網羅外部的專家，然後成立一個以上的小組來擬定一些策略。這些解決問題的策略要根據已確定的需求、問題和目標、理論是否健全、以及可行性等準則加以評估。

輸入成本評鑑的結果，是用來決定是否有充分可行的解決策略。如果沒有，工作人員要重新考慮這個變革是否很重要，如果是，則必須再找一個新的解決策略。如果已發現可行策略，則要決定實行這個策略時是否要進一步的測試。如果已經很了解這個策略，對於實施方面也沒有顧慮的話，那麼工作人員就可以把這個改變併入機構的例行活動中，而不必做進一步評鑑。

如果這個策略需要進一步的測試，就要進行策略的實地檢測。此時要從事過程及成果評鑑，以檢驗並修正實施的程序，並達到預期的表現和足夠的水準。假如此方案的表現不理想或成本太高，那麼機構領導者可能判斷此方案不值得繼續實施，便將它廢止；或者就算沒有滿意的表現，但仍有繼續努力的價值，就再回到測試階段，重新實地檢驗可行的策略。如表圖 2 的右下方，如果一個方案的實施相當成功，機構領導者認為之前需要改變的地方已不存在，就會終止這個方案；如果這個方案雖然成功，但解決問題的策略仍有持續實行的必要，那麼這個機構就要正式採用這個已經證明

為適用的方案，並定期評鑑。

由 CIPP 評鑑模式的流程圖看來，其評鑑流程從規律的系統化運作，經過定期的背景評鑑，到特定的輸入成本評鑑、過程評鑑和成果評鑑，以至決定是否要終止方案，或納入原有體系繼續運作。這種具系統改良觀點的評鑑方法，有下列特徵：

1. 評鑑是機構正規方案的一部份，而不只是改革方案的特別行動。CIPP 或其他評鑑方法的實行，只是機構正式或非正式評鑑活動的一部份。

2. 評鑑在刺激和計畫變遷方面扮演重要的角色。

3. CIPP 模式中每個評鑑類型的運用，是根據現有需求的資訊，而不是根據各評鑑類型實行的價值而定。換句話說，背景、輸入成本、過程、成果評鑑只是一個機構評鑑活動的一部份，它們最重要的功能在於滿足機構對於評鑑資訊的需求。

4. 新方案的發展應提供未來方案設置所需使用的評鑑。

5. 將評鑑的資訊加以記錄保存並提供公眾查閱，那麼方案不只能提供機構解決問題的方針，也能為特定方案的存續決定提供判斷的標準。

CIPP 模式並沒有提供特定形式和檢測的假設，它是透過背景、輸入成本、和過程的資訊，提供豐富的背景資料來說明和瞭解成果。

CIPP 模式在醫學教育評鑑之應用^[20,21]

現有醫學教育評鑑方式，主要是針對醫學生臨床技能做測驗，關於教學內容亦使用自製評量表或學生評量表；評鑑涵蓋元素最廣者即是 COME 品質評估，然其乃針對社區裡進行之醫學教育，且仍在發展測試中；TMAC 的成立為我國醫學院評鑑開創了新的展望，評鑑內容

大都是翻譯自國外既有資料，其內容固然豐富多元，亦考量各項品質水準的要求，且符合國際標準，但評鑑對象鎖定為醫學院；至於整體醫學教育的評鑑，我們仍缺乏一個具宏觀架構性的評鑑模式。是以在此提出 CIPP 模式，期能整合各項醫學教育評鑑元素及方法，找出完善的評鑑模式。

就前述 CIPP 評鑑模式的精神與內涵而言，CIPP 適用於醫學教育領域的主要特質至少可歸納為以下幾點：

(一)系統化的評鑑過程

醫學教育是一個整體性、連貫性的過程，影響其成效的因素廣泛且環環相扣；CIPP 可以透過背景、輸入、過程、成果的評鑑，將可能影響醫學教育效果的要素都納入評估，因此這種系統的評鑑過程很適合醫學教育，不但可以有效掌握醫學教育的全貌，更可以顧全各部分的關聯。

(二)兼重形成性與總結性的功能

CIPP 不僅可作為證明績效責任、選擇方法的總結性評估工具，更具有改良及發展的形成性功能，有助於啟發醫學教育的研究發展，更能診斷問題。醫學教育評鑑在我國正處於萌芽期，需要更多資訊與方法作為改進的參考，透過 CIPP 模式可以奠定深厚的基礎。

(三)適用於各層級的醫學教育

承上所述醫學教育可分為：(1)醫療（臨床醫學），(2)公共衛生（預防醫學）及(3)生物醫學科學（基礎醫學）等領域，而醫學生接受相關訓練的單位除了學校外，尚有醫學中心、醫院、社區基層醫療單位等不同層級的場所，故在評鑑時必須考量評鑑對象領域及層級的不同，CIPP 可以針對方案、機構、母群體、服務



對象，甚至個人做評鑑，沒有人或事的侷限，適用於各領域或層級的醫學教育。

四、適用自評與外部評鑑

CIPP 模式強調改良導向，且需要提供資訊來輔助評鑑，評鑑者要協助決定方案的計劃、結構、執行與考核，此時評鑑人員是內部評鑑者；但如果評鑑要提供績效責任時，也需要使用外部的評鑑人員。我國醫學教育評鑑委員會運作的形式，通常會請受評醫學院先行自評，再派專家學者前往評鑑考察，是一種需要內部評鑑與外部評鑑的過程，與 CIPP 評鑑模式的特質不謀而合。

至於如何運用 CIPP 評鑑模式來進行醫學

教育的評鑑項目，作者擬提供下列幾項作為參考：

(一)背景評鑑：應包括規劃層面的標準，醫學院成立的宗旨與目的為何、組織架構、歷史沿革、師生概況、現況如何、有何特色，問題或困難為何等。

(二)輸入評鑑：評估為達成醫學教育目標所投入的資源及運用情形。輸入評鑑時首重方案或機構之需求，它可以四個 W 和一個 H 方式來達成：

1. Who：管理者／董事會背景？教授及老師資歷？醫學生入學資格？行政人員有哪些？師生人數、比例多少？

2. What：醫學生需要學習什麼？醫學院可

表 3. CIPP 評鑑模式應用於醫學生臨床技能測驗方案

	一般資料需求	評鑑所需蒐集之資料	蒐集資料方法
背景評鑑	1. 醫學生基本資料 2. 教學者或評鑑者資料 3. 欲達成之臨床醫學技能，如：X 光判讀 4. 測驗方式，如：OSCE 或 SP	1. 醫學生年級、性別、經歷等 2. 教學或評鑑者資歷 3. 測驗技能之應用模式有哪些 4. 測驗方式之標準流程	1. 校方提供 2. 課程大綱 3. 教學者或評鑑者訂定
輸入評鑑	1. 授課時數 2. 學生練習機會與時間 3. 教學所費設備或經費 4. 測驗所需設備與經費	1. 課程準備項目 2. 課程進度表 3. 預算表 4. 測驗標準流程	1. 教學者提供 2. 可利用設備列表 3. 相關文獻研究
過程評鑑	1. 授課內容與進度 2. 授課方式 3. 學生出席率 4. 學生提問或討論率 5. 學生自行查閱相關資料時數或比率 6. 教學或學習困難	1. 授課紀錄 2. 課程綱要與教學完成率 3. 運用圖書館、網路或其他資源 4. 小考成績或期中教學狀況檢討 5. 問題解決	1. 期中、期末測驗 2. 教師教學紀錄 3. 課堂或學校日誌 4. 學生分組紀錄 5. 學生筆記 6. 資源使用紀錄表
成果評鑑	1. 學生學習滿意度 2. 教學完成度 3. OSCE 成績 4. SP 評估結果 5. 非預期效益（如：小組討論之啟發）	1. 學生自我期待與實際達成率 2. 教師預期目標與實際達成率 3. 課程經費結算表或收支明細 4. 課程優缺點與改進建議	1. 課程實施前後學生技能改變狀況 2. 與去年及前年之授課成果比較 3. 教師自評紀錄 4. 學生對課程之評量表 5. 其他測驗或臨床技能相關研究報告

以提供哪些課程？教授及老師應規劃什麼樣的課程內容？學校可運用資源有哪些？需要多少經費及設備？

3. How：董事會或校方管理者如何訂定目標？如何做決策？學校如何招收學生？如何聘僱教師及行政人員？學生如何學習？學習方式及管道有哪些？課程或方案如何執行？

4. Where：學習環境及授課地點在哪裡？是否便利實用？需要什麼空間或特殊地點？

5. When：董事會或校方協調會議召開時間頻率？教師每週開課時數？學生每週上課、實習時數？師生各花多少時間備課？課程安排順序等。

經由以上的分析，更能瞭解醫學院師生的需求，以及如何提供良好醫學教育制度以滿足其需求。另外要掌握評鑑經費，許多機構經由評鑑了解各種方案是否達成預期目標與效果，故應明列預算、經費來源、支出項目及明細等。再者是要有良好的組織架構與人力配置，行政、計劃、執行等工作人員，最好邀請專家學者或相關人員參與規劃。

(三)過程評鑑：過程評鑑可定期提供回饋，同時亦能評估工作任務之掌握與推動情形。包括醫學教育的內容、方式、目標等。要掌握實施過程中是否按學年計劃及課程大綱進行，如細部的活動、工作任務、教材、配套措施、時間進度的掌控、階段性目標達成率、校園方案或課程推動之困難等等。

(四)成果評鑑：評估工作目標的達成情形。像是：學生學習成效、教師教學滿意度、行政問題解決的程度、醫學院整體的改變等等。

要言之，CIPP 評鑑模式是以敘事方式說明評鑑的計劃，重點置於結果的分析與報導。它是以「多種」的研究法達到評量的目的。表二是將 CIPP 評鑑模式應用於醫學教育之醫學生臨床技能測驗方案之項目一覽表。作者期待經

由 CIPP 模式能協助判斷評估各項醫學教育措施的成效，並期待受評鑑者能內化為自我評鑑的習慣。

誌 謝

本報告承蒙衛生署提供經費補助，以及湯茹雲、鄭智文兩位助理協助文書整理，謹致謝忱。

參考文獻

1. 宋瑞樓：台灣醫學教育改進方向之我見。台灣醫學 2003；1：68。
2. 宋瑞樓：醫學教育改進之研議。醫學教育 1998；4：369。
3. 謝博生：現代醫學在台灣。台大醫學院：金名圖書。2001 年 7 月初版。
4. 劉克明：改進國家執照考試方式以評量台灣醫學教育之目的。醫學教育 2002；4：379。
5. 劉克明：參加 2002 年歐洲醫學教育學會會議：迎向優質教學。醫學教育 2002；3：357-8。
6. 劉競明：21 世紀醫學教育的新趨勢：醫學教育評鑑與小組互動式教學。台灣醫界 2002；3：153。
7. 陳晶瑩：以實際表現為基礎之評估方式。醫學教育 1999；4：407-10。
8. 王維典：臨床醫學教育—以標準病人為教、學及評量工具。醫學教育 1998；4：378-83。
9. 劉克明，黃裕勝：應用英美醫學生臨床技能測驗模式落實台灣醫學生臨床技能測驗評估。醫學教育 2003；1：16。
10. 劉文俊：社區導向醫學教育的教學品質評



- 估。醫學教育 1999；3：.290-1。
11. 劉文俊，梁繼權，王維典，邱泰源，陳慶餘，呂碧鴻：醫學生社區醫學教育學習成效之評估。醫學教育 2002；4：425-31。
 12. 賴明亮：歐洲醫學教育聯盟 2001 年會：世代改變中之醫學教育及標準。醫學教育 2001；5：254-55。
 13. 朱宗信：教學品質的評估。醫學教育 1999；3：197-99。
 14. 高美英，李明濱，呂碧鴻，何弘能，黃天祥，謝博生：醫學系教學評量結果之三年追蹤分析。醫學教育 1999；1：71。
 15. 黃崑巖：美國教育部國外醫學教育與認可委員會－會後報告，2001。
 16. Questionnaire Standards and Processes used by Countries to Evaluate Medical Education. National Committee on Foreign Medical Education and Accreditation. March, 1998.
 17. 唐雅蘭：CIPP 評鑑模式應用在特殊教育推展成效評估之研究－以花東地區啟智教育為例。碩士論文，2000:14-6
 18. Madaus, GF Scriven, MS Stufflebeam DL: The CIPP Model For Program Evaluation, Evaluation Models: Viewpoints on Educational and Human Services Evaluation. Boston: Kluwar-Nijhoff Publication, 1993:118-22
 19. 游進年：CIPP 模式在台灣省國民中學訓輔工作評鑑應用之研究－以宜蘭縣為例。碩士論文，1999。
 20. 呂春嬌：從 CIPP 評鑑模式談圖書館的評鑑。大學圖書館 1999; 4: 15-28。
 21. 黃慧文：以 CIPP 評鑑模式建構我國遠距高等教育評鑑項目之研究。1999，碩士論文。

Feasibility of Using the CIPP Model to Assess Medical Education

Tsung-Tsair Yang, M.D., Ph.D.¹, Ming-Been Lee, M.D.², Yin-Chang Wu, Ph.D.³

The purpose of medical education is to educate students concerning: (1) clinical medicine; (2) public health; and (3) biomedicine. Since 1992, the medical schools in Taiwan have begun a series of educational reforms that include the implementation of small group tutorials, integrated courses in basic science and clinical medicine, innovative teaching strategies for the medical humanities and medical ethics, etc.. However there has been very little research with respect to outcome evaluation of these programs or particular projects. Adequate assessment or accreditation methods are one of the most important ways to provide feedback and thus improve the quality of program implementation. In the past, most evaluations have centered on individual student competency. Popular evaluation methods include the Standardized Patient (SP), the Objective Structure Clinical Examination (OSCE), the Clinical Practice Examination (CPX), and Community-Oriented Medical Education (COME). Most of these methods are used to evaluate clinical competency, and we still lack a method applicable to program evaluation in the educational fields of biomedicine and public health.

An evaluation system plays an important role in improving the quality of medical education. In 1998, the National Committee on Foreign Medical Education & Accreditation (NCFMEA) of the United States were

critical of medical education in Taiwan because there was no accepted method that was used to assess medical education in Taiwan, and they judged Taiwan's medical education "non-comparable" to the American system. At that point, the Ministry of Education (Taiwan) requested the National Health Research Institutes to establish the 'Taiwan Medical Accreditation Council (TMAC)'. TMAC put forward a new vision for the evaluation of Taiwan medical schools, but it still lacked a macro and structural model for evaluation. Therefore, the authors aim to introduce the CIPP model and discuss its applicability to accreditation of medical education. The CIPP model was developed for evaluating the institutes of education in the late 1960s, and is directed at concrete goals, tests, and experimental designs. The model is composed of the following four parts: context evaluation (C), input evaluation (I), process evaluation (P), and product evaluation (P). It is a comprehensive evaluation model that integrates crucial evaluation issues and methods of medical education. The authors think the CIPP model is applicable to medical education because of the following characteristics: 1) the process of evaluation is systematized; 2) it has both formative and summative functions; 3) it can be applied at each level of medical education; and 4) it can be used both internally by staff or externally by experts. (Full text in Chinese)

Key words: evaluation, medical education, CIPP model
(J Med Education 2003; 7:218 ~33)

Department of Psychiatry, Cardinal Tien Hospital and Fu-Jen University School of Medicine, Taiwan¹ Department of Psychiatry, National Taiwan University College of Medicine and Taipei City Psychiatric Center² Department of Psychology, National Taiwan University³.

Received: July 1, 2003; Accepted: August 28, 2003.

Correspondence to: Ming-Been Lee, M.D., Taipei City Psychiatric Center, 309 Sungde Rd., Taipei, Taiwan 110.