



對畢業前醫學教育核心 EPAs 的審思

前言

今年3月號 *Academic Medicine* 刊登一篇在撰文時仍在哈佛醫學院擔任評量中心主任的 Krupat E. 教授 (PhD) 所發表的個人觀感——“Critical Thoughts About the Core Entrustable Professional Activities in Undergraduate Medical Education”。作者在文中透露出他對可信賴專業活動 (entrustable professional activity, EPA) 不以為然之處，其所持理由部分頗有道理，但由於 Krupat 只有 PhD 學位，故沒有臨床教學和評估醫學生能力的實戰經驗，其對能力評估的感受與領悟當然會有隔靴搔癢的情形，誤解的地方也不少。但無論如何，在能力導向醫學教育 (competency-based medical education, CBME) 當道，里程碑 (milestones) 和 EPAs 被奉為醫學教育的神聖教條的此刻，有人對 EPAs 提出批判，其論點當然值得一讀，更值得深思。

內容摘要

作者在本文首段即開門見山，說出心裡的話：「我不喜歡核心 EPAs」(I do not like the Core EPAs)。作者以美國醫學院學會 (American Association of Medical Colleges, AAMC) 所提出的 13 項核心 EPAs 為例，對 EPAs 作出批判。他將 EPAs 的問題歸納為三類——內容與焦點 (content and focus)、測量與評估 (measurement and assessment)，以及其他問題。分述如下：

一、內容與焦點

1. 教學內容的問題：

- (1) 作者引述 Tekian A. 於 2017 年初在 *Medical Teacher* 撰文批評 AAMC 的 13 項核心 EPAs 的論述“Are all EPAs really EPAs” (見文獻來源 4)。該文指出 13 項核心 EPAs 並非全部都是真的 EPAs，其中三項不是臨床上獨立執行的任務 (如協調跨領域團隊)，也有項目較像是教育目標而非任務 (如形成臨床問題及搜尋證據 [EPA 7])。
- (2) 核心 EPAs 的範圍歸屬頗有爭議，例如病史詢問和身體診察應屬於兩項 EPAs 卻被結合為一項 (EPA 1)，而文件記錄 (EPA 5) 和口頭報告 (EPA 6) 則應可視為一項 EPA——「對臨床互動過程作出摘要」。
- (3) 深淺程度有很大的落差，例如交班 (EPA 8) 和認知病人需要緊急照護 (EPA 10) 應非醫學生的必修項目。另外，一位有經驗的畢業後醫學教育的教師告訴作者，識別系統錯誤 (EPA 13) 是第二年住院醫師的課題。

2. 教學焦點的爭議

- (1) 作者引述 ten Cate 所言，將 EPA 概念自畢業後醫學教育 (graduate medical education, GME) 推展至畢業前醫學教育 (undergraduate medical education, UME)，是以醫師被許可做什麼事來重新定義「醫師」。他擔心的是未來會出現「唯 EPAs 獨尊」的醫學教育 (筆者按：EPAs 是臨床上可被觀察和可被測量的工作單元，除了 EPAs 之外，醫學教育還須包含許多不能被



觀察且不易評量的重要元素)，並以圖形-背景關係 (figure-ground relationship) 來比喻 EPAs 和醫學教育其他內容的關係。他對此提出質疑，因為「評估驅動學習」是可以預見的，在以「EPAs 至上」的醫學教育課程中，將會產出只熟知如何滿足照護所需的技術和有能力完成某些使醫院有效運作的任務的醫師 (筆者按：就是所謂的「醫匠」)。在執行以 EPAs 為主軸的醫學教育時，教育者必須思考如何將建立醫病間的照護關係視為獨立行為單位來觀察和評估，這恐怕不會比形成臨床問題和搜尋證據 (EPA 7) 來得容易。作者建議使用一些既有的框架來訂出操作元素，如四習慣模式 (Four habits Model)，使能評估照護關係的可信賴狀況。

- (2) AAMC 推動核心 EPAs 是為了要銜接 UME 和 GME，使專科訓練計畫主持人能以 EPAs 的表現作為錄取住院醫師的依據，並在住院醫師報到的第一天就能掌握住院醫師已具備什麼能力。然而，研究證據顯示，專科訓練計畫主持人最關心的是住院醫師的個人素質 (例如與他人合作的能力、可靠性和誠實等)，而不是他們是否已對這些核心 EPAs 獲得信賴。

二、測量與評估

以 EPAs 為主軸的醫學教育，對能力評估的度量標準是從根本作出改變，在 EPAs 問世以前各種評估方法所做的判斷 (judgment) 均被「信賴」(entrustment) 所取代。從正向思考，信賴的概念在直覺上很有吸引力，大多數醫師都很熟悉並且習慣於思考受訓者是否能完成某些目標和任務，因此接受度並不會低。然而，EPAs 的評估在許多方面都不符合社會科學 (筆者按：指教育學) 測量理論中許多有關優質測量的既定原則。EPAs 在測量與評估的問題如下：

1. 欠缺合理的評分系統 (rating system)。
 - (1) 以觀察行為來判斷是否勝任並無任何標準。
 - (2) 因不同情境對行為有不同的需求和影響，如何從行為來產生信任和決定需要的監督，很難訂出一致的標準。
 - (3) 可信度判斷 (entrustability judgments) 最終必超越目前 EPAs 所要求的知識和技能，還須評估真誠、責任心和辨別力等專業特質。
2. 測量欠缺明確標的 (end points) 或支點 (anchors)。
 - (1) 醫學生所須學習的 EPAs 之中有不少是以五個監督層次中的「間接監督」(indirect supervision) 作為測量標準，但「間接監督」卻是十分模糊而難以有一致的定義。
 - (2) 另一個有爭議的監督層次是「可監督他人執行」，因監督能力與執行能力是在不同的層面，故違反了測量的基本規則——同一項測量的各評分點須在同一面向之上。

三、其他問題

1. UME 及 GME 需要達成不同監督水平所相應的能力，在銜接上仍有問題。
2. 雖然 CBME 強調「對能力的要求恆定，訓練時間須有彈性」。但目前推行核心 EPAs 的訓練計畫幾乎所有學員都是同步達標，意味著對學員的要求乃是



「最低標準」，所以大家才會同時變成「可信賴」(entrustable)。

3. EPAs 的評估十分主觀，充滿偏誤 (biases)。

最後，作者對核心 EPAs 作出下列的結論、建議和結語：

一、結論

1. EPA 導向的教育系統並未增加評分的信度和判斷的效度，亦無助於造就性回饋 (formative feedback) 或總結性決定 (summative decision)。
2. 過於著重平凡的技术性技巧而扭曲了醫師的角色。
3. 違反多項優質評量的基本原則，無意中造成許多偏誤。

二、建議

1. 採用單一混合模式，稱之為醫療任務 (tasks of medicine, TOMs)。
2. 為 TOMs 發展出涵蓋全程專業發展 (自醫學生至專科醫師) 的一套能力描述符 (descriptors)。
3. 對於能力要有客觀的評估程序，建議將執行過程攝錄成影帶，供不知學生背景和表現的教師評估，以求客觀。

三、結語

EPAs 像一列已啟動的火車，不但已駛出車站，而且載著相當的動能勇往直前。然而，由於存在著許多問題，應該遵從“slow down when we should”的守則，暫緩對 EPAs 的推動並重新評估，待釐清疑慮之後再重新出發。

評論

在本文投稿之後，ACGME 的「里程碑發展與評量」資深副理事長 Holmboe ES 醫師立即撰文為 EPAs 政策辯護 (見文獻來源 3)。文中並未對 Krupat 教授的質疑直接回應，僅提到 CBME 是必要的改革，也是一項不容易的任務等等，內容之空泛令人失望。不但如此，他在文中一段企圖說明為何大多數醫學生或醫師都可同步取得「可信賴」的說辭呈中現出錯誤的觀念：

On the assessment side, the balance needs to shift from the 20th century's heavy reliance on high-stakes "proxy" assessments, such as multiple-choice question (MCQ) examinations and objective structured clinical examinations (OSCEs), to work-based assessment in the clinical space, especially in GME. That is not to say that MCQ exams, for example, do not have value—they do—but at some point we in the educational community need to recognize that despite decades of advances in highstakes MCQ exams and OSCEs, we have the health system performance we have.

觀念錯在誤以為 workplace-based assessment 才是評估臨床能力最好的地方 (這是許多竭力宣揚 CBME 或 EPAs 的人常有的謬誤)，事實上不同的評估方式有不同的角色，如醫病溝通、人際技巧、急症處置、團隊默契，這些重要但不易在日常工作時測試的能力都有賴 OSCE 等在模擬狀況下的評估方法，因為這些情況在工作場所所有取樣問題 (George Miller 在他的演說和文章中提到)，甚至根本不容許在工作場所中進行這些評估。另一方面，像 mini-CEX 或 DOPS 等在工作場所的客觀性皆不如 OSCE。因此，過度強調工作場所是評估臨床能力最重要的場域，並不合宜。至此，筆者建議可以忽略 Holmboe 的回應。

言歸正傳，以下為筆者對 Krupat 教授撰文的評論。

首先要注意的是作者的背景，包括提出 EPAs 的 Olle ten Cate、在 2017 年批

判 AAMC 核心 EPAs 的 Ara Tekian，以及本文作者 Edward Krupat，他們都僅有 PhD 學位而不具 MD 資格，由於毫無實戰經驗，這些學者在評論臨床醫學教育都是靠學理判斷，易生謬誤是可以理解和預期的。

Krupat 在文章之初批判核心 EPAs 主要是引用 Tekian 的文章。Krupat 似乎完全支持 Tekian 錯誤的論點，理由如前所述，就是因為不了解臨床醫療和教學的情境。在 13 個核心 EPAs 之中，Tekian 誤以為 EPA 6(對臨床案例作口頭報告)不是真正的 EPA，是因為他不知道 case presentation 是醫師的例行專業活動；他誤以為 EPA 7(形成臨床問題並蒐尋證據以推動病人照護)不是真正的 EPA，是因為他不知道實證醫學(evidence-based medicine, EBM)的方法學；他誤以為 EPA 9(善作跨領域團隊成員)不是真正 EPA，是因為他不知道 TeamSTEPS 等課程；他誤以為 EPA 13(識別系統失誤及促進病安文化)不是真正 EPA，是因為他不知道醫品病安的制度和活動。這四個項目理應可以列為 EPAs，但由於 Tekian 和 Krupat 對相關知識知道的不多，加上筆者猜測他們沒有仔細閱讀 AAMC 的出版物(見文獻來源 4)，所以才會產生誤解。至於 Krupat 建議 EPA 1(病史詢問及身體診察)須分為二項，筆者頗為認同，但要將文件記錄(EPA 5)和口頭報告(EPA 6)視為一項 EPA 則純屬門外漢的胡言，因為製作病歷和口頭報告案例應是兩項不同的能力。

至於在教學焦點的爭議和測量與評估等陳述的各項問題，筆者則十分贊同。也是近年來筆者對 CBME 過度強調只有可被觀察和被測量的能力(例如以 EPAs 來呈現)才是教育和決定是否能信賴的依據，感到不妥和擔憂的地方。

對於作者提出的三個結論筆者原則是贊同的，特別是“slow down when we should”的概念更是值得肯定。然而，接下來作者的三個建議則甚有問題：(1) 推出所謂的 TOMs 並沒有比 EPAs 好，反而會令大家更加混淆；(2) 發展出涵蓋全程專業發展的一套能力描述並不符實際需要；(3) 主觀性評估(subjective assessment) 主要問題常在評分者的信度而非是否欠缺盲性，且醫院環境及資源無法做到也不宜去做評估過程的現場錄影。

總括而言，Krupat 這篇個人觀點很值得閱讀，但卻切勿照單全收！

文獻來源

1. Krupat E. Critical Thoughts About the Core Entrustable professional Activities in Undergraduate Medical Education. *Acad Med* 2018; 93: 371–376.
2. Takian A. Are all EPAs really EPAs? *Medical Teacher*. 2017; 39: 32–233
3. Holmboe ES. Competency-based Medical Education and the Ghost of Kuhn: Reflections on the Messy and meaningful Work of Transformation. *Acad Med* 2018; 93: 350–353.
4. AAMC. Core Entrustable Professional Activities for Entering Residency. Curriculum Developers' Guide. 2014.