教學經驗與分享

醫學教育研究與測量工具選擇及發展

◎文／徐祥清、馮瑞鶯

前言

首先，很感謝眾多師長的支持與鼓勵讓我有機會能夠獲得 109 學年度教學特優教師的殊榮，我相信在成功大學醫學院與附設醫院中的每一位臨床教師，不僅在醫療環境中照顧患者，在此同時也肩負臨床教學的責任，這是我們身為臨床人員的工作與使命，在此同時我們也彼此分享對於教育的熱誠與成果。我想藉著這個機會，分享護理系馮瑞鶯老師對於醫學教育研究多年來的心得與做研究的方法，希望能夠在此困難的時期彼此互相成長。

醫學教育研究是推動醫學教育精進與改革重要的一環，而其研究範圍包羅萬象，舉凡對於教育相關的議題都是不斷被探討的主題，但議題的熱門度及重要性隨著時代及科技人文思維的變革而不同。我們從 social network analysis 來分析從過去到現在有關醫學教育研究的主題，可以發現在 1997-2005 年最常被探討的前三大主題是 internet, evidence-based medicine, distant medicine，

而自2006-2015年的熱門主題則是young adult, quality improvement, social media，從技術層面轉至學習者文化與更好的教學及照護品質，由此可見，醫學教育的主題是會隨著社會環境轉變而不斷的與時俱進。如果從topic cluster上來看，許多主題是彼此互相關聯的，而從最大的族群來看則包括有clinical competencies, internship and residency, education medical graduate, questionnaires, educational measurement是最被常探討的，其中問卷與測量被提及的機率增加，精確測量與品質改善的關係不容小覷。再從另一個觀點來看，在教學過程中的，從老師、學生、知識技能態度的養成、醫病關係、到預後的分析，當中每一個階段都是能夠被探討。

在強調醫學教育的成效同時，也會發現容易被詬病不夠嚴謹，這也使得教育成效是否能夠被客觀量測、能否被複製或與其他文獻結果做科學性比較成為焦點。過去發表的文獻常見的困境包含缺乏使用一致的測量工具來評值類似的成果、缺乏使用具有信效度的工具，或者有的研究雖然使用現存工具，但使用時卻只針對該單篇研究目的或內容，挑選該測量工具的一部分或感興趣的題項，忽略該測量工具已建立好的信效度之概念完整性，因此使得知識建構變得殘缺。因此在做研究之前需要思考，欲探討的主題的重要性是什麼、在這個主題中有何未知的部分、對於未來的思考性等問題，而對於過程與成果則須考

量測量的過程、研究結果的可信度等相關問題。

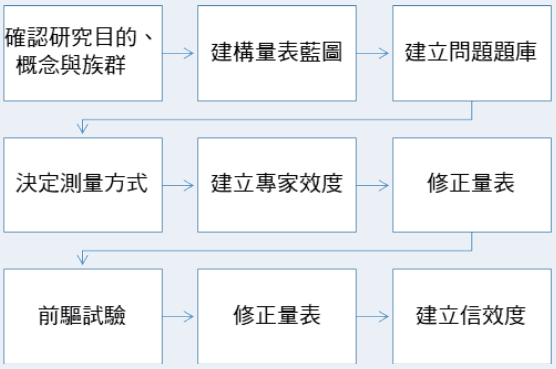
因此，正確、精確測量是知識建構的一個重要步驟，要了解測量工具的選擇與發展之前，必須要先了解什麼是測量及工具發展。測量 (measurement) 就是量化特殊現象的方法，或者是利用標準化的規則，將物件轉化為客觀數值的過程。工具發展 (instrumentation) 指的是以特定規則發展工具，並藉由可信證據評價研究結果。這個過程越嚴謹，所發展出來的測量工具就越具有代表性，對於各個科學領域的知識發展是很重要的。常見的測量工具很多，包括問卷、篩檢量表、量尺或生物檢驗法等，都是常用來作為滿意度調查、教育成效評值或者學術研究。而檢視這些工具發展及信效度驗證過程，則是選擇工具的重要原則。

通常採用現成的測量工具是比較快速且方便的，且因為它可能已經被多次測試，可能的錯誤或誤差也已被處理，加上使用相同量表的研究結果在科學文獻上也可以被比較。然而，也會有些困境，包括工具是否有授權、有中文版、或是否必須要付費，同時須考量原量表所施測的對象是否跟與現行研究類似或相同，內容與可讀性是否需要修改與驗證，如果無法找到適合的量測工具，就得發展新的量測工具。

好的量測工具得具備有良好的信效度，才能真實反映要量測的變項，研究結果之證據才能被採信，無論研究設計再如何嚴謹，臨床試驗控制或盲化做得再徹底，使用不良的量測工具會產生所謂資料的 GIGO (garbage in, garbage out)。如果將選用或者評價現存工具，必須注意以下幾方面，(1) 使用該工具的目的與目標：每個工具發展時都有其相對的目標，包括描述、診斷、篩檢、選擇、或預測等；(2) 對象族群：是否符合測量對象的背景與條件；(3) 使用環境：工具本身是否再在特定的場合中被發展與使用；(4) 時間觀點：受測者回憶的時間點或者時間點中的事件；(5) 概念基礎：工具的測量概念與使用者的觀點是否一致、或現存概念與工具發展的符合度；(6) 信效度。

如果選擇發展新的測量工具，我們必須認識一般測量工具發展的流程。在發展前，須清楚確定研究或測量的目的，例如是為了要診斷、篩檢或者是了解此現象在群體中的狀況，例如，調查性研究用來瞭解狀態改變的程度或差異，也就是

要先能夠清楚地釐清要測量的是什麼，量表中的內涵如態度、知識、技能、意圖、行為、感受、認知等，然後透過理論與文獻查證已確立概念的是否正確。其次，對於受測的族群的特質，例如；種族、性別、年齡等，其語言與閱讀能力。第三，建構量表藍圖，包括兩個面向，及測量的內容與能力，也就是範圍與層次。第四，建立問題的題庫，通常這些問題必須符合測量目標且可反映預測概念，問題最好呈現過度飽和的狀態，並以不同的切入點來測量現象；這些題庫的來源可以參考相關量表，或根據研究理論自行訂定。問項的建立須注意，閱讀的程度以國小五年級到國中一年級的閱讀程度為依據，注意語法與文法、避免難懂的話語，問項須完整，避免過長、縮寫、雙重含義、模糊不清或模稜兩可，用詞須避免偏見，避免使用否定、被動式與代名詞。第五，測量反應格式，可以有語意差異量表 (semantic differential)、李克特量表 (Likert scale)、視覺類比量表 (visual analog scale)、與二分量表 (dichotomous)。最後，建立專家內容效度；專業的專家可以是研究學者或臨床專家，目的是檢閱量表的初稿、整體格式進行評估是否有代表性、清晰性與完整性，整合專家意見後進行修改與刪減內容。然後，我們可以進行約 30-50 人的前驅測試，評估量表的明確性、指引的有效性、答題的完整性與完成問卷所約的時間，並瞭解可能產生的問題與誤差，同時執行初步的信度評估，校正量表內容。最後，再以更大的樣本驗證其包含信效度之心理計量。



結論

以上只是大略的介紹醫學教育研究工具的選擇與發展，我們在發展工具的同時，也須瞭解醫學科學與教育兩者是會隨著時代而改變，沒有什麼研究的結果是絕對的，狹義的來講教育是種制度，但廣義來說則是傳播人類的文化、知識、技能與經驗的活動紀錄，能夠探究醫學教育的過程並從中以科學的角度來探討與改進，是我們推動醫學教育進步的動力。