

หลักการการประเมินผลทางแพทยศาสตร์ ศึกษา ประสบการณ์จากการร่วมการประชุม วิชาการแพทยศาสตร์ศึกษานานาชาติของ สมาคมแพทยศาสตร์ศึกษายุโรป “Association of Medical Education in Europe (AMEE) 2014”

พศ.ดร.พญ.ฉันทนา สิริธัญญ

ภาควิชาสรีรวิทยา, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร 10700.

จากการร่วมการประชุมแพทยศาสตร์ศึกษานานาชาติ ของสมาคมแพทยศาสตร์ศึกษายุโรป (association of medical education in Europe; AMEE 2014) ณ เมืองมิลาน ประเทศอิตาลี ระหว่างวันที่ 30 สิงหาคม พ.ศ.2557 ถึงวันที่ 3 กันยายน พ.ศ.2557 ซึ่งเป็นการประชุมเกี่ยวกับการเรียนการสอนด้านแพทยศาสตร์ในด้านต่าง ๆ โดยหัวข้อที่ข้าพเจ้าสนใจและได้เข้าร่วมการจัดประชุมเชิงปฏิบัติการคือ ทักษะสำคัญในการประเมินผลทางแพทยศาสตร์ศึกษา (essential skills in medical education assessment; ESMEA) ซึ่งเป็นการอบรมที่จัดขึ้นเพื่อให้อาจารย์ในวิชาชีพแพทย์และวิชาชีพเกี่ยวกับการดูแลสุขภาพ เข้าใจเกี่ยวกับหลักการพื้นฐานและความรู้ที่ใช้เกี่ยวกับการประเมินผลในทางแพทยศาสตร์ศึกษาในปัจจุบัน โดยการจัดอบรมเชิงปฏิบัติการนี้จะจัดการเรียนการสอนเป็นกลุ่มเล็ก ๆ การนำเสนอผลงานสั้น ๆ เนื้อหาจะประกอบด้วย หลักการพื้นฐานของการประเมินผล (principles of assessment)

การเขียนแผนผังการสอบ (blueprinting) การออกแบบการประเมินผล (designing assessments) การออกโจทย์ที่มีหลายตัวเลือกให้ดีขึ้น (writing better multiple-choice questions for your exams) การทดสอบทางคลินิกโดยใช้วัตถุประสงคเป็นโครงสร้าง (objective structured clinical examination, OSCE) และการตั้งค่ามาตรฐาน (standard setting)

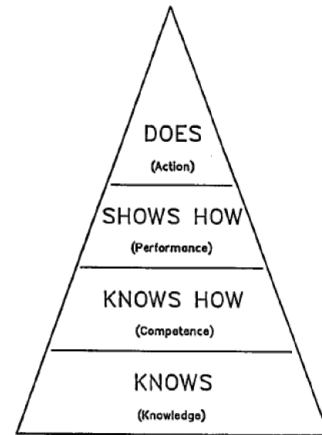
การประเมินผล (assessment) คือ การกระทำอย่างตั้งใจหรือเป็นทางการ เพื่อที่จะได้รับข้อมูลเกี่ยวกับความสามารถ (competence) และสมรรถภาพ (performance) ของนักศึกษาซึ่ง assessment มักจะใช้ในบริบทของการเรียนรู้ของนักศึกษา ส่วน evaluation มักจะใช้ในบริบทของโปรแกรมทางการศึกษา (educational program)¹ โดยการประเมินผลเป็นการวัด ความรู้ (knowledge) และทักษะ (skill) ซึ่งบอกถึง ความสามารถทางคลินิก (clinical competence) โดย J Cohen ได้กล่าวไว้ในการประชุมวิชาการแพทยศาสตร์ศึกษานานาชาติ

ของสมาคมแพทยศาสตรศึกษายุโรป AMEE 2006 ว่า “Students always want to know: Will this be on test? Students and residents don't respect what you expect; they respect what you inspect” นักศึกษามักต้องการที่จะทราบว่า สิ่งนี้จะอยู่ในแบบทดสอบหรือไม่ นักศึกษาและแพทย์ประจำบ้านไม่ได้เคารพในสิ่งที่คุณคาดหวัง แต่พวกเขาเคารพในสิ่งที่คุณตรวจสอบ นั่นก็แสดงให้เห็นถึงการประเมินผลเป็นสิ่งสำคัญในการกำหนดการเรียนรู้ของนักศึกษา

หลักการพื้นฐานของการประเมินผล (principles of assessment) มีคำถามที่สำคัญ 4 คำถามคือ 1. ทำไมต้องประเมินผล (Why are you doing the assessment?) 2. จะประเมินอะไร (What are you going to assess?) 3. จะประเมินอย่างไร (How are you going to assess it?) และ 4. การทดสอบนั้นดีหรือไม่ อย่างไร (How well is the assessment working?)

1. ทำไมต้องประเมินผล สำหรับการประเมินผลมีจุดประสงค์ 2 อย่าง คือ อย่างแรกคือการประเมินผลระหว่างการจัดการเรียนรู้ (formative evaluation) เพื่อประเมินผลความก้าวหน้า ปรับปรุงและพัฒนาการจัดการเรียนรู้ ซึ่งจะเป็นการป้อนกลับระหว่างเรียน (in course/ in training feedback) อย่างที่สองคือการประเมินผลเมื่อสิ้นสุดการจัดการเรียนการสอน (summative evaluation) เพื่อตัดสินคุณภาพผู้เรียน และจัดการเรียนรู้ เพื่อการจบการศึกษาหรือการได้รับประกาศนียบัตร (graduation/ post-graduate certification)

2. จะประเมินอะไร สำหรับการประเมินความสามารถทางคลินิก (clinical competence) สามารถประเมินได้ 3 องค์ประกอบคือ ความรู้ (knowledge) ทักษะ (skills) และเจตคติ (attitude) โดยการประเมินความรู้ ประกอบด้วย ความรู้ที่เป็นข้อเท็จจริง (factual knowledge) และความรู้ที่เป็นการประยุกต์เหตุผลทางคลินิก (applied: clinical reasoning) การประเมินทักษะประกอบด้วย ทักษะการสื่อสาร (communication) และทักษะทางคลินิก (clinical) ส่วนการประเมิน



รูปที่ 1. Miller's pyramid

นำมาจาก Miller GE. The assessment of clinical skills/ competence/ performance. Acad Med. 1990 Sep;65 (9 Suppl):S63-7.²

ทางเจตคติ จะเป็นการประเมินพฤติกรรมวิชาชีพ (professional behavior)

3. จะประเมินอย่างไร การประเมินผลจำเป็นต้องประเมินให้ถูกจุดประสงค์กับสิ่งที่ต้องการวัด² ดังรูปที่ 1

การวัดที่เป็นฐานคือ ต้องการวัดว่ารู้หรือไม่ (knows) ให้วัดความรู้ (knowledge) สามารถใช้วัดความรู้ที่เป็นข้อเท็จจริง โดยใช้แบบทดสอบที่ให้เลือกคำตอบที่ดีที่สุดข้อเดียว (single best answer, SBA) หรือเป็นคำถามที่ให้ตอบสั้น ๆ (short answer questions, SAQ) หรือ คำถามจับคู่แบบขยายตัวเลือก (extended matching questions, EMQ) การวัดระดับสูงขึ้นมาคือ รู้ว่าทำอย่างไร (knows how) โดยการวัดความสามารถ (competence) สามารถวัดโดยใช้บริบททางคลินิก (clinical context based) ซึ่งสามารถทดสอบได้โดยวิธี SBA, SAQ หรือ EMQ ได้เช่นกัน ในระดับที่สูงขึ้นมาอีก ให้แสดงว่าทำอย่างไร (shows how) โดยให้ประเมินการปฏิบัติ (performance assessment) ซึ่งใช้การทดสอบโดยวัดความรู้ทางด้านทักษะและหัตถการทางคลินิก ซึ่งเป็นการทดสอบทางคลินิกโดยใช้วัตถุประสงค์เป็นโครงสร้าง (objective structured clinical

examination, OSCE) ส่วนการประเมินผลระดับสูงสุด คือ ให้กระทำ (does) โดยการประเมินการปฏิบัติหรือ การกระทำ (action) อาจทำได้โดย การประเมินจากสถานที่ ปฏิบัติงาน (workplace based assessment; WBA) เช่น การประเมินทางคลินิกในสภาพการณ์และการปฏิบัติงานจริง (mini-clinical evaluation exercise; Mini-CEX) จากการสังเกตโดยตรงจากทักษะกระบวนการ (directly observed procedural skills; DOPs), การประเมิน พฤติกรรมการทำงานเป็นทีม (team assessment behaviour; TBA)

4. การทดสอบนั้นดีหรือไม่ อย่างไร ซึ่งการจะ ตอบคำถามนี้ ต้องพิจารณาจากลักษณะของเครื่องมือ ในการประเมินผล (the characteristics of assessment instruments) ซึ่งได้แก่ ความถูกต้อง (validity, V) ความน่าเชื่อถือ (reliability, R), ผลกระทบต่อการศึกษา (educational impact, E) การยอมรับ (acceptability, A) และค่าใช้จ่าย (cost, C)

ความถูกต้อง (validity) นั้น จะประกอบด้วย 1. เนื้อหา (content) ซึ่งเนื้อหาของการทดสอบนั้น ขึ้นตอนที่สำคัญที่ควรมีคือ การเขียนแผนผังการสอบ (blueprinting) เพื่อกำหนดหัวข้อการทดสอบไว้ตาม ผลการเรียนรู้ไว้ตั้งแต่ต้น และเนื้อหาที่ออกสอบควรเป็น

ไปตามแบบที่วางไว้ตามผลเรียนรู้ที่กำหนด 2. กระบวนการสอบ (response process) เช่น การใช้กระดาษ หรือ รูปแบบอิเล็กทรอนิกส์ การกำหนดขอบเขตความผิดพลาด (scope for error) เช่น ในการให้คะแนน 3. โครงสร้าง ภายใน (internal structure) เช่น จำนวนข้อสอบ รูปแบบ ข้อสอบ ลักษณะของตัวเลือก ระบบการให้คะแนน ความน่าเชื่อถือ (reliability) ซึ่งความน่าเชื่อถือเป็น ปัจจัยด้านหนึ่งของความถูกต้อง และกำหนดมาตรฐาน (standard setting) 4. ความสัมพันธ์กับปัจจัยอื่น ๆ เช่น ความสัมพันธ์ระหว่างส่วนที่แตกต่างกันของ การทดสอบ หรือระหว่างแบบทดสอบที่ต่างกัน 5. ผลที่ ตามมา (consequences) เช่น ผลตกหรือผ่าน ที่มีผล กับนักศึกษาหรือสถาบัน หรือมีต่อการตั้งมาตรฐาน ของวิชาชีพ (professional standards)

ความน่าเชื่อถือ ซึ่งเป็นส่วนหนึ่งของความถูกต้อง (validity) ประกอบด้วย 1. ความแม่นยำในการทำซ้ำ (reproducibility) เมื่อมีความแตกต่างกันของผู้ให้คะแนน คำถาม กรณีศึกษา และโอกาสต่าง ๆ 2. ความสามารถ แยกแยะระหว่างเด็กเก่งและเด็กอ่อน (capability of differentiating consistently between good and poor students) 3. สามารถวัดได้ เช่น วัดโดย Cronbach's alpha coefficient³

ตารางที่ 1. ตัวอย่างของการทำ blueprinting ของการสอบ OSCE ของหลักสูตรตามระบบ

ระบบ	ซีกประวัติ	อธิบาย	การตรวจ	กระบวนการ
หัวใจและหลอดเลือด	เจ็บหน้าอก	การให้ยา	ระบบหัวใจ และหลอดเลือด	การวัดความดัน
หายใจ	ไอเป็นเลือด	ผลของการสูบบุหรี่	ระบบหายใจ	การตรวจสมรรถภาพปอด ด้วยการวัดอัตราการไหลสูงสุด
ทางเดินอาหาร	ปวดท้อง	สาเหตุของดีซ่าน	ช่องท้อง	การตรวจทางทวารหนัก
สืบพันธุ์	ไม่มีประจำเดือน	ผลการป้ายมูกปากมดลูกที่ผิดปกติ	การป้ายมูกจากปากมดลูก	การตรวจภายใน
ประสาท	ปวดศีรษะ	อาการที่เกิดจากเยื่อหุ้มสมองอักเสบ	ระบบประสาท	
กล้ามเนื้อ	กล้ามเนื้ออ่อนแรง	วินิจฉัยโรครูมาตอยด์		
ทั่วไป	การประเมินก่อนการผ่าตัด	การขอแสดง ความยินยอม ของผู้ป่วย		การให้เลือด

การสุมตัวอย่างในวงกว้างจำเป็นสำหรับความน่าเชื่อถือที่เพียงพอ การสุมตัวอย่างข้ามเนื้อหา เช่น การมีช่วงของกรณีหรือสถานการณ์ต่าง ๆ โดยใช้ blueprinting ในการช่วยทำให้สามารถทำแผนที่ในการประเมินต่อผลของการเรียนรู้ของหลักสูตรหรือวัตถุประสงค์การเรียนรู้ตัวอย่างการทำ blueprinting ของการสอบ OSCE ของ หลักสูตรตามระบบ (system based course) ได้แสดงไว้ดังตารางที่ 1

สำหรับหลักการการประเมินผลนั้น ไม่มีการประเมินผลใดที่สมบูรณ์แบบ จำเป็นจะต้องมีการประนีประนอม (compromise) ซึ่งจะขึ้นอยู่กับเนื้อหาของ การประเมินผล โดยคุณภาพของการประเมินผล เป็นผลของ การออกแบบของการประเมินผลทั้งหมด คุณภาพของเครื่องมือแต่ละชนิด และการรวบรวมและกลั่นกรองหลักฐานความถูกต้องทั้งหมด

แลกเปลี่ยนประสบการณ์ ความคิดเห็น อภิปราย และสรุป

แลกเปลี่ยนประสบการณ์ จากการเรียน ESMEA course ที่ได้เข้าร่วม มีจัดการเรียนการสอนแบบกลุ่มย่อย มี interactive lecture และ small group discussion ได้แลกเปลี่ยนความรู้กับผู้ดูแลหลักสูตร แพทยศาสตรบัณฑิตจากสถาบันต่างประเทศหลายสถาบัน และพบว่า คณะแพทยศาสตรศิริราชพยาบาล ได้นำหลักการประเมินผลที่ถูกต้องมาประยุกต์ใช้บ้างแล้ว และเสนอให้มีการพัฒนาระบบการประเมินผลให้ดียิ่ง ๆ ขึ้นไป เพื่อสามารถวัดความรู้ความสามารถของนักศึกษาได้อย่างถูกต้อง เหมาะสม และสามารถนำไปชี้แนะและพัฒนาการเรียนรู้ของนักศึกษาได้

ความคิดเห็น การประเมินผลเป็นสิ่งที่สำคัญมาก นอกจากจะมีผลในการตัดสินใจการเรียนรู้นักศึกษา การได้ตก หรือการจบการศึกษาแล้ว สิ่งที่สำคัญอีกประการหนึ่ง คือ การกำหนดวิธีการเรียนรู้ของนักศึกษาจากการสอบถาม พบปะ พูดคุย กับนักศึกษา ถ้าการประเมินผลนั้นเป็นการประเมินผลที่เน้นความคิดเห็น รวบรวม (concept) การประยุกต์ความรู้ (application) นักศึกษาก็จะเรียนรู้โดยพยายามเข้าใจถึงหลักที่สำคัญ

และนำความรู้ไปประยุกต์ในสถานการณ์ต่าง ๆ แต่ถ้าเป็นการประเมินผลที่เน้นรายละเอียดของข้อมูล ไม่เน้นกระบวนการคิด ไม่เน้นการประยุกต์ใช้ นักศึกษาจะเกิดการเรียนรู้จากการประเมินผลครั้งแรก และจะมีการปรับตัวในการเรียนรู้ โดยจะเรียนรู้แบบท่องจำ เน้นรายละเอียด ไม่เน้นกระบวนการคิดและการประยุกต์ใช้ ดังนั้นการประเมินผลที่ดี จึงเป็นการกำหนดการเรียนรู้ของนักศึกษาที่แท้จริง สิ่งที่จะกำหนดว่า การประเมินผลนั้นควรเป็นแบบใด ขึ้นอยู่กับ จุดประสงค์ว่าเราต้องการให้มีผลลัพธ์การเรียนรู้แบบใด (outcome based) เช่น ถ้าจุดประสงค์คือ ต้องการให้นักศึกษามีความรู้ จำได้ การประเมินผลควรเป็นการถามความรู้นั้น ถ้าจุดประสงค์คือ ต้องการให้นักศึกษาคิดเป็น ประยุกต์เป็น เน้นกระบวนการคิด การประเมินผลควรเป็น การออกโจทย์ที่เน้นกระบวนการคิดและประยุกต์ใช้ ถ้าจุดประสงค์คือ ต้องการให้นักศึกษาสามารถทำได้ การประเมินผลควรเป็นการแสดง หรือวัดกระบวนการทำ หรือการทดสอบทางคลินิกโดยใช้วัตถุประสงค์เป็นโครงสร้าง เป็นต้น

อภิปรายและสรุป การประเมินผลจะเป็นการตัดสินผลการเรียนรู้ และเป็นตัวกำหนดการเรียนรู้ของนักศึกษา ดังนั้นผู้จัดหลักสูตรต้องกำหนดผลลัพธ์การเรียนรู้อย่างชัดเจน การประเมินผลที่ดีต้องมีความถูกต้อง เชื่อถือได้ เมื่อประเมินจากผู้ประเมินต่างคนกัน และต้องสามารถแยกความแตกต่างระหว่างผู้ที่มีความรู้ความสามารถและไม่มีความรู้ความสามารถได้ วัดสิ่งที่ต้องการวัดอย่างเหมาะสม มีการกำหนดแผนผังการสอบ blueprinting ว่าประเมินอะไร อย่างไร ไม่ได้ประเมินไปตามรายชื่อข้อสอบ แต่ต้องประเมินตามผลลัพธ์การเรียนรู้ที่เราต้องการให้นักศึกษาได้เรียนรู้ เพื่อพัฒนาให้นักศึกษาได้มีความรู้ ทักษะ และความสามารถ ในการประกอบวิชาชีพต่อไปในอนาคต

เอกสารอ้างอิง

1. Tejinder S. Principles of Assessment in Medical Education: Jaypee Brothers Pvt Ltd; 2012.
2. Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990 Sep;65(9 Suppl):S63-7.
3. Downing SM. Reliability: on the reproducibility of assessment data. Med Educ. 2004 Sep;38(9):1006-12.