

การกำหนดเกณฑ์ผ่านทักษะทางคลินิกโดยการสอบ Objective Structured Clinical Examination (OSCE) เปรียบเทียบระหว่างวิธี Angoff's Method และ The Borderline Regression Method

เจริญพร แก้วละเอียด และ นันทสัณห์ อธิรปัญญากร

ภาควิชาอายุรศาสตร์ คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ จังหวัดสงขลา 90000

kjarernp@medicine.psu.ac.th

บทคัดย่อ

การประเมินทักษะทางคลินิกโดยการสอบ Objective Structured Clinical Examination หรือ OSCE เป็นวิธีการจำลองสถานการณ์ทางคลินิกที่นิยมใช้เพื่อตรวจสอบความสามารถทางคลินิกของนักศึกษาแพทย์ เป็นวิธีการประเมินผลที่มีความน่าเชื่อถือ และครอบคลุมการวัดทั้งระดับความรู้ ทักษะ เจตคติและการตัดสินใจ ซึ่งถือเป็นการประเมินตามสภาพจริงจากการปฏิบัติ (Direct Observation) เพื่อให้การตัดสินใจผลการประเมินมีความตรง (Validity) จึงขึ้นอยู่กับวิธีการเลือกใช้เกณฑ์ผ่าน (Cut off Score) หรือเกณฑ์มาตรฐาน (Standard Setting) ที่เหมาะสม การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อ (1) ศึกษาทักษะทางคลินิกของนักศึกษาแพทย์โดยการสอบ OSCE จำแนกตามทักษะการเรียนรู้ (2) ศึกษาความสอดคล้องผลการประเมินเมื่อกำหนดเกณฑ์ผ่านวิธี Angoff's Method กับ Borderline Regression Method ทำการศึกษากับนักศึกษาแพทย์ชั้นปีที่ 4 ที่หมุนเวียนมาปฏิบัติงานรายวิชาสุขภาพและโรคของผู้ใหญ่และผู้สูงอายุ 2 (Health and Diseases of Adults and Elderly II) ปีการศึกษา 2561 จำนวน 139 คน ทำการสอบด้วยข้อสอบ OSCE จำนวน 20 สถานี ครอบคลุมทุกทักษะการเรียนรู้ ทำการวิเคราะห์ด้วยสถิติภาคบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย (Mean) ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) การเปรียบเทียบคะแนนในแต่ละทักษะโดยการวิเคราะห์ความแปรปรวนทางเดียว (One way ANOVA) หากพบความแตกต่างทำการเปรียบเทียบรายคู่โดยวิธีการ Scheffe การวิเคราะห์การถดถอยอย่างง่าย (Simple Regression Analysis) และวิเคราะห์ความสอดคล้องผลการตัดสินใจผ่าน-ตก ทั้ง 2 วิธี โดยสถิติแคปปา (Kappa Statistics)

ผลการศึกษา พบว่า (1) นักศึกษาแพทย์มีทักษะทางคลินิกโดยภาพรวมมีคะแนนเฉลี่ยการประเมิน 69.22 คะแนน จากคะแนนเต็ม 100 มีทักษะการเรียนรู้แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 (p -value < .01) (2) การตัดสินใจผลการสอบผ่าน-ตก โดยวิธี Angoff's Method และ Borderline Regression Method ให้ผลการประเมินสอดคล้องกัน อย่างมีนัยสำคัญทางสถิติ (p -value < 0.01) ทุกสถานีสอบ โดยมีขนาดความสอดคล้องโดยภาพรวมอยู่ในระดับดี มีค่าเฉลี่ยสัมประสิทธิ์ความสอดคล้อง (Kappa) เท่ากับ 0.74 (range : 0.22 - 1.00)

คำสำคัญ: การกำหนดเกณฑ์ผ่าน, OSCE, Borderline Regression Method

บทนำ

การประเมินทักษะทางคลินิกโดยการสอบ Objective Structured Clinical Examination หรือ OSCE เป็นวิธีการจำลอง สถานการณ์ทางคลินิกที่นิยมใช้เพื่อตรวจสอบความสามารถทางคลินิก (Mitchell et al. 2009) เป็นวิธีประเมินผลรูปแบบหนึ่งซึ่งช่วยให้การประเมินทางคลินิกมีความน่าเชื่อถือที่สูงขึ้น และครอบคลุมทักษะทางคลินิกที่สำคัญ เนื่องจากเป็นเครื่องมือประเมินทักษะทางคลินิกที่มีรายงานการศึกษาถึงความเที่ยงตรงและความน่าเชื่อถืออยู่ในเกณฑ์ดี (Gormely, 2011) สามารถแยกแยะระดับความรู้ของนักศึกษารายบุคคลได้อย่างแท้จริง ครอบคลุมการวัดทุกระดับความรู้ ทั้งในด้านทักษะ เจตคติและการตัดสินใจ ซึ่งถือเป็นการประเมินตามสภาพจริง จากการปฏิบัติจริง (Direct Observation) การตัดสินใจผลการสอบส่วนใหญ่ใช้การตัดสินใจแบบอิงเกณฑ์ เพื่อให้มีมาตรฐานตามเกณฑ์ที่กำหนดไว้ โดยการตั้งเกณฑ์ผ่าน (Cut off Score) หรือเกณฑ์มาตรฐาน (Standard Setting) ซึ่งเป็นเกณฑ์ยอมรับได้ว่าความสามารถน้อยที่สุดที่ยอมรับให้ผู้เรียนผ่านการประเมินไปได้ การกำหนดคะแนนจุดตัด (setting cut scores) เป็นการดำเนินการกำหนดคะแนนเพื่อใช้เป็นขีดจำกัด (Threshold) ในการแบ่งคะแนนสอบตามระดับความสามารถของผู้สอบ

งานวิจัยที่ผ่านมาส่วนใหญ่เป็นการศึกษาเพื่อประเมินทักษะความสามารถทางคลินิกโดยพิจารณาคะแนนสอบและอัตราการสอบผ่านในแต่ละทักษะ เช่น สาริต คลังสิน และคณะ (2559) ศึกษาเปรียบเทียบคะแนนสอบ Objective Structured Clinical Examination (OSCE) และอัตราการสอบผ่านแต่ละทักษะของนักศึกษาแพทย์ชั้นปีที่ 6 ที่หมุนเวียนและปฏิบัติงานภาควิชาสูติศาสตร์และนรีเวชวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ พบว่าคะแนนเฉลี่ยทักษะการซักประวัติ การตรวจร่างกาย การแปลผลทางห้องปฏิบัติการ หัตถการ และทักษะการสื่อสารการให้คำปรึกษา มีคะแนนแต่ละทักษะแตกต่างกันอย่างมีนัยสำคัญ อัตราการสอบผ่านร้อยละ 71.5, 78.4, 53.2, 74.1 และ 46.2 ตามลำดับ และแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($p\text{-value} < 0.01$) นอกจากนี้การศึกษารายชื่อของ เกษร สายธนู และคณะ (2559) ศึกษาการใช้รูปแบบการสอบ OSCE เพื่อประเมินสมรรถนะการปฏิบัติการสร้างสัมพันธ์ภาพเพื่อการบำบัดของนักศึกษาพยาบาล คณะพยาบาลศาสตร์ มหาวิทยาลัยอุบลราชธานี พบว่า นักศึกษาพยาบาลส่วนใหญ่สอบผ่านตามเกณฑ์การประเมินผล ร้อยละ 81.82 ค่าเฉลี่ยคะแนนความคิดเห็นต่อการสอบ OSCE ของนักศึกษาพยาบาลอยู่ในระดับดี ส่วนการศึกษากำหนดเกณฑ์ผ่านของข้อสอบ OSCE พบได้ในการศึกษาของ Thananit Sangkomakamhang (2014) ทำการศึกษาในกลุ่มนักศึกษาแพทย์ชั้นปีที่ 6 คณะแพทยศาสตร์มหาวิทยาลัยขอนแก่น พบว่าการกำหนดวิธีการกำหนดเกณฑ์ผ่านระหว่างวิธี Angoff's Method, Borderline Regression Method, Borderline Group ส่งผลให้จำนวนผู้สอบผ่านในแต่ละวิธีแตกต่างกัน

จากความสำคัญดังกล่าว ทำให้ผู้วิจัยสนใจศึกษาความสามารถทางคลินิกของนักศึกษาแพทย์ โดยการสอบ OSCE จำแนกตามทักษะการเรียนรู้ และศึกษาความสอดคล้องผลการประเมินเมื่อกำหนดเกณฑ์ผ่าน โดยวิธี Angoff's Method กับ Borderline ความสามารถทางคลินิกจำแนกตามทักษะการเรียนรู้ และผลการกำหนดเกณฑ์ผ่านที่ต่างกัน โดยพิจารณาจากค่าความสอดคล้องในการตัดสินผลการประเมินกำหนดสมมติฐานการวิจัย คือ นักศึกษามีทักษะการเรียนรู้แตกต่างกัน การกำหนดเกณฑ์ผ่านข้อสอบ OSCE โดยวิธีต่างกัน ให้ผลการประเมินสอดคล้องกัน การศึกษาครั้งนี้จะนำไปสู่การให้ข้อมูลสะท้อนกลับแก่ผู้เรียน และเป็นประโยชน์ในการวางแผนการพัฒนาระบบการจัดการเรียนการสอน การวัดและประเมินผลการสอบ OSCE ในรายวิชาให้มีประสิทธิภาพมากยิ่งขึ้น

วิธีดำเนินการวิจัย

การศึกษานี้ใช้รูปแบบการวิจัยเชิงพรรณนา (Descriptive Research)

ประชากร

นักศึกษาแพทย์ชั้นปีที่ 4 ที่ลงทะเบียนรายวิชาสุขภาพและโรคของผู้ใหญ่และผู้สูงอายุ 2 (Health and Disease of Adults and Elderly II) ซึ่งหมุนเวียนปฏิบัติงาน ปีการศึกษา 2561 จำนวน 2 กลุ่ม ประกอบด้วย นักศึกษาจากคณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ จำนวน 78 คน ศูนย์แพทยศาสตร์ศึกษาโรงพยาบาลหาดใหญ่ จำนวน 42 คน ศูนย์แพทยศาสตร์ศึกษาโรงพยาบาลยะลา จำนวน 19 คน รวมทั้งสิ้น จำนวน 139 คน ซึ่งแต่ละสถาบันมีรูปแบบการจัดการสอน การวัดประเมินผลในลักษณะเดียวกัน และใช้ข้อสอบร่วมกัน ในการศึกษานี้ใช้ประชากรเป็นกลุ่มตัวอย่างในการศึกษา

ตัวแปรในการศึกษา

กำหนดตัวแปรในการศึกษาสำหรับการตอบวัตถุประสงค์ที่ 1 การเปรียบเทียบทักษะการเรียนรู้ทางคลินิก ตัวแปรอิสระ (Independent Variable) ได้แก่ ทักษะการเรียนรู้ทางคลินิก ประกอบด้วยทักษะการซักประวัติ ทักษะการตรวจร่างกาย การแปลผลทางห้องปฏิบัติการ การทำหัตถการ การสื่อสาร/การให้คำปรึกษา ตัวแปรตาม (Dependent Variable) คือคะแนนสอบในแต่ละทักษะ การตอบวัตถุประสงค์ข้อที่ 2 กำหนดตัวแปรอิสระ ได้แก่รูปแบบการกำหนดเกณฑ์ผ่านวิธี Angoff's Method และ The Borderline Regression Method ตัวแปรตาม คือ ความสอดคล้องผลการประเมินผ่าน-ตก เมื่อกำหนดเกณฑ์ผ่านวิธี Angoff's Method กับ Borderline Regression Method

การเก็บรวบรวมข้อมูล

ขั้นตอนที่ 1 การวิเคราะห์ก่อนสอบ ดำเนินการโดยคณะกรรมการรายวิชาทั้ง 3 สถาบัน จำนวน 5 ท่าน ตรวจสอบคุณภาพด้านความตรงเชิงเนื้อหา (Content Validity) และความเหมาะสมของข้อคำถาม พร้อมทั้งกำหนดเกณฑ์ผ่านข้อสอบโดยวิธี Angoff's Method โดยพิจารณาร่วมกันในแต่ละข้อว่านักศึกษาที่มีความสามารถระดับคาบเส้น (Borderline Examiness) มีความรู้ความสามารถอยู่ใน

ระดับใดที่สามารถยอมรับได้ว่ามีทักษะทางคลินิกเพียงพอ ตามหัวข้อเรื่อง และประเภทข้อสอบ (Taxonomy) ที่กำหนดตามผังการออกข้อสอบ (Table of Specification) กลุ่มละ 20 สถานี ประกอบด้วย ทักษะการซักประวัติ การตรวจร่างกาย การแปลผลทางห้องปฏิบัติการ หัตถการ และการสื่อสาร/การให้คำปรึกษา

ขั้นตอนที่ 2 การวิเคราะห์หลังสอบ

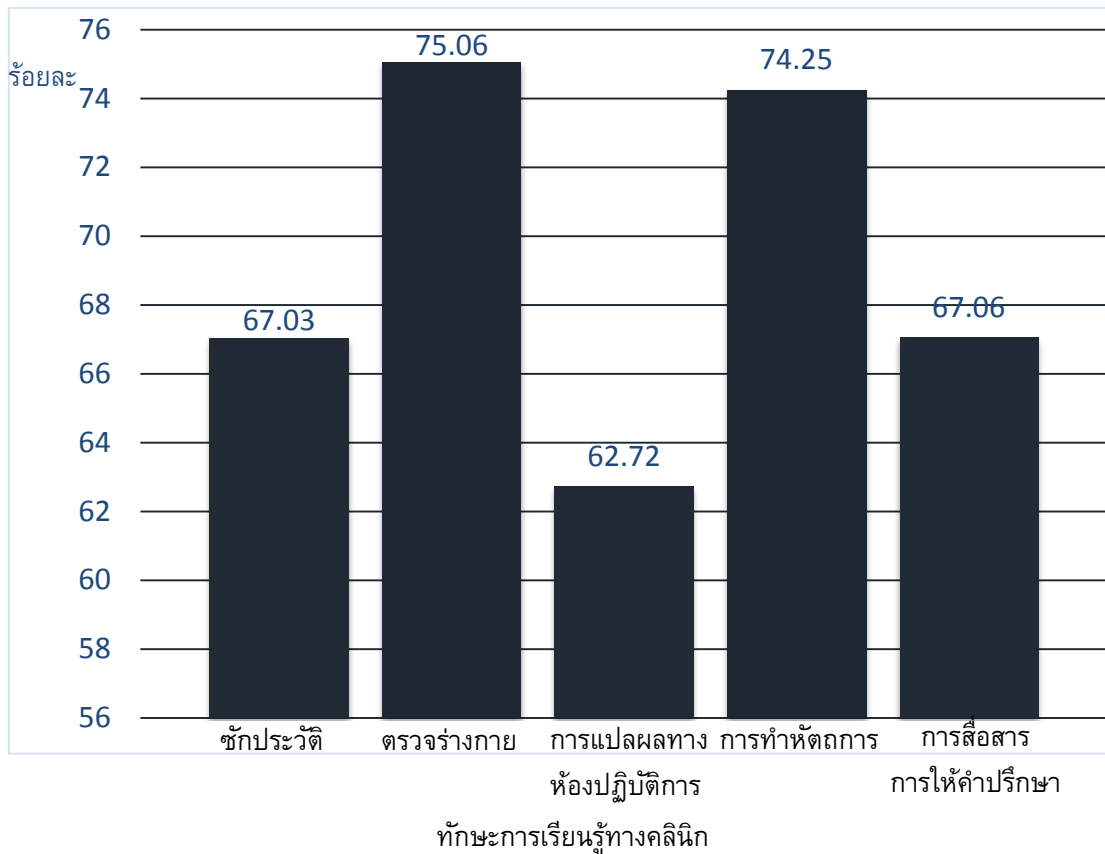
นำข้อมูลที่ได้จากการสอบมาวิเคราะห์หาเกณฑ์ผ่านโดยการวิเคราะห์การถดถอยเพื่อหาค่าเกณฑ์ผ่านโดยวิธี Borderline Regression Method (BRM) กำหนดให้กลุ่ม Borderline มีค่าอยู่ในระดับ 2 ตามเกณฑ์การให้คะแนน (Check List) ความสามารถในภาพรวม (Overall Performance) แต่ละสถานีสอบ ซึ่งกำหนดไว้ 4 ระดับ (4 = Excellent 3 = Pass 2= Borderline 1 = Fail)

สถิติที่ใช้ในการวิจัย

1. ทำการวิเคราะห์ด้วยสถิติภาคบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย (Mean) ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation S.D.)
2. วิเคราะห์เปรียบเทียบความสามารถในแต่ละทักษะทางคลินิก โดยการวิเคราะห์ความแปรปรวนทางเดียว (One Way ANOVA) หากพบความแตกต่างทำการเปรียบเทียบรายคู่โดยวิธีการของ Scheffe
3. การวิเคราะห์การถดถอยอย่างง่าย (Simple Regression Analysis) เพื่อหาเกณฑ์ผ่านโดยวิธี Borderline Regression Method
4. การวิเคราะห์ความสอดคล้องผลการตัดสินจำนวนผู้ผ่าน-ไม่ผ่านเกณฑ์ โดยใช้สถิติแคปปา (Kappa Statistics) กำหนดเกณฑ์การพิจารณาระดับความสอดคล้องตามแนวทางของ Landis and Koch (1977) อ้างอิงจาก ประสพชัย พสุนนท์ (2558) ดังนี้ ความสอดคล้องดีมาก (Almost Perfect: 0.81-1.00) มีความสอดคล้องดี (Substantial: 0.61 -0.80) มีความสอดคล้องปานกลาง (Moderate: 0.41-0.60) มีความสอดคล้องพอใช้ (Fair: 0.21- 0.40) มีความสอดคล้องเล็กน้อย (Slight: 0.00- 0.20) ไม่มีความสอดคล้อง (Poor: น้อยกว่า 0.00)

ผลและอภิปรายผล

1. นักศึกษาแพทย์ชั้นปีที่ 4 ปีการศึกษา 2561 ที่ลงเรียนรายวิชาสุขภาพและโรคของผู้ใหญ่และผู้สูงอายุ 2 ทั้ง 3 สถาบัน จำนวน 139 คน พบว่ามีคะแนนการสอบโดยภาพรวมเฉลี่ย 69.22 คะแนน จากคะแนนเต็ม 100 การตรวจร่างกายมีคะแนน 75.06 \pm 9.44 (rang: 43.83- 100.00) การทำหัตถการมีคะแนน 74.25 \pm 7.69 (rang: 52.32-93.00) การสื่อสาร/การให้คำปรึกษา มีคะแนน 67.06 \pm 11.62 (rang: 30.50-90.50) การซักประวัติมีคะแนน 67.03 \pm 8.48 (rang: 35.50-84.17) การแปลผลทางห้องปฏิบัติการ มีคะแนน 62.72 \pm 9.37 (rang: 37.50-80.63) ตามลำดับ เมื่อทำการเปรียบเทียบทักษะการเรียนรู้ทางคลินิก และทำการเปรียบเทียบรายคู่โดยวิธีการของ Scheffe พบว่า นักศึกษามีทักษะการซักประวัติสูงกว่าการแปลผลทางห้องปฏิบัติการ ทักษะการตรวจร่างกายสูงกว่าการซักประวัติ การแปลผลทางห้องปฏิบัติการ การสื่อสาร/การให้คำปรึกษา นักศึกษามีทักษะการสื่อสาร/การให้คำปรึกษาสูงกว่าทักษะการแปลผลทางห้องปฏิบัติการ มีทักษะการทำหัตถการสูงกว่าทักษะการซักประวัติ การสื่อสาร/การให้คำปรึกษา การแปลผลทางห้องปฏิบัติการ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ($p < .01$) ดังรูปที่ 1



รูปที่ 1 แสดงทักษะการเรียนรู้ทางคลินิก

2. ความสอดคล้องผลการประเมินโดยใช้เกณฑ์การตัดสินที่ต่างกันพบว่า การตัดสินผลการสอบวิธี Angoff's Method และ Borderline Regression Method ให้ผลการประเมินสอดคล้องกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ($p < 0.01$) ทุกสถานีสอบ โดยมีขนาดความสอดคล้องโดยภาพรวมอยู่ในระดับดี มีค่าเฉลี่ยเท่ากับ 0.74 (range : 0.22-1.00) มีขนาดความสอดคล้องดีมาก (Kappa 0.81-1.00) จำนวน 11 สถานีสอบ (ร้อยละ 52.38) มีความสอดคล้องดี (Kappa 0.61 -0.80) จำนวน 5 สถานีสอบ (ร้อยละ 23.80) มีความสอดคล้องปานกลาง (Kappa 0.41- 0.60) จำนวน 2 สถานีสอบ (ร้อยละ 9.52) มีความสอดคล้องพอใช้ (Kappa 0.21- 0.40) จำนวน 3 สถานีสอบ (ร้อยละ 14.28) เมื่อจำแนกตามทักษะการประเมินในแต่ละสถานีสอบพบว่า (1) ทักษะการซักประวัติ (History Taking) มีดัชนีความสอดคล้องจากมากไปหาน้อย คือ Edema , Chest Pain , Major Depression , Head Brain, Jaundice ตามลำดับ (2) ทักษะการทำหัตถการ (Procedure) มีค่าดัชนีความสอดคล้องจากมากไปหาน้อย คือ Basic CPR, Percussion(Test1), IV Insertion, Chest PT (test 2), Stool Examination, Intravenous, Hand Washing ตามลำดับ (3) ทักษะการตรวจร่างกาย (Physical Examination) มีค่าดัชนีความสอดคล้องจากมากไปหาน้อย คือ COPD, Cardio Vascular, ตรวจเท้าเบาหวาน, Neuro, Cerebellar Signs ตามลำดับ (4) การสื่อสารการให้คำปรึกษา (Communication skill) มีค่าดัชนีความสอดคล้องจากมากไปหาน้อย คือ Essential Hypertension, Chest Pain, Lumbar Puncture, HIV Infection ตามลำดับ ดังตารางที่ 1

ตารางที่ 1 การเปรียบเทียบคะแนนจุดผ่าน อัตราการสอบผ่าน และความสอดคล้อง การกำหนดเกณฑ์ผ่านวิธี Angoff's Method และ Boderline Regression Method

สถานี (Stations)	mean	SD	n	Standard Setting Method				ดัชนีความสอดคล้อง (Kappa)
				Angoff 's Method		Boderline Regression Method (BRM)		
				Cut score	Pass Rate (%)	Cut score	Pass Rate (%)	
การซักประวัติ (History Taking)								
Major depression	68.41	11.98	139	60	77.00	63.24	66.9	.75** (SE=0.06) 95%CI (0.69,0.81)
Jaundice	60.42	12.35	65	65	27.69	52.69	75.38	.22** (SE =0.06) 95%CI (0.16,0.28)
chest pain	72.27	9.64	65	65	80.00	65.42	75.38	.86** (SE =.07) 95%CI (0.79,0.93)
Head brain	60.37	14.57	74	60	58.10	54.01	74.32	.64** (SE =.08) 95%CI (0.56,0.72)
Edema	72.27	13.75	74	60	83.78	60.90	82.43	.94** (SE =.04) 95%CI (0.90,0.98)
การทำหัตถการ (Procedure)								
Intravenous	59.88	16.18	65	60	55.38	55.85	61.53	.87 ** (SE =0.06) 95%CI (0.81,0.93)
Hand washing	91.53	11.41	65	70	96.92	70.86	89.23	.41** (SE =0.20) 95%CI (0.21,0.61)
Percussion(Test1)	59.41	13.67	65	60	56.92	55.99	58.46	.96** (SE =0.03) 95%CI (0.93,0.99)
Basic CPR	90.81	9.75	139	80	89.20	81.01	88.5	.96 ** (SE=0.03) 95%CI (0.93,0.99)
Chest PT (test 2)	55.13	19.93	74	60	56.75	53.82	59.45	.94** (SE =0.03) 95%CI (0.91,0.97)
Stool examination	77.55	11.54	74	70	77.02	67.11	81.08	.87** (SE =0.06) 95%CI (0.81,0.93)
IV insertion	71.38	17.29	74	60	78.37	56.00	79.72	.95** (SE =0.04) 95%CI (0.91,0.99)
การตรวจร่างกาย (Physical Examination)								
ตรวจเท้าเบาหวาน	35.18	9.85	65	60	3.07	44.20	10.76	.41** (SE =.20) 95%CI (0.21,0.61)
Neuro	91.95	12.75	65	60	98.46	75.99	92.30	.31** (SE =0.24) 95%CI (.07,.55)
COPD	73.07	15.28	139	60	79.90	58.72	79.90	1.00** (SE=0.00) 95%CI (1.00,1.00)
Cardio vascular	88.20	12.22	74	65	90.54	71.61	91.89	.91** (SE =0.08) 95%CI (0.83,0.99)
Cerebellar signs	85.83	16.73	74	60	66.21	55.46	27.02	.31** (SE =0.07) 95%CI (0.24,0.38)
การสื่อสาร-การให้คำปรึกษา (Communication skill)								
Essential Hypertension	75.43	7.62	65	65	90.76	65.08	90.76	1.00** (SE =0.00) 95%CI (1.00,1.00)
chest pain	75.58	10.73	65	65	86.15	66.35	80.00	.78** (SE =.10) 95%CI (0.68,0.88)

สถานี (Stations)	Standard Setting Method							ดัชนีความสอดคล้อง (Kappa)
				Angoff 's Method		Boderline Regression Method (BRM)		
	mean	SD	n	Cut score	Pass Rate (%)	Cut score	Pass Rate (%)	
HIV infection	48.64	10.66	74	44	70.27	49.34	58.10	.74** (SE =0.07) 95%CI (0.67,0.81)
Lumbar Puncture	70.62	12.36	74	60	79.72	56.63	86.48	.76** (SE =0.10) 95%CI (0.66,0.86)

อภิปรายผล

1. การประเมินทักษะและหัตถการทางคลินิก ด้วยการสอบ OSCE ของนักศึกษาแพทย์ชั้นปีที่ 4 ปีการศึกษา 2561 รายวิชา สุขภาพและโรคของผู้ใหญ่และผู้สูงอายุ 2 ทั้ง 3 สถานี พบว่านักศึกษามีทักษะทางคลินิกของนักศึกษามีคะแนนเฉลี่ยเท่ากับ 69.22 คะแนน จากคะแนนเต็ม 100 มีทักษะการซักประวัติสูงกว่าการแปลผลทางห้องปฏิบัติการ ทักษะการตรวจร่างกายสูงกว่าการซักประวัติ การสื่อสาร/การให้คำปรึกษา การแปลผลทางห้องปฏิบัติการ ส่วนทักษะการสื่อสาร/การให้คำปรึกษาสูงกว่าการแปลผลทางห้องปฏิบัติการ นักศึกษามีทักษะการทำหัตถการสูงกว่าการซักประวัติ การสื่อสาร/การให้คำปรึกษา การแปลผลทางห้องปฏิบัติการ ซึ่งสอดคล้องกับสมมติฐานข้อที่ 1 การศึกษาครั้งนี้สะท้อนให้เห็นถึงผลการจัดการศึกษาในรายวิชา ที่ควรเน้นการจัดกิจกรรมเพื่อให้ให้นักศึกษามีประสบการณ์ในการฝึกทักษะการแปลผลทางห้องปฏิบัติการให้มากขึ้น แต่อย่างไรก็ตามข้อจำกัดในการสอบ OSCE ทั้งในเรื่องกลุ่มตัวอย่างในการศึกษาที่มีจำนวนน้อยในการจัดสอบในแต่ละครั้ง คะแนนที่มีการกระจายตัวมาก นอกจากนี้ยังมีปัจจัยอื่น เช่น ผู้ป่วยจำลองที่ใช้ในการสอบมีการเปลี่ยนแปลงไปในแต่ละครั้ง ซึ่งอาจแก้ไขโดยการกำหนดมาตรฐานของผู้ป่วยจำลองให้มีมาตรฐาน และจากการศึกษาของ Chesser และคณะ (2004) ที่พบว่าการจัดลำดับสถานีสอบส่งผลต่อคะแนนการสอบ OSCE ผลการศึกษาครั้งนี้ไม่สอดคล้องกับการศึกษาของ สาริต คลังสิน และคณะ (2559) ที่พบว่านักศึกษาแพทย์ชั้นปี 6 ที่ปฏิบัติงานในภาควิชาสูติศาสตร์และนรีเวช คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ ที่ผ่านการสอบ OSCE มีทักษะทางคลินิกด้านการสื่อสาร และด้านการให้คำปรึกษาน้อยกว่าทักษะอื่นๆ

2. การกำหนดเกณฑ์ผ่านข้อสอบ OSCE ส่วนใหญ่มักใช้กำหนดเกณฑ์ผ่านแบบตัดสินแบบอิงเกณฑ์ เพื่อแสดงออกถึงมาตรฐานของวิชาชีพแพทย์ เพื่อให้สังคมได้รับการบริการทางการแพทย์ที่มีคุณภาพ จากการศึกษาครั้งนี้พบว่ากำหนดเกณฑ์ผ่านทั้ง 2 วิธี ให้ผลการตัดสินมีความสอดคล้องกัน โดยมีขนาดความสอดคล้องโดยภาพรวมอยู่ในระดับดี ซึ่งสอดคล้องกับสมมติฐานข้อที่ 2 ที่กำหนดไว้ ทั้งนี้ตามแนวคิดการกำหนดเกณฑ์ผ่านจากการพิจารณาโดยใช้ดุลยพินิจของผู้เชี่ยวชาญในเนื้อหา (Judgmental Methods) การกำหนดจุดตัดวิธีนี้จะได้จุดตัดที่คงที่ไม่แปรเปลี่ยนไปตามคะแนนของผู้สอบ แต่ขึ้นกับประสบการณ์ผู้กำหนดเป็นหลักต้องมีประสบการณ์และเข้าใจคุณลักษณะของผู้เรียนกลุ่มคาบเส้นเป็นอย่างดีจึงจะสามารถกำหนดเกณฑ์ได้อย่างเหมาะสม ส่วนการกำหนดจุดตัดโดยใช้ข้อมูลจากการสอบของกลุ่มผู้สอบ (Empirical Methods) จุดตัดที่ได้จากวิธีนี้มีค่าแปรเปลี่ยนหรือสัมพันธ์กับคะแนนการสอบของกลุ่มผู้สอบ (เชิดศักดิ์ ไรรมณีรัตน์, 2006) ผลการศึกษาครั้งนี้ไม่สอดคล้องกับการศึกษาของ Thananit Sangkomakamhang (2014) พบว่าการกำหนดวิธีการกำหนดเกณฑ์ผ่านข้อสอบ OSCE ระหว่างวิธี Angoff 's Method, Boderline Regression, Borderline group ส่งผลให้จำนวนผู้สอบผ่านในแต่ละวิธีแตกต่างกัน ทั้งนี้อาจเนื่องมาจากปัจจัยทางด้านคุณภาพของข้อสอบเช่นคุณภาพด้านความเชื่อมั่นของข้อสอบ หรือการคัดเลือกผู้ที่ทำหน้าที่กำหนดเกณฑ์ผ่านย่อมส่งผลต่อการศึกษา ประกอบกับการเลือกใช้วิธีการกำหนดเกณฑ์ผ่านวิธี Borderline group ซึ่งคำนวณเกณฑ์ผ่านจากจำนวนผู้เข้าสอบที่มีความสามารถอยู่ในระดับปานกลางซึ่งกรณีที่มีสอบในกลุ่มย่อยตามรูปแบบการสอบ OSCE อาจมีตัวแทนจำนวนที่น้อยเกินไปแม้จะนำค่ามัธยฐานมากำหนดเป็นจุดผ่านก็ตาม

ข้อดีของการศึกษาครั้งนี้คือขาดการวิเคราะห์คุณภาพข้อสอบด้านความเชื่อมั่น (Reliability) แต่มีการตรวจสอบคุณภาพด้านความตรง (Content Validity) จากผู้เชี่ยวชาญในเนื้อหาเป็นอย่างดี มีรายงานการศึกษาที่ยืนยันคุณภาพของข้อสอบ OSCE ทั้งความเที่ยงตรงและความน่าเชื่อถืออยู่ในเกณฑ์ดี เช่นการศึกษาของสถาบันพระบรมราชชนก (2556) ที่กล่าวว่าเป็นการจัดการเรียนการสอนด้วยการประเมินผลที่สามารถประเมินการปฏิบัติจริง มีศักยภาพสูงในการพัฒนาความสามารถผู้เรียนทั้งทางด้านความรู้ ความคิด ทักษะ เจตคติ และการตัดสินใจ และจากการศึกษาของ Woragidpoonpol and Yarmsuoy (2006) ที่พัฒนารูปแบบการ

ประเมิน ผลการปฏิบัติการพยาบาล ในกระบวนการวิชาการฝึกปฏิบัติการพยาบาลแม่และเด็ก 2 โดยการสอบแบบ OSCE พบว่าเป็นการทดสอบที่วัดความสามารถในการปฏิบัติทางคลินิกและความรู้ ของบุคคลได้จริงครอบคลุมทุกด้านทั้งการปฏิบัติและทฤษฎีรวมทั้งจากการสะท้อนคิดและประเมินผลนักศึกษาผดุงครรภ์มหาวิทยาลัย Swansea ต่อการใช้ OSCE ในการประเมินทักษะทางคลินิกพบว่าเป็นการเพิ่มโอกาสในการฝึกปฏิบัติที่เพียงพอมีความปลอดภัยทำให้เกิดทักษะทางคลินิกที่ดีและสามารถฝึกการคิดอย่างมีวิจารณญาณ (Einion, 2013)

สรุปผลการวิจัย

นักศึกษาแพทย์ชั้นปีที่ 4 ที่ลงทะเบียนรายวิชาสุขภาพและโรคของผู้ใหญ่และผู้สูงอายุ 2 มีทักษะทางคลินิกเรื่องการตรวจร่างกายมีการทำหัตถการ การสื่อสาร/การให้คำปรึกษา การซักประวัติ และการแปลผลทางห้องปฏิบัติการ ตามลำดับ การศึกษาครั้งนี้ชี้ให้เห็นว่าทางรายวิชาสุขภาพและโรคของผู้ใหญ่และผู้สูงอายุ ทั้ง 3 สถาบัน ควรมีการจัดประสบการณ์การเรียนรู้ด้านการแปลผลทางห้องปฏิบัติการให้มากขึ้นอาจเป็นการจัดชั่วโมงให้นักศึกษาค้นคว้าด้วยตนเอง (SDL) หรือจัดชั่วโมงเพิ่มเติมให้นักศึกษามีโอกาสเรียนรู้ทักษะการแปลผลทางห้องปฏิบัติการจากอาจารย์ผู้สอนโดยตรงมากขึ้น เพราะการอ่านและแปลผลทางห้องปฏิบัติการถือเป็นทักษะ (Skill) ที่ต้องอาศัยการปฏิบัติซ้ำ ๆ จึงจะสามารถปฏิบัติได้อย่างถูกต้อง จากการศึกษาการกำหนดเกณฑ์ผ่านการสอบ OSCE โดยวิธี Angoff's Method และ Borderline Regression Method พบว่าการกำหนดเกณฑ์ผ่านทั้ง 2 วิธีให้ผลการประเมินสอดคล้องกันอยู่ในระดับสูง โดยเฉพาะทักษะการทำหัตถการทางการแพทย์ ที่ให้ค่าความสอดคล้องอยู่ในระดับสูงมากทั้งนี้เนื่องจากการประเมินตามสภาพจริงและประเมินผู้เรียนจากการปฏิบัติจริง การศึกษาครั้งนี้แนะนำให้ใช้การกำหนดเกณฑ์ผ่านวิธี Angoff's Method ซึ่งเป็นการกำหนดเกณฑ์โดยผู้เชี่ยวชาญ ซึ่งจะไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ทั้งนี้ตามลักษณะรูปแบบการสอบ OSCE เป็นการประเมินผู้เรียนกลุ่มเล็ก มีการบริหารจัดการสอบที่ค่อนข้างยุ่งยากเนื่องจากต้องมีการจัดเตรียมทั้งอุปกรณ์ สถานที่ และผู้ป่วยจำลอง จึงต้องดำเนินการจัดสอบให้มีมาตรฐาน การนำเทคนิคการกำหนดเกณฑ์ผ่านโดยวิธี Borderline Regression Method จำเป็นต้องอาศัยข้อมูลจากกลุ่มผู้ถูกประเมินในระดับคาบเส้นมากขึ้นเพื่อไม่ให้ละเมิดข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอยพหุคูณ โดยเฉพาะการแจกแจงของข้อมูลที่จำเป็นต้องมีการแจกแจงเป็นโค้งปกติ เพื่อให้การวิเคราะห์ได้ผลการศึกษาน่าเชื่อถือมากยิ่งขึ้น

ข้อเสนอแนะ

1. ควรพัฒนาโจทย์สถานการณ์ของการสอบ Objective Structured Clinical Examination (OSCE) ให้มีลักษณะเป็นโจทย์สถานการณ์ที่ครอบคลุมเนื้อหาตามเกณฑ์มาตรฐานของแพทย์สภาที่กำหนดไว้ เพื่อให้การคัดเลือกข้อสอบมาใช้ในการสอบแต่ละกลุ่มมีความยากง่ายที่ใกล้เคียงกัน และพัฒนาจัดทำเป็นคลังข้อสอบต่อไป
2. ควรมีการจัดทำมาตรฐานของผู้ป่วยจำลอง (Standardized Patient) เพราะการให้ข้อมูลในแต่ละครั้งของผู้ป่วยต้องมีความคงเส้นคงวา สามารถให้ข้อมูลเหมือนกันกับนักศึกษาผู้เข้าสอบทุกคน เพราะการให้ข้อมูลตามบทผู้ป่วยจำลองที่มีการเตรียมข้อมูลไว้ล่วงหน้ามีผลโดยตรงต่อคะแนนการประเมิน
3. ควรรวบรวมข้อมูลผลการใช้ข้อสอบ OSCE ให้มากขึ้น ซึ่งจะเป็นการเพิ่มโอกาสจำนวนกลุ่มศึกษากลุ่มคาบเส้น ซึ่งส่งผลโดยตรงต่อการกำหนดเกณฑ์ผ่านด้วยวิธีวิเคราะห์การถดถอยพหุคูณ
4. ควรทำการศึกษาการกำหนดเกณฑ์ผ่าน ในรายวิชาอื่นๆ ที่มีการจัดสอบ OSCE เพื่อประเมินทักษะทางคลินิก ภายใต้อาจารย์แพทย์ศาสตร์ มหาวิทยาลัยสงขลานครินทร์ เพื่อให้การกำหนดเกณฑ์ผ่านมีข้อมูลที่จะนำมาใช้ในการกำหนดเกณฑ์ในรูปแบบเดียวกัน เพื่อให้มาตรฐานการวัดและกำหนดเกณฑ์ผ่านข้อสอบ OSCE เหมือนกันในทุกรายวิชา
5. การคัดเลือกผู้ที่ทำหน้าที่กำหนดเกณฑ์ผ่านของข้อสอบ OSCE โดยวิธี Angoff's Method ย่อมส่งผลต่อการกำหนดเกณฑ์ผ่าน ดังนั้นควรมีระบบคัดเลือกผู้ที่ทำหน้าที่ดังกล่าวอย่างเหมาะสม และควรเป็นผู้ที่มีประสบการณ์ในการสอนและออกข้อสอบ OSCE มาก่อน

บรรณานุกรม

- Chesser AM, Laing MR, Miedzybrodzka ZH, et al. (2004). Factor analysis can be a useful standard setting tool in a high stakes OSCE assessment. *Medical Education*, 38: 825 - 31.
- Einion, A. (2013). OSCE assessment for emergency scenarios in midwifery education: A reflection and evaluation. *British Journal of Midwifery*, 21(12), 893-897.
- Gormley, G. (2011). Summative OSCEs in Undergraduate Medical Education. *Nurse Education in Practice*, 80: 127-132.
- Mitchell, M. L., Henderson, A., Groves, M., Dalton, M. & Nulty, D. (2009). The Objective Structured Clinical Examination (OSCE): Optimizing its Value in the Undergraduate Nursing Curriculum. *Nurse Education Today*, 29(4): 398-404.
- Thananit Sangkomkamhang. (2014). The Standard setting method to evaluation OSCE pass-scores of sixth year medical students. (Abstract) Presented at 16th Ottawa conference, 25-29 April, 2014, Ontario, Canada.
- Woragidpoonpol, P., & Yarmsuoy, A. (2006). Development of Nursing Practicum Assessment Model in Maternal and Newborn Nursing Practicum 2 Course Using OSCE in its Instruction. *Journal of Nursing Education*, 17, 25-33. [in Thai]
- เกษร สายธนู และคณะ. (2559). ศึกษาการใช้รูปแบบการสอบ OSCE เพื่อประเมินสมรรถนะการปฏิบัติการสร้างสัมพันธภาพเพื่อการบำบัดของนักศึกษาพยาบาล คณะพยาบาลศาสตร์ มหาวิทยาลัยอุบลราชธานี. การประชุมวิชาการและนำเสนอผลงานวิจัยระดับชาติ ราชธานีวิชาการ ครั้งที่ 1.
- เชิดศักดิ์ ไอรมนรัตน์. (2006). Iramaneerat C. Passing standard: Part III (Thai). *Medical Education Pamphlet 2006*; 2(3): 1.
- ประสพชัย พสุนนท์. (2558). การประเมินความเชื่อมั่นระหว่างผู้ประเมินโดยใช้สถิติแคปปา วารสารวิชาการศิลปศาสตร์ประยุกต์ มกราคม-มิถุนายน 2558.
- สาธิต คลังสิน และคณะ. (2559). เปรียบเทียบทักษะด้านต่างๆ ในการสอบ Obstructive Structured Clinical Examination ของนักศึกษาแพทย์ชั้นปีที่ 6 ในภาควิชาสูติศาสตร์และนรีเวชวิทยา สงขลานครินทร์เวชสาร ปีที่ 34 ฉบับที่ 6 พ.ย.-ธ.ค. 2559