



การพัฒนารูปแบบการสอบทางคลินิกเชิงโครงสร้างแบบปรนัยในการวัดทักษะทางคลินิก ของนักเรียนปฏิบัติการฉุกเฉินทางการแพทย์

A development of an objective structured clinical examination in
performance assessment for Advance emergency medical technician

นุชปียา ทองโชติ¹

Nuchpiya Thongchot

กรมแพทย์ทหารเรือ

Naval Medical Department

(Received 12/04/2021, Revised 1/06/2021, Accepted 4/06/2021)

บทคัดย่อ

การสร้างและพัฒนารูปแบบการประเมินผลภาคปฏิบัติ มีวัตถุประสงค์เพื่อวัดทักษะทางคลินิก โดยการสอบทางคลินิกเชิงโครงสร้างแบบปรนัย ประกอบด้วย 4 ขั้นตอน ได้แก่ 1) การสร้างผังข้อสอบ 2) การพัฒนาโจทย์สถานการณ์ 3) การฝึกอบรมครูผู้ประเมิน 4) การตรวจสอบคุณภาพเครื่องมือ จัดสอบจำนวน 3 สถานการณ์ สถานะ 10 นาที กลุ่มตัวอย่าง คือ นักเรียนโรงเรียนนาวิกเวชกิจ จำนวน 20 คน สุ่มอย่างง่ายด้วยวิธีจับสลาก เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบวัดทางคลินิกเชิงโครงสร้างแบบปรนัย ผลการวิจัยมีดังนี้ 1. ตรวจสอบคุณภาพเครื่องมือรายข้อ ได้แก่ 1) การวิเคราะห์หาดัชนีความยากของงาน พบว่า ผู้ตรวจคนที่ 1 มีค่าดัชนีความยากของสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.69, 0.71 และ 0.69 ตามลำดับ และ ผู้ตรวจคนที่ 2 ค่าดัชนีความยากของสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.70, 0.73 และ 0.69 ตามลำดับ 2) ค่าดัชนีอำนาจจำแนกของผู้ตรวจคนที่ 1 ในสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.74, 0.26 และ 0.74 ตามลำดับ ค่าดัชนีอำนาจจำแนกของผู้ตรวจคนที่ 2 ในสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.74, 0 และ 0.38 ตามลำดับ 2. การตรวจสอบคุณภาพเครื่องมือทั้งฉบับ ได้แก่ 1) ตรวจสอบความตรงตามเนื้อเรื่อง เท่ากับ 0.66 2) ความเที่ยงระหว่างผู้ประเมิน โดยหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันในสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.88, 0.98 และ 0.97 ตามลำดับ และหาค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด PXSR design พบว่า ค่าความน่าเชื่อถือของการวัดสำหรับการตัดสินใจเชิงสัมพัทธ์ ($\rho^{\wedge} s^2$) มีค่าเท่ากับ 0.49 และความน่าเชื่อถือของการวัดสำหรับการตัดสินใจเชิงสัมบูรณ์ ($\rho^{\wedge} \Delta^2$) มีค่าเท่ากับ 0.48 ถ้าต้องการความเที่ยงอย่างน้อย 0.70 ในการทดสอบควรมีผู้ประเมิน จำนวน 5 คน

คำสำคัญ : การประเมินผลภาคปฏิบัติ, การสอบทางคลินิกเชิงโครงสร้างแบบปรนัย, นักเรียนเวชกิจฉุกเฉิน

¹ ผู้ช่วยครูแผนกศึกษา โรงเรียนนาวิกเวชกิจ ศูนย์วิทยาการ กรมแพทย์ทหารเรือ, E-mail: nuchpiya29@yahoo.com



Abstract

The performance development aims at clinical skill qualification for advance emergency medical technician (AEMT) students. The objective of this study is to develop and validate objective structured clinical examination (OSCE). Regarding the study, the logical development steps are modified for experiment with the following four steps which are 1) blueprints and mapping, 2) scenario development, 3) examiner training and 4) validity and reliability testing. For an OSCE, there are 3 situations and each situation has 10 minutes. 20 AEMT students from Naval operational medical school are target group. The results were summarized as following.

For item-level analysis focus on task difficulty index, the first rater has difficulty index 0.69 for the first situation, 0.71 for the second situation and 0.69 for the third situation whereas the second rater has difficulty index 0.70 for the first situation, 0.73 for the second situation and 0.69 for the third situation. For item-level analysis focus on discrimination index, the first rater has difficulty index 0.74 for the first situation, 0.26 for the second situation and 0.74 for the third situation whereas the second rater has index 0.74 for the first situation, 0.00 for the second situation and 0.38 for the third situation

Test level analysis focus on item objective congruence (IOC) is 0.66. Focusing on inter-rater reliability, the Pearson product-moment correlation coefficient of situation shows the result as 0.88 for the first situation, 0.98 for the second situation and 0.97 for the third situation. Generalizability Coefficient (G-coefficient) PXSXR design is applied for relative decision. The results show that (ρ^2) is about 0.49 and G-coefficient for absolute decision (ρ^2) is 0.48. For reliability less than 0.7, it is recommended five persons for target group.

Key words: Performance assessment, objective structured clinical examination, Advance emergency medical technician student

1. ความเป็นมาและความสำคัญของปัญหา

การพัฒนาศักยภาพของผู้ปฏิบัติการฉุกเฉินทางการแพทย์นั้นมีความสำคัญอย่างยิ่งในการจัดการศึกษาให้มีความสอดคล้องกับความต้องการผู้ใช้งานจริงและมีสมรรถนะตามที่สถาบันการแพทย์ฉุกเฉิน (สพฉ.) กำหนด นอกเหนือจากสิ่งที่องค์กรวิชาชีพได้กำหนดมาตรฐานของนักเรียนหลักสูตรผู้ปฏิบัติการฉุกเฉินทางการแพทย์ (Advance emergency medical technician: AEMT) การปรับปรุงการจัดการเรียนการสอนให้มีประสิทธิภาพจะนำมาซึ่งคุณภาพชีวิตที่ดีของผู้ป่วย การคงไว้ซึ่งทักษะการปฏิบัติ อันจะนำมาสู่คุณค่าของวิชาชีพนั้น ๆ จากการศึกษาพบว่า การประเมินทักษะภาคปฏิบัติด้วยการสอบทางคลินิกเชิงโครงสร้างแบบปรนัย (Objective structured clinical examination: OSCE) จะสามารถทำนายการคงไว้ซึ่งทักษะที่ประเมินในระยะเวลา 1 ปี (Lee, Ho, Yu, & Chao, 2020) นอกจากนี้ยังพบว่า การวิเคราะห์เวลาที่ใช้ในการปฏิบัติการจะช่วยลดผลกระทบจากการทำงานที่ล่าช้า เช่น เวลาที่เริ่มต้นทำการกดหน้าอก (Time to First Compression) เวลาที่เริ่มช็อกไฟฟ้า (time to first shock) ในการปฏิบัติการกู้ชีพตามแนวทางการกู้ชีพ (AmericanHeartAssociation, 2017) การคัดกรองและปฏิบัติงานตามแนวทางการดูแลผู้ป่วยโรคหัวใจ เช่น



เวลาที่ใช้ในการประเมินผู้ป่วยที่ห้องตรวจโรคฉุกเฉินจนกระทั่งส่งไปทำหัตถการสวนหัวใจ (Door to Balloon) การดูแลผู้ป่วยโรคหลอดเลือดสมอง เช่น เวลาเริ่มต้นในการให้ยา (Door to Drug) (Shinnick & Woo, 2018)

การพัฒนาและประเมินความสามารถทางคลินิกนับว่าเป็นกิจกรรมที่มีความท้าทายอย่างยิ่งในการจัดองค์ประกอบของการศึกษาให้เป็นไปตามมาตรฐานขององค์การวิชาชีพ แม้ว่าปัจจุบันจะมีวิธีการประเมินผลความสามารถทางคลินิกหลายวิธี เช่น การสอบปากเปล่า (Oral Examination) การสังเกตจากการปฏิบัติจริง (Direction Observer) แต่วิธีดังกล่าวอาจมีข้อจำกัด ได้แก่ การสอบปากเปล่าใช้เวลานานและอาจไม่ครอบคลุมเนื้อหาทั้งหมด การประเมินโดยการสังเกตอาจเกิดความลำเอียงจากผู้ประเมินแต่ละคน ส่วนการประเมินความสามารถทางคลินิกด้วย (OSCE) เป็นวิธีการจำลองสถานการณ์ทางคลินิกที่ถูกนำมาใช้เพื่อการเรียนรู้อย่างต่อเนื่องของนักเรียน เป็นวิธีการประเมินผลรูปแบบหนึ่ง ที่ช่วยให้การประเมินทางคลินิกมีความน่าเชื่อถือสูง และครอบคลุมทักษะทางคลินิกที่สำคัญและได้กลายเป็นวิธีที่ได้รับการยอมรับอย่างกว้างขวาง (รุ่งนภา นาคะภาร, 2540)

การจัดสอบ OSCE ตามมาตรฐาน มีจุดเด่นดังนี้ 1) ลดความแปรปรวนของผู้ป่วยและผู้ประเมิน เนื่องจากใช้ผู้ป่วยและผู้ประเมินคนเดียวกัน 2) รูปแบบการประเมินมีความเที่ยง (Reliability) เนื่องจากมีการวางแผนล่วงหน้าอย่างมีหลักเกณฑ์แน่นอน โดยการสร้าง Checklist สำหรับการให้คะแนนในสถานการณ์จำลอง 3) รูปแบบการประเมินมีความตรงตามเนื้อหา (Content Validity) เนื่องจากมีการออกข้อสอบที่มีเนื้อหาครอบคลุม แบ่งข้อสอบออกเป็นข้อย่อย ๆ และมีเกณฑ์การให้คะแนนที่ชัดเจน 4) สามารถใช้การประเมินทักษะและการตัดสินใจทางคลินิกที่มีความสำคัญในการดูแลผู้ป่วยฉุกเฉิน แต่อาจไม่พบบ่อยในการฝึกปฏิบัติงานจริง รวมถึงสามารถประเมินความรู้ขั้นสูงด้วยการประยุกต์ใช้การคิดอย่างมีวิจารณญาณและการตัดสินใจทางคลินิก 5) การประเมิน OSCE ทำให้ผู้สอนสามารถบอกจุดอ่อนของนักเรียนได้ 6) การสอบ OSCE สะท้อนให้เห็นถึงคุณภาพและประสิทธิภาพการสอนของผู้สอน ซึ่งจะนำไปสู่การปรับปรุงการเรียนการสอนต่อไป 7) ประเมินความรู้ความสามารถที่ไม่สามารถประเมินด้วยวิธีอื่น ๆ เช่น ทักษะการให้คำแนะนำแก่ผู้ป่วย ทักษะในการสื่อสารกับผู้ป่วย (ศิริวรรณ ชูกำเนิด และสุระพร บุษเจริญ, 2560) แม้ว่าจุดเด่นของการประเมินแบบ OSCE จะมีจุดเด่นจำนวนมากแต่มีข้อจำกัดในการประเมินเช่นกัน เช่น การประเมินอยู่ในระดับ Knows How เท่านั้น (Miller, 1990) และการประเมินแบ่งความสามารถทางคลินิกออกเป็นส่วนย่อย ๆ ซึ่งไม่ตรงกับชีวิตจริงที่ผู้ปฏิบัติการฉุกเฉินทางการแพทย์จะต้องผสมผสานความสามารถทางคลินิกหลาย ๆ อย่างร่วมกันในการดูแลผู้ป่วย

โรงเรียนนาวิกเวชกิจเป็นสถานศึกษาในสังกัดกระทรวงกลาโหม มีภารกิจในการผลิตบุคลากรด้านการแพทย์ฉุกเฉินในระดับประกาศนียบัตรวิชาชีพชั้นสูง สาขาปฏิบัติการฉุกเฉินทางการแพทย์ให้มีความรู้ความสามารถตามขอบเขต บทบาท หน้าที่ ตามที่สถาบันการแพทย์ฉุกเฉินกำหนด เพื่อตอบสนองกรมแพทย์ทหารเรือ โดยมีปรัชญาในการจัดการศึกษาให้นักเรียนสามารถปฏิบัติงานร่วมกับหน่วยงานที่เกี่ยวข้องโดยใช้กระบวนการคิดวิเคราะห์ ในการแก้ปัญหาอย่างมีวิจารณญาณ ให้การดูแลผู้ป่วยฉุกเฉินอย่างมีศาสตร์และศิลป์อย่างเป็นองค์รวม (Holistic Care) ในการวิจัยครั้งนี้ได้ทำการพัฒนาพัฒนาการสอบทางคลินิกเชิงโครงสร้างแบบปรนัยในการวัดทักษะทางคลินิกในรายวิชาการดูแลรักษาผู้บาดเจ็บฉุกเฉิน (Emergency Medical Care of Trauma Patients) ซึ่งมีจุดมุ่งหมายให้นักเรียนมีความรู้ความเข้าใจเกี่ยวกับการประเมินผู้บาดเจ็บฉุกเฉินที่มีความผิดปกติในสถานการณ์ต่าง ๆ ตลอดจนให้การดูแลรักษา ณ จุดเกิดเหตุ และระหว่างนำส่งโรงพยาบาลได้อย่างถูกต้อง รวดเร็ว ตามขอบเขตของเจ้าพนักงานฉุกเฉินทางการแพทย์ โดยใช้กระบวนการคิดอย่างมีวิจารณญาณและการตัดสินใจทางคลินิก รวมถึงตระหนักถึงความสำคัญ มีเจตคติที่ดี และคุณธรรมจริยธรรมในการดูแลรักษาผู้บาดเจ็บฉุกเฉิน จากการวิเคราะห์องค์ประกอบของการจัดการทดสอบแบบ OSCE



ในรายวิชาดังกล่าว ในส่วนของการพัฒนาด้านกระบวนการ ผู้วิจัยได้นำแนวทางกระบวนการสืบเสาะความตรง (Validity Inquiry process) (Conn, Bohan, Pieper, & Musumeci, 2020) มาใช้ในการประชุมครู ผู้ประเมิน เพื่อตรวจสอบการประเมินภาคปฏิบัติและสร้างการอ้างอิงอย่างสมเหตุสมผล ประกอบด้วยการวิเคราะห์เนื้อหา (Content Analysis) รูปแบบการสืบเสาะความตรง (Validity Inquiry Form) เมตารูบรีค (Meta Rubric) การสำรวจนักเรียน (Student Survey) ดำเนินการทบทวนความเที่ยง (Conducting Review) รวมถึงการวิเคราะห์องค์ประกอบของการจัดสถานการณ์ OSCE เพื่อเป็นข้อมูลประกอบการสนทนากลุ่ม เพื่อหาแนวทางในการประเมินที่เป็นแนวปฏิบัติที่ดีที่สุด อันจะนำไปสู่การประเมินทักษะผู้เรียนในการดูแลผู้ป่วยฉุกเฉินที่มีคุณภาพมาตรฐาน มีความสอดคล้องกับการทดสอบในระดับประเทศของสถาบันการแพทย์ฉุกเฉินแห่งชาติ

2. วัตถุประสงค์การวิจัย

- 2.1 เพื่อสร้างและพัฒนารูปแบบการสอบทางคลินิกเชิงโครงสร้างแบบปรนัยในรายวิชาการดูแลผู้ป่วยฉุกเฉิน
- 2.2 เพื่อศึกษาคุณภาพของแบบสอบทางคลินิกเชิงโครงสร้างแบบปรนัยในรายวิชาการดูแลผู้ป่วยฉุกเฉิน

3. วิธีดำเนินการวิจัย

- 3.1 รูปแบบการวิจัย เป็นการวิจัยและพัฒนา (Research and Development)
- 3.2 ประชากร คือ นักเรียนหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง โรงเรียนนาวิกเวชกิจ ชั้นปีที่ 1 ปีการศึกษา 2563 จำนวน 60 คน ที่เคยผ่านการเรียนรายวิชาการดูแลผู้ป่วยฉุกเฉิน
- 3.3 กลุ่มตัวอย่างคือ นักเรียนหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง โรงเรียนนาวิกเวชกิจ ชั้นปีที่ 1 ปีการศึกษา 2563 จำนวน 20 คน คัดเลือกกลุ่มตัวอย่างด้วยวิธีสุ่มอย่างง่ายด้วยวิธีจับสลาก
- 3.4 เนื้อหาวิชาที่ใช้ในการสร้างแบบสอบเป็นการประเมินทักษะปฏิบัติทางคลินิกในรายวิชาการดูแลผู้ป่วยฉุกเฉิน ประกอบด้วย ผู้บาดเจ็บฉุกเฉินที่มีความผิดปกติในสถานการณ์ต่าง ๆ ตลอดจนให้การดูแลรักษา จุดเกิดเหตุ และระหว่างนำส่งโรงพยาบาลได้อย่างถูกต้อง รวดเร็ว ตามความรับผิดชอบของเจ้าพนักงานฉุกเฉินทางการแพทย์
- 3.5 การวิจัยครั้งนี้ใช้ขั้นตอนในการพัฒนาของ Kamran Khan et al. (2013) มี 4 ขั้นตอน ได้แก่ 1) การสร้างผังข้อสอบ (Blueprint and Mapping) 2) การพัฒนาโจทย์สถานการณ์ 3) การฝึกอบรมครูผู้ประเมิน 4) การตรวจสอบคุณภาพเครื่องมือ (Kamran Khan, Kathryn Gaunt, Sankaranarayanan Ramachandran, & Pushkar, 2013) โดยกำหนดองค์ประกอบในการจัดสถานการณ์ โดยแบ่งเป็น 2 ระยะ ได้แก่ ระยะที่ 1 การพัฒนาแบบวัดทักษะทางคลินิกโดยการสอบทางคลินิกเชิงโครงสร้างแบบปรนัย ระยะที่ 2 การตรวจสอบคุณภาพของแบบวัดทักษะทางคลินิกโดยการสอบทางคลินิกเชิงโครงสร้างแบบปรนัย ดังแสดงในตารางที่ 1

ตารางที่ 1 แสดงองค์ประกอบในการจัดสถานการณ์

ประเภท	เนื้อเรื่อง
งานของนักเรียน	สถานการณ์โรคของผู้ป่วย, ข้อคำถามในการประเมิน, เวลาที่ใช้ในการประเมิน
แนวทางปฏิบัติของผู้ป่วยมาตรฐาน	คำถามที่ใช้ในการประเมิน, การตอบสนองต่อคำถามและการปฏิบัติของนักเรียน
แนวทางสำหรับครูผู้ประเมิน	คำถามในการประเมิน, การแก้ไขปัญหาเมื่อเกิดสถานการณ์ที่ไม่คาดคิด
มาตรฐานสำหรับการประเมิน	มาตรฐานการใช้มาตรวัด, คำนิยามของความสำเร็จอย่างสมบูรณ์, ความสำเร็จเป็นบางส่วนและการตกอย่างสมบูรณ์
แบบตรวจสอบรายการ	ข้อคำถามในแบบตรวจสอบรายการและการตัดสินผล
การเตรียมอุปกรณ์	อุปกรณ์และจำนวนคนที่ใช้ในการประเมิน



ระยะที่ 1 การพัฒนาวัดทักษะทางคลินิกโดยการสอบทางคลินิกเชิงโครงสร้างแบบปรนัย

1. การพัฒนาผังข้อสอบภาคปฏิบัติ (Table Specification) โดยการตรวจสอบความครอบคลุมเนื้อหา ความเป็นตัวแทน ของโจทย์สถานการณ์ พบว่า ยังไม่ครอบคลุมเนื้อหาในการจัดการเรียนการสอน โดยก่อนการสร้างเครื่องมือมีความครอบคลุมใน 6 เนื้อเรื่อง และหลังสร้างเครื่องมือมีความครอบคลุม 10 เนื้อเรื่อง โดยโจทย์เดิมมีความซ้ำซ้อนวัดทักษะการดูแลผู้ป่วยฉุกเฉินที่มีความคล้ายคลึงกันทั้ง 3 สถานการณ์ จะทำให้เกิดการวัดเนื้อหาในเชิงลึกแต่ขาดความกว้างที่ครอบคลุมองค์ความรู้ในการดูแลผู้ป่วยฉุกเฉิน นอกจากนี้พบว่า การระบุโรคไม่สอดคล้องกับสถานการณ์ที่กำหนดไว้ในโจทย์ เช่น ระบุโรคที่มีกระดูกหัก แต่เนื้อหาไม่มีสถานการณ์ที่เป็นกระดูกหักจะทำให้ขาดการประเมินเนื้อหาในส่วนดังกล่าว ดังนั้นผู้วิจัย จึงสร้างโจทย์ใหม่ที่มีสถานการณ์ที่มีความครอบคลุมทุกหัวข้อ เพื่อให้การประเมินทักษะการปฏิบัติมีความครอบคลุมองค์ความรู้และตรงตามจุดมุ่งหมายของการจัดการเรียนการสอน ดังตารางที่ 2

ตารางที่ 2 แสดงผังข้อสอบภาคปฏิบัติ

เนื้อเรื่อง	วัตถุประสงค์ เชิงพฤติกรรม	OSCE 1		OSCE 2		OSCE 3	
		ก่อน	หลัง	ก่อน	หลัง	ก่อน	หลัง
1. Trauma overview	2.2.2						
2. Bleeding Traumatic shock & Resuscitation	2.3.2	✓	✓				✓
	2.4.1						
3. Soft tissue Injury	3.3	✓	✓	✓	✓	✓	
4. Orthopedic trauma	4.1		✓				
5. Head, Facial, neck & spinal trauma	4.5				✓		
6. Nervous system				✓	✓		
7. Chest trauma				✓	✓	✓	
8. Abdominal& Genitourinary						✓	✓
9. Special consideration in Trauma							✓
10. Mechanism of injury		✓	✓	✓	✓	✓	✓
11. Environmental emergency					✓		

2. ดำเนินการประชุม ในส่วนของการพัฒนาด้านกระบวนการ ผู้วิจัยได้นำแนวทางกระบวนการสืบเสาะความจริง (Validity Inquiry Process) (Conn et al., 2020) มาใช้ในการประชุมครูผู้สอน เพื่อตรวจสอบการประเมินภาคปฏิบัติและสร้างการอ้างอิงอย่างสมเหตุสมผล ประกอบด้วยการวิเคราะห์เนื้อหา (Content Analysis) รูปแบบการสืบเสาะความจริง (Validity Inquiry Form) เมตารูบรีค (Meta Rubric) การสำรวจนักเรียน (Student Survey) ดำเนินการทบทวนความเที่ยง (Conducting review) รวมถึงการวิเคราะห์องค์ประกอบของการจัดสถานการณ์ OSCE เพื่อเป็นข้อมูลประกอบการสนทนากลุ่มเพื่อหาแนวทางในการประเมินที่เป็นแนวปฏิบัติที่ดีที่สุด อันจะนำไปสู่การประเมินทักษะผู้เรียนในการดูแลผู้ป่วยฉุกเฉินที่มีคุณภาพมาตรฐาน มีความสอดคล้องกับการทดสอบในระดับประเทศของสถาบันการแพทย์ฉุกเฉินแห่งชาติ จากมติที่ประชุมสรุปผลการประชุมได้ดังนี้

2.1 การวิเคราะห์เนื้อหา (Content Analysis) กำหนดจุดมุ่งหมายของการวัดและประเมินทักษะการดูแลผู้ป่วยฉุกเฉินเพื่อประเมินทักษะการดูแลผู้ป่วยฉุกเฉินของนักเรียนจำหนวเรือชั้นปีที่ 1 ว่าสามารถดูแลผู้ป่วยได้ถูกต้องตามหลักการการดูแลรักษาผู้ป่วยตั้งแต่ก่อนจนถึงโรงพยาบาล (Pre-Hospital Trauma



Life Support: PHTLS) ตามเกณฑ์มาตรฐานของสถาบันการแพทย์ฉุกเฉินแห่งชาติ นำโจทย์ที่สร้างขึ้นมา ประชุมกับครูผู้สอนหลังจากนั้นนำไปตรวจสอบความตรงตามเนื้อเรื่องกับผู้เชี่ยวชาญ

2.2 รูปแบบการสืบเสาะความตรง (Validity Inquiry Form) กำหนดกรอบการประเมินการดูแล ผู้บาดเจ็บ (Emergency Care of Trauma Patient) สำหรับนักเรียนหลักสูตรปฏิบัติการฉุกเฉินทางการแพทย์ ชั้นปีที่ 1 แบ่งออกเป็น 8 องค์ประกอบ ซึ่งเป็นรูปแบบมาตรฐานของของสถาบันการแพทย์ฉุกเฉินแห่งชาติ ที่ใช้ในการปฏิบัติงานและการสอบเพื่อขอรับใบอนุญาตประกอบวิชาชีพซึ่งมีรายละเอียดดังนี้ 1) การประเมิน สถานการณ์ (Scene size-up) 2) การประเมินสภาพผู้บาดเจ็บและผู้ป่วยฉุกเฉิน ณ จุดเกิดเหตุ (Initial assessment) 3) การประเมินการบาดเจ็บอย่างรวดเร็ว (Rapid Scan) 4) การประเมินซ้ำระหว่างนำส่ง (Reassessment) 5) การวัดสัญญาณชีพ การประเมินระดับความรู้สึกตัวผู้บาดเจ็บทางสมองและระดับน้ำตาล ในเลือด (V/S & GCS & DTX) 6) การซักประวัติตามหลัก SAMPLE 7) ประเมินแบบละเอียด Detailed Physical Assessment 8) การส่งมอบผู้บาดเจ็บ MIST กำหนดเกณฑ์ผ่านการประเมินทั้งหมด คือ นักเรียน จะต้องได้คะแนนทักษะการดูแลผู้บาดเจ็บ ไม่น้อยกว่า 17 คะแนน (ร้อยละ 60) ในการสอบผู้ป่วยจำลอง ทั้งหมด 3 สถานการณ์

2.3 เมตารูบริก (Meta Rubric) ทบทวนการใช้เกณฑ์รูบริกที่ใช้ในการทดสอบของประเทศไทย พบว่า สถาบันการแพทย์ฉุกเฉินซึ่งเป็นหน่วยงานหลักของประเทศจัดการทดสอบใช้เกณฑ์รูบริก ดังนั้น ในการ จัดการเรียนการสอนควรมีการพัฒนาเกณฑ์ในรายวิชาการดูแลผู้บาดเจ็บฉุกเฉินให้เป็นเกณฑ์รูบริก และ กำหนดประเด็นที่ต้องการประเมินทักษะการดูแลผู้บาดเจ็บ เน้นกระบวนการปฏิบัติงาน (Process) และ ผลงาน (Process) กล่าวคือ มีการประเมินคุณภาพของการประเมินสถานการณ์และการให้การรักษา ทุกขั้นตอนและคุณภาพของการดูแลผู้บาดเจ็บเชิงปริมาณ คือ จำนวนการดูแลผู้บาดเจ็บที่ผ่านเกณฑ์ การประเมินทักษะการดูแลผู้บาดเจ็บ

2.4 การสำรวจนักเรียน (Student Survey) กำหนดงานและสถานการณ์ที่ให้ผู้รับการประเมิน ปฏิบัติเป็นงานที่เป็นเหตุการณ์ (Event Tasks) โดยมีผู้ป่วยมาตรฐานจำนวน 3 ราย เพื่อให้ นักเรียนหลักสูตร ปฏิบัติการฉุกเฉินทางการแพทย์ให้การดูแลผู้บาดเจ็บโดยมีภาพเหตุการณ์สถานการณ์จริงโดยประยุกต์ใช้กับ โปรแกรม Power Point โดยอาจารย์ผู้สอนร่วมกันพิจารณาโดยใช้เกณฑ์การประเมินทักษะการดูแลผู้บาดเจ็บ 3 ระดับ รวมถึงพิจารณาความยากง่ายให้สอดคล้องกับระดับการศึกษาในชั้นปีที่ 1

2.5 ดำเนินการทบทวนความเที่ยง (Conducting Review) การจัดการเรียนการสอนในรายวิชา ผู้บาดเจ็บฉุกเฉินยังไม่เคยตรวจสอบความเที่ยง และยังไม่มีการตรวจสอบคุณภาพเครื่องมือการประเมิน ภาควิชาปฏิบัติ

2.6 ตรวจสอบความตรงของเกณฑ์การประเมิน โดยพิจารณาจากค่าดัชนีความสอดคล้องของ วัตถุประสงค์กับแบบประเมินภาควิชาปฏิบัติการดูแลผู้บาดเจ็บฉุกเฉินโดยผู้เชี่ยวชาญจำนวน 3 ท่าน และ ปรับปรุงแก้ไขตามคำแนะนำ ผศ.อภรณ์ศรีประเสริฐและผศ.อภรณ์ศรีประเสริฐ

2.7 อธิบายแนวทางการประเมินและมอบหมายให้นักเรียนทำแบบฝึกทักษะการดูแลผู้บาดเจ็บ ฉุกเฉินโดยใช้แผนการสอน

2.8 จัดสอบเพื่อประเมินทักษะการดูแลผู้บาดเจ็บฉุกเฉินและนำเกณฑ์ไปใช้ประเมินภาระงาน โดยผู้ประเมินจำนวน 2 คน ที่เป็นอิสระต่อกัน สถานีละ 10 นาที จำนวน 3 สถานการณ์

ระยะที่ 2 การตรวจสอบคุณภาพของแบบวัดทักษะทางคลินิก โดยการสอบทางคลินิกเชิงโครงสร้างแบบปรนัย

การตรวจสอบคุณภาพเครื่องมือ ได้แก่ การวิเคราะห์ค่าสถิติพื้นฐาน การวิเคราะห์คุณภาพเครื่องมือรายข้อ ได้แก่ ดัชนีความยากของงาน ดัชนีอำนาจจำแนก การวิเคราะห์คุณภาพเครื่องมือทั้งฉบับ ได้แก่ ความเที่ยงระหว่างผู้ประเมิน

เครื่องมือที่ใช้ในการวิจัย

1. แบบประเมินทักษะภาคปฏิบัติวิชาการดูแลผู้บาดเจ็บฉุกเฉิน จำนวน 3 สถานการณ์ ใช้เกณฑ์รูบริก 3 ระดับ โดยกำหนดเกณฑ์การให้คะแนนดังนี้ ให้ 2 คะแนน เมื่อปฏิบัติได้ถูกต้องครบถ้วน, ให้ 1 คะแนน เมื่อปฏิบัติได้ถูกต้องบางส่วน, ให้ 0 คะแนน เมื่อปฏิบัติไม่ถูกต้องตามที่กำหนดหรือไม่ปฏิบัติ



รูปที่ 1 การประเมินโดยครูคนที่ 1



รูปที่ 2 การประเมินโดยครูคนที่ 2

4. สรุปผลการวิจัย

ผลการวิเคราะห์ข้อมูลของแบบประเมินทักษะภาคปฏิบัติในวิชาการดูแลผู้บาดเจ็บฉุกเฉิน จำนวน 3 สถานการณ์ สามารถนำเสนอเป็น 2 ระยะ โดยมีรายละเอียดดังนี้

ระยะที่ 1 การพัฒนาวัดทักษะทางคลินิกโดยการสอบทางคลินิกเชิงโครงสร้างแบบปรนัย

1. การวิเคราะห์ความตรงตามเนื้อเรื่อง (Content Validity) การตรวจสอบความตรงตามเนื้อเรื่องในการปรับเกณฑ์รูบริก โดยการนำเสนองานในชั้นเรียนกับอาจารย์ผู้สอนและนิสิตภายในชั้นเรียนรายวิชา การประเมินภาคปฏิบัติ จุฬาลงกรณ์มหาวิทยาลัย และตรวจสอบความตรงตามเนื้อเรื่องกับผู้เชี่ยวชาญจำนวน 3 ท่าน จากนั้นนำคะแนนที่ได้มาคำนวณหาค่าดัชนีความสอดคล้องระหว่างมิติหรือองค์ประกอบของเครื่องมือวัด (IOC) เท่ากับ 0.66

ระยะที่ 2 การตรวจสอบคุณภาพของแบบวัดทักษะทางคลินิกโดยการสอบทางคลินิกเชิงโครงสร้างแบบปรนัย

1. การวิเคราะห์ค่าสถิติพื้นฐาน การประเมินการปฏิบัติในสถานการณ์การทดสอบทั้งหมด 3 สถานการณ์ ของผู้สอบจำนวน 20 คน ที่ให้คะแนนโดยผู้ประเมินจำนวน 2 คน สามารถแสดงรายละเอียดผลคะแนนได้ดังตารางที่ 3



ตารางที่ 3 ค่าสถิติพื้นฐานที่ได้จากการวิเคราะห์ข้อมูล

ค่าสถิติ	คะแนนผู้ประเมินคนที่ 1			คะแนนผู้ประเมินคนที่ 2		
	สถานการณ์ที่ 1	สถานการณ์ที่ 2	สถานการณ์ที่ 3	สถานการณ์ที่ 1	สถานการณ์ที่ 2	สถานการณ์ที่ 3
คะแนนสูงสุด	23	23	23	22	23	22
คะแนนต่ำสุด	14	16	14	16	17	15
คะแนนเฉลี่ย	19.27	19.90	19.27	19.55	20.30	19.45
ส่วนเบี่ยงเบนมาตรฐาน	2.41	2.12	2.41	2.02	1.77	1.90
ความแปรปรวน	5.82	4.49	5.83	4.08	3.14	3.63

จากตาราง 3 พบว่า ค่าสถิติพื้นฐานของคะแนนในการประเมินการปฏิบัติจากผู้ประเมินคนที่ 1 พบว่า ในสถานการณ์ที่ 1, 2 และ 3 มีคะแนนสูงสุดเท่ากับ 23 เช่นเดียวกัน คะแนนต่ำสุดเท่ากับ 14, 16 และ 14 ตามลำดับ คะแนนเฉลี่ยเท่ากับ 19.27, 19.90 และ 19.27 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 2.41, 2.12 และ 2.41 ตามลำดับ ค่าความแปรปรวนเท่ากับ 5.82, 4.49 และ 5.83 ตามลำดับ และเมื่อทำการวิเคราะห์ค่าสถิติพื้นฐานของคะแนนในการประเมินการปฏิบัติ จากผู้ประเมินคนที่ 2 พบว่า ในสถานการณ์ที่ 1, 2 และ 3 มีคะแนนสูงสุดเท่ากับ 22, 23 และ 22 ตามลำดับ คะแนนต่ำสุดเท่ากับ 16, 17 และ 15 ตามลำดับ คะแนนเฉลี่ยเท่ากับ 19.27, 19.90 และ 19.27 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 2.02, 1.77 และ 1.90 ตามลำดับ ค่าความแปรปรวนเท่ากับ 4.08, 3.14 และ 3.63 ตามลำดับ

2. การวิเคราะห์คุณภาพเครื่องมือวิจัย

2.1 ดัชนีความยากของงาน (Task Difficulty Index) โดยคำนวณจากสัดส่วนของค่าเฉลี่ยของคะแนนรายข้อหรืองานแต่ละงานของผู้รับการประเมินกับคะแนนที่เป็นไปได้สูงสุดในแต่ละข้อ สามารถแสดงผลการวิเคราะห์ที่ได้ดังตารางที่ 4

ตารางที่ 4 แสดงดัชนีความยากของงานของผู้ตรวจคนที่ 1 และผู้ตรวจคนที่ 2

สถานการณ์	ผู้ประเมินคนที่ 1			ผู้ประเมินคนที่ 2		
	คะแนนเฉลี่ยแต่ละงาน	คะแนนสูงสุดที่เป็นไปได้	ดัชนีความยากของงาน	คะแนนเฉลี่ยแต่ละงาน	คะแนนสูงสุดที่เป็นไปได้	ดัชนีความยากของงาน
สถานการณ์ที่ 1	19.27	28	0.69	19.55	28	0.70
สถานการณ์ที่ 2	19.90	28	0.71	20.30	28	0.73
สถานการณ์ที่ 3	19.27	28	0.69	19.45	28	0.69

จากตารางที่ 4 พบว่า ข้อมูลจากผู้ประเมินคนที่ 1 จะได้ค่าดัชนีความยากของสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.69, 0.71 และ 0.69 ตามลำดับ และข้อมูลจากผู้ประเมินคนที่ 2 จะได้ค่าดัชนีความยากของสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.70, 0.73 และ 0.69 ตามลำดับ โดยค่าดัชนีความยากควรมีค่าอยู่ระหว่าง 0.3 ถึง 0.9 (กมลวรรณ ตังธนากานนท์, 2563) แสดงว่าค่าดัชนีความยากนี้อยู่ในระดับที่สามารถยอมรับได้



2.2 ดัชนีอำนาจจำแนก (Discrimination Index) โดยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์แบบพอยต์ไปซีเรียล (Point Biserial Correlation Coefficient) ในสถานการณ์ที่ 1, 2 และ 3 ของผู้ตรวจคนที่ 1 และผู้ตรวจคนที่ 2 สามารถสรุปได้ดังตารางที่ 5

ตารางที่ 5 ดัชนีอำนาจจำแนก (Discrimination index)

ผู้ตรวจ	ดัชนีอำนาจจำแนก		
	สถานการณ์ที่ 1	สถานการณ์ที่ 2	สถานการณ์ที่ 3
คนที่ 1	0.74	0.26	0.74
คนที่ 2	0.74	0	0.38

จากตารางที่ 5 ค่าดัชนีอำนาจจำแนกของผู้ประเมินคนที่ 1 ในสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.74, 0.26 และ 0.74 ตามลำดับ ค่าดัชนีอำนาจจำแนกของผู้ประเมินคนที่ 2 ในสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.74, 0 และ 0.38 ตามลำดับ ซึ่งจะเห็นว่าสถานการณ์การทดสอบที่ 1, 3 มีค่าอำนาจจำแนกมากกว่า 0.40 แสดงว่าการประเมินการปฏิบัติในสถานการณ์ที่ 1 และสถานการณ์ที่ 3 มีอำนาจจำแนกดีมาก ในขณะที่สถานการณ์ที่ 2 มีค่าอำนาจจำแนกเท่ากับ 0.26 จากผู้ตรวจคนที่ 1 แสดงว่ามีอำนาจจำแนกพอใช้ได้ ในขณะที่มีอำนาจจำแนกเท่ากับ 0 จากผู้ตรวจคนที่ 2 แสดงว่าควรมีการปรับปรุง

3. การตรวจคุณภาพเครื่องมือทั้งฉบับ

การตรวจคุณภาพเครื่องมือทั้งฉบับ ได้แก่ ความเที่ยงระหว่างผู้ประเมิน (Inter-rater Reliability) โดยค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันและความเที่ยงระหว่างผู้ประเมิน (Inter-rater Reliability) โดยหาค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Coefficient)

3.1 การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันในแต่ละสถานการณ์สามารถแสดงรายละเอียดดังตารางที่ 6

ตารางที่ 6 ค่าสัมประสิทธิ์ความเที่ยงระหว่างผู้ประเมิน

สถานการณ์	สัมประสิทธิ์ความเที่ยงระหว่างผู้ประเมิน
สถานการณ์ที่ 1	0.88
สถานการณ์ที่ 2	0.98
สถานการณ์ที่ 3	0.97

จากตาราง 6 ค่าความเที่ยงระหว่างผู้ประเมินในสถานการณ์ที่ 1, 2 และ 3 มีค่าเท่ากับ 0.88, 0.98 และ 0.97 ตามลำดับ ซึ่งค่าที่ได้มีค่าเข้าใกล้ 1 แสดงว่ามีความสัมพันธ์กันสูง นั่นคือ มีความเที่ยงระหว่างผู้ประเมินค่อนข้างมาก

3.2 การหาความเที่ยงระหว่างผู้ประเมิน (Inter-rater Reliability) โดยหาค่าสัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Coefficient) PXSXR design นั่นคือ ผู้เรียนจำนวน 20 คน (Person: P) เข้ารับการทดสอบคนละ 3 สถานการณ์ (Situation: S) และมีผู้ประเมินให้คะแนนแต่ละสถานการณ์ (Rater: R) จำนวน 2 คน ด้วยโปรแกรม EduG สามารถทำการวิเคราะห์ได้ดังตารางที่ 7



ตารางที่ 7 แสดงผลการวิเคราะห์ความแปรปรวนและองค์ประกอบความแปรปรวนในการศึกษา G
สำหรับการออกแบบ PXSXR

Source of Variation	Sum of Square	df	Mean Square	Estimated Error Component	Percentage of Error Variance
P	59.12500	19	3.11184	1.39583	13.7
S	4.36250	2	2.18125	0.04989	0.5
R	0.00833	1	0.00833	0.03838	0.4
PS	112.88750	38	2.97072	-2.87697	0.0
PR	9.32500	19	0.49079	-2.74463	0.0
SR	11.87917	2	5.93958	-0.13925	0.0
PSR	331.53750	38	8.72467	8.72467	85.5
Total	529.12500	119			100

เมื่อ P แทน ผู้สอบ S แทน สถานการณ์ที่ประเมิน R แทน ผู้ประเมิน
 PS แทน ปฏิสัมพันธ์ระหว่างผู้สอบและสถานการณ์ทดสอบ
 PR แทน ปฏิสัมพันธ์ระหว่างผู้สอบและผู้ประเมิน
 IR แทน ปฏิสัมพันธ์ระหว่างสถานการณ์ทดสอบและผู้ประเมิน
 PSR แทน ปฏิสัมพันธ์ระหว่างผู้สอบ สถานการณ์ทดสอบ และผู้ประเมิน

จากตารางที่ 7 นำเสนอผลการวิเคราะห์ความแปรปรวนขององค์ประกอบจากแหล่งความแปรปรวน 7 แหล่ง พบว่า ความแปรปรวนของคะแนนจริง (σ_p^2) มีค่ามากที่สุด คิดเป็นร้อยละ 13.7 นอกนั้นเป็นความแปรปรวนของความคลาดเคลื่อนจากแหล่งอื่น ๆ จำนวน 6 แหล่ง ได้แก่ ความแปรปรวนของคะแนนเฉลี่ยระหว่างสถานการณ์สอบ (σ_s^2) คิดเป็นร้อยละ 0.5 ความแปรปรวนของคะแนนเฉลี่ยระหว่างผู้ประเมิน (σ_r^2) คิดเป็นร้อยละ 0.4 ความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบและสถานการณ์ทดสอบ จริง (σ_{ps}^2) คิดเป็นร้อยละ 0.0 ความแปรปรวนของปฏิสัมพันธ์ระหว่างระหว่างผู้สอบและผู้ประเมิน (σ_{pr}^2) คิดเป็นร้อยละ 0.0 ความแปรปรวนของปฏิสัมพันธ์ระหว่างสถานการณ์การสอบและผู้ประเมิน (σ_{sr}^2) คิดเป็นร้อยละ 0.0 จะเห็นได้ว่าความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้สอบ ข้อสอบ และผู้ประเมิน คิดเป็นร้อยละ 85.5 แสดงให้เห็นว่ามีความคลาดเคลื่อนที่ยังไม่สามารถวัดได้ค่อนข้างมาก

เมื่อทำการวิเคราะห์ความน่าเชื่อถือของผลการวัดในการศึกษา D หรือ D Study สำหรับการออกแบบ (PXSXR) เมื่อผู้ตรวจมีจำนวนเท่ากับ 1, 2, 3, 4, 5 และ 6 ตามลำดับ โดยมีรายละเอียดดังตารางที่ 8

ตารางที่ 8 ผลการวิเคราะห์ความน่าเชื่อถือของผลการวัดในการศึกษา D สำหรับการออกแบบ PXSXR เมื่อ R เท่ากับ 1, 2, 3, 4, 5 และ 6 ตามลำดับ

	n_s'	3	3	3	3	3	3
	n_r'	1	2	3	4	5	6
Error Variance	$\sigma^{\wedge} \delta^2$	2.90822	1.45411	0.96941	0.72706	0.58164	0.48470
	$\sigma^{\wedge} \Delta^2$	2.96323	1.48993	0.99883	0.75328	0.60595	0.50773
G-Coefficient	$\rho^{\wedge} \delta^2$	0.32431	0.48978	0.59014	0.65752	0.70587	0.74225
	$\rho^{\wedge} \Delta^2$	0.32021	0.48370	0.58289	0.64949	0.697259	0.73327



จากตารางที่ 8 พบว่า ในการทดสอบที่มีจำนวนผู้สอบ 20 คน ผู้สอบแต่ละคนเข้ารับการทดสอบทั้งหมด 3 สถานการณ์ และในแต่ละสถานการณ์มีผู้ประเมิน 2 คน จะได้ค่าความน่าเชื่อถือของการวัดสำหรับการตัดสินใจเชิงสัมพัทธ์ ($\rho^{\wedge s^2}$) มีค่าเท่ากับ 0.48978 และความน่าเชื่อถือของการวัดสำหรับการตัดสินใจเชิงสัมบูรณ์ ($\rho^{\wedge \Delta^2}$) มีค่าเท่ากับ 0.48370 ซึ่งหากมีการจัดการทดสอบที่ต้องการความน่าเชื่อถือไม่ต่ำกว่า 0.7 ทั้งในการตัดสินใจเชิงสัมพัทธ์ และการตัดสินใจเชิงสัมบูรณ์ ควรมีการจัดการทดสอบสำหรับผู้สอบจำนวน 20 คน ที่ทำการทดสอบจำนวน 3 สถานการณ์ ที่มีผู้ประเมินจำนวน 5 คน จึงได้ความน่าเชื่อถือตามที่ต้องการ

ตารางที่ 9 แสดงผลการประเมินการปฏิบัติในแต่ละสถานการณ์

ผู้สอบ (คนที่)	คะแนนของผู้ประเมินคนที่ 1			คะแนนของผู้ประเมินคนที่ 2		
	สถานการณ์ที่ 1	สถานการณ์ที่ 2	สถานการณ์ที่ 3	สถานการณ์ที่ 1	สถานการณ์ที่ 2	สถานการณ์ที่ 3
1	21.5	21	18.5	20.5	21.5	18
2	19.5	23	18.5	21	23	19
3	22	17.5	19	21.5	18	19
4	21.5	21.5	18	20.5	21.5	18.5
5	16.5	21	21.5	16	21	21
6	19	17.5	20.5	20.5	18	20.5
7	20	20.5	17	21	21	17.5
8	14	21	21	16	21.5	21
9	21	16	23	21.5	17	22
10	19	20.5	14.5	19	20.5	15
11	14.5	20.5	19	16	21	19
12	23	16	21	22	17	21
13	21	20	14	20.5	21	16
14	17	20.5	20	17.5	21	21
15	20.5	17.5	19	20.5	19	19.5
16	21.5	21	16.5	21	21	17.5
17	18	21.5	21.5	17.5	21.5	20.5
18	19	17.5	22	18	18	21.5
19	18.5	23	19.5	21.5	22	20.5
20	18.5	21	21.5	19	21.5	21

จากตารางที่ 9 คะแนนเต็มในแต่ละสถานการณ์มีค่าเท่ากับ 28 คะแนน ซึ่งในการประเมินการปฏิบัตินี้มีการกำหนดเกณฑ์การผ่านเกณฑ์ขั้นต่ำอยู่ที่ร้อยละ 60 นั่นคือ ผู้เรียนจะต้องได้คะแนนไม่ต่ำกว่า 17 คะแนน จึงจะผ่านเกณฑ์การทดสอบ เมื่อทำการวิเคราะห์คะแนนที่ได้จากผู้ประเมินคนที่ 1 พบว่าในสถานการณ์ที่ 1 มีผู้ผ่านเกณฑ์จำนวน 17 คน และมีผู้ไม่ผ่านเกณฑ์จำนวน 3 คน ได้แก่ ผู้สอบคนที่ 5, 8 และ 11 ตามลำดับ ในสถานการณ์ที่ 2 มีผู้ผ่านเกณฑ์จำนวน 18 คน และมีผู้ไม่ผ่านเกณฑ์จำนวน 2 คน ได้แก่ ผู้สอบคนที่ 9 และ 12 และในสถานการณ์ที่ 3 มีผู้ผ่านเกณฑ์จำนวน 17 คน และมีผู้ไม่ผ่านเกณฑ์จำนวน 3 คน ได้แก่ ผู้สอบคนที่ 10, 13 และ 16 ตามลำดับ ดังนั้น หากพิจารณาผู้ผ่านการประเมินการปฏิบัติทั้งหมด 3 สถานการณ์ จากคะแนนของผู้ตรวจคนที่ 1 พบว่า มีผู้ผ่านการประเมินทุกสถานการณ์ จำนวน 12 คน คิดเป็นร้อยละ 60 ของผู้เข้ารับการประเมินการปฏิบัติทั้งหมด



เมื่อทำการวิเคราะห์คะแนนที่ได้จากผู้ประเมินคนที่ 2 พบว่าในสถานการณ์ที่ 1 มีผู้ผ่านเกณฑ์จำนวน 17 คนและมีผู้ไม่ผ่านเกณฑ์จำนวน 3 คน ได้แก่ ผู้สอบคนที่ 5, 8 และ 11 ตามลำดับ ในสถานการณ์ที่ 2 ผู้สอบผ่านเกณฑ์การทดสอบทุกคน ในขณะที่ในสถานการณ์ที่ 3 มีผู้ผ่านเกณฑ์จำนวน 18 คนและมีผู้ไม่ผ่านเกณฑ์จำนวน 2 คน ได้แก่ ผู้สอบคนที่ 10 และ 13 ดังนั้น หากพิจารณาผู้ผ่านการประเมินการปฏิบัติทั้งหมด 3 สถานการณ์ จากคะแนนของผู้ประเมินคนที่ 2 พบว่า มีผู้ผ่านการประเมินทุกสถานการณ์จำนวน 15 คน คิดเป็นร้อยละ 75 ของผู้เข้ารับการประเมินการปฏิบัติทั้งหมด

5. อภิปรายผลการวิจัย

ในการจัดการเรียนการสอนภาคปฏิบัติควรมีการทบทวนองค์ประกอบในการฝึกประสบการณ์และการวัดและประเมินผลให้มีความครอบคลุม เนื่องจากการดูแลผู้บาดเจ็บฉุกเฉินมีความซับซ้อนและต้องการความรวดเร็วที่คำนึงถึงความปลอดภัย โจทย์สถานการณ์ที่มีการฝึกทักษะที่ซ้ำกันอาจทำให้นักเรียนมีทักษะในเชิงลึกเฉพาะเรื่องนั้น ๆ แต่ถ้าหากครูไม่ได้ประเมินโดยใช้ผังข้อสอบจะทำให้การวัดและประเมินผลมีจุดอ่อนโดยขาดความเป็นตัวแทนของทักษะที่ต้องการวัดที่กำหนดไว้ตามจุดมุ่งหมายของหลักสูตร นอกจากนี้ การตรวจสอบความตรงตามเนื้อเรื่อง ควรเพิ่มผู้เชี่ยวชาญที่เป็นผู้ใช้บัณฑิตเพื่อให้ได้สถานการณ์ที่มีความสอดคล้องกับสถานการณ์จริง หรือสถานการณ์ที่ต้องการทักษะที่จำเป็นแต่อาจไม่พบบ่อย การตรวจให้คะแนนที่มีความสอดคล้องกับสถาบันการแพทย์ฉุกเฉินแห่งชาติ จะทำให้มีความเป็นมาตรฐานและสามารถทราบถึงสมรรถนะภาคปฏิบัติเมื่อนักเรียนต้องเข้ารับการทดสอบเพื่อขอรับใบอนุญาตประกอบวิชาชีพจาก สพด.

จากการตรวจสอบคุณภาพเครื่องมือผลการวิเคราะห์ข้อมูลที่ได้จะเห็นว่าในส่วนของดัชนีความยาก อำนาจจำแนก ความตรง และความเที่ยงระหว่างผู้ประเมินในภาพรวมของรายงานนี้อยู่ในระดับที่ยอมรับได้ เนื่องจากครูผู้สอนและผู้ประเมินมีความเชี่ยวชาญในด้านการดูแลผู้บาดเจ็บฉุกเฉิน สอดคล้องกับการพัฒนาแบบสอบแบบมีโครงสร้างเชิงปรนัยในการจัดการทดสอบสมรรถนะทางคลินิกของนักเรียนพยาบาลก่อนสำเร็จการศึกษาของประเทศได้หวั่น พบว่า การคัดเลือกผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อเรื่อง และควรกำหนดประสบการณ์การสอนทางคลินิกของผู้ประเมินไม่ควรต่ำกว่า 5 ปี ส่วนครูผู้จัดการทดสอบ ควรมีประสบการณ์การสอนทางคลินิกไม่ควรต่ำกว่า 2 ปี (Lee et al., 2020) ในส่วนของกระบวนการสืบเสาะความตรง (Validity Inquiry Process) ในการประเมินภาคปฏิบัติในครั้งนี้ ยังไม่ได้เชิญผู้ใช้บัณฑิตและครูพี่เลี้ยงมาร่วมประชุม จึงอาจทำให้ดัชนีความสอดคล้องระหว่างองค์ประกอบกับวัตถุประสงค์ไม่สูงมากซึ่งเป็นโอกาสในการพัฒนาการจัดการเรียนการสอนครั้งต่อไป และดำเนินการติดตามผลหลังนักเรียนสำเร็จการศึกษาเพื่อหาความตรงตามสภาพ (Concurrent Validity) ซึ่งเวลาเป็นข้อจำกัดในการดำเนินการวิจัยในครั้งนี้ นอกจากนี้ในการศึกษาครั้งต่อไปควรเพิ่มจำนวนผู้ประเมินในแต่ละสถานการณ์ให้มากขึ้น และควรตรวจสอบความลำเอียงจากผู้ประเมิน เนื่องจากเป็นปัจจัยที่ส่งผลต่อคุณภาพของเครื่องมือวัดและประเมินทักษะปฏิบัติ (กมลวรรณ ตังธานนท์, 2563)

การวัดและประเมินผลทางการศึกษาเป็นการวัดทางสังคมศาสตร์ ซึ่งในทุก ๆ ครั้งที่มีการวัดผลจะมีความคลาดเคลื่อนเกิดขึ้น สำหรับการวิจัยในครั้งนี้ได้นำเครื่องมือไปใช้ทำการวิเคราะห์ความเที่ยงด้วยทฤษฎีการสุปร้องอ้างอิงความน่าเชื่อถือของผลการวัด ซึ่งมีความแตกต่างจากการทดสอบแบบดั้งเดิม โดยการวิเคราะห์แยกส่วนความแปรปรวนของคะแนนความคลาดเคลื่อนออกเป็นความคลาดเคลื่อนอย่างเป็นระบบ อันเนื่องมาจากการเงื่อนไขต่าง ๆ ของการวัด กับความคลาดเคลื่อนสุ่มจึงทำให้ทราบแหล่งความคลาดเคลื่อนที่สำคัญทำให้สามารถควบคุมความคลาดเคลื่อนได้ตรงประเด็น อันจะนำไปสู่การตัดสินใจในการเลือกใช้เงื่อนไขหรือสถานการณ์การวัดในระดับที่ต้องการ (ศิริชัย กาญจนวาสี, 2555) พบว่า หากมีการจัดการทดสอบที่ต้องการ



ความน่าเชื่อถือไม่ต่ำกว่า 0.7 ทั้งในการตัดสินใจเชิงสัมพัทธ์ และการตัดสินใจเชิงสัมบูรณ์ ควรมีการจัดการทดสอบสำหรับผู้สอบจำนวน 20 คน ที่ทำการทดสอบจำนวน 3 สถานการณ์ ที่มีผู้ตรวจจำนวน 5 คน จึงได้ความน่าเชื่อถือตามที่ต้องการ

6. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

- 6.1 เป็นแนวทางในการวัดและประเมินผลภาคปฏิบัติของโรงเรียนนาวิกเวทีก
- 6.2 ครูผู้สอนทราบคุณภาพเครื่องมือที่ใช้ในการประเมินภาคปฏิบัติ

บรรณานุกรม

- กมลวรรณ ตั้งชนกานนท์. (2563). **การวัดและประเมินทักษะการปฏิบัติ**. กรุงเทพฯ: สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- รุ่งนภา นาคะภากร. (2540). **การพัฒนาการสอบทางคลินิกเชิงโครงสร้างแบบปรนัยในการวัดทักษะทางคลินิกสำหรับนักศึกษาพยาบาล**. (วิทยานิพนธ์ครุศาสตรมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ.
- ศิริชัย กาญจนวาสี. (2555). **ทฤษฎีการทดสอบแนวใหม่**. พิมพ์ครั้งที่ 4. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริวรรณ ชูกำเนิด และสุระพร ปู่เจริญ. (2560). The use of Objective Structured Clinical Examination in Nursing Education. **วารสารเครือข่ายวิทยาลัยพยาบาลและการสาธารณสุขภาคใต้**, 4(3), 240.
- American Heart Association. (2017). **Get with the guidelines-resuscitation overview**. Retrieved 14 December 2020, from <https://www.heart.org/en/professional/quality-improvement/get-with-the-guidelines/get-with-the-guidelines-resuscitation/get-with-the-guidelines-resuscitation-overview#.WPefZ4jyv6Q>.
- Conn, C. A., Bohan, K. J., Pieper, S. L., & Musumeci, M. (2020). Validity inquiry process: Practical guidance for examining performance assessments and building a validity argument. **Studies in Educational Evaluation**, 65. doi:10.1016/j.stueduc.2020.100843.
- Kamran Khan, Kathryn Gaunt, Sankaranarayanan Ramachandran, & Pushkar, P. (2013). The Objective Structured Clinical Examination (OSCE) AMEE Guide No.81. Part2: Organisation & Administration. **MedicalTeacher**, 35,e1447-e1463. doi:<https://doi.org/10.3109/0142159X.2013.818635>.
- Lee, K. C., Ho, C. H., Yu, C. C., & Chao, Y. F. (2020). The development of a six-station OSCE for evaluating the clinical competency of the student nurses before graduation: A validity and reliability analysis. **Nurse Educ Today**, 84, 104247. doi:10.1016/j.nedt.2019.104247



Miller, G. (1990). The assessment of clinical skill/competence/performance. **Academic Medical education**, 65, S63-S67.

Shinnick, M. A., & Woo, M. A. (2018). Validation of time to task performance assessment method in simulation: A comparative design study. **Nurse Educ Today**, 64, 108-114. doi:10.1016/j.nedt.2018.02.011.

