

การสร้างและทดสอบแบบวัดการตัดสินใจร่วมด้วยราชสีโมเดล

สุรัส ลีลาศ¹, พิสิษฐ์ เวชกามา², นื่องเล็ก คุณวรดิษฐ์³, แสง วัชรธนกิจ³

¹ฝ่ายเภสัชกรรม โรงพยาบาลสิรินธร จังหวัดอุบลราชธานี

²แผนกอายุรกรรม โรงพยาบาลสรรพสิทธิประสงค์ จังหวัดอุบลราชธานี

³กลุ่มวิชาเภสัชกรรมปฏิบัติ คณะเภสัชศาสตร์ มหาวิทยาลัยอุบลราชธานี

บทคัดย่อ

วัตถุประสงค์: เพื่อสร้างและทดสอบแบบวัดการตัดสินใจร่วมระหว่างผู้ป่วยและบุคลากรทางการแพทย์จากมุมมองของผู้ป่วย **วิธีการ:** แบบสอบถามฉบับร่างสร้างขึ้นจากขั้นตอนของกระบวนการตัดสินใจร่วม และตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ ดัชนีความสอดคล้องของข้อคำถามกับวัตถุประสงค์มีค่าอยู่ระหว่าง 0.67-1.00 แบบสอบถามฉบับสุดท้ายมีคำถาม 17 ข้อ การศึกษาเก็บข้อมูลด้วยแบบวัดที่สร้างขึ้นในผู้ป่วยจำนวน 240 รายจาก 4 แผนกในโรงพยาบาล ดังนี้ 1) แผนกผู้ป่วยนอก 2) แผนกผู้ป่วยใน 3) แผนกอุบัติเหตุ-ฉุกเฉิน และ 4) คลินิกโรคไม่ติดต่อเรื้อรัง หลังจากนั้นวิเคราะห์ข้อมูลโดยใช้ราชสีโมเดล **ผลการศึกษา:** แบบสอบถามมีความเป็นเอกมิติและมีความเป็นอิสระของข้อคำถาม การทดสอบความเที่ยงพบว่า person reliability, person separation และสัมประสิทธิ์แอลฟา มีค่า 0.86, 2.45, 0.92 ตามลำดับ ซึ่งอยู่ในเกณฑ์เป็นที่ยอมรับ item reliability มีค่า 0.86 และ item separation มีค่า 2.52 การทดสอบความเหมาะสมรายข้อ พบว่า คำถาม 15 ข้อ (ร้อยละ 88.24) มีค่าความเหมาะสมรายข้ออยู่ระหว่างเกณฑ์ 0.60-1.40 หลังจากปรับปรุงรายการเลือกคำตอบเหลือ 3 ตัวเลือกตอบ และตัดกลุ่มตัวอย่างที่ไม่เหมาะสมกับราชสีโมเดลจำนวน 28 ราย แล้วทดสอบความเที่ยงพบว่า item reliability และ item separation มีค่าเพิ่มขึ้น แต่ยังต่ำกว่าเกณฑ์ที่ยอมรับเล็กน้อย โดยพบว่าค่า person reliability, person separation และค่าสัมประสิทธิ์แอลฟามีค่า 0.85, 2.39, 0.92 ตามลำดับ ค่า item reliability มีค่า 0.88 และ item separation มีค่า 2.74 ความเหมาะสมรายข้อจำนวน 12 ข้อ (ร้อยละ 70.59) มีค่าอยู่ระหว่าง 0.60-1.27 **สรุป:** แบบวัดการตัดสินใจร่วมมีความเป็นเอกมิติและความเป็นอิสระของข้อคำถาม มีความตรงและความเที่ยงที่ดี แต่อย่างไรก็ตามในการศึกษาครั้งต่อไป ควรทดสอบแบบวัดในผู้ป่วยจำนวนมากขึ้น

คำสำคัญ: การตัดสินใจร่วม แบบสอบถาม ราชสีโมเดล

รับต้นฉบับ: 18 ก.ค. 2561, รับลงตีพิมพ์: 24 ก.ย. 2561

ผู้ประสานงานบทความ: แสง วัชรธนกิจ กลุ่มวิชาเภสัชกรรมปฏิบัติ คณะเภสัชศาสตร์ มหาวิทยาลัยอุบลราชธานี อ่าเภอวารินชำราบ จังหวัดอุบลราชธานี 34190 E-mail: sawaeng.w@ubu.ac.th

Development and Validation of the Shared Decision Making Questionnaire Using Rasch Model

Suwaros Leeras¹, Phisitt Vejakama², Nonglek Khunawaradisai³, Sawaeng Watcharathanakij³

¹Department of Pharmacy, Sirinthorn Hospital, Ubon Ratchathani

²Department of Internal Medicine, Sunpasithiprasong Hospital, Ubon Ratachathani

³Department of Pharmacy Practice, Faculty of Pharmaceutical Sciences, Ubon Ratchathani University

Abstract

Objective: To develop and validate the questionnaire measuring shared decision making (SDM) between patients and healthcare professional from patient's perspective. **Methods:** The first draft of the SDM questionnaire was developed based on steps of SDM process and its face validity was assessed using experts' opinion. Index of item objective congruence ranged 0.67-1.00. The final version with 17 items was administered in 240 patients from 4 medical clinics including 1) outpatient department 2) inpatient department 3) accident and emergency department and 4) chronic non-communicable disease clinic. Subsequently, data were analyzed using the Rasch model. **Results:** The questionnaire showed unidimensionality and local independence. The reliability tests unveiled that person reliability, person separation index, and alpha reliability were 0.86, 2.45 and 0.92 respectively, which were regarded as acceptable. Item reliability was 0.86 and item separation was 2.52. Test of item fit statistics revealed that 15 items (88.24%) had item fit statistics ranged between 0.60-1.40. After combining 5 response categories to 3 categories and excluding 28 mis-fitted persons (11.67%) from the Rasch model, the reliability tests unveiled item reliability and item separation increased, but were marginally below the acceptable cut-off value. Person reliability, person separation index, and alpha reliability were 0.85, 2.39 and 0.92 respectively. Item reliability was 0.88 and item separation was 2.74. Item fit statistics of 12 items (70.59%) ranged between 0.60-1.27. **Conclusion:** The SDM questionnaire has unidimensionality and local independence with high validity and reliability. However, it should be further tested in a larger number of patients.

Keywords: shared decision making, questionnaire, Rasch model

บทนำ

การตัดสินใจร่วม (shared decision making) เป็นวิธีสร้างความสัมพันธ์ที่ดีตั้งแต่แรก และเป็นแนวคิดที่พัฒนาขึ้นเพื่อช่วยให้การรักษาเป็นไปอย่างราบรื่น กระบวนการตัดสินใจร่วมเป็นการปฏิบัติงานร่วมกันระหว่างผู้ป่วยและแพทย์ในการเลือกการทดสอบและการรักษา โดยอาศัยหลักฐานทางคลินิกและการยินยอมในการรักษาของผู้ป่วย การตัดสินใจร่วมเป็นการให้ข้อมูลหลักฐานทางคลินิกในแง่ทางเลือกของการรักษา ผลลัพธ์และความไม่แน่นอนในการรักษา ร่วมกับการให้คำปรึกษาเพื่อสนับสนุนการตัดสินใจ และการบันทึกการให้คำปรึกษาของผู้ป่วย (1) ลักษณะของการตัดสินใจร่วมเกี่ยวข้องกับผู้ที่มีส่วนร่วมอย่างน้อย 2 ส่วน ได้แก่ ผู้เชี่ยวชาญทางด้านสุขภาพซึ่งส่วนใหญ่หมายถึงแพทย์ และผู้ป่วย บ่อยครั้งมีผู้มีส่วนร่วมในการตัดสินใจที่หลากหลาย ซึ่งได้แก่ผู้ป่วย และบุคลากรทางการแพทย์สาขาต่าง ๆ (2) เนื่องจากการปฏิบัติงานร่วมกันของบุคลากรทางการแพทย์ช่วยให้ผลลัพธ์ในการดูแลผู้ป่วยดีขึ้น เช่น ลดอาการไม่พึงประสงค์จากยาที่สามารถป้องกันได้ ทำให้ผู้ป่วยได้รับยาในขนาดที่เหมาะสม ลดอัตราป่วยและอัตราตาย (3) เพิ่มศักยภาพของกระบวนการนำผู้ป่วยเข้ามามีส่วนร่วมในการตัดสินใจ และเพิ่มคุณภาพในการตัดสินใจ (4) ดังนั้นในงานวิจัยนี้ การตัดสินใจร่วมหมายถึง การมีส่วนร่วมในการตัดสินใจของผู้ป่วยและบุคลากรทางการแพทย์

เครื่องมือวัดการตัดสินใจร่วมได้รับการสร้างและพัฒนาในหลายประเทศและหลากหลายภาษา (5) เช่น OPTION, (6) perceived involvement in care scale (PICS) (7) และ 9-item shared decision-making questionnaire (SDM-Q-9) (8) เป็นต้น หากพิจารณาจากโครงสร้างของเครื่องมือวัดการตัดสินใจร่วมเหล่านี้ จะสามารถจำแนกประเภทออกเป็น 3 ประเภทหลักดังนี้ 1) เครื่องมือวัดสิ่งที่เกิดขึ้นก่อนการตัดสินใจ (decision antecedents) 2) เครื่องมือวัดสิ่งที่เกิดระหว่างการตัดสินใจ (decision process) และ 3) เครื่องมือวัดผลลัพธ์ของการตัดสินใจ (decision outcomes) (5)

อย่างไรก็ตามลักษณะของการประเมินการตัดสินใจร่วมที่ดียังเป็นข้อถกเถียงในประเด็นของจุดมุ่งหมายที่จะประเมิน เดิมมีความเห็นว่ามีคามจำเป็นที่

ต้องพิจารณาทั้งกระบวนการและผลลัพธ์ของการตัดสินใจ แต่ผลลัพธ์ของกระบวนการการตัดสินใจร่วมไม่ได้เป็นตัวบ่งชี้คุณภาพในการวัดที่เหมาะสม เนื่องจากผลจากการรักษาอาจจะขึ้นอยู่กับหลายภายนอกอื่น ๆ อีกหลายปัจจัย (9) นอกจากนี้การตัดสินใจควรประเมินเกี่ยวกับวิธีในการตัดสินใจของแต่ละบุคคล และการวัดประสิทธิภาพของการตัดสินใจร่วมต้องแยกผลของการตัดสินใจออกจากกระบวนการของการตัดสินใจ (10) สำหรับในประเทศไทยยังไม่พบการสร้างและพัฒนาเครื่องมือวัดการตัดสินใจร่วม ดังนั้นการวิจัยนี้มีวัตถุประสงค์เพื่อสร้างเครื่องมือประเมินกระบวนการการตัดสินใจร่วมพร้อมตรวจสอบคุณสมบัติการวัดทางจิตวิทยาด้านความตรง (validity) และความเที่ยง (reliability) ของเครื่องมือวัดที่สร้างขึ้น

วิธีการวิจัย

ตัวอย่าง

ตัวอย่าง คือ ผู้ป่วย 240 รายที่มารับบริการที่โรงพยาบาลสิรินธร อำเภอสรินธร จังหวัดอุบลราชธานี ที่มีอายุ 18 ปีขึ้นไป สามารถอ่านและเข้าใจภาษาไทย และเป็นผู้ที่ยินดีเข้าร่วมการศึกษาวิจัย จำนวนตัวอย่างที่แนะนำสำหรับการทดสอบแบบสอบถามที่มีหลายตัวเลือกตอบโดยใช้ราสช์โมเดล (Rasch model) คือ ควรมีอย่างน้อย 10 คนต่อ 1 รายการคำตอบ (11)

เครื่องมือ

เครื่องมือที่ใช้ในการศึกษาค้างนี้ คือ แบบสอบถามซึ่งแบ่งออกเป็น 2 ส่วน ได้แก่ ส่วนที่ 1 เป็นคำถาม 6 ข้อที่วัดข้อมูลส่วนบุคคล และส่วนที่ 2 เป็นคำถาม 17 ข้อที่วัดการมีส่วนร่วมในการตัดสินใจของผู้ป่วยและบุคลากรทางการแพทย์ (shared decision making: SDM) ดังแสดงในตารางที่ 3 แบบวัดเป็นมาตราส่วนประมาณค่า 5 ระดับ คำถามสร้างโดยอ้างอิงจากแนวคิดของกระบวนการมีส่วนร่วมในการตัดสินใจที่ปรับปรุงจากของ Makoul และ Clayman (12) ซึ่งประกอบด้วย 7 ด้านตามองค์ประกอบ คือ 1. การระบุปัญหาหรือประเด็นที่ต้องมีการตัดสินใจ (ข้อ 1 ในตารางที่ 3) 2. การระบุทางเลือก หรือ ความเสี่ยงที่มี (ข้อ 2) 3. การพิจารณาข้อดี และข้อเสีย (ข้อ 3) 4. การค้นหาความต้องการหรือความคาดหวังของผู้ป่วย (ข้อ 4-9) 5. การให้ข้อมูลเกี่ยวกับทางเลือกที่ผู้ป่วยพึงพอใจ (ข้อ 10-12) 6. การเจรจาต่อรองระหว่างผู้ป่วยและบุคลากรการแพทย์

(ข้อ 13-14) และ 7. การตัดสินใจเลือกทางเลือกของการรักษาร่วมกัน (ข้อ 15-17)

การตรวจสอบคุณภาพของเครื่องมือ

การวิจัยตรวจสอบความตรงเชิงเนื้อหาโดยใช้ผู้เชี่ยวชาญจำนวน 3 ท่าน ได้แก่ แพทย์ พยาบาล และเภสัชกรที่มีประสบการณ์การทำงาน 10 ปีขึ้นไป พิจารณาความสอดคล้องของคำถามที่ใช้วัดระดับการตัดสินใจร่วมระหว่างผู้ป่วยและบุคลากรทางการแพทย์ว่าตรงตามคุณลักษณะหรือองค์ประกอบที่กำหนดหรือไม่ การศึกษาคำนวณค่าความตรงเชิงเนื้อหาจากค่าดัชนีความสอดคล้องหรือ item-objective congruence index (IOC) ค่า IOC ที่ ≥ 0.50 (13) ถือว่า คำถามสอดคล้องกับคุณลักษณะหรือองค์ประกอบที่ต้องการวัด

การเก็บรวบรวมข้อมูล

การศึกษานี้เก็บข้อมูลโดยให้ผู้ป่วยรายงานข้อมูลด้วยตัวเองหลังจากมารับบริการตรวจรักษา การเลือกตัวอย่างใช้วิธีการสุ่มตัวอย่างแบบเป็นระบบ คือ สุ่มตัวอย่างทุกลำดับที่ 5 ของผู้ป่วยที่เข้ารับบริการจาก 4 แผนกในโรงพยาบาล ดังนี้ แผนกผู้ป่วยนอก แผนกผู้ป่วยใน แผนกอุบัติเหตุ-ฉุกเฉิน และคลินิกโรคไม่ติดต่อเรื้อรัง ทั้งนี้เพื่อให้เป็นตัวแทนของผู้ป่วยที่มารับบริการในโรงพยาบาล เนื่องจากเครื่องมือวัดการตัดสินใจร่วมที่สร้างขึ้นมีวัตถุประสงค์เพื่อใช้ในกลุ่มผู้ป่วยทั่วไป และเพื่อให้ครอบคลุมบริบทของการดูแลรักษาที่อาจมีความแตกต่างกันในเรื่องการตัดสินใจร่วม

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลส่วนบุคคลของกลุ่มตัวอย่างใช้สถิติเชิงพรรณนา การศึกษาใช้โปรแกรม WINSTEPS version 3.9.1.0 วิเคราะห์ความตรงของแบบสอบถามด้วยวิธี Rasch model โดยมีรายละเอียดดังนี้ เส้นโค้งการเลือกรายการคำตอบ (category response curves or CRC) เป็นแผนภาพแสดงความน่าจะเป็นในการเลือกรายการคำตอบของตัวอย่างที่มีระดับคุณลักษณะแฝงแตกต่างกัน นั่นคือ CRC แสดงความน่าจะเป็นในการเลือกรายการคำตอบที่สัมพันธ์กับการตัดสินใจร่วมระหว่างผู้ป่วยและบุคลากรทางการแพทย์ที่แตกต่างกัน

fit statistics แบ่งเป็น person fit เป็นการตรวจจบริูปแบบของคะแนน (item-score) ที่ไม่เป็นไปตาม Rasch model ของตัวอย่าง (14) และ item fit ซึ่งเป็นการทดสอบความเหมาะสมของคำถามรายข้อโดย Rasch model ซึ่งเป็น

รูปแบบการทดสอบแบบลอจิสต์ หรือที่เรียกกันว่า one parameter logistic model การทดสอบนี้ได้ค่าสถิติที่ระบุว่าข้อมูลรายข้อนั้นสอดคล้องกับโมเดลที่กำหนดไว้หรือไม่ ค่าสถิติที่ใช้คือ OUTFIT MNSQ (outfit mean square) และ INFIT MNSQ (infit mean square) ทั้งสองค่านี้อยู่บนฐานแนวคิดของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนมาตรฐานระหว่างผลการตอบข้อคำถามและโมเดลการวัดนั้น (15) งานวิจัยนี้ใช้ค่า OUTFIT MNSQ และ INFIT MNSQ ระหว่าง 0.60-1.40 (16-17) เป็นเกณฑ์ในการตัดสินใจคุณภาพของข้อคำถาม

การทดสอบความเที่ยง เป็นการทดสอบความคงเส้นคงวาของการวัด เกณฑ์ที่ยอมรับคือ ค่า person reliability > 0.80 (เทียบเท่ากับสัมประสิทธิ์แอลฟา) item reliability > 0.90 , person separation > 2 และ item separation > 3 (18)

การทดสอบความเป็นเอกมิติ (unidimensionality) ทำโดยการทดสอบ principal components analysis (PCA) ซึ่งเป็นการวิเคราะห์เศษเหลือ (residuals) หลังจากการประเมินมิติหลัก (19) โดยเกณฑ์ในการพิจารณาลักษณะความเป็นเอกมิติมี 3 ข้อ ได้แก่ 1) ค่าความแปรปรวนที่อธิบายโดยมิติของการวัด (variance explained by the measurement dimension) ต้องไม่น้อยกว่าร้อยละ 40 (20) 2) ค่าความแปรปรวนที่อธิบายได้โดยองค์ประกอบแรกของเศษเหลือ (variance explained by first principal component of the residuals) ต้องไม่เกินร้อยละ 15 และ 3) อัตราส่วนของความแปรปรวนในข้อ 1 และข้อ 2 ต้อง $\geq 3:1$ (21)

ความเป็นอิสระของข้อคำถาม (local independence) หมายความว่า การตอบข้อสอบข้อใดข้อหนึ่งถูก จะไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ (การเลือกตอบข้อคำถามใด ๆ จะไม่มีผลต่อการเลือกตอบข้อคำถามอื่น ๆ) ซึ่งสามารถพิจารณาได้จากค่า intra-item correlations ของ standardized residuals ของแบบสอบถาม ซึ่งควรมีค่า < 0.70 (22)

แผนภาพบุคคลและข้อคำถาม (item-person map) แสดงความสัมพันธ์ของระดับความสามารถของผู้ทดสอบ (person ability) หรือคุณลักษณะแฝงภายในบุคคล (ระดับของการตัดสินใจร่วมในงานวิจัยนี้) และความยาก-ง่ายของข้อคำถาม (item hierarchies)

การทำหน้าที่ต่างกันของข้อคำถาม (differential item functioning or DIF) หมายถึงคำถามทำให้ผลการตอบ

คำถามระหว่างกลุ่มของกลุ่มตัวอย่างที่มีความสามารถหรือคุณลักษณะแฝงภายในบุคคลระดับเดียวกัน มีโอกาสในการตอบข้อคำถามได้แตกต่างกัน เช่น กรณีที่กลุ่มตัวอย่างที่มีความสามารถหรือคุณลักษณะแฝงภายในบุคคลในระดับเดียวกัน แต่อยู่กลุ่มต่างกัน มีความน่าจะเป็นในการเลือกรายการคำตอบแตกต่างกัน ถือว่าข้อคำถามนั้นมีการทำหน้าที่ต่างกัน การวิเคราะห์ DIF ช่วยในการตรวจสอบว่าข้อคำถามมีความแตกต่างกันระหว่างกลุ่มของกลุ่มตัวอย่างหรือไม่ เกณฑ์ที่ใช้ในการตัดสิน คือ ถ้ามีความแตกต่างของผลตอบสนองต่อข้อคำถามระหว่างกลุ่มของกลุ่มตัวอย่าง ≥ 0.50 logit ถือว่าข้อคำถามนั้นมีการทำหน้าที่ต่างกัน (17)

ผลการวิจัย

ลักษณะของตัวอย่าง

การศึกษานี้มีตัวอย่างตอบแบบสอบถาม 240 ราย การวิจัยตัดตัวอย่าง 28 รายออกจากการวิเคราะห์เนื่องจากรูปแบบของคะแนนไม่เหมาะสม (person misfit) กับรหัสโมเดล ตารางที่ 1 แสดงคุณลักษณะของตัวอย่างที่เหลือ ตัวอย่าง 212 รายที่คงเหลือส่วนใหญ่เป็นเพศหญิง (ร้อยละ 59.90) มีอายุระหว่าง 51-60 ปี (ร้อยละ 30.20) รองลงมา มีอายุระหว่าง 41-50 ปี (ร้อยละ 26.40) จบการศึกษาระดับประถมศึกษา (ร้อยละ 68.90) ประกอบอาชีพเกษตรกรกรรม (ร้อยละ 39.30) รองมาเป็นแม่บ้าน (ร้อยละ 25.70) รายได้

ตารางที่ 1. ลักษณะของกลุ่มตัวอย่าง (n=212 ราย)

ลักษณะทั่วไป	จำนวน (ร้อยละ)
เพศ	
หญิง	127 (59.90)
ชาย	85 (40.10)
อายุ (ปี)	
18-30 ปี	15 (7.10)
31-40 ปี	35 (16.50)
41-50 ปี	56 (26.40)
51-60 ปี	64 (30.20)
60 ปีขึ้นไป	42 (19.80)
การศึกษา	
ประถมศึกษา	146 (68.90)
มัธยมศึกษา	49 (23.10)
อนุปริญญา	5 (2.40)
ปริญญาตรี	12 (5.60)
อาชีพ	
เกษตรกรกรรม	84 (39.30)
แม่บ้าน	55 (25.70)
รับจ้าง/ลูกจ้าง	52 (24.30)
ค้าขาย/ธุรกิจส่วนตัว	17 (7.90)
รับราชการ/รัฐวิสาหกิจ	6 (2.80)

ตารางที่ 1. ลักษณะของกลุ่มตัวอย่าง (n=212 ราย) (ต่อ)

ลักษณะทั่วไป	จำนวน (ร้อยละ)
รายได้	
ต่ำกว่า 10,000 บาท	158 (74.50)
10,001-30,000 บาท	46 (21.70)
30,001-40,000 บาท	5 (2.40)
มากกว่า 40,000 บาท	3 (1.40)
สาเหตุที่มาพบแพทย์	
โรคเรื้อรัง เช่น เบาหวาน ความดัน	100 (47.20)
ปวดกล้ามเนื้อ/ข้อ	25 (11.80)
อุบัติเหตุ	19 (9.00)
ปวดท้อง	14 (6.60)
ไข้หวัด	11 (5.20)
ปวดหัว	10 (4.70)
โรคผิวหนัง	8 (3.80)
อื่น ๆ	25 (11.70)
แผนกที่เข้ารับบริการ	
ผู้ป่วยนอก	62 (29.25)
คลินิกพิเศษ	62 (29.25)
อุบัติเหตุ/ฉุกเฉิน	43 (20.28)
ผู้ป่วยใน	45 (21.22)

ต่ำกว่า 10,000 บาท (ร้อยละ 74.50) สาเหตุที่มาพบแพทย์ คือ ป่วยด้วยโรคเรื้อรัง เช่น เบาหวาน ความดันโลหิตสูง (ร้อยละ 47.20) ผู้ป่วยเข้ารับบริการในแผนกผู้ป่วยนอก (ร้อยละ 29.25) และแผนกคลินิกพิเศษ (ร้อยละ 29.25)

โครงสร้างของการเลือกรายการคำตอบ

ตารางที่ 2 แสดงผลการวิเคราะห์โครงสร้างของการเลือกรายการคำตอบ (rating scale response structures) ผลการวิเคราะห์แสดงความถี่การเลือกตัวเลือกที่ไม่เท่าเทียม สะท้อนให้เห็นถึงความไม่สม่ำเสมอของการกระจายการเลือกคำตอบ แต่ไม่พบตัวเลือกใดที่มีความถี่เป็นศูนย์ ตัวเลือกที่มีการตอบมากที่สุดคือ “เห็นด้วย” (ร้อยละ 72) และรายการคำตอบ “ไม่เห็นด้วย”, “ไม่แน่ใจ”, “เห็น

ด้วยอย่างยิ่ง” และ “ไม่เห็นด้วยอย่างยิ่ง” มีผู้ตอบร้อยละ 13, 9, 6 และ <1 ตามลำดับ รายการคำตอบ “ไม่เห็นด้วยอย่างยิ่ง” มีผู้เลือกน้อยที่สุดจากรายการคำตอบ 5 ระดับ โดยพบเฉพาะคำถามข้อที่ 11 (ท่านทราบทางเลือกที่ดีที่สุดในการรักษาของท่าน)

เมื่อปรับรายการตัวเลือกให้เหลือ 3 รายการพบว่าตัวเลือกที่มีการตอบมากที่สุดคือ “เห็นด้วย”, “ไม่แน่ใจ” “ไม่เห็นด้วย” และ “ไม่เห็นด้วยอย่างยิ่ง” และ “เห็นด้วยอย่างยิ่ง” คิดเป็นร้อยละ 72, 22 และ 6 ตามลำดับ ดังตารางที่ 2

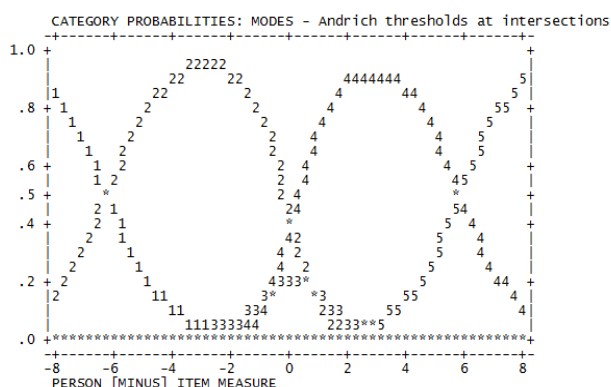
เส้นโค้งการเลือกรายการคำตอบ

จากเส้นโค้งการเลือกรายการคำตอบ (CRC) จำนวน 5 รายการคำตอบของตัวอย่าง พบว่าความน่าจะเป็น

ตารางที่ 2. การกระจายของการเลือกรายการคำตอบ

	5 รายการคำตอบ					3 รายการคำตอบ		
	ไม่เห็นด้วยอย่างยิ่ง (1)	ไม่เห็นด้วย (2)	ไม่แน่ใจ (3)	เห็นด้วย (4)	เห็นด้วยอย่างยิ่ง (5)	ไม่เห็นด้วย-ไม่แน่ใจ (1)	เห็นด้วย (2)	เห็นด้วยอย่างยิ่ง (3)
ความถี่	2	528	362	2,955	233	892	2,955	233
ร้อยละ	<1	13	9	72	6	22	72	6

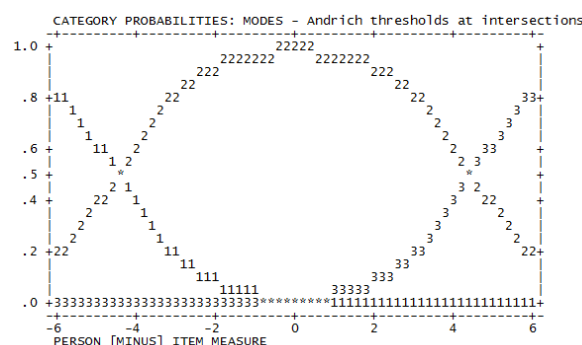
เป็นในการเลือกรายการคำตอบลำดับที่ 3 (ไม่แน่ใจ) มีค่าต่ำกว่ารายการคำตอบอื่น ๆ โดยโอกาสสูงสุดเพียงร้อยละ 20 และมีความน่าจะเป็นในการเลือกรายการคำตอบลำดับที่ 3 ซ้อนทับกับตัวเลือกที่ 2 และ 4 เป็นส่วนใหญ่ ดังแสดงในรูปที่ 1 ดังนั้นจึงปรับปรุงรายการคำตอบโดยรวมรายการคำตอบลำดับที่ 1, 2 และ 3 เข้าด้วยกัน และปรับเกณฑ์ให้คะแนน ดังนี้ “ไม่แน่ใจ, ไม่เห็นด้วย, ไม่เห็นด้วยอย่างยิ่ง” =1 คะแนน “เห็นด้วย”=2 คะแนน และ “เห็นด้วยอย่างยิ่ง”=3 คะแนน



รูปที่ 1. เส้นโค้งการเลือกรายการคำตอบของคำถามที่มีตัวเลือกคำตอบ 5 ระดับ

เส้นโค้งการเลือกรายการคำตอบ เมื่อปรับให้เหลือ 3 รายการคำตอบ พบว่า ตัวอย่างที่มีระดับการมีส่วนร่วมในการตัดสินใจในระดับต่ำที่ -6.00 logit มีโอกาสตอบว่า “ไม่แน่ใจ, ไม่เห็นด้วย, ไม่เห็นด้วยอย่างยิ่ง” มากที่สุด ประมาณร้อยละ 80 ทั้งนี้พบโอกาสที่จะเลือกตอบ “เห็นด้วย” ประมาณร้อยละ 20 และไม่พบโอกาสที่จะเลือกตอบว่า “เห็นด้วยอย่างยิ่ง” สำหรับกลุ่มตัวอย่างที่มีระดับการมีส่วนร่วมในการตัดสินใจในระดับสูง ที่ระดับคะแนน 6.00 logit พบโอกาสที่จะเลือกตอบว่า “เห็นด้วยอย่างยิ่ง” มากที่สุด ประมาณร้อยละ 80 พบโอกาสที่จะเลือกตอบ “เห็นด้วย”

ประมาณร้อยละ 20 และไม่พบโอกาสที่จะเลือกตอบ “ไม่แน่ใจ, ไม่เห็นด้วย, ไม่เห็นด้วยอย่างยิ่ง” ดังแสดงในรูปที่ 2



รูปที่ 2. เส้นโค้งการเลือกรายการคำตอบของข้อคำถามที่มีตัวเลือกคำตอบ 3 ระดับ

ความเหมาะสมของบุคคล

ความเหมาะสมของบุคคล (person fit) เป็นการตรวจจึบรูปแบบของคะแนน (item-score) ที่ไม่เป็นไปตามรหัสโมเดลของกลุ่มตัวอย่าง ซึ่งพบว่ามีความสอดคล้องในระดับบุคคลที่ไม่เหมาะสมกับรหัสโมเดลจำนวน 28 รายการ จากกลุ่มตัวอย่างจำนวน 240 ราย (ร้อยละ 11.67) โดยมีค่า outfit mean-squares อยู่ระหว่าง 1.56-5.03 และถูกตัดออกจากการวิจัยครั้งนี้

ความเหมาะสมของข้อคำถาม

จากตารางที่ 3 การทดสอบความเหมาะสมรายข้อ (item fit) พบว่า ก่อนการตัดตัวอย่างที่ไม่เหมาะสมกับรหัสโมเดล คำถาม 15 ข้อ มีค่า infit mean-squares และ outfit mean-squares อยู่ในเกณฑ์เหมาะสม (0.60-1.40) ยกเว้นคำถามข้อที่ 8 มีค่า outfit mean-squares = 1.60 และข้อที่ 16 มีค่า infit mean-squares = 0.55 และ Outfit mean-squares = 0.44

ค่า item fit หลังจากตัดตัวอย่างที่ไม่เหมาะสมกับรหัสโมเดลจำนวน 28 รายการออก พบว่า มีคำถามที่มีค่า infit

mean-squares และ outfit mean-squares ที่อยู่ในเกณฑ์เหมาะสม จำนวน 12 ข้อ (ร้อยละ 70.59) และมีคำถามที่ไม่เหมาะสมกับบรรทัดโมเดลจำนวน 5 ข้อ ได้แก่ ข้อคำถามลำดับที่ 1, 8, 9, 10 และ 16 ซึ่งมีค่า outfit mean-squares = 1.41, 1.64, 0.31, 0.53 และ 0.44 ตามลำดับ โดยมีค่า infit mean-squares และ outfit mean-squares อยู่ระหว่าง 0.31-1.63 และยังคงไว้ในแบบสอบถาม

ความเป็นเอกมิติ

การทดสอบความเป็นเอกมิติซึ่งในงานวิจัยครั้งนี้หมายถึง การวัดเจตคติของผู้ป่วยเรื่องการมีส่วนร่วมในการ

ตัดสินใจรักษาของข้อมูลว่ามีเพียงมิติเดียวหรือไม่ โดยใช้ข้อมูลก่อนตัดตัวอย่างที่ไม่เหมาะสมกับบรรทัดโมเดลมาทดสอบด้วย PCA ตารางที่ 4 แสดงผลการทดสอบความเป็นเอกมิติ ความแปรปรวนที่อธิบายได้โดยมิติของการวัดคือ ร้อยละ 46.90 ค่าความแปรปรวนของเศษเหลือที่อธิบายได้โดยองค์ประกอบแรก มีค่าเท่ากับร้อยละ 8.40 และอัตราส่วนของความแปรปรวนทั้งสอง เท่ากับ 5.58 : 1

ผลการทดสอบหลังจากตัดตัวอย่างที่ไม่เหมาะสมกับบรรทัดโมเดล 28 รายการออก มิติของการวัดสามารถอธิบายค่าความแปรปรวนได้ร้อยละ 43.30 ค่าความแปรปรวนของ

ตารางที่ 3. ความเหมาะสมรายข้อ item fit (infit mean-squares และค่า outfit mean-squares)

คำถาม	MNSQ ¹ (n=240)		MNSQ ¹ (n=212)	
	5 รายการคำตอบ		3 รายการคำตอบ	
	infit	outfit	infit	outfit
1. บุคลากรทางการแพทย์ได้อธิบายปัญหาสุขภาพ ที่ต้องมีการตัดสินใจเลือกวิธีการรักษาให้ท่านทราบ	1.18	1.09	1.63*	1.41*
2. บุคลากรทางการแพทย์ได้อธิบายทางเลือกต่างๆ ในการรักษาให้ท่านทราบ	1.03	0.68	0.92	0.61
3. บุคลากรทางการแพทย์ได้อธิบาย ข้อดีและข้อเสีย ของทางเลือกต่างๆ ในการรักษาให้ท่านทราบ	0.94	0.77	0.92	0.60
4. บุคลากรทางการแพทย์รับทราบความคาดหวังเกี่ยวกับวิธีการรักษาที่ท่านต้องการ	1.06	1.04	1.24	1.12
5. บุคลากรทางการแพทย์รับทราบความกังวลเกี่ยวกับปัญหาด้านสุขภาพของท่าน	0.99	0.72	0.90	0.62
6. บุคลากรทางการแพทย์ได้สอบถามถึงทางเลือกในการรักษาที่ท่านพึงพอใจ	0.86	0.92	0.96	0.60
7. บุคลากรทางการแพทย์มีเวลาเพียงพอให้ท่านซักถามปัญหาด้านสุขภาพของท่าน	1.14	1.14	0.99	1.18
8. ท่านมีโอกาสในการซักถามปัญหาด้านสุขภาพของท่าน	1.29	1.60*	1.12	1.64*
9. ท่านมีเวลาเพียงพอสำหรับซักถามปัญหาด้านสุขภาพของท่าน	0.85	0.74	0.55*	0.31*
10. ท่านมีความเข้าใจข้อมูลเกี่ยวกับปัญหาสุขภาพของท่าน	0.99	0.74	0.81	0.53*
11. ท่านทราบทางเลือกที่ดีที่สุดในการรักษาของท่าน	1.21	0.90	1.10	0.80
12. ท่านทราบถึงข้อดีและข้อเสียของทางเลือกในการรักษาของท่าน	0.87	0.76	0.97	0.74
13. บุคลากรทางการแพทย์และท่านร่วมกันพิจารณาทางเลือกต่างๆ ในการรักษา	0.92	0.85	0.82	0.69
14. ท่านสามารถหารือกับบุคลากรทางการแพทย์กรณีที่มีความต้องการทางเลือกในการรักษาที่แตกต่างกัน	0.77	0.85	1.27	1.03
15. บุคลากรทางการแพทย์นำเอาความคิดเห็นของท่านมาใช้เพื่อเลือกวิธีการรักษา	1.39	1.32	1.12	0.88
16. ในระหว่างการให้คำปรึกษา ท่านรู้สึกมีส่วนร่วมในการตัดสินใจในการรักษา	0.55*	0.44*	0.71	0.44*
17. บุคลากรทางการแพทย์ และท่าน ร่วมกันเลือกวิธีการรักษาด้วยกัน	0.80	0.64	0.89	0.59

1: MNSQ=mean square

*: คำถามที่มีค่า infit mean-squares และ outfit mean-squares ไม่อยู่ในเกณฑ์เหมาะสม

ตารางที่ 4. การทดสอบความเป็นเอกมิติ (unidimensionality)

ค่าสถิติ	เกณฑ์ (ร้อยละ)	MNSQ ¹ (240 ราย)	MNSQ ¹ (212 ราย)
1. ค่าความแปรปรวนที่อธิบายโดยมิติของการวัด	≥ 40	46.9	43.30
2. ค่าความแปรปรวนของเศษเหลือที่สามารถอธิบายได้โดยองค์ประกอบแรก	< 15	8.4	9.20
3. อัตราส่วนของความแปรปรวนในข้อ 1 และ 2	3 : 1	5.58 : 1	4.60 : 1

1: MNSQ=mean square

เศษเหลือที่สามารถอธิบายได้โดยองค์ประกอบแรก เท่ากับ ร้อยละ 9.20 และอัตราส่วนของความแปรปรวนทั้งสอง คือ 4.60 : 1 ซึ่งให้เห็นว่า แบบวัด SDM มีความเป็นเอกมิติ

ความเป็นอิสระของข้อคำถาม

ความเป็นอิสระของข้อคำถามพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากตาราง largest residual correlations for items ที่แสดงค่า intra-item correlations ของ standardized residuals ของคำถาม หลังจากตัดตัวอย่างที่ไม่เหมาะสมกับราสซิมเดล มีค่าสหสัมพันธ์ดังกล่าวมีค่าระหว่าง -0.23 ถึง 0.53 ซึ่งแสดงให้เห็นว่าคำถามมีอิสระต่อกัน เนื่องจากมีค่าสหสัมพันธ์ < 0.70

ความเที่ยง

หลังจากตัดตัวอย่างที่ไม่เหมาะสมกับราสซิมเดล พบว่า person reliability มีค่า 0.85 และมี person separation มีค่า 2.39 และค่าสัมประสิทธิ์แอลฟา 0.92 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ (ตารางที่ 5) ส่วนค่า item reliability มีค่า 0.88 และมีค่า item separation มีค่า 2.74 ซึ่งต่ำกว่าเกณฑ์ที่ยอมรับได้เล็กน้อย ทั้งนี้อาจเป็นเพราะตัวอย่างอาจมีขนาดไม่มากพอที่จะยืนยันความแม่นยำของตำแหน่งของคำถามต่อตัวแปรแฝงที่ต้องการวัด (18)

แผนภาพบุคคลและคำถาม

แผนภาพบุคคลและคำถาม (item-person map) แสดงอยู่ในรูปที่ 3 โดยด้านซ้าย (person) แสดงระดับความสามารถของตัวอย่าง (person ability) หรือความเห็นต่อ SDM ส่วนด้านขวา (item) แสดงลำดับความยาก-ง่ายของข้อคำถาม (item hierarchies) ทั้งนี้โปรแกรม winsteps จัดเรียงให้ ด้านล่างซ้ายของภาพแสดงถึงจำนวนตัวอย่างที่ไม่เห็นด้วยต่อ SDM (กลุ่มที่ตอบไม่เห็นด้วยหรือไม่เห็นด้วยอย่างยิ่ง) และด้านล่างขวาแสดงลำดับข้อคำถามที่มีตัวอย่างเลือกตอบเห็นด้วยหรือเห็นด้วยอย่างยิ่งค่อนข้างบ่อย ส่วนด้านบนซ้ายของรูปแสดงถึงจำนวนตัวอย่างที่เห็นด้วยต่อ SDM (ตัวอย่างตอบเห็นด้วยหรือเห็นด้วยอย่างยิ่ง) และด้านบนขวาแสดงลำดับข้อคำถามที่ตัวอย่างเลือกตอบไม่เห็นด้วยหรือไม่เห็นด้วยอย่างยิ่งค่อนข้างบ่อย ด้านขวาใช้เครื่องหมาย "#" แทนจำนวนตัวอย่าง จำนวน 5 ราย และเครื่องหมาย "." แทนกลุ่มตัวอย่าง จำนวน 1-4 ราย และด้านซ้ายเป็นลำดับข้อคำถาม ค่า "M" หมายถึงค่าเฉลี่ยของตัวอย่างและข้อคำถาม ค่า "S" หมายถึง หนึ่งเท่าของส่วนเบี่ยงเบนมาตรฐานจากค่าเฉลี่ยของตัวอย่าง และข้อคำถาม ค่า "T" หมายถึง สองเท่าของส่วนเบี่ยงเบนมาตรฐานจากค่าเฉลี่ยของตัวอย่างและข้อคำถาม

ตารางที่ 5. person reliability, item reliability และ separation statistics

การทดสอบความเที่ยง	เกณฑ์	240 ราย, 17 ข้อคำถาม, 5 รายการคำตอบ	212 ราย, 17 ข้อคำถาม, 3 รายการคำตอบ
person reliability	≥ 0.80	0.86	0.85
alpha coefficient	≥ 0.80	0.92	0.92
item reliability	≥ 0.90	0.86*	0.88*
person separation	≥ 2.00	2.45	2.39
item separation	≥ 3.00	2.52*	2.74*

*หมายถึงมีค่าต่ำกว่าเกณฑ์ที่ยอมรับ

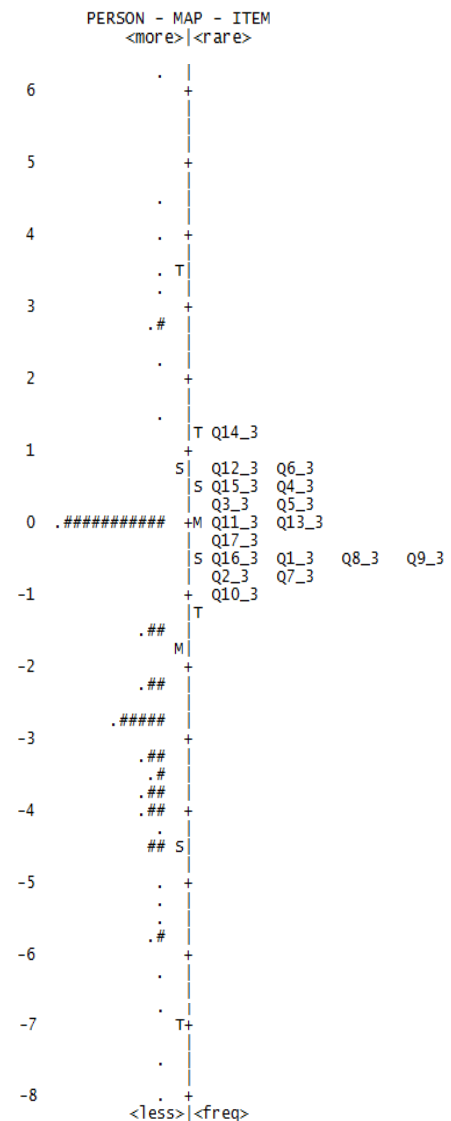
จากรูปที่ 3 คำถามส่วนใหญ่มีการกระจายอยู่ด้านบนเหนือกลุ่มตัวอย่าง ค่าเฉลี่ยของคำถาม (item mean) สูงกว่าค่าเฉลี่ยกลุ่มตัวอย่าง (person mean) เท่ากับ 1.90 logit (แสดงด้วยตำแหน่ง "M" บนสเกลด้าน ขวาและซ้ายตามลำดับ) ซึ่งหมายถึงข้อคำถามค่อนข้างยากเมื่อเทียบกับความสามารถของกลุ่มตัวอย่าง แสดงว่า ตัวอย่างส่วนใหญ่มีส่วนร่วมในการตัดสินใจรักษา ร่วมกับบุคลากรทางการแพทย์เพียงบางองค์ประกอบของกระบวนการตัดสินใจร่วม คำถามลำดับที่ 10 "ท่านมีความเข้าใจข้อมูลเกี่ยวกับปัญหาสุขภาพของท่าน" เป็นข้อที่ตัวอย่างเลือกตอบ "เห็นด้วย" มากที่สุด ส่วนข้อคำถามลำดับที่ 14 "ท่านสามารถหรือกับบุคลากรทางการแพทย์กรณีที่มีความต้องการทางเลือกในการรักษาที่แตกต่างกัน" เป็นข้อที่กลุ่มตัวอย่างเลือกตอบ "เห็นด้วย" น้อยที่สุดเมื่อเทียบกับข้อคำถามอื่น และเลือกตอบ "ไม่เห็นด้วย" มากกว่าข้อคำถามอื่น ๆ

การทำหน้าที่ต่างกันของข้อคำถาม

การทดสอบการทำหน้าที่ต่างกันของข้อคำถามระหว่างกลุ่มตัวอย่าง (DIF contrast) ที่จำแนกตามเพศ พบคำถาม 6 ข้อ มี DIF contrast ระหว่างกลุ่มตัวอย่าง < 0.50 logit ได้แก่ ข้อ 2, 8, 9, 10, 16 และ 17 ดังแสดงในรูปที่ 4 สำหรับการทดสอบ DIF ของข้อคำถามระหว่างกลุ่มตัวอย่างที่จำแนกตามอายุ ระดับการศึกษา อาชีพ รายได้ อาการหรือสาเหตุที่มาพบแพทย์ และแผนกที่เข้ารับบริการ พบว่าคำถามส่วนใหญ่มี DIF contrast ≥ 0.50 logit ซึ่งอาจจากตัวอย่างในแต่ละกลุ่มย่อยมีขนาดไม่มากพอที่จะยืนยันความแม่นยำของ DIF contrast ระหว่างกลุ่มตัวอย่าง จำนวนตัวอย่างเพื่อประเมิน DIF contrast ระหว่างกลุ่มตัวอย่างที่มีขนาด effect size เท่ากับ 0.50 ที่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ $p \leq 0.05$ และมีค่าอำนาจการทดสอบเท่ากับ 0.80 จะต้องมีจำนวนตัวอย่างอย่างน้อย 64 รายในแต่ละกลุ่มย่อย (23)

การอภิปรายผล

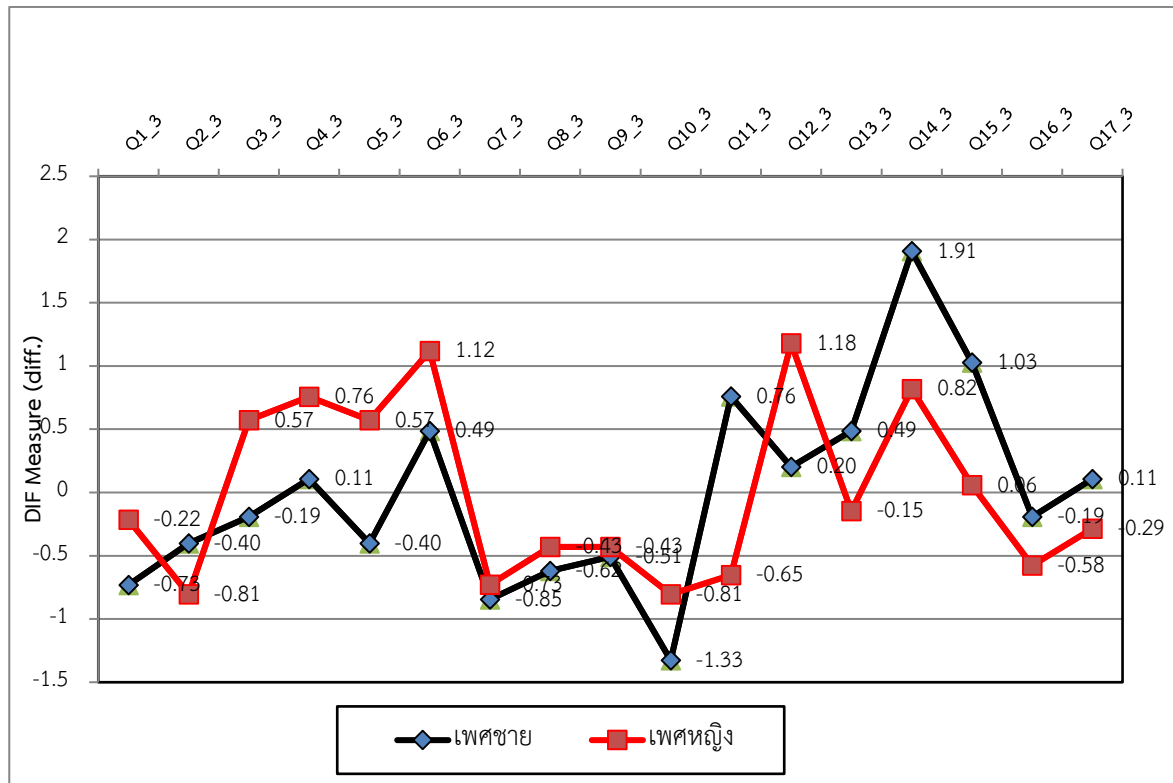
แบบวัด SDM ที่สร้างขึ้นมีความเที่ยงที่ดีเมื่อเทียบกับเกณฑ์ที่ระบุว่าไม่ควรน้อยกว่า 0.80 (24) แบบสอบถามแสดงความเป็นเอกมิติและมีความเป็นอิสระของข้อคำถามโดยค่า person reliability มีค่าเท่ากับ 0.86 และค่า person



รูปที่ 3. แผนภาพบุคคลและข้อคำถาม

separation มีค่าเท่ากับ 2.39 ซึ่งเทียบเท่าค่าสัมประสิทธิ์แอลฟาที่ 0.92 สำหรับค่า item fit statistics ของข้อมูลดิบพบว่าคำถามที่มีค่า infit mean-squares และ outfit mean-squares อยู่ในเกณฑ์เหมาะสม (0.60-1.40) จำนวน 15 ข้อ (ร้อยละ 88.24) หลังจากตัดตัวอย่างที่ไม่เหมาะสมกับราศีโมเดลออก พบว่ามีคำถามที่มีค่า infit mean-squares และค่า outfit mean-squares จัดอยู่ในเกณฑ์ที่เหมาะสมจำนวน 12 ข้อ (ร้อยละ 70.59)

งานวิจัยมีการปรับปรุงรายการคำตอบโดยรวมตัวเลือก "ไม่แน่ใจ", "ไม่เห็นด้วย" และ "ไม่เห็นด้วยอย่างยิ่ง" เข้าด้วยกัน เนื่องจากการวิเคราะห์การเลือกคำตอบแสดงความถี่การเลือกรายการคำตอบที่ไม่เท่าเทียม พบความถี่



รูปที่ 4. การทำหน้าที่ต่างกันของข้อคำถามจำแนกตามเพศ

การตอบ “ไม่เห็นด้วยอย่างยิ่ง” ต่ำที่สุด โดยพบเฉพาะคำถามข้อที่ 11 จาก CRC พบความน่าจะเป็นในการเลือกคำตอบ “ไม่แน่ใจ” มีค่าต่ำกว่าตัวเลือกอื่น ๆ และมีความน่าจะเป็นในการเลือกคำตอบซ้อนทับกับ “ไม่เห็นด้วย” และ “เห็นด้วย” เป็นส่วนใหญ่ นอกจากนี้ พบปัญหา item category: ability mean ของข้อคำถามที่ 10 และ 11 ในการวิจัยครั้งต่อไปอาจพิจารณาลดจำนวนตัวเลือกของคำถามลง ซึ่งทำให้สอดคล้องกับความเป็นจริงที่ว่าผู้ป่วยมีการรับรู้ได้หรือไม่ได้รับบริการตามองค์ประกอบของกระบวนการตัดสินใจร่วม

ข้อคำถามทั้ง 17 ข้อถูกคงไว้ในแบบวัด แม้ว่ามีคำถามที่ไม่เหมาะสมกับรหัสโมเดลจำนวน 5 ข้อ ได้แก่คำถามลำดับที่ 1, 8, 9, 10 และ 16 ซึ่งมีค่า outfit mean-squares = 1.41, 1.64, 0.31, 0.53 และ 0.44 ตามลำดับ ซึ่งข้อคำถามที่มีค่า outfit mean-squares < 0.50 หมายถึงคำถามที่สามารถคาดการณ์ได้สูง ทำให้มีประสิทธิผลน้อยสำหรับการวัด ซึ่งอาจเกิดได้จากการที่เนื้อหาซ้ำซ้อนกับคำถามข้ออื่น สำหรับข้อคำถามที่มีค่า outfit mean-squares > 1.50 หมายถึงข้อคำถามไม่เหมาะสมสำหรับโครงสร้างการวัด แต่คำถามไม่ได้ผิดเพี้ยนออกไปจากโครงสร้างการ

วัด (not degrading) (16) แต่ยังคงข้อคำถามเหล่านี้ในแบบสอบถาม เนื่องจากเมื่อตัดข้อคำถามทิ้งไปจะมีผลต่อความเที่ยงของแบบสอบถาม การคงข้อคำถามเหล่านี้ไว้ก็ไม่ทำให้เกิดผลเสียหายต่อผลลัพธ์ที่ประเมินได้จากข้อคำถามอื่น ๆ

แม้ว่าค่า item reliability มีค่าเท่ากับ 0.88 และมีค่า item separation มีค่าเท่ากับ 2.74 ซึ่งมีค่าต่ำกว่าเกณฑ์ที่ยอมรับเล็กน้อย แสดงถึงขนาดตัวอย่างที่ไม่มากพอที่จะยืนยันความแม่นยำของตำแหน่งของข้อคำถาม (item hierarchy) ต่อระดับความสามารถของผู้ทดสอบ (person ability) ที่ต้องการวัด (17) ขนาดตัวอย่างเพื่อการประมาณค่าที่แน่นอนเท่ากับ 50 รายต่อ 1 ข้อคำถาม (25) หรืออาจแก้ไขโดยการเพิ่มจำนวนข้อคำถามให้มากขึ้น และข้อคำถามที่เพิ่มเข้าไปควรสามารถวัดระดับความสามารถหรือระดับการตัดสินใจร่วมที่สูงขึ้น

การตรวจสอบ DIF contrast ระหว่างกลุ่ม พบว่าคำถามส่วนใหญ่มี DIF contrast เนื่องจากกลุ่มตัวอย่างที่มีอาชีพรับราชการมีจำนวนผู้ทดสอบเพียงร้อยละ 2.80 (n=6) โดยจำนวนผู้ทดสอบเพื่อประเมิน DIF contrast ต้องมีจำนวนอย่างน้อย 64 รายในแต่ละกลุ่มย่อย (26) ดังนั้น

การศึกษาครั้งต่อไปควรควบคุมจำนวนตัวอย่างในแต่ละกลุ่มย่อยให้เท่าเทียม ควบคุมความแตกต่างของกระบวนการรักษา และความแตกต่างของภาวะสุขภาพของกลุ่มตัวอย่าง หากพบ DIF contrast สูงกว่า ≥ 0.50 logit อาจต้องตัดข้อคำถามเหล่านี้ทิ้งไปสำหรับการวัดผลลัพธ์ที่ต่างกัน เนื่องจากความแตกต่างของภาวะสุขภาพที่อาจนำไปสู่การสร้างเครื่องมือวัด SDM สำหรับใช้ในผู้ป่วยทั่วไป (generic instrument) และเครื่องมือวัดการตัดสินใจร่วมสำหรับใช้เฉพาะโรค (specific instrument) สำหรับข้อดีของแบบสอบถามทั่วไป คือสามารถใช้งานสำหรับผู้ป่วยทุก ๆ โรค แต่มีข้อเสียคือแปรผลยากและใช้เปรียบเทียบระหว่างโรคไม่ได้ และอาจพลาดบางองค์ประกอบที่น่าสนใจสำหรับประชากรบางกลุ่ม ในขณะที่เครื่องมือเครื่องมือวัด SDM สำหรับใช้เฉพาะโรค อาจมีความไวต่อการเปลี่ยนแปลงได้ดีกว่า

สรุปผล

แบบสอบถามวัดระดับ SDM แบ่งเป็น 7 องค์ประกอบตามกระบวนการตัดสินใจ มีคำถามจำนวน 17 ข้อ ตรวจสอบคุณสมบัติการวัดด้วยการวิเคราะห์ราสช์โมเดล พบว่าแบบวัด SDM มีความตรงและความเที่ยงที่ดี มีความเป็นความเป็นเอกมิติ และความเป็นอิสระของข้อคำถาม แต่อย่างไรก็ตามควรมีการทดสอบคุณสมบัติทางจิตวิทยาโดยสมบูรณ์ของแบบสอบถามในผู้ป่วยจำนวนมากต่อไป

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณเจ้าหน้าที่ของโรงพยาบาลสิรินธร ที่ให้ความช่วยเหลือในการเก็บข้อมูล รวมทั้งผู้ป่วยทุกท่าน ที่ให้ความร่วมมือในการเก็บข้อมูล

เอกสารอ้างอิง

1. Coulter A, Collins A. Making shared decision-making a reality: no decision about me, without me. London: The King's Fund; 2011.
2. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: What does it mean? (or it takes at least two to tango). Soc Sci Med 1997; 44: 681-92.

3. Brennan B, Holly M. Interprofessional collaboration in health care: Lessons to be learned from competitive sports. Can Pharm J 2015; 148: 176-9.
4. Haggerty JL, Reid RJ, Freeman GK, Starfield BH, Adair CE, McKendry R. Continuity of care: a multidisciplinary review. BMJ 2003; 327: 1219-21.
5. Scholl I, Koelewijn-van Loon M, Sepucha K, Elwyn G, Légaré F, Härter M, et al. Measurement of shared decision making—a review of instruments. Z Evid Fortbild Qual Gesundheitswes 2011; 105: 313-24.
6. Elwyn G, Edwards A, Wensing M, Hood K, Atwell C, Grol R. Shared decision-making: developing the OPTION scale for measuring patient involvement. Qual Saf Health Care 2003; 12: 93-9.
7. Lerman CE, Brody DS, Caputo GC, Smith DG, Lazaro CG, Wolfson HG. Patient's perceived involvement in care scale: Relationship to attitudes about illness and medical care. J Gen Intern Med 1990; 5: 29-33.
8. Kriston L, Scholl I, Hölzel L, Simon D, Loh A, Härter M. The 9 - item Shared Decision Making Questionnaire (SDM-Q-9): Development and psychometric properties in a primary care sample. Patient Educ Couns 2010; 80: 94-9.
9. Elwyn G, Miron-Shatz T. Deliberation before determination: The definition and evaluation of good decision making. Health Expect 2010 Jun; 13: 139-47.
10. NHS Organisations. Measuring shared decision making: A review of research evidence [online]. 2012 (cited Apr 1, 2015). Available from: www.england.nhs.uk/wp-content/uploads/2013/08/7sdm-report.pdf.
11. Linacre JM. Understanding Rasch measurement: Optimizing rating scale category effectiveness. J Appl Meas 2002; 3: 85-106.
12. Makoul G, Clayman ML. An integrative model of shared decision making in medical encounters. Patient Educ Couns 2006; 60: 301-12.

13. Guntayoung C, Chinchai S. The content validity and test-retest reliability of the developmental visual perception test (DTVP-2) in Thai children. *International Journal of Medicine and Pharmaceutical Sciences* 2013; 3: 1-6.
14. Meijer RR, Sijtsma K. Methodology review: Evaluating person fit. *Appl Psychol Meas* 2001; 25: 107-35.
15. DeMars C. Measuring higher education outcomes with a multidimensional Rasch model. *J Appl Meas* 2004; 5: 350-61.
16. Linacre JM. Fit diagnosis: Infit outfit mean-square standardized [online]. 1999 (cited Apr 1, 2018).
17. Bond T, Fox C. Applying the Rasch model: Fundamental measurement in the human sciences. 2nd ed. Mahwah: Lawrence Erlbaum Associates; 2007.
18. Linacre JM. Reliability and separation of measures [online]. 1999 (cited May 12, 2016). Available from: www.winsteps.com/winman/reliability.htm.
19. Smith EV. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 2002; 3: 205-31.
20. Linacre JM. Data variance explained by Rasch measures. *Rasch Measurement Transactions* 2006 ; 20: 1045.
21. Embetson SE, Reise SP. Item response theory for psychologists. Mahwah: New Jersey, L. Erlbaum Associates, 2000.
22. Wang Z, Zhou J, Luo X, Xu Y, She X, Chen L, et al. Rasch analysis of the Adult Strabismus Quality of Life Questionnaire (AS-20) among Chinese adult patients with strabismus. *PLOS ONE* 2015; 10: e0142188. doi:10.1371/journal.pone.0142188.
23. Tellez A, Cadena CHG, Corral-Verdugo V. Effect size, confidence intervals and statistical power in psychological research. *Psychology in Russia: State of the Art* 2015; 8: 27-46.
24. Fox CM, Jones JA. Uses of Rasch modeling in counseling psychology research. *J Couns Psychol* 1998; 45: 30-45.
25. Linacre JM. Small sample size, sample size and item calibration (or person measure) stability [online]. 1998 (cited Sep 1, 2018). Available from: www.rasch.org/rmt/rmt74m.htm.