

คุณสมบัติการวัดทางจิตวิทยาของแบบประเมินความแตกฉานด้านสุขภาพสำหรับ ชาวไทยชนิตรายการค้าที่มีคำถามทดสอบความเข้าใจในผู้ป่วยชาวไทยมุสลิม

นุรไอนา ดารามาลย์¹, สวงน ลือเกียรติวัฒน์², วรณช แสงเจริญ³

¹กลุ่มงานเภสัชกรรม โรงพยาบาลยะหริ่ง จังหวัดปัตตานี

²ภาควิชาบริหารเภสัชกิจ คณะเภสัชศาสตร์ มหาวิทยาลัยสงขลานครินทร์

³ภาควิชาเภสัชกรรมคลินิก คณะเภสัชศาสตร์ มหาวิทยาลัยสงขลานครินทร์

บทคัดย่อ

วัตถุประสงค์: เพื่อศึกษาความตรงและความเที่ยงของแบบประเมินความแตกฉานด้านสุขภาพสำหรับชาวไทยชนิตรายการค้าที่มีคำถามทดสอบความเข้าใจหรือ Thai Health Literacy Assessment Using Word List with Extended Questions to Test Comprehension (THLA-W+) ในผู้ป่วยชาวไทยมุสลิมและหาจุดตัดคะแนนที่เหมาะสมของแบบวัด วิธีการ: การศึกษานี้ใช้แบบวัด THLA-W+ ที่พัฒนาโดยผดุง จันชูโต ที่ประกอบด้วยคำจำนวน 48 คำและมี 4 ตัวเลือกในแต่ละคำเพื่อทดสอบความเข้าใจของตัวอย่างต่อคำในแบบวัด การคำนวณคะแนนมี 3 แบบ คือ 1) จำนวนจากการทดสอบการอ่าน (THLA-W+R: reading test) 2) จำนวนจากการทดสอบความเข้าใจ (THLA-W+C หรือ comprehension test) และ 3) จำนวนจากการทดสอบการอ่านร่วมกับการทดสอบความเข้าใจ (THLA-W+RC: reading and comprehension test) ตัวอย่าง คือ ผู้ป่วยนอกที่เป็นชาวไทยมุสลิม 767 คนของโรงพยาบาลยะหริ่ง จังหวัดปัตตานี การทดสอบความตรงของ THLA-W+ ทำโดยหาความสัมพันธ์กับตัวชี้วัดความแตกฉานด้านสุขภาพ (health literacy: HL) ระดับการศึกษา และแบบวัด Thai Health Literacy Assessment Using Nutrition Label (THLA-N8) การหาจุดตัดและความสามารถในการทำนายใช้การวิเคราะห์ receiver operating curve ผลการวิจัย: THLA-W+R, THLA-W+C และ THLA-W+RC มี Cronbach Alpha 0.98, 0.97, และ 0.97 ตามลำดับ THLA-W+C และ THLA-W+RC สามารถแยกตัวอย่างที่มีระดับการศึกษาที่แตกต่างกันได้ดี และดีกว่า THLA-W+R ค่าสัมประสิทธิ์สหสัมพันธ์ (r) ระหว่างคะแนน THLA-W+ ทั้ง 3 แบบกับตัวชี้วัด HL ต่าง ๆ (เช่น ความสามารถในการอ่าน ความเข้าใจเอกสารทางสุขภาพ) เป็นบวกและมีนัยสำคัญทางสถิติ ($r=0.22-0.82$; $P<0.05$) ซึ่ง THLA-W+C และ THLA-W+RC มีค่าสัมประสิทธิ์สหสัมพันธ์ที่สูงกว่า THLA-W+R คะแนนจาก THLA-W+C และ THLA-W+RC มีความสัมพันธ์กันสูงโดย $r=0.99$ หากใช้ความสามารถในการอ่าน ความเข้าใจเอกสารทางสุขภาพ และ THLA-N8 เป็นตัววัดมาตรฐาน พบว่า THLA-W+R มีพื้นที่ใต้โค้ง (area under the curve; AUC) เท่ากับ 0.64-0.66 (ขึ้นกับตัววัดมาตรฐาน) ซึ่งน้อยกว่า AUC ของ THLA-W+C และ THLA-W+RC ที่เท่ากับ 0.82-0.83 และ 0.82-0.83 ตามลำดับ จุดตัดคะแนนที่เหมาะสมของ THLA-W+R, THLA-W+C และ THLA-W+RC คือ 47, 37 และ 37 ตามลำดับ ความไวของแบบวัด คือ ร้อยละ 41.90-46.52, 68.68-78.61 และ 70.74-80.75 ตามลำดับ THLA-W+C และ THLA-W+RC มีค่าความไวมากกว่า THLA-W+R ส่วนค่าความจำเพาะของแบบวัดจากการคำนวณคะแนนทั้ง 3 แบบมีค่าใกล้เคียงกัน คือ 76.21-88.10, 67.96-83.33 และ 66.02-80.95 ตามลำดับ สรุป: THLA-W+ มีความเที่ยงและความตรงในผู้ป่วยชาวไทยมุสลิมในการทดสอบควรใช้ THLA-W+C เพราะมีคุณสมบัติการวัดที่ดีกว่า THLA-W+R และไม่ต้องใช้บุคลากรในการตัดสินการอ่านออกเสียงคำ THLA-W+C มีความไวและความจำเพาะเป็นที่น่าพอใจ ผู้ที่ได้คะแนนน้อยกว่าหรือเท่ากับ 37 ถือว่ามี HL ไม่เพียงพอ

คำสำคัญ: ความแตกฉานด้านสุขภาพ แบบประเมินความแตกฉานด้านสุขภาพ เกณฑ์คะแนน ชาวไทยมุสลิม

รับต้นฉบับ: 12 ต.ค. 2561 ได้รับบทความฉบับปรับปรุง: 14 ธ.ค. 2561 รับลงตีพิมพ์: 25 ธ.ค. 2561

ผู้ประสานงานบทความ: นุรไอนา ดารามาลย์ กลุ่มงานเภสัชกรรม โรงพยาบาลยะหริ่ง อำเภอยะหริ่ง จังหวัดปัตตานี 94150 E-mail: nuraina056@gmail.com

Psychometric Properties of the Thai Health Literacy Assessment Using Word List with Extended Questions to Test Comprehension in Thai Muslim Patients

Nuraina Daraman¹, Sanguan Lerkiatbundit², Woranuch Saengcharoen³

¹Pharmacy Department, Yaring Hospital, Pattani

²Department of Pharmacy Administration, Faculty of Pharmaceutical Sciences, Prince of Songkla University

³Department of Clinical Pharmacy, Faculty of Pharmaceutical Sciences, Prince of Songkla University

Abstract

Objectives: To test validity and reliability of the Thai Health Literacy Assessment Using Word List with Extended Questions to Test Comprehension (THLA-W+) in Thai Muslim patients and to determine its cut-off. **Methods:** This study used the THLA-W+ developed by Phadoong Chanchuto. The scale consisted of 48 words with 4 corresponding choices to test the comprehension of word meaning. Scoring of the scale was conducted in 3 methods: 1) calculating the scores from reading test (THLA-W+R), 2) calculating the scores from comprehension test (THLA-W+C) and 3) calculating the scores from both reading and comprehension test (THLA-W+RC). Subjects were 767 Muslim outpatients at Yaring hospital in Pattani province. Validity testing of THLA-W+ was conducted by determining its correlation with various indicators of health literacy (HL) and the Thai Health Literacy Assessment Using Nutrition Label (THLA-N8). Cut-off point and predictive ability were determined by the analysis of receiver operating curve. **Results:** Cronbach's Alphas of THLA-W+R, THLA-W+C and THLA-W+RC were 0.98, 0.97 and 0.97, respectively. The THLA-W+C and THLA-W+RC could well discriminate the subjects with different levels of education, and better than the THLA-W+R did. Correlation coefficients (r) between all three THLA-W+ scores and HL indicators (e.g. reading ability, understanding of health documents) were positive and statistically significant ($r=0.222-0.816$; $P<0.05$). The THLA-W+C and THLA-W+RC showed a higher correlation coefficient than the THLA-W+R did. The THLA-W+C and the THLA-W+RC highly correlated with $r=0.987$. When using reading ability, understanding of health documents and THLA-N8 as gold standards, THLA-W+R had an area under the curve (AUC) of 0.641-0.660 (depending on gold standards), which was less than those of the THLA-W+C and THLA-W+RC, which were 0.820-0.830 and 0.826-0.832, respectively. The cut-off values of the THLA-W+R, THLA-W+C and THLA-W+RC were 47, 37 and 37 respectively. Sensitivities of the scales were 41.90-46.52, 68.68-78.61 and 70.74-80.75, respectively. THLA-W+C and THLA-W+RC were more sensitive than the THLA-W+R did. Specificities of three measures were similar at 76.21-88.10, 67.96-83.33 and 66.02-80.95, respectively. **Conclusion:** The THLA-W+ was reliable and valid for using in Muslim patients. The scale should be administered as THLA-W+C because it provides the scores with a better psychometric property than THLA-W+R, and needs no staff to judge the pronunciation of the subjects. The THLA-W+C shows a satisfactory sensitivity and specificity with the score of 37 or lower indicating inadequate HL.

Keywords: health literacy, health literacy assessment tool, cut-off, Thai Muslims

บทนำ

ความแตกฉานด้านสุขภาพ (health literacy: HL) ประกอบด้วยความรู้ ความต้องการ และความสามารถที่บุคคลจำเป็นต้องมีเพื่อให้สามารถเข้าถึง เข้าใจ ประเมิน และใช้ข้อมูลสุขภาพสำหรับตัดสินใจในชีวิตประจำวันในเรื่องเกี่ยวกับการรักษาพยาบาล การป้องกันโรค และการสร้างเสริมสุขภาพเพื่อเพิ่มคุณภาพชีวิต (1) HL ที่ต่ำสัมพันธ์กับผลลัพธ์ทางสุขภาพที่ไม่ดีและการไม่ได้รับบริการสุขภาพที่สมควรดี (2)

การวัด HL เพื่อประเมินสภาพปัญหาจึงมีความสำคัญทั้งในระดับบุคคลและประเทศ เครื่องมือวัด HL มี 3 ชนิดที่นิยมใช้กันมากที่สุดในต่างประเทศ คือ Rapid Estimate of Adult Literacy in Medicine (REALM) (3), Test of Functional Health Literacy in Adults (TOFHLA) (4) และ Newest Vital Sign (NVS) (5) REALM ถูกพัฒนาเพื่อใช้ประเมินทักษะในการอ่านคำ 125 คำที่คัดเลือกจากคำที่ปรากฏในเอกสารที่ใช้ในสถานบริการปฐมภูมิ (3) ต่อมาจึงพัฒนาเป็น REALM-S หรือแบบวัดฉบับสั้นเพื่อลดเวลาในการทดสอบและทำให้การใช้สะดวกมากขึ้น ส่วน TOFHLA วัด HL ขั้นพื้นฐาน (functional health literacy) ได้แก่ ทักษะการอ่าน เขียน และการคำนวณ (4) แบบวัดมีความยาวมาก จึงมีการพัฒนาแบบวัดฉบับสั้น 40 ข้อหรือ S-TOFHLA แต่ยังไม่สะดวกในการใช้งาน เนื่องจากยังคงมีจำนวนข้อคำถามมากและใช้เวลาในการทดสอบนาน ส่วน NVS เป็นคำถาม 6 ข้อที่ทดสอบทักษะการอ่าน การแปลความหมาย และการคำนวณโดยอาศัยข้อมูลบนฉลากโภชนาการของไอศกรีม การทดสอบใช้การสัมภาษณ์ทำให้มีข้อจำกัดในการใช้และทำให้ผู้ถูกทดสอบกังวลในการตอบ จึงอาจทำให้ผลการประเมินไม่สอดคล้องกับความจริง (5)

ในประเทศไทย มีผู้พัฒนาแบบวัด HL เฉพาะกลุ่มโรคหรืออาการ เช่น แบบวัดความรู้ด้านสุขภาพตามหลัก 3อ2ส สำหรับคนไทยกลุ่มเสี่ยงโรคเบาหวานและความดันโลหิตสูง (6) เป็นต้น แต่ในงานวิจัยครั้งนี้มุ่งทดสอบแบบวัด HL ชนิดทั่วไปที่ไม่เจาะจงโรคหรืออาการ ซึ่งสามารถใช้ในผู้มารับบริการในสถานพยาบาล จึงไม่อาจใช้แบบวัดที่กล่าวมาข้างต้นได้

สำหรับแบบวัด HL ทั่วไปฉบับภาษาไทยนั้น บังอร ศรีจินดาวงศ์ (7) ได้แปลแบบวัด REALM เป็นภาษาไทยซึ่งมีผลทำให้ระดับความยากง่ายในการอ่านของคำเปลี่ยนไปจากต้นฉบับภาษาอังกฤษ อีกทั้งแบบวัดนี้ไม่ได้

ผ่านการทดสอบความตรงโดยการหาความสัมพันธ์กับความสามารถอื่น ๆ ที่บ่งชี้ถึง HL นอกจากนี้ยังพบการกระจุกตัวของคะแนนในช่วงคะแนนสูง (ceiling effect) ผู้ที่มีการศึกษาในระดับที่ต่างกันสามารถอ่านคำได้ถูกต้องใกล้เคียงกัน จึงอนุมานว่า แบบวัดอาจไม่สามารถแยกแยะผู้ที่มี HL ต่างกันได้ดีนัก แบบวัด S-TOFHLA บางส่วนได้รับการแปลเป็นไทยเช่นเดียวกัน (7) แบบวัดจำนวนมากมีเนื้อหาตามระบบสาธารณสุขของประเทศสหรัฐอเมริกา ทำให้ไม่อาจใช้กับชาวไทยได้

ต่อมาวิทยา พันธุ์ทอง (8) ได้สร้างแบบวัด Thai Health Literacy Assessment Using Word List หรือ THLA-W โดยยึดแนวคิดของ REAL-M แบบวัดประกอบด้วยรายการคำ 48 คำที่ใช้ในเอกสารต่าง ๆ ในโรงพยาบาลระดับปฐมภูมิ ระดับ HL คือ จำนวนคำที่ตัวอย่างตามถูก การทดสอบแบบวัดในผู้ป่วยนอก 502 รายของโรงพยาบาลปางเบซาร์ จังหวัดสงขลา พบว่า แบบวัดมี Cronbach's alpha=0.91 และมีความตรง เนื่องจากสามารถจำแนกคะแนนในผู้ที่มีการศึกษาในระดับที่ต่างกันได้อีกทั้งยังมีความสัมพันธ์กับตัวแปรที่บ่งชี้ HL (เช่น ความเข้าใจในฉลากยา) อย่างไรก็ตาม ผลการศึกษาพบว่า ยังคงมี ceiling effect จากนั้นผดุง จันชูโต (9) ปรับ THLA-W โดยเพิ่มคำถามชนิดมีตัวเลือกเพื่อทดสอบความเข้าใจในคำที่ใช้ทดสอบเพื่อลด ceiling effect และลดข้อวิจารณ์ที่ว่า THLA-W ประเมินเฉพาะทักษะการอ่าน จนได้แบบวัด Thai Health Literacy Assessment Using Word List with Extended Questions หรือ THLA-W+ วิธีการดังกล่าวประยุกต์โดยใช้แนวคิดจาก SAHLSA (Short Assessment of Health Literacy for Spanish-Speaking Adults) (10) การทดสอบ THLA-W+ ในผู้ป่วยนอก 313 รายของโรงพยาบาลสะเดา จังหวัดสงขลา พบว่า คะแนนที่ได้มีการกระจายตัวดีกว่า THLA-W แบบวัดมี Cronbach's alpha=0.91 และมีความตรง เนื่องจากสามารถจำแนกคะแนนในผู้ที่มีการศึกษาในระดับที่ต่างกันได้ และมีความสัมพันธ์กับตัวแปรที่บ่งชี้ HL (เช่น ความเข้าใจในบัตรนัดของโรงพยาบาล)

หลังจากนั้น ปรีนา ณ พัทลุง (11) ได้สร้างแบบวัด Thai Health Literacy Assessment Using Nutrition Label (THLA-N8) ซึ่งเป็นแบบวัด 8 ข้อที่ใช้ฉลากโภชนาการเป็นเนื้อหาในคำถามตามหลักการของ NVS และเพิ่มคำถามคัดกรอง HL ชนิดรายงานด้วยตนเอง การทดสอบในผู้ป่วย

นอก 497 รายของโรงพยาบาลตราง จังหวัดตราง พบว่า แบบวัดมี Cronbach's $\alpha=0.70$ และมีความตรงโดยผู้ที่มีระดับการศึกษาสูงกว่ามีคะแนน THLA-N8 มากกว่า คะแนนจากแบบวัดยังสัมพันธ์กับตัวแปรที่บ่งชี้ HL (ความเข้าใจในเอกสารสุขภาพและความสามารถในการอ่านที่ประเมินด้วยแบบวัด ซึ่งดัดแปลงจาก S-TOFHLA) ในการวิเคราะห์โค้ง ROC พบว่า แบบวัดสามารถแยกแยะผู้ที่มี HL ที่เพียงพอและไม่เพียงพอได้ดี คะแนน THLA-N8 ที่น้อยกว่าหรือเท่ากับ 4.83 บ่งบอก HL ที่ไม่เพียงพอ แบบวัดมีความไว 0.71-0.75 และความจำเพาะ 0.67-0.73 ขึ้นกับตัววัดมาตรฐาน

การวิจัยนี้เป็นการทดสอบแบบวัด THLA-W+ ในชาวไทยมุสลิม เนื่องจาก THLA-W+ ถูกสร้างและพัฒนาในตัวอย่างที่เป็นชาวไทยพุทธมากกว่าร้อยละ 90 ซึ่งใช้ภาษาไทยเป็นภาษาหลัก แม้ว่า THLA-W+ จะมีความตรง ความเที่ยง และมีคุณสมบัติการวัดที่ดีในชาวไทยพุทธ (8, 9) แต่ไม่สามารถเชื่อมั่นได้ว่า แบบวัดจะมีคุณสมบัติการวัดที่ดีในชาวไทยมุสลิมในสามจังหวัดชายแดนใต้ที่มีวัฒนธรรม การสื่อสารที่แตกต่างจากชาวไทยพุทธ โดยมีการใช้ภาษามลายูเป็นภาษาหลักในการสื่อสาร ลักษณะสำคัญของ THLA-W+ คือ การทดสอบการอ่านและความเข้าใจคำภาษาไทย ความสามารถทางภาษาไทยที่น้อยกว่าของประชาชนในสามจังหวัดชายแดนใต้เป็นที่ทราบกันและพบได้ในผู้ที่ไม่ใช่ผู้สูงอายุ ดังเห็นได้จากผลการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ในปีพ.ศ. 2560 พบว่า นักเรียนระดับชั้นมัธยมศึกษาชั้นปีที่ 6 ในจังหวัดยะลา ปัตตานี และนราธิวาส ซึ่งมีประชากรเป็นชาวมุสลิมเป็นส่วนใหญ่ ได้คะแนนทดสอบวิชาภาษาไทย 36.41, 34.74 และ 35.38 ตามลำดับ ในขณะที่ค่าเฉลี่ยคะแนนของประเทศ คือ 49.25 และค่าเฉลี่ยคะแนนของจังหวัดสงขลา คือ 47.93 (12) จากการที่ชาวมุสลิมบางส่วนอาจมีความสามารถทางภาษาไทยที่น้อยและมีวัฒนธรรมที่แตกต่างจากชาวพุทธ จึงอาจทำให้มี HL น้อย เนื่องจากการไม่เข้าใจข้อมูลที่เป็นภาษาไทย ทั้งยังสามารถส่งผลต่อความตรงและความเที่ยงของแบบวัด THLA-W+ ได้ด้วย

การศึกษาเลือกทดสอบ THLA-W+ เพราะมีคุณสมบัติการวัดที่ดีกว่า THLA-W และมีแนวคิดที่ใช้วัด HL คือ การทดสอบความสามารถในการอ่านและเข้าใจคำภาษาไทย ซึ่งน่าจะเป็นวิธีการที่เหมาะสมกับกลุ่มที่มีทักษะเหล่านี้ น้อย ขณะที่ THLA-N8 ใช้หลักการในการประเมิน

HL ที่ต่างออกไป (เน้นทักษะการค้นหาข้อมูลและการคำนวณ) ผู้วิจัยจึงเลือกใช้แบบวัด THLA-W+ มาศึกษาในงานวิจัยครั้งนี้ อีกทั้งการทดสอบคุณสมบัติของแบบวัด THLA-W+ ยังทำในกลุ่มประชากรที่ไม่หลากหลายนัก

งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อทดสอบความตรงและความเที่ยงของ THLA-W+ ในชาวไทยมุสลิม รวมทั้งหาเกณฑ์การแปลผลระดับ HL ในประชากรกลุ่มนี้ การศึกษานี้คำนวณคะแนน THLA-W+ ใน 3 แบบ คือ 1) คะแนนจากการอ่านคำถูกต้อง (reading test) หรือ THLA-W+R 2) คะแนนจากตอบคำถามทดสอบความเข้าใจถูกต้อง (comprehension test) หรือ THLA-W+C และ 3) คะแนนจากการอ่านคำถูกต้องควบคู่กับการตอบคำถามทดสอบความเข้าใจถูกต้อง (reading and comprehension tests) หรือ THLA-W+RC) การศึกษาเปรียบเทียบคุณสมบัติของคะแนนจากการคำนวณทั้ง 3 แบบ

วิธีการวิจัย

การวิจัยนี้ผ่านการอนุมัติจากคณะกรรมการพิจารณาจริยธรรมการวิจัย คณะเภสัชศาสตร์ มหาวิทยาลัยสงขลานครินทร์ (ศร 0521.1.07/1870) รวมทั้งได้รับการอนุญาตให้ใช้แบบวัด THLA-W+ และแบบวัด THLA-N8 แล้ว

การทดสอบความตรง-ความเที่ยง

สถานที่ทำวิจัยและตัวอย่าง

สถานที่ทำวิจัย คือ โรงพยาบาลยะหริ่ง ซึ่งเป็นโรงพยาบาลชุมชนขนาด 60 เตียง ในจังหวัดปัตตานี ประชากรในอำเภอยะหริ่งมีจำนวน 86,546 คน โดยประชากรร้อยละ 95.72 นับถือศาสนาอิสลาม (อีกร้อยละ 4.28 นับถือศาสนาพุทธ) ตัวอย่าง คือ ผู้ป่วยนอกซึ่งมารับบริการที่โรงพยาบาลยะหริ่งและมีคุณสมบัติตามเกณฑ์คัดเข้า ดังนี้ 1) เป็นชาวไทยมุสลิมที่มีอายุอย่างน้อย 20 ปี 2) สามารถพูดสื่อสารและอ่านภาษาไทยได้ 3) ยินดีเข้าร่วมการวิจัย และ 4) ไม่จบการศึกษาหรือประกอบอาชีพในสาขาที่เกี่ยวข้องสาธารณสุข การศึกษาครั้งนี้เก็บข้อมูลใน 5 คลินิก คือ คลินิกโรคทั่วไป คลินิกโรคไม่ติดต่อเรื้อรัง คลินิกทันตกรรม คลินิกโรคหืดและปอดอุดกั้นเรื้อรัง และคลินิกฝากครรภ์ การเลือกตัวอย่างผู้ป่วยใช้วิธีแบบตามสะดวกจากผู้ที่มีคุณสมบัติตามเกณฑ์

การคำนวณขนาดตัวอย่างใช้สูตรการเปรียบเทียบตัวแปรเชิงปริมาณ (คะแนนจาก THLA-W+) ระหว่างกลุ่มที่

มีระดับการศึกษาต่างกันด้วยการวิเคราะห์ความแปรปรวนทางเดียว โดยใช้โปรแกรม G*power (13) ในการคำนวณ มีการกำหนดความคลาดเคลื่อนชนิดที่ 1 ที่ 0.05 อำนาจการทดสอบที่ 0.80 จำนวนกลุ่มเปรียบเทียบ คือ 4 กลุ่ม (ประถมศึกษา มัธยมศึกษาตอนต้น มัธยมศึกษาตอนปลาย/ประกาศนียบัตรวิชาชีพ (ปวช.) และประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.)/ปริญญาตรีหรือสูงกว่า) ในที่นี้กำหนดให้ขนาดอิทธิพล (effect size) เท่ากับ 0.15 ขนาดตัวอย่างที่คำนวณได้ คือ 492 ราย ดังนั้นต้องเก็บข้อมูลอย่างน้อยประมาณ 125 คนในแต่ละระดับการศึกษา การศึกษาครั้งนี้เก็บตัวอย่างได้ทั้งหมด 767 ราย

วิธีการเก็บข้อมูล

เมื่อพบตัวอย่างที่มีคุณสมบัติตามเกณฑ์ ผู้วิจัยขอให้ตัวอย่างตอบแบบสอบถามที่วัดข้อมูลทั่วไป (เพศ อายุ อาชีพ การศึกษา และการประเมินความสามารถในการอ่านของตนเอง) ต่อมาผู้วิจัยยื่นแบบวัด THLA-W+ (9) ให้ตัวอย่าง แบบวัด THLA-W+ ประกอบด้วยรายการคำจำนวน 48 คำ และแต่ละคำมี 4 ตัวเลือก ผู้วิจัยกล่าวกับตัวอย่างว่า “ลองอ่านคำหมายเลข 1 แบบออกเสียงดูซิคะ” พร้อมทั้งชี้คำด้วยปลายปากกา หลังจากตัวอย่างอ่านให้กล่าวว่า “คำไหนเหมือนหรือคล้ายกับคำที่เพิ่งอ่านไปมากที่สุด หากไม่รู้หรือไม่แน่ใจให้เลือกข้อ ง ได้คะ” การตัดสินใจว่า ตัวอย่างอ่านออกเสียงถูกหรือผิดใช้วิธีการเดียวกับงานวิจัยในอดีต (9) ถ้าตัวอย่างอ่านคำไม่ได้ 3 คำติดต่อกัน ผู้วิจัยขอให้ตัวอย่างดูรายการคำที่ยังไม่อ่านทั้งหมดและถามว่า อ่านคำไหนได้บ้างและให้เลือกตัวเลือกของคำนั้น ๆ

หลังจากนั้นให้ตัวอย่างตอบแบบสอบถามที่วัดตัวบ่งชี้ถึง HL ดังนี้ ก) คำถามปลายปิดจำนวน 5 ข้อที่วัดความสามารถในการอ่านที่ดัดแปลงมาจากแบบวัด S-TOFHLA (4) แบบวัดมีลักษณะเป็น cloze test คือ เป็นประโยคที่มีข้อความบางส่วนหายไป ผู้ตอบต้องเลือกคำที่เหมาะสมจากตัวเลือกที่กำหนดให้ เพื่อให้ประโยคสมบูรณ์ ข) คำถามปลายเปิด 6 ข้อที่วัดความเข้าใจเอกสารทางสุขภาพ เช่น จากฉลากยาที่ระบุว่า “รับประทานครั้งละ 2 เม็ด ทุก 4-6 ชั่วโมง เฉพาะเวลาปวดหรือมีไข้” หากท่านปวดและรับประทานยาแก้ปวดเวลา 10.00 น. แต่อาการปวดไม่ลดลง ท่านสามารถรับประทานยาแก้ปวดซ้ำได้อีกครั้งหนึ่งในเวลาใด และ ค) แบบวัด THLA-N8 จำนวน 8 ข้อ โดย 3 ข้อแรกเป็นคำถามคัดกรองผู้ที่มี HL ไม่เพียงพอโดยมีตัวเลือก 5 ระดับ และอีก 5 ข้อเป็นคำถามทดสอบความ

เข้าใจฉลากโภชนาการ คะแนนของ THLA-N8 ได้จากค่าเฉลี่ยของคำถาม 3 ข้อสำหรับคัดกรองผู้ที่มี HL ไม่เพียงพอ (พิสัย คือ 1-5) บวกกับคะแนนรวมจากคำถามเกี่ยวกับการใช้ข้อมูลจากฉลากโภชนาการ 5 ข้อ (พิสัย คือ 0-5) คะแนนมีพิสัยที่เป็นไปได้ คือ 1-10 (11)

การวิเคราะห์ข้อมูล

การวิเคราะห์ความเที่ยงของ THLA-W+ ใช้ Cronbach's alpha การทดสอบความตรงทำโดย 1) เปรียบเทียบคะแนน THLA-W+ ระหว่างผู้ที่มีระดับการศึกษาต่างกัน หากผู้ที่มีระดับการศึกษาสูงกว่าได้คะแนน THLA-W+ มากกว่า แสดงว่า THLA-W+ มีความตรง การเปรียบเทียบความเท่าเทียมกันของความแปรปรวนในแต่ละกลุ่มการศึกษาด้วย Levene test พบว่า ความแปรปรวนแตกต่างกัน การทดสอบความแปรปรวนทางเดียวจึงใช้ Brown-Fortsynthe test ส่วนการเปรียบเทียบรายคู่ใช้ Games-Howell test 2) หาความสัมพันธ์ระหว่างคะแนน THLA-W+ กับตัวชี้วัด HL โดยใช้สัมประสิทธิ์สหสัมพันธ์ (r) ตัวชี้วัด HL ประกอบด้วย ก) ความสามารถในการอ่านที่ดัดแปลงจาก S-TOFHLA ข) ความสามารถในการเข้าใจเอกสารสุขภาพ ค) แบบคัดกรองผู้ที่มี HL ไม่เพียงพอซึ่งเป็นส่วนหนึ่งของ THLA-N8 และ ง) คะแนนรวมจากแบบวัด THLA-N8 ค่าสัมประสิทธิ์สหสัมพันธ์ที่สูง บ่งชี้ความตรงของแบบวัด THLA-W+

ความสามารถในการทำนาย

ตัววัดมาตรฐาน

การหาเกณฑ์คะแนนที่เหมาะสมเพื่อใช้ตัดสินระดับ HL ใช้การวิเคราะห์โค้ง receiver operating characteristic (ROC) ตัววัดมาตรฐาน (gold standard: GS) ที่ใช้ประเมินความเพียงพอของ HL คือ ความสามารถพื้นฐานที่บุคคลจำเป็นต้องมีเพื่อให้สามารถประกอบกิจกรรมสุขภาพเบื้องต้นได้ GS ในการศึกษาครั้งนี้ มี 3 ชนิด คือ ก. GS1 หรือคะแนนรวมของความสามารถในการอ่าน (5 ข้อ) และความเข้าใจเอกสารทางสุขภาพ (6 ข้อ) พิสัยของคะแนน คือ 0-11 ในที่นี้ถือว่าผู้ที่ได้คะแนนเต็ม 11 เป็นผู้ที่มี HL เพียงพอ ข. GS2 หรือความเข้าใจเอกสารทางสุขภาพ (6 ข้อ) พิสัยของคะแนน คือ 0-6 ในที่นี้ถือว่า ผู้ที่ได้คะแนนเต็ม 6 เป็นผู้ที่มี HL เพียงพอ และ ค. GS3 หรือ THLA-N8 ซึ่งงานวิจัยของปริณา ณ พัทลุงพบว่า คะแนน THLA-N8 ที่มากกว่า 4.83 ถือว่ามี HL เพียงพอ

การวิเคราะห์โค้ง ROC

ในการวิเคราะห์โค้ง ROC เกณฑ์คะแนนของ THLA-W+ ที่เหมาะสม คือ คะแนนที่ทำให้มีความไวบวก ความจำเพาะมีค่าสูงสุด (14) ค่าพื้นที่ใต้โค้ง ROC (area under the curve: AUC) มากกว่า 0.9, 0.7-0.9 และ 0.5-0.7 แสดงว่า แบบวัดสามารถแยกแยะระหว่างผู้ที่มี HL ที่เพียงพอ/ไม่เพียงพอออกจากกันได้ดี ปานกลาง และน้อย ตามลำดับ (15) ส่วนแบบวัดที่มี LR (likelihood ratio) มากกว่า 10.0/น้อยกว่า 0.10, 5-10/0.10-0.20 และ 2.0-5.0/0.20-0.50 สามารถให้ข้อมูลที่เป็นประโยชน์ต่อการตัดสินใจทางคลินิกมาก ปานกลาง และน้อย (แต่อาจมีความสำคัญทางคลินิก) ตามลำดับ ส่วน LR ระหว่าง 0.50-2.0 ให้ข้อมูลที่ไม่เป็นประโยชน์ (16)

ผลการวิจัย

คุณลักษณะของตัวอย่าง

ตัวอย่างทั้งหมด 767 คน เป็นเพศชายจำนวน 384 คน (ร้อยละ 50.1) ตัวอย่างส่วนใหญ่มีอายุอยู่ในช่วง 36-50 ปี (ร้อยละ 55.60) มีสถานะสมรส (ร้อยละ 63.4) ประกอบอาชีพค้าขาย (ร้อยละ 37.9) รองลงมา คือ รับจ้าง (ร้อยละ 28.4) และรับราชการ/รัฐวิสาหกิจ (ร้อยละ 15.3) การศึกษาส่วนใหญ่อยู่ในระดับ ปวส./ปริญญาตรีหรือสูงกว่า (ร้อยละ 41.7) รองลงมา คือ ระดับมัธยมศึกษาปลาย/ปวช. (ร้อยละ 23.7) ตัวอย่างร้อยละ 71.7 ศึกษาในโรงเรียนสายสามัญ ร้อยละ 23.9 ศึกษาในโรงเรียนเอกชนสอนศาสนา และร้อยละ 4.4 ศึกษาในสถานศึกษาปอเนาะที่สอนเฉพาะความรู้ทางศาสนา ตัวอย่างร้อยละ 68.7 ใช้ภาษามลายูในการสื่อสารในชีวิตประจำวัน และร้อยละ 28.7 ใช้ทั้งภาษามลายูและภาษาไทยในการสื่อสาร ตัวอย่างร้อยละ 2.0 เห็นว่าตนเองมีความสามารถในการอ่านไม่ดี เมื่อถามด้วยคำถามคัดกรองผู้ที่มี HL ไม่เพียงพอพบว่า ตัวอย่างร้อยละ 4.3 รายงานว่าไม่เข้าใจเกี่ยวกับอาการของตนเองจากการอ่านเอกสารที่ได้รับ ร้อยละ 3.8 ไม่มั่นใจในการกรอกแบบฟอร์มต่าง ๆ ของโรงพยาบาล และร้อยละ 46.9 เคยขอให้ผู้อื่นช่วยอ่านเอกสารของโรงพยาบาล ตัวอย่างร้อยละ 32.7 ในการศึกษาครั้งนี้เป็นผู้ป่วยในคลินิกโรคไม่ติดต่อเรื้อรัง

คะแนน THLA-W+ และความเที่ยง

ผลการคำนวณคะแนน THLA-W+ ทั้ง 3 แบบพบว่า การกระจายคะแนนของ THLA-W+R มีลักษณะเบ้

ซ้าย โดยคะแนนกระจุกอยู่ทางด้านคะแนนมาก ส่วนคะแนน THLA-W+C และ THLA-W+RC มีการกระจายตัวมากกว่า อย่างเห็นได้ชัด THLA-W+R มีคะแนนเฉลี่ยมากกว่า THLA-W+C และ THLA-W+RC โดยมีคะแนนเฉลี่ยเท่ากับ 42.47 ± 11.04 , 27.48 ± 13.95 และ 26.89 ± 13.96 ตามลำดับ (จากคะแนนเต็ม 48) แสดงว่าการเพิ่มคำถามทดสอบความเข้าใจทำให้แบบทดสอบมีความยากขึ้น ส่งผลให้คะแนนมีการกระจายตัว หรือ SD มากขึ้น ซึ่งบ่งบอกว่า การทดสอบความเข้าใจทำให้สามารถแยกผู้ที่มี HL แตกต่างกันได้ดีมากขึ้น สำหรับเวลาที่ใช้ในการทดสอบพบว่า ตัวอย่างร้อยละ 91.4 ใช้เวลาน้อยกว่า 8 นาที เวลาที่ใช้เฉลี่ย คือ 6.07 ± 2.09 นาที

THLA-W+ ทั้ง 3 แบบมีค่า Cronbach's alpha อยู่ในช่วง 0.97-0.98 ซึ่งถือว่ามีความเที่ยงสูง ส่วน THLA-N8 มีค่า Cronbach's alpha 0.70 ถือว่ามีความเที่ยงเป็นที่น่าพอใจ

THLA-W+ ระหว่างระดับการศึกษาที่ต่างกัน

ผลการเปรียบเทียบคะแนน THLA-W+ ระหว่างระดับการศึกษาที่ต่างกัน 4 ระดับ แสดงในตารางที่ 1 คะแนน THLA-W+ ในแต่ละระดับการศึกษาไม่ได้มีความเท่าเทียมกันจากการคำนวณคะแนนทั้ง 3 แบบ ($P < 0.001$) เมื่อเปรียบเทียบรายคู่พบว่า THLA-W+R สามารถจำแนกเฉพาะกลุ่มประถมศึกษาหรือต่ำกว่า ออกจากกลุ่มการศึกษาอื่น ๆ ได้ แต่ไม่สามารถจำแนกกลุ่มมัธยมศึกษาปีที่ 1-3 มัธยมศึกษาตอนปลาย/ปวช. และ ปวส./ปริญญาตรีหรือสูงกว่าออกจากกันได้ ส่วนคะแนน THLA-W+C และ THLA-W+RC สามารถจำแนกผู้ที่มีระดับการศึกษาทั้ง 4 ระดับออกจากกันได้ดีกว่า THLA-W+R ผลการศึกษานี้เป็นหลักฐานชิ้นหนึ่งแสดงถึงความตรงของแบบวัด THLA-W+ โดยเฉพาะอย่างยิ่ง THLA-W+C และ THLA-W+RC

ความตรงของ THLA-W+

THLA-W+C และ THLA-W+RC มีความสัมพันธ์ในระดับที่สูงมาก ($r = 0.99$) บ่งบอกว่าสามารถใช้ THLA-W+C แทน THLA-W+RC ได้ นั่นคือ สามารถใช้แบบทดสอบนี้โดยให้ตัวอย่างตอบคำถามทดสอบความเข้าใจเพียงอย่างเดียวและไม่ต้องทดสอบการอ่านคำ ซึ่งจะทำให้การใช้แบบวัดสะดวกมากขึ้น ส่วน THLA-W+R มีความสัมพันธ์กับ THLA-W+C และ THLA-W+RC โดย $r =$

ตารางที่ 1. การวิเคราะห์ความแปรปรวนทางเดียวเพื่อเปรียบเทียบคะแนน THLA-W+ ในระดับการศึกษาต่าง ๆ

ระดับการศึกษา	จำนวน	ค่าเฉลี่ย±SD (คะแนนเต็ม 48 คะแนน) ³		
		THLA-W+R	THLA-W+C	THLA-W+RC
ประถมศึกษาหรือต่ำกว่า	125	20.97±11.68 ^a	9.28±12.23 ^a	7.72±10.35 ^a
มัธยมศึกษาปีที่ 1-3	141	46.30±5.00 ^b	15.28±8.16 ^b	14.98±8.0 ^b
มัธยมศึกษาปีที่ 4-6/ปวช. ¹	182	46.26±3.76 ^b	32.54±7.98 ^c	32.05±8.32 ^c
ปวส. ² /ปริญญาตรีหรือสูงกว่า	319	47.03±1.94 ^b	37.12±6.38 ^d	36.71±6.38 ^d
Levene test		F=114.04, df=3, 763 P<0.001	F=21.67, df=3, 76 P<0.001	F=10.98, df=3, 763 P<0.001
Brown-Fortsynthe test		F=468.12, df=3, 200 P<0.001	F=382.36, df=3, 389 P<0.001	F=464.86, df=3, 46 P<0.001

1: ปวช. คือ ประกาศนียบัตรวิชาชีพ

2: ปวส. คือ ประกาศนียบัตรวิชาชีพชั้นสูง

3: ค่าเฉลี่ยที่มีอักษรตัวย่อเหมือนกันหมายถึงไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติเมื่อทดสอบ Games-Howell

0.64 และ 0.68 ตามลำดับ ซึ่งถือว่ามีความสัมพันธ์ในระดับที่สูง เป็นการสะท้อนว่าวิธีการทดสอบทั้งสอง (การทดสอบการอ่านและการทดสอบความเข้าใจ) วัดตัวแปรตัวเดียวกันหรือใกล้เคียงกัน

ค่าสัมประสิทธิ์สหสัมพันธ์ (r) ระหว่าง THLA-W+ กับตัวบ่งชี้ HL ต่าง ๆ เป็นบวกสอดคล้องกับการทำนายทางทฤษฎีและมีนัยสำคัญทางสถิติ (ตารางที่ 2) ซึ่งบ่งชี้ถึงความตรงของแบบวัด โดยเฉพาะอย่างยิ่ง ค่า r ระหว่าง

THLA-N8 และ แบบคัดกรองผู้ที่มี HL ไม่เพียงพอกับ THLA-W+C และ THLA-W+RC ซึ่งมีค่าสูงเกินกว่า 0.70

ค่า r ระหว่าง THLA-W+C และ THLA-W+RC กับตัวบ่งชี้ HL ต่าง ๆ มีค่าใกล้เคียงกันมาก บ่งบอกความตรงในระดับเดียวกัน ส่วนค่า r ของ THLA-W+R มีค่าต่ำกว่า นั่นคือ การทดสอบความเข้าใจทำให้แบบวัดมีความตรงมากกว่าการทดสอบการอ่านเพียงอย่างเดียว

ตารางที่ 2. ค่าสัมประสิทธิ์สหสัมพันธ์ของ THLA-W+R, THLA-W+C และ THLA-W+RC กับตัวแปรต่าง ๆ ที่บ่งชี้ HL¹ โดยเปรียบเทียบกับผลการศึกษาของวิทยา พันธุ์ทอง (8) และผดุง จันชูโต (9)

แบบวัด	ความสามารถในการอ่าน	ความเข้าใจเอกสารสุขภาพ	THLA-N8	แบบคัดกรองผู้ที่มี HL ไม่เพียงพอ
การศึกษาครั้งนี้				
THLA-W+R	0.24	0.22	0.46	0.55
THLA-W+C	0.31	0.31	0.73	0.79
THLA-W+RC	0.31	0.32	0.74	0.82
THLA-W จากงานของวิทยา พันธุ์ทอง (8)	0.57	-	-	0.54
การศึกษาของผดุง จันชูโต (9)				
THLA-W+R	-0.039 (P=0.497)	-	-	0.42
THLA-W+C	0.22	-	-	0.46
THLA-W+RC	0.20	-	-	0.48

1: ค่าสัมประสิทธิ์สหสัมพันธ์ทุกตัวมีค่า P<0.05 ยกเว้นที่ระบุค่า P ในตาราง

ตารางที่ 3. คุณสมบัติของ THLA-W+R, THLA-W+C และ THLA-W+RC ที่จุดตัดคะแนน 47, 37 และ 37 ตามลำดับ

ตัววัดมาตรฐานและแบบวัด	AUC	ความไว	ความจำเพาะ	ความถูกต้อง	PPV	NPV	LR+	LR-
GS1 เป็นตัววัดมาตรฐาน								
THLA-W+R	0.66	41.90	87.18	44.20	98.39	7.44	3.27	0.67
THLA-W+C	0.82	68.68	82.05	69.36	98.62	12.31	3.83	0.38
THLA-W+RC	0.83	70.74	79.49	71.19	98.47	12.70	3.45	0.37
GS2 เป็นตัววัดมาตรฐาน								
THLA-W+R	0.66	42.07	88.10	44.59	98.39	8.10	3.53	0.66
THLA-W+C	0.82	68.97	83.33	69.75	98.62	13.46	4.14	0.37
THLA-W+RC	0.83	71.03	80.95	71.58	98.47	13.93	3.73	0.36
GS3 เป็นตัววัดมาตรฐาน								
THLA-W+R	0.64	46.52	76.21	54.50	84.19	34.35	1.96	0.70
THLA-W+C	0.83	78.61	67.96	75.75	86.98	53.85	2.45	0.31
THLA-W+RC	0.83	80.75	66.02	76.79	86.62	55.74	2.38	0.29

1: GS1 คือ คะแนนรวมของความสามารถในการอ่านและความเข้าใจเอกสารทางสุขภาพ รวม 11 ข้อ ผู้ที่ได้คะแนนเต็ม 11 ถือเป็น ผู้ที่มี HL เพียงพอ; GS2 คือ คะแนนความเข้าใจเอกสารทางสุขภาพ รวม 6 ข้อ ผู้ที่ได้คะแนนเต็ม 6 ถือเป็น ผู้ที่มี HL เพียงพอ; GS3 คือ THLA-N8 ผู้ที่มีคะแนน > 4.83 ถือเป็น ผู้ที่มี HL เพียงพอ

พื้นที่ใต้โค้ง ROC และจุดตัดคะแนน

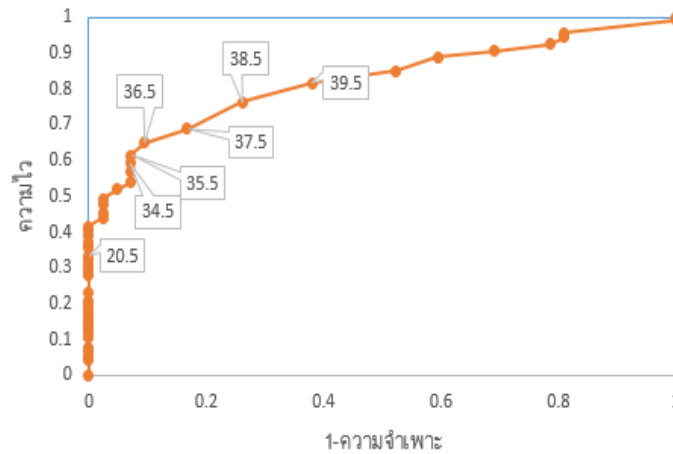
การวิเคราะห์โค้ง ROC โดยใช้ GS1, GS2 และ GS3 เป็นตัววัดมาตรฐานพบว่า AUC ซึ่งบ่งบอกถึงความสามารถในการจำแนกผู้ที่มี HL ที่เพียงพอและไม่เพียงพอออกจากกัน มีค่าเท่ากับ 0.64-0.66 (ตารางที่ 3) สำหรับ THLA-W+R (ขึ้นกับตัววัดมาตรฐาน) ซึ่งถือว่าอยู่ในระดับน้อยถึงปานกลาง (15) THLA-W+C และ THLA-W+RC มี AUC ที่มากกว่า คือ 0.82-0.83 และ 0.82-0.83 ตามลำดับ (ขึ้นกับตัววัดมาตรฐาน) ซึ่งอยู่ในระดับปานกลางถึงดี ผลการวิจัยบ่งบอกว่า THLA-W+C และ THLA-W+RC มีความสามารถในการจำแนกระดับ HL ดีกว่า THLA-W+R

จุดตัดคะแนนที่เหมาะสมของ THLA-W+R เมื่อใช้ GS1, GS2 และ GS3 เป็นตัววัดมาตรฐาน คือ 47 (จากคะแนนเต็ม 48) ส่วนจุดตัดคะแนนของ THLA-W+C คือ 38, 38 และ 36 เมื่อใช้ GS1, GS2 และ GS3 เป็นตัววัดมาตรฐาน ตามลำดับ ดังนั้นจึงเลือก 37 เป็นจุดตัดคะแนน เพราะเป็นค่าระหว่างกลาง (รูปที่ 1) สำหรับ THLA-W+RC จุดตัดคะแนน คือ 37, 37 และ 36 เมื่อใช้ GS1, GS2 และ

GS3 เป็นตัววัดมาตรฐาน ตามลำดับ จึงเลือกคะแนน 37 เป็นจุดตัดคะแนน ดังนั้น จึงถือว่าตัวอย่างที่ได้คะแนน THLA-W+C หรือ THLA-W+RC \geq 37 มีระดับ HL ที่ไม่เพียงพอ

ความไวและความจำเพาะ

ความไว คือ สัดส่วนของผู้ที่มี HL ไม่เพียงพอที่ได้คะแนน THLA-W+ \leq ค่าจุดตัดคะแนน (true positive rate) ส่วนความจำเพาะ คือ สัดส่วนของผู้ที่มี HL เพียงพอที่ได้คะแนน THLA-W+ > ค่าจุดตัดคะแนน (true negative rate) จากตารางที่ 3 THLA-W+R ที่จุดตัดคะแนน 47 มีความไว น้อย คือ ร้อยละ 41.90-46.52 (ขึ้นกับตัววัดมาตรฐาน) ส่วน THLA-W+C และ THLA-W+RC มีความไวที่ดี (ร้อยละ 68.68-78.61 และร้อยละ 70.74-80.75 ตามลำดับ) และสูงกว่า THLA-W+R แต่ความจำเพาะของ THLA-W+R, THLA-W+C และ THLA-W+RC ไม่ต่างกันนัก คือ ร้อยละ 76.21-88.10, 67.96-83.33 และ 66.02-80.95 ตามลำดับ) หากบุคลากรทางการแพทย์ต้องการค้นหาผู้ป่วยที่มี HL ไม่



รูปที่ 1. โค้ง ROC ของ THLA-W+C เมื่อใช้ GS2 เป็นตัววัดมาตรฐาน

หมายเหตุ: โค้ง ROC ของ THLA-W+C และ THLA-W+RC ในทั้งสามตัววัดมาตรฐานแทบจะทับซ้อนกันเป็นเส้นเดียวกัน ในที่นี้จึงแสดงเฉพาะโค้งจาก THLA-W+C เมื่อใช้ GS2 เป็นตัววัดมาตรฐาน

เพียงพอ เพื่อให้คำแนะนำที่เหมาะสมกับผู้รับบริการนั้น ๆ การเลือกแบบวัดในวัตถุประสงค์นี้ ควรให้ความสำคัญกับความไว

ความสามารถในการทำนาย

ความถูกต้องในการทำนาย (accuracy) เป็นค่าที่แสดงว่า แบบวัดทำนายการมี HL เพียงพอหรือไม่เพียงพอถูกต้องมากน้อยเพียงไร จากตารางที่ 3 พบว่า THLA-W+R มีความถูกต้องในการทำนายร้อยละ 44.20-54.50 ขึ้นกับตัววัดมาตรฐาน ซึ่งน้อยกว่าความถูกต้องในการทำนายของ THLA-W+C และ THLA-W+RC (ร้อยละ 69.36-75.75 และ 71.19-76.79 ตามลำดับ)

positive predictive value (PPV) คือ ความถูกต้องในการทำนายเมื่อผลการวัดโดย THLA-W+ ระบุว่า ตัวอย่างมี HL ไม่เพียงพอ การคำนวณคะแนน THLA-W+ ทั้งสามรูปแบบมี PPV ที่สูง (ร้อยละ 84.19-98.62) ส่วน negative predictive value (NPV) คือ ความถูกต้องในการทำนายเมื่อผลการวัดโดย THLA-W+ ระบุว่า ตัวอย่างมี HL เพียงพอ THLA-W+ ที่คำนวณจากคะแนนทั้ง 3 แบบมี NPV ต่ำ (ร้อยละ 7.44-55.74)

THLA-W+C และ THLA-W+RC มีค่า LR+ เกือบทั้งหมดมากกว่า 2 แต่น้อยกว่า 5 และมี LR- มีค่าน้อยกว่า 0.5 ซึ่งบ่งชี้ว่า แบบวัดอาจมีประโยชน์ต่อการตัดสินใจในทางคลินิก สำหรับ THLA-W+R พบว่า LR- มีค่ามากกว่า 0.5 และ LR+ บางตัวน้อยกว่า 2 ซึ่งบ่งชี้ว่า แบบวัดไม่มีประโยชน์ต่อการตัดสินใจในทางคลินิก

การอภิปรายผล

การวิจัยในอดีตที่คล้ายกับการศึกษาครั้งนี้ คือ การศึกษาของวิทยา พันธุ์ทอง ที่ทดสอบแบบวัด THLA-W โดยการอ่านคำเพียงอย่างเดียวในผู้ป่วยของโรงพยาบาลป่าดงเบงชาร์ท จังหวัดสงขลา (8) โดย และการศึกษาของผดุง จันชูโต (9) ที่ใช้แบบวัด THLA-W+ เช่นเดียวกับการศึกษาครั้งนี้ โดยทดสอบในผู้ป่วยนอกของโรงพยาบาลสะเดา จังหวัดสงขลา การศึกษาครั้งนี้ในชาวมุสลิมและในอดีต (9) ได้ผลตรงกันว่า THLA-W+R มีการกระจายแบบเบ้ซ้าย โดยคะแนนกระจุกตัวอยู่ในช่วงคะแนน 46-48 คะแนน (จากคะแนนเต็ม 48) คะแนนเฉลี่ยของ THLA-W+R ในการศึกษานี้เท่ากับ 42.47 ± 11.04 ซึ่งใกล้เคียงกับที่พบในงานวิจัยของวิทยา พันธุ์ทอง (8) (คะแนนเฉลี่ย 43.17 ± 5.99) และงานวิจัยของผดุง จันชูโต (9) (คะแนนเฉลี่ย 46.04 ± 4.19) ตัวอย่างในงานวิจัยของวิทยา พันธุ์ทอง (8) ได้คะแนนการอ่านน้อยกว่างานวิจัยของผดุง จันชูโต (9) ถึงแม้ว่าจะเป็นตัวอย่างในจังหวัดสงขลาเหมือนกันและมีลักษณะของตัวอย่างที่คล้ายกัน ซึ่งอาจเนื่องมาจากงานวิจัยของวิทยา พันธุ์ทอง (8) เป็นการทดสอบการอ่านคำเพียงอย่างเดียวอย่างต่อเนื่อง จึงคล้ายกับการทดสอบในเวลาจำกัดหรือ speed test ทำให้ตัวอย่างอ่านออกเสียงผิดได้ง่าย แต่งานวิจัยของผดุง จันชูโต (9) ทดสอบการอ่านคำและทดสอบความเข้าใจในแต่ละคำควบคู่กันไป ทำให้จังหวะในการทดสอบการอ่านช้าลง จึงทำให้ตัวอย่างมีคะแนนการทดสอบการอ่านคำที่สูงขึ้น ตัวอย่างในการศึกษานี้ผ่านกระบวนการทดสอบเหมือนกับการวิจัยของผดุง จันชูโต (9)

แต่ได้คะแนนน้อยกว่า ซึ่งอาจเกิดจากการมีระดับ HL ที่น้อยกว่าหรือมีทักษะทางภาษาไทยที่น้อยกว่าเนื่องจากชาวมุสลิมจำนวนมากใช้ภาษามลายูสื่อสารในชีวิตประจำวัน

การเพิ่มคำถามทดสอบความเข้าใจทำให้แบบวัดมีความยากขึ้น THLA-W+C และ THLA-W+RC ในการศึกษาครั้งนี้มีค่าเฉลี่ยคือ 27.48 ± 13.95 และ 26.89 ± 13.96 ตามลำดับ ซึ่งน้อยกว่า THLA-W+R (42.47 ± 11.04) การศึกษาของผดุง จันชูโต (9) มีค่าเฉลี่ยของ THLA-W+C และ THLA-W+RC สูงกว่าในการศึกษาคครั้งนี้ คือ 35.02 ± 7.78 และ 34.56 ± 7.95 ตามลำดับซึ่ง อาจเกิดจากตัวอย่างชาวมุสลิมในการศึกษาคครั้งนี้มีทักษะทางภาษาที่น้อยกว่าหรือมีระดับ HL ที่น้อยกว่า

เวลาที่ใช้ในการทดสอบด้วย THLA-W ในงานวิจัยของวิทยา พันธุ์ทอง (8) ที่ทดสอบการอ่านคำ คือ 1.41 ± 0.73 นาที ซึ่งใกล้เคียงกับแบบประเมิน REALM-S ที่ทดสอบการอ่านคำจำนวน 66 คำที่ใช้เวลาเพียง 1-2 นาที (3) ส่วน THLA-W+C และ THLA-W+RC ในงานวิจัยของผดุง จันชูโต (9) ใช้เวลา 7.86 ± 2.12 นาที ส่วนในการศึกษาคครั้งนี้ใช้เวลา 6.07 ± 2.09 ซึ่งใกล้เคียงกัน

ความเที่ยงของ THLA-W+ จากการคำนวณคะแนนทั้ง 3 แบบของการศึกษาคครั้งนี้มีค่า 0.97-0.98 ซึ่งใกล้เคียงกับงานวิจัยของผดุง จันชูโต (0.89-0.92) และวิทยา พันธุ์ทอง (0.91) จึงสรุปได้ว่า ความเที่ยงของ THLA-W+ ในทุกรูปแบบของการคำนวณมีค่าสูงในระดับที่ยอมรับได้เมื่อทดสอบในชาวไทยมุสลิม

ความตรงของแบบวัดในชาวไทยมุสลิมเห็นได้จากการที่ผู้ที่มีระดับการศึกษาสูงกว่ามีคะแนน THLA-W+ ที่มากกว่า (ในทุกรูปแบบของการคำนวณคะแนน) THLA-W+C และ THLA-W+RC มีความตรงมากกว่า THLA-W+R เนื่องจากสามารถจำแนกความแตกต่างของผู้ที่อยู่ใน 4 ระดับการศึกษาที่ต่างกันได้ ส่วน THLA-W+R ไม่สามารถแยกกลุ่มที่มีระดับการศึกษามากกว่าชั้นประถมศึกษาออกจากกันได้

เมื่อพิจารณาค่า r ของ THLA-W+ กับตัวชี้วัด HL ต่าง ๆ (ความสามารถในการอ่าน ความเข้าใจในเอกสารทางสุขภาพ คำถามคัดกรอง HL และ THLA-N8) มีค่าเป็นบวกและมีนัยสำคัญทางสถิติ ($P < 0.05$) ซึ่งบ่งชี้ถึงความตรงของแบบวัด โดยเฉพาะอย่างยิ่ง THLA-W+C และ THLA-W+RC มีความสัมพันธ์กับคำถามคัดกรอง HL และ THLA-N8 สูงมาก ($r > 0.70$) จากตารางที่ 2 พบว่า r ของ THLA-

W+ ทั้ง 3 แบบกับตัวชี้วัด HL ในการศึกษาครั้งนี้โดยรวมไม่ต่างจากงานวิจัยของผดุง จันชูโต (9) มากนัก แต่ THLA-W+R ในการศึกษาครั้งนี้มีความตรงในระดับที่น้อยกว่าในการศึกษาของวิทยา พันธุ์ทอง (r ของความสามารถในการอ่านกับแบบวัด คือ 0.24 และ 0.57 ตามลำดับ) การสูญเสียลักษณะของ speed test ใน THLA-W+R อาจอธิบายถึงผลการศึกษานี้ได้

ผู้ที่ตอบคำถามวัดการอ่านได้ถูกมีคะแนน THLA-W+R, THLA-W+C และ THLA-W+RC มากกว่าผู้ที่ตอบผิด ลักษณะดังกล่าวพบในคำถามจำนวน 4 ข้อจากทั้งหมด 5 ข้อ ผู้ที่ตอบคำถามวัดความเข้าใจในเอกสารสุขภาพถูกมีคะแนน THLA-W+R, THLA-W+C และ THLA-W+RC มากกว่าผู้ที่ตอบผิด ลักษณะดังกล่าวพบในคำถามจำนวน 5, 3 และ 4 ข้อจากทั้งหมด 6 ข้อ ตามลำดับ ผลการศึกษานี้คล้ายกับที่พบในงานวิจัยของผดุง จันชูโต (9)

ในการศึกษาคครั้งนี้และงานวิจัยของผดุง จันชูโต (9) พบว่า THLA-W+C และ THLA-W+RC มีความตรงมากกว่า THLA-W+R เนื่องจากการเพิ่มคำถามทดสอบความเข้าใจในคำทำให้ความสามารถในการจำแนกความแตกต่างของ HL ดีมากขึ้น อย่างไรก็ตาม THLA-W+C และ THLA-W+RC ในการศึกษาครั้งนี้ มีความตรงมากกว่าในงานวิจัยของผดุง จันชูโต (9) ทั้งนี้เนื่องจากหลักการสำคัญของแบบวัด THLA-W+ คือ การทดสอบการอ่านคำภาษาไทยและการทดสอบความเข้าใจในคำภาษาไทย แบบวัดจึงมีความตรงหรือใช้ได้ดีในผู้ที่มีความสามารถจำกัดในภาษาไทย

ผลการวิจัยสรุปได้ว่า THLA-W+C มีความตรงมากกว่า THLA-W+R ส่วน THLA-W+C และ THLA-W+RC มีค่า r สูงถึง 0.987 ดังนั้น จึงสามารถใช้ THLA-W+C แทน THLA-W+RC ได้ในการประเมิน HL ในผู้ป่วยชาวไทยมุสลิม

การวิเคราะห์โค้ง ROC ในการศึกษาครั้งนี้พบว่า AUC ของแบบวัด THLA-W+R มีค่าน้อยกว่า AUC ของ THLA-W+C และ THLA-W+RC ซึ่งสอดคล้องกับผลการวิจัยของผดุง จันชูโต (9) แต่อย่างไรก็ตาม ไม่อาจเปรียบเทียบขนาดของ AUC จากการศึกษาครั้งนี้กับงานวิจัยของผดุง จันชูโต (9) ได้เนื่องจากใช้ตัววัดมาตรฐานในลักษณะที่แตกต่างกัน

จุดตัดคะแนนของ THLA-W+R, THLA-W+C และ THLA-W+RC ในการศึกษาครั้งนี้ คือ 47, 37, และ 37 คะแนน

ตามลำดับ ทั้งยังได้ค่าใกล้เคียงกับจุดตัดคะแนนในงานวิจัยของผดุง จันชูโต (9) คือ 4537, และ 37 คะแนน ตามลำดับ การศึกษาครั้งนี้และงานวิจัยของผดุง จันชูโต (9) พบว่า THLA-W+R มีความไวน้อยกว่า แต่มีความจำเพาะมากกว่า THLA-W+C และ THLA-W+RC อย่างไรก็ตาม ไม่อาจเปรียบเทียบความตรงและความจำเพาะของการศึกษาี้กับงานวิจัยของผดุง จันชูโต (9) หรือวิทยา พันธุ์ทอง (8) ได้ เนื่องจากใช้ตัววัดมาตรฐานในลักษณะที่แตกต่างกัน

THLA-W+ มุ่งเน้นใช้เพื่อค้นหาผู้ที่มี HL ไม่เพียงพอ ดังนั้น จึงควรให้น้ำหนักกับความไวมากกว่าความจำเพาะ นั่นคือ การค้นหาผู้ที่มี HL ไม่เพียงพอให้พบนั้นมีความสำคัญในการดูแลรักษาผู้ป่วย แต่การทำนายผู้ที่มี HL เพียงพออย่างถูกต้อง (ทำนายผู้ป่วยว่ามี HL เพียงพออย่างถูกต้อง) นั้น เกิดผลกระทบเชิงลบต่อผู้ป่วยน้อยกว่า ดังนั้นจึงอาจยอมให้มีค่าความจำเพาะต่ำได้บ้าง ดังนั้นการคำนวณคะแนนที่แนะนำ คือ THLA-W+C

แบบวัด THLA-W+ มีความยาวถึง 48 ข้อ การวิจัยในอนาคตควรพัฒนาแบบวัดฉบับย่อโดยตัดคำถามข้อที่ตัวอย่างเกือบทุกคนตอบได้ออก การวิจัยครั้งนี้ศึกษาในตัวอย่างชาวไทยมุสลิมในอำเภอยะหริ่ง จังหวัดปัตตานี ดังนั้นผลที่ได้จากการวิจัยครั้งนี้ เช่น เกณฑ์คะแนน ความตรง และความเที่ยง อาจมีความจำเพาะกับกลุ่มประชากรที่วิจัยเท่านั้น ดังนั้นจึงควรเพิ่มพื้นที่ในการศึกษาเพื่อยืนยันผลการวิจัย

สรุปผล

การทดสอบ THLA-W+R, THLA-W+C และ THLA-W+RC ในชาวไทยมุสลิมพบว่า มีค่าความเที่ยงสูงและอยู่ในระดับใกล้เคียงกัน THLA-W+C คำนวณคะแนนจากการทดสอบความเข้าใจเพียงอย่างเดียวโดยไม่ต้องทดสอบการอ่าน เป็นทางเลือกที่น่าสนใจ เนื่องจากทำให้การทดสอบสะดวกจากการที่ไม่ต้องใช้ผู้ตัดสินความถูกต้องของการออกเสียง และยังช่วยลดระยะเวลาของการทดสอบได้อีกด้วย THLA-W+C มีความตรง ความไว ความจำเพาะ AUC, accuracy, PPV, NPV, LR+ และ LR ในระดับเดียวกับ THLA-W+RC ทั้ง THLA-W+C และ THLA-W+RC มีคุณสมบัติการวัดและการทำนายที่ดีกว่า THLA-W+R ดังนั้น จึงสรุปได้ว่า การทดสอบ THLA-W+ ในผู้ป่วยชาวไทยมุสลิมควรทำเฉพาะการทดสอบความเข้าใจเพียงอย่างเดียว

เอกสารอ้างอิง

1. Sørensen K, den Broeck SV, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health* 2012; 25: 12: 80.
2. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Viera A, Crotty K, et al. Health literacy interventions and outcomes: an updated systematic review. Evidence Report/ Technology Assessment No. 199. Rockville, MD. Agency for Healthcare Research and Quality; 2011.
3. Davis TC, Crouch M, Wills G, Abdehou D. Rapid assessment of literacy levels of adult primary care patients. *Fam Med* 1991; 23: 433-55.
4. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *J Gen Intern Med* 1995; 10: 537-41.
5. Weiss BD, Mays MZ, Martz W, Casto KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. *Ann Fam Med* 2005; 3: 514-22.
6. Health Education Division, Department of Health Service Support, Ministry of Public Health. Guidelines for the development of health literacy to change behavior and reduce risk. Nonthaburi: Ministry of Public Health, 2013.
7. Jindawong B. Validity and reliability of the Thai version of health literacy screening tools for patients at Srinagarind Hospital, Khon Kaen, Thailand. [master thesis]. Khon Kaen: Khon Kaen University; 2013.
8. Phantong W. Development of the Thai Health Literacy Assessment using Word List (THLA-W). [master thesis]. Songkla: Prince of Songkla University; 2016.
9. Chanchuto P. Development of the Thai Health Literacy Assessment Using Word List with extended questions to test comprehension (THLA-W+): testing

- in Sadao Hospital. [master thesis]. Songkla: Prince of Songkla University; 2017.
10. Lee SY, Bender DE, Ruiz RE, Cho YI. Development of an easy-to-use Spanish health literacy test. *Health Serv Res* 2006; 41: 1392-412.
 11. Na Phatthalung P, Lerkiatbundit S. Development of Thai health literacy assessment based on the assessment of ability to use nutrition label. *Thai Journal of Pharmacy Practice* 2019; 11: 659-77.
 12. Office of the Private Schools within 5 Southern Border Provinces. Comparison of the average score on ordinary national educational test [online]. 2017 [cited Apr 16, 2018]. Available from: www.skprivate.go.th/uploads/information/9a33fc62f981f0a441b0cac91da055c6.xlsx.
 13. Faul GF, Erdfelder E, Buchner A and Lang AG. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods* 2009; 41:1149-60.
 14. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950; 3: 32-5.
 15. Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988; 240:1285-93.
 16. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*. 1994; 271: 703-7.