

การพัฒนาแบบวัดความแตกฉานทางสุขภาพสำหรับชาวไทยโดยประเมินจากทักษะที่สำคัญ

พิชญกร วัชรตะมกุล¹, สงวน ลือเกียรติบัณฑิต², วรณช แสงเจริญ³

¹กลุ่มงานเภสัชกรรม โรงพยาบาลหาดใหญ่

²ภาควิชาบริหารเภสัชกิจ คณะเภสัชศาสตร์ มหาวิทยาลัยสงขลานครินทร์

³ภาควิชาเภสัชกรรมคลินิก คณะเภสัชศาสตร์ มหาวิทยาลัยสงขลานครินทร์

บทคัดย่อ

วัตถุประสงค์: เพื่อพัฒนาแบบวัดความแตกฉานทางสุขภาพสำหรับชาวไทยโดยประเมินจากทักษะที่สำคัญ (Thai Health Literacy Assessment based on Skills: THLA-S) และหาเกณฑ์แปลผลระดับความแตกฉานทางสุขภาพ (health literacy: HL) ของแบบวัด **วิธีการ:** การพัฒนาแบบวัดประกอบด้วย 5 ขั้นตอน คือ 1) การวิเคราะห์ทักษะที่ประเมินโดยแบบวัด HLSI (Health Literacy Skills Instrument) ซึ่งเป็นต้นแบบ 2) การจัดทำแบบวัด THLA-S ฉบับเริ่มต้นที่วัดทักษะที่ใกล้เคียงกับ HLSI 3) การพิจารณาแบบวัดโดยผู้เชี่ยวชาญ 6 ท่านและการปรับปรุงตามข้อเสนอแนะ 4) การทดสอบความตรงและความเที่ยงในตัวอย่างขนาดใหญ่ในโรงพยาบาลศูนย์ โดยตัวอย่างเป็นผู้ป่วยนอกและญาติจำนวน 522 คนที่เลือกมาแบบตามสะดวก ตัวอย่างตอบแบบวัด THLA-S และแบบสอบถามที่วัดระดับของ HL (การรับรู้ทักษะของตนเองในการทำกิจกรรมที่จำเป็นต่อการมี HL ความสามารถในการอ่านจาก cloze test คะแนนจากแบบคัดกรองผู้มี HL ไม่เพียงพอ และความสามารถในการใช้ฉลากโภชนาการ) และ THLA-N (Thai Health Literacy Assessment Using Nutrition Label) และ 5) การหาเกณฑ์คะแนนด้วยการวิเคราะห์โค้ง receiver operating characteristic (ROC) โดยใช้คะแนนจาก THLA-N เป็น gold standard **ผลการวิจัย:** แบบวัด THLA-S มีค่า Cronbach's alpha เท่ากับ 0.81 แบบวัดมีความตรง เพราะ 1) ในการประเมินความสามารถของตนเองในการทำ 6 กิจกรรม (จากจำนวน 8 กิจกรรมที่จำเป็นต่อการมี HL ผู้ที่ตอบว่า ง่ายได้คะแนน THLA-S สูงกว่าผู้ที่ตอบว่ายาก 2) ผู้ที่มีระดับการศึกษาสูงกว่าได้คะแนน THLA-S มากกว่า 3) ความสัมพันธ์ของ THLA-S เป็นบวกและมีนัยสำคัญทางสถิติ ($P < 0.05$) กับกับตัวชี้วัด HL ต่อไปนี้ การรับรู้ความสามารถของตนเองในการทำกิจกรรมที่จำเป็นต่อการมี HL ความสามารถในการอ่านจาก cloze test คะแนนจากแบบคัดกรองผู้มี HL ไม่เพียงพอ ความสามารถในการใช้ฉลากโภชนาการ และ THLA-N โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ 0.34, 0.40, 0.16, 0.63 และ 0.61 ตามลำดับ การวิเคราะห์โค้ง ROC พบพื้นที่ใต้โค้ง ROC 0.84 คะแนน THLA-S ≤ 13 ถือว่ามี HL ไม่เพียงพอ ณ จุดตัดนี้ ได้ค่าความไว ร้อยละ 75.0 ค่าความจำเพาะ ร้อยละ 78.4 ค่า PPV ร้อยละ 79.12 ค่า NPV ร้อยละ 74.19 ค่า LR+ และค่า LR- ได้ค่า 3.47 และ 0.32 ตามลำดับ **สรุป:** แบบวัด THLA-S มีความตรง ความเที่ยง ความไว ความจำเพาะ และความสามารถในการจำแนกระดับ HL ได้ดีเป็นที่น่าพอใจ อย่างไรก็ตาม แบบวัดที่ได้นี้จำเป็นต้องพัฒนาต่อไป เพราะแบบวัดมีจำนวนคำถามมาก

คำสำคัญ: แบบวัด THLA-S ความตรง ความเที่ยง ความไว ความจำเพาะ ความแตกฉานทางสุขภาพ

รับต้นฉบับ: 10 พ.ย. 2561, ได้รับบทความฉบับปรับปรุง: 15 ธ.ค. 2561, รั้งลงตีพิมพ์: 13 ม.ค. 2562

ผู้ประสานงานบทความ: พิชญกร วัชรตะมกุล กลุ่มงานเภสัชกรรม โรงพยาบาลหาดใหญ่ อำเภอหาดใหญ่ จังหวัด.สงขลา 90110

E-mail: wawaphicha16@gmail.com

Development of the Thai Health Literacy Assessment Based on the Testing of Essential Skills

Phitchayakon Warotamakul¹, Sanguan Lerkiatbundit², Woranuch Saengcharoen³

¹Pharmacy Department, Hatyai Hospital

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

Objective: To develop the Thai Health Literacy Assessment based on Skills: THLA-S), and to determine the cut-off value to interpret the level of health literacy (HL). **Methods:** There were 5 steps in development of this measure: 1) analysis of skills assessed by the HLSI (Health Literacy Skills Instrument), the measurement model of THLA-S, 2) preparation of the first draft of the THLA-S for measuring the skills equivalent to those assessed by the HLSI, 3) review of the measure by 6 experts and revision of the measure accordingly, 4) validity and reliability testing in large samples in a regional hospital with 522 outpatients and their relatives chosen by accidental sampling. The subjects completed the THLA-S and questionnaire on various indicators of HL (self-perception on skills for activities necessary for HL, reading ability from cloze test, HL screening test, and ability to use nutritional label), and THLA-N (Thai Health Literacy Assessment Using Nutrition Label), 5) determination of cut-off of the THLA-S using THLA-N as gold standard. **Results:** The THLA-S showed a Cronbach's alpha of 0.81. The measure was valid because 1) In assessing self-perception on skills for activities necessary for HL, those reporting 6 activities (from all 8 activities) as easy had a higher level of the THLA-S than those reporting the tasks difficult 2) those with a higher education level showed a higher level of the THLA-S 3) relationships between the THLA-S scores with the following HL indicators were positive and statistically significant ($P < 0.05$)--self-perception on skills for activities necessary for HL, reading ability as measured by the cloze test, scores from screening tests for inadequate HL, ability to use nutrition labels and the THLA-N with correlation coefficients of 0.34, 0.40, 0.16, 0.63 and 0.61, respectively. Analysis of the ROC curve using the THLA-N as a gold standard showed an AUC of 0.84. The THLA-S scores less than or equal to 13 indicated inadequate HL. At this cut-off, the scale had a sensitivity of 75.0 percent, specificity of 78.4 percent, PPV of 79.12 percent and NPV of 74.19 percent, with LR+ and LR- of 3.47 and 0.32 respectively. **Conclusion:** The THLA-S is valid, reliable, sensitive, and specific with satisfying discriminating power of HL. However, the THLA-S still needs further improvement because of its large number of items.

Keywords: THLA-S, validity, reliability, sensitivity, specificity, health literacy

บทนำ

“ความแตกฉานทางสุขภาพ” (health literacy: HL) หมายถึง สิ่งที่ยोगกับความแตกฉานทั่วไป (literacy) ของบุคคลซึ่งประกอบด้วย ความรู้ ความต้องการ และความสามารถที่จำเป็นต้องมีเพื่อให้สามารถเข้าถึง เข้าใจ ประเมิน และใช้ข้อมูลสุขภาพเพื่อการตัดสินใจในชีวิตประจำวันในเรื่องที่เกี่ยวกับการรักษาพยาบาล การป้องกันโรค และการสร้างเสริมสุขภาพเพื่อเพิ่มคุณภาพชีวิต (1) ระดับ HL ที่ต่ำสัมพันธ์กับผลลัพธ์ทางสุขภาพที่ไม่ดีและการไม่ได้รับบริการทางสุขภาพที่เพียงพอ เช่น ผู้ที่มีระดับ HL ต่ำมีการเจ็บป่วย การนอนพักรักษาตัวในโรงพยาบาล การเข้ารับบริการฉุกเฉินของโรงพยาบาล และมีค่าใช้จ่ายทางการแพทย์ที่สูงกว่าผู้ที่มี HL ระดับปานกลางถึงสูง (2, 3) ซึ่งในกลุ่มผู้สูงอายุจะเห็นผลได้อย่างชัดเจน (4)

รายงานฉบับสำคัญชื่อ health literacy: a prescription to end confusion ของ Institute of Medicine (5) ให้คำแนะนำว่า ควรพัฒนาแบบประเมิน HL ขึ้นใหม่ให้เหมาะสมกับวัฒนธรรมและสามารถใช้สำรวจในประชากรได้ด้วย (6-8) โดยสนับสนุนให้มีการพัฒนาวิธีการวัดที่ครอบคลุม HL ในมิติต่าง ๆ มากขึ้น Health Literacy Skills Instrument (HLSI) (9) ถูกพัฒนาขึ้นตามคำแนะนำข้างต้น โดยมีเนื้อหาการวัดครอบคลุมทักษะ 4 ด้าน คือ ทักษะด้านสิ่งพิมพ์ (print literacy) ทักษะด้านตัวเลข (numeracy skills) ทักษะด้านการพูดและการฟัง (oral literacy skills) และทักษะการแสวงหาข้อมูล การประเมินใช้การทดสอบในรูปแบบข้อเขียนและไม่ใช้ข้อเขียน เช่น การใช้เสียงเป็นคำถามเพื่อทดสอบทักษะการฟังข้อมูลสุขภาพ การทดสอบการสืบค้นข้อมูลจากอินเทอร์เน็ต เป็นต้น นอกจากนี้เนื้อหาของคำถามใน HLSI ยังครอบคลุมประเด็นสุขภาพที่ประชาชนทั่วไปพบเจอได้ตลอดช่วงชีวิต ทั้งประเด็นการส่งเสริมสุขภาพ การป้องกันโรค การดูแลและรับการรักษาเมื่อเจ็บป่วย และการเข้าใช้ระบบบริการสุขภาพ

ประเทศไทยได้มีผู้พัฒนาแบบวัด HL เฉพาะกลุ่มโรคหรืออาการ เช่น แบบวัดความรู้ด้านสุขภาพตามหลัก 3อ 2ส สำหรับคนไทยกลุ่มเสี่ยงโรคเบาหวานและความดันโลหิตสูง (10) อย่างไรก็ตาม การวิจัยนี้มุ่งพัฒนาแบบวัด HL ชนิดทั่วไปเพื่อนำไปใช้ประเมินผู้ป่วยหรือประชาชนทั่วไปแบบไม่เจาะจงกลุ่มโรคหรืออาการ จึงไม่สามารถนำมาแบบวัดข้างต้นมาพัฒนาต่อยอดได้ แบบวัด HL ชนิดทั่วไปฉบับภาษาไทยในปัจจุบันส่วนหนึ่งพัฒนามาจาก

แบบวัดของต่างประเทศ เช่น การแปล Rapid Estimate of Adult Literacy in Medicine (REALM) เป็นภาษาไทยและทดสอบในผู้ป่วยโรงพยาบาลศรีนครินทร์จังหวัดขอนแก่น (11) แบบวัดประกอบด้วยคำเรียงตามจำนวนพยางค์จากน้อยไปมาก ซึ่งบ่งชี้ความยากในการออกเสียง คะแนน HL วัดจากจำนวนคำที่ผู้ถูกทดสอบอ่านได้ แบบวัดมีข้อดี คือ สามารถประเมิน HL ภายในเวลา 2-3 นาที อย่างไรก็ตามเมื่อแปลรายการคำใน REALM เป็นภาษาไทย ทำให้ความยากง่ายในการอ่านเปลี่ยนไป ทั้งยังพบ ceiling effect คือ ผู้ที่มีระดับการศึกษาที่ต่างกัน (สิ่งบ่งชี้ HL) ล้วนได้คะแนนในแบบวัดสูงและไม่ต่างกันมาก อีกทั้งแบบวัดฉบับภาษาไทยยังไม่ได้ผ่านการทดสอบความตรงโดยการหาความสัมพันธ์กับความสามารถของผู้ป่วยที่บ่งชี้ถึง HL เช่น การคำนวณขนาดยา การอ่านฉลากยา เป็นต้น นอกจากนี้ แบบวัดดังกล่าวยังประเมินเพียงความสามารถในการอ่านเท่านั้น จึงยังไม่มีข้อมูลความตรงและความเที่ยงที่มากพอ และยังไม่มีการกำหนดจุดตัดคะแนนที่ใช้ตัดสินระดับ HL

จากข้อดีของ REALM ที่ใช้งานง่าย และใช้เวลาทดสอบสั้น ทำให้มีการพัฒนาแบบคัดกรอง HL สำหรับชาวไทย ชนิตรายการคำ หรือ Thai Health Literacy Assessment Using Word List (THLA-W) (12) โดยอาศัยหลักการของ REALM จากการทดสอบในผู้ป่วยนอกของโรงพยาบาลชุมชน 502 รายพบว่า แบบวัดมี Cronbach's alpha 0.91 แบบวัดมีความตรง เพราะสามารถจำแนกผู้ที่มีการศึกษาในระดับที่ต่างกันได้ และมีความสัมพันธ์กับตัวแปรที่บ่งชี้ HL (เช่น ความเข้าใจในฉลากยา) อย่างไรก็ตาม THLA-W มีจุดอ่อนเหมือน REALM คือ ทดสอบทักษะการอ่านเพียงด้านเดียว โดยไม่ได้ประเมินความเข้าใจในคำที่อ่าน (13) อีกทั้งยังพบเกิดการกระจุกตัวของคะแนนในระดับที่สูง (ceiling effect) ต่อมา มีแบบวัด Thai Health Literacy Assessment Using Word List with Extended Questions to Test Comprehension หรือ THLA-W+ (14) ซึ่งต่อยอด THLA-W โดยการเพิ่มคำถามชนิดมีตัวเลือกเพื่อทดสอบความเข้าใจในคำของตัวอย่าง โดยเป็นการประยุกต์วิธีการของแบบวัด SAHLSA (Short Assessment of Health Literacy for Spanish-Speaking Adults) (15) ผลการวัดด้วย THLA-W+ ทำให้ได้คะแนนที่มีการกระจายตัวที่ดีกว่า THLA-W การคำนวณคะแนนจากการทดสอบความเข้าใจ คำพบว่า แบบวัดมี Cronbach's alpha 0.91 แบบวัดมีความตรง เพราะสามารถจำแนกผู้ที่มีการศึกษาในระดับที่ต่างกัน

ได้ และมีความสัมพันธ์กับตัวแปรที่บ่งชี้ HL (เช่น ความเข้าใจในบัตรนัดของโรงพยาบาล) ในระดับ 0.20 ถึง 0.48 ซึ่งเป็นไปตามทฤษฎีและทิศทางที่คาดไว้ แต่ข้อจำกัดของแบบวัดนี้ คือ เนื้อหาการทดสอบมีเพียงการทดสอบการอ่านและการเข้าใจคำ แต่ไม่ครอบคลุมเนื้อหาของ HL ในด้านอื่น ๆ

สำหรับแบบวัด Thai Health Literacy Assessment using Nutrition Label (THLA-N) (16) ประยุกต์วิธีการวัด HL จากแบบวัด Newest Vital Sign (NVS) (17) และคำถาม set of brief screening questions (SBSQ) หรือชุดคำถามสั้น ๆ (18) คำถามใน THLA-N วัดความสามารถในการใช้ฉลากโภชนาการและการประเมินตนเองของตัวอย่างตามคำถามคัดกรอง HL แบบวัดมี Cronbach's alpha 0.71 ผู้ที่มีระดับการศึกษาต่างกันมีคะแนน THLA-N ต่างกัน THLA-N มี $r=0.20$ และ 0.47 กับความเข้าใจในเอกสารสุขภาพและความสามารถในการอ่าน ตามลำดับ แบบวัดมีความไว 0.71 และ 0.75 และความจำเพาะ 0.69 และ 0.75 ขึ้นกับ gold standard ที่ใช้ แม้แบบวัดมีความสะดวกในการใช้ เพราะเป็นแบบวัดชนิดตอบเอง แต่ไม่ครอบคลุมเนื้อหาของ HL อื่น ๆ เช่น ทักษะด้านการพูด-ฟัง ทักษะด้านการค้นหาข้อมูลทางอินเทอร์เน็ต เป็นต้น

ผู้วิจัยจึงต้องการพัฒนาแบบวัด HL ชนิดทั่วไปที่มีเนื้อหาครอบคลุมทักษะที่สื่อถึง HL ที่มากกว่าแบบวัดในอดีต การศึกษานี้พัฒนาแบบวัดขึ้นใหม่ โดยไม่ใช้การแปลแบบวัดในอดีต แต่สร้างขึ้นโดยอาศัยแนวคิดของ Health Literacy Skills Instrument (HLSI) และพัฒนาให้มีความตรง ความเที่ยง และมีความเหมาะสมกับสภาพสังคมไทย การแปลแบบวัด HLSI เป็นภาษาไทยมีข้อจำกัดหลายประการ คือ 1) HLSI ใช้คำถามที่ไม่เหมาะกับบริบทของไทย เช่น การทดสอบที่ใช้เอกสารกำกับยาของประเทศสหรัฐอเมริกาซึ่งมีรูปแบบที่แตกต่างจากข้อกำหนดของสำนักงานคณะกรรมการอาหารและยาของไทย หรือการทดสอบที่ใช้บัตรบันทึกการรับประทานยาซึ่งไม่ได้ใช้ในประเทศไทย และ 2) ลิขสิทธิ์ของแบบวัดฉบับแปลยังคงเป็นของ Research Triangle Institute (RTI) International ทำให้ผู้วิจัยคนไทยต้องขออนุญาตใช้แบบวัดจากสถาบันดังกล่าวทุกครั้งในการวิจัย แม้ว่าการใช้ HLSI โดยบุคคลทั่วไปซึ่งไม่ใช่การใช้ในทางพาณิชย์สามารถทำได้โดยไม่มีค่าใช้จ่ายก็ตาม งานวิจัยนี้จึงอาศัยเพียงแนวคิดของ HLSI และผู้ที่สนใจแบบวัดฉบับนี้สามารถนำไปใช้ได้โดยไม่มี

ค่าใช้จ่าย ข้อมูลจากการประเมินด้วยแบบวัดทำให้บุคลากรทางการแพทย์ซึ่งมีข้อจำกัดทั้งกำลังคนและเวลาสามารถให้ความดูแลเป็นพิเศษกับกลุ่มผู้ที่มี HL ไม่ดี ซึ่งมีแนวโน้มในการเกิดปัญหาในการรักษาพยาบาล ทำให้นำไปสู่การรักษาที่เน้นผู้ป่วยเป็นศูนย์กลางและผลการรักษาที่ดีขึ้น

วิธีการวิจัย

การวิจัยนี้ได้ผ่านการพิจารณาของคณะกรรมการจริยธรรมการวิจัยในมนุษย์ คณะเภสัชศาสตร์ มหาวิทยาลัยสงขลานครินทร์ และได้รับการอนุญาตให้ทำวิจัยจากผู้อำนวยการของโรงพยาบาลที่เป็นสถานที่วิจัย ขั้นตอนในการพัฒนาแบบวัดประกอบด้วย 1) การวิเคราะห์ทักษะที่ประเมินโดยแบบวัด HLSI 2) การจัดทำแบบวัดของไทย ฉบับเริ่มต้น 3) การพิจารณาแบบวัดโดยผู้เชี่ยวชาญและบุคคลที่มี HL สูง 4) การทดสอบความตรงและความเที่ยงในตัวอย่างขนาดใหญ่ และ 5) การหาเกณฑ์คะแนนที่ใช้ตัดสินระดับ HL ในคนไทย

การวิเคราะห์แบบวัด HLSI

การวิเคราะห์แบบวัด HLSI ทำเพื่อใช้เป็นกรอบในการพัฒนาแบบวัด HL ฉบับภาษาไทย ทำให้แบบวัดที่ได้มีหลักการวัดเหมือน HLSI แต่ถูกออกแบบให้มีเนื้อหาเหมาะสมกับบริบทของไทย แบบวัด HLSI ประกอบด้วยคำถาม 25 ข้อที่ทดสอบทักษะที่แสดงถึง HL ใน 3 ด้าน จำแนกตามสื่อที่ใช้ทดสอบ ได้แก่ 1) ทักษะด้านสิ่งพิมพ์ 2) ทักษะด้านการพูด-ฟัง และ 3) ทักษะด้านการค้นหาข้อมูลทางอินเทอร์เน็ต ทักษะด้านสิ่งพิมพ์จำแนกออกเป็นทักษะเกี่ยวกับความเรียง (prose literacy) ทักษะเกี่ยวกับเอกสาร (document literacy) และทักษะเชิงตัวเลข (quantitative literacy) ส่วนคำถามในแบบวัด HLSI สามารถจำแนกตามชนิดของความสามารถเป็น 5 รูปแบบ คือ 1) การค้นและการเข้าใจข้อความเกี่ยวกับสุขภาพ 2) การแปลข้อมูลในรูปของตาราง แผนภาพ ภาพนิ่ง ภาพเคลื่อนไหว แผ่นผัง หรือเครื่องหมาย 3) การคำนวณ 4) การหาข้อสรุปจากข้อมูลหรือประยุกต์ใช้ข้อมูลกับสถานการณ์หนึ่ง ๆ และ 5) การใช้อินเทอร์เน็ตในการแสวงหาข้อมูลสุขภาพ แบบวัด HLSI มีเนื้อหาครอบคลุม 3 ประเด็นใหญ่เกี่ยวกับสุขภาพที่บุคคลทั่วไปสามารถพบได้ตลอดช่วงชีวิต คือ การสร้างเสริมสุขภาพ-การป้องกันโรค การรักษาและใช้บริการสาธารณสุข และการเข้าถึงบริการในระบบสุขภาพ

การจัดทำแบบวัดของไทยฉบับเริ่มต้น

ผู้วิจัยไม่ได้แปลแบบวัด HLSI เป็นภาษาไทย แต่พัฒนาแบบวัดฉบับภาษาไทยขึ้นใหม่ทั้งฉบับ โดยใช้แนวคิดของ HLSI แบบวัดภาษาไทยที่ได้มีชื่อว่า แบบวัดความแตกฉานทางสุขภาพสำหรับชาวไทยโดยประเมินจากทักษะที่สำคัญ (Thai Health Literacy Assessment based on Skills: THLA-S) ฉบับเริ่มแรก ซึ่งถูกออกแบบให้มีความยาว 26 ข้อ คำถามสามารถจำแนกตามสื่อที่ใช้ทดสอบ (3 สื่อ) ความสามารถที่ทดสอบ (5 ทักษะ) และเนื้อหาของคำถามที่ใช้ (3 เนื้อหา) คือ การสร้างเสริมสุขภาพ-การป้องกันโรค (17 ข้อ) การรักษาและใช้บริการสาธารณสุข (5 ข้อ) และการเข้าถึงบริการในระบบสุขภาพ (4 ข้อ)

เกณฑ์ในการพัฒนาคำถาม

เกณฑ์ในการพัฒนาคำถามมีดังนี้ 1) คำถามถามในประเด็นสุขภาพที่เกี่ยวข้องกับประชาชนจำนวนมาก นั่นคือ เป็นประเด็นหรือโรคที่พบได้ทั้งชายและหญิง โดยไม่ใช่โรคที่พบยากหรืออาการที่พบในคนกลุ่มเล็ก ๆ 2) เป็นประเด็นสุขภาพที่เหมาะสมกับคนไทย คือ เป็นสิ่งที่คนไทยพบเห็นได้ง่าย 3) เป็นประเด็นสุขภาพที่สำคัญต่อคนไทย และมีความเห็นพ้องในหมู่ผู้เชี่ยวชาญต่อประเด็นที่อยู่ในคำถามและคำตอบ (ไม่เป็นประเด็นที่ยังเป็นที่โต้เถียงกันทางวิทยาศาสตร์หรือยังไม่มีข้อยุติ) 4) คำถามใช้สื่อประกอบที่หลากหลายทั้งข้อความ เสียงในระบบให้ข้อมูลอัตโนมัติทางโทรศัพท์ ภาพเคลื่อนไหว และเว็บไซต์ 5) เนื้อหาที่ใช้ประกอบคำถามแต่ละข้อใช้เวลาอ่านหรือศึกษาไม่เกิน 1 นาที 6) คำถามและสื่อประกอบพัฒนาโดยนักวิจัย โดยไม่ได้คัดลอกจากแหล่งใดเพื่อป้องกันปัญหาเรื่องลิขสิทธิ์หรือได้จากแหล่งที่มีการเผยแพร่ต่อสาธารณะเพื่อให้ใช้ประโยชน์โดยทั่วไป (เช่น เว็บไซต์ของ สสส.) และ 7) การตอบคำถามสามารถทำได้อย่างถูกต้องโดยใช้ข้อมูลที่อยู่ในสื่อประกอบคำถาม ซึ่งผู้ถูกทดสอบไม่จำเป็นต้องมีความรู้ในเรื่องที่ถามมาก่อน

การพิจารณาแบบวัดเบื้องต้น

การตรวจสอบความเหมาะสมของ THLA-S ทำโดยผู้เชี่ยวชาญ 6 ท่าน ประกอบด้วย 1) นักวิจัย 1 ท่านที่มีความเชี่ยวชาญในเรื่องการสร้างเครื่องมือวัดและประเมินผล รวมทั้งมีประสบการณ์การวิจัยมากกว่า 20 ปี 2) แพทย์เภสัชกร และพยาบาลวิชาชีพ สาขาละ 1 ท่าน และ 3) เจ้า

หน้าสาธารณสุข ซึ่งมีประสบการณ์ในการให้คำแนะนำปรึกษาแก่ผู้ป่วยมากกว่า 10 ปี จำนวน 2 ท่าน

ผู้วิจัยนำแบบวัดที่ปรับตามคำแนะนำของผู้เชี่ยวชาญแล้ว ไปตรวจสอบความเหมาะสมโดยบุคคลที่มี HL สูง ซึ่งคือ พยาบาลที่จบการศึกษาขั้นต่ำในระดับปริญญาตรี จำนวน 28 คน เพื่อค้นหาความเข้าใจผิดอันเกิดจากการออกแบบวัดที่ไม่ดี (ไม่ใช่การเข้าใจผิดจากระดับ HL น้อย) การทดสอบใช้เทคนิคการคิดออกเสียง (think aloud) (19) ผู้วิจัยให้พยาบาลทำแบบทดสอบครั้งละ 1 คน โดยให้ผู้ถูกทดสอบอ่านคำถามแบบออกเสียงพร้อมทั้งบรรยายความคิดที่เกิดขึ้นขณะทำแบบทดสอบออกมา ผู้วิจัยบันทึกความคิดของตัวอย่าง โดยเฉพาะอย่างยิ่งในประเด็นที่แสดงให้เห็นถึงความเข้าใจผิดอันเนื่องมาจากแบบทดสอบที่ไม่ชัดเจน หลังจากนั้นผู้วิจัยปรับแก้แบบวัดตามความเห็นที่ได้รับก่อนจะเริ่มทดสอบในตัวอย่างคนต่อไป กระบวนการทดสอบและการปรับแก้แบบวัดจะทำสลับกันไปเรื่อย ๆ จนไม่เจอข้อผิดพลาดหลัก

การทดสอบในตัวอย่างขนาดใหญ่

ตัวอย่าง

ตัวอย่างในการศึกษานี้ คือ ผู้ป่วย ญาติผู้ป่วย หรือผู้ติดตามผู้ป่วยที่มาใช้บริการจากหน่วยบริการหรือคลินิกบริการผู้ป่วยนอกของโรงพยาบาลศูนย์ที่เป็นสถานที่วิจัยในช่วงวันที่ 1-30 เมษายน พ.ศ. 2560 การเลือกคลินิกหรือหน่วยบริการในการเก็บข้อมูลแต่ละวันทำแบบสุ่ม ผู้รับการทดสอบต้อง 1) เป็นคนไทยที่มีตั้งแต่อายุ 20 ปีขึ้นไป นั่นคือบรรลุนิติภาวะตามกฎหมายและสามารถให้คำยินยอมในการเข้าร่วมการวิจัยโดยไม่ต้องได้รับความยินยอมจากผู้แทนโดยชอบธรรม 2) สามารถพูดสื่อสารด้วยภาษาไทยและอ่านภาษาไทยได้ 3) มีความสามารถในการฟังเป็นปกติ (ทดสอบโดยการพูดโต้ตอบกับนักวิจัยได้ และ 4) ไม่ได้จบการศึกษาหรือประกอบอาชีพในสาขาที่เกี่ยวกับสาธารณสุข การศึกษาใช้การเลือกตัวอย่างแบบบังเอิญ

เกณฑ์คัดออก คือ 1) ผู้รับบริการใช้บริการจาก 1) หน่วยให้บริการที่ตั้งอยู่นอกพื้นที่ของโรงพยาบาล 2) หน่วยบริการพิเศษที่จัดตั้งเฉพาะสำหรับกลุ่มผู้รับบริการ เช่น แรงงานต่างด้าว แรงงานตามสถานประกอบการ หน่วยบริการที่ผู้ป่วยอายุน้อยกว่า 20 ปี และ 3) คลินิกที่มีผู้ป่วยเปราะบางหรือที่มีโอกาสถูกตีตราทางสังคมสูง เช่น ผู้ป่วยมะเร็ง วัณโรค เอชไอวี เป็นต้น

ขนาดตัวอย่าง

ผู้วิจัยใช้สูตรการคำนวณขนาดตัวอย่างสำหรับการวิเคราะห์ความแปรปรวนทางเดียวเพื่อเปรียบเทียบตัวแปรเชิงปริมาณ (คะแนนจาก THLA-S) ระหว่างกลุ่มที่มีระดับการศึกษาต่างกันโดยใช้โปรแกรม G*power (20) ความคลาดเคลื่อนชนิดที่ 1 ถูกกำหนดที่ 0.05 อำนาจการทดสอบที่ 0.80 จำนวนกลุ่มเปรียบเทียบ คือ 4 กลุ่ม (ประถมศึกษา มัธยมศึกษาตอนต้น มัธยมศึกษาตอนปลาย/ประกาศนียบัตรวิชาชีพ (ปวช.) และประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.)/ปริญญาตรีหรือสูงกว่า) ในที่นี้กำหนดให้ขนาดอิทธิพล (effect size) เท่ากับ 0.15 (ค่า 0.10 และ 0.25 แสดงถึง ความแตกต่างระหว่างกลุ่มขนาดน้อยและปานกลาง ตามลำดับ) (20) ขนาดอิทธิพลที่ 0.15 คือ ความแตกต่างระหว่างกลุ่มที่มีขนาดน้อยถึงปานกลาง ขนาดตัวอย่างที่คำนวณได้ คือ 492 คน ดังนั้นควรเก็บข้อมูลอย่างน้อยประมาณ 125 คนในแต่ละระดับการศึกษา

กระบวนการทดสอบ

แบบวัดที่เก็บข้อมูลมีทั้งหมด 4 ตอน ดังนี้ ตอนที่ 1 ประกอบด้วยคำถามสำหรับข้อมูลทั่วไปของผู้ตอบ 9 ข้อ และชุดคำถาม 3 ข้อที่ดัดแปลงมาจากแบบคัดกรองผู้ที่มี HL ไม่เพียงพอ (18) เกี่ยวกับการอ่านเอกสารของโรงพยาบาล การกรอกแบบฟอร์มทางการแพทย์ และการทราบอาการของตนเองจากการอ่านเอกสาร ผู้ป่วยต้องประเมินตนเองในคำถามทั้งสามบนสเกล 5 ระดับ (ไม่เคยเลย/ไม่แน่ใจเลย ถึง ทุกครั้ง/มั่นใจมากที่สุด) คำถามสามข้อนี้เป็นส่วนหนึ่งของแบบวัด THLA-N (Thai Health Literacy Assessment Using Nutrition Label) ถัดมาเป็นชุดคำถาม 8 ข้อซึ่งให้ตัวอย่างประเมินตนเองในเรื่องความสามารถต่าง ๆ ที่จำเป็นต้องการมี HL ได้แก่ การคิดเลข (คำตอบมี 4 ระดับ จากเห็นด้วยอย่างยิ่ง ถึง ไม่เห็นด้วยอย่างยิ่ง) การจดจำข้อมูลที่อ่านเองและที่คนอื่นบอก การเข้าใจข้อมูลที่อ่านเอง และที่คนอื่นบอก การอธิบายเรื่องสุขภาพให้แพทย์ฟัง รวมทั้งการค้นหาข้อมูลทางสุขภาพบนอินเทอร์เน็ตและที่ต้องการเอง (คำตอบ 4 ระดับ คือ ยากมาก ยาก ค่อนข้างง่าย และง่ายมาก)

ตอนที่ 2 เป็นแบบวัดจำนวน 4 ข้อที่วัดความสามารถในการอ่านแบบ cloze test ซึ่งปรับปรุงมาจากแบบวัด S-TOFHLA ฉบับภาษาไทย (11) โดยปรับให้มีเนื้อหาที่เหมาะสมกับคนไทย ผู้ตอบต้องเลือกตัวเลือกที่ดีที่สุดเพื่อเติมคำในช่องว่างของประโยคที่เว้นว่างไว้ในคำถาม

ตอนที่ 3 เป็นแบบวัด THLA-S 26 ข้อที่พัฒนาขึ้นในการวิจัยนี้ ซึ่งผู้ตอบสามารถตอบแบบวัดได้ด้วยตนเอง คำถามชุดแรกมีการใช้ข้อความหรือเอกสารประกอบการถาม-ตอบ จำนวน 20 ข้อ และคำถามชุดที่ 2 มีการใช้สื่อที่ไม่ใช่ข้อความ/เอกสารประกอบการถามจำนวน 6 ข้อ ได้แก่ การใช้คลิปเสียงเป็นสื่อประกอบ 2 ข้อ การใช้ภาพเคลื่อนไหวเป็นสื่อประกอบ 2 ข้อ และการใช้อินเทอร์เน็ตในการค้นหาข้อมูลหรือคำนวณ (เช่น คำนวณหาค่าพลังงานที่ร่างกายต้องการ (BMR)) 2 ข้อ โดยในการเก็บข้อมูลจากคำถามชุดที่ 2 จะทำการทดสอบเมื่อตัวอย่างตอบแบบสอบถามส่วนอื่น ๆ เสร็จแล้ว ผู้วิจัยจึงจะเปิดแฟ้มเสียงหรือภาพเคลื่อนไหวจากคอมพิวเตอร์ให้ชมหรือฟัง หรือเปิดหน้าจอของเว็บไซต์ที่ใช้ประกอบให้ตัวอย่าง

ตอนที่ 4 เป็นคำถาม 5 ข้อที่เป็นส่วนหนึ่งของแบบวัด THLA-N ผู้ตอบต้องใช้ข้อมูลบนฉลากโภชนาการในการตอบคำถาม THLA-N ทดสอบทักษะการอ่าน การแปลความหมาย และการคำนวณที่เกี่ยวข้องกับการใช้ฉลากโภชนาการ เช่น จากฉลากโภชนาการ ถ้าท่านต้องการพลังงานวันละ 2,000 กิโลแคลอรี รับประทานมันฝรั่งทอดกรอบชนิดนี้ 30 กรัม ให้พลังงานคิดเป็นกี่เปอร์เซ็นต์ของพลังงานที่ท่านต้องการต่อวัน

การวิเคราะห์ข้อมูล

การวิเคราะห์คำถามรายข้อ: การศึกษาตัดคำถามข้อที่มีผู้ตอบถูกจำนวนมากหรือตอบผิดจำนวนมาก เพราะเป็นคำถามที่ไม่สามารถแยกผู้ที่มี HL สูงและต่ำออกจากกันได้ และตัดคำถามที่มีค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient; r) กับคะแนนรวมของ THLA-S น้อย (น้อยกว่า 0.3) เพราะแสดงให้เห็นว่า คำถามข้อนั้นวัดในสิ่งที่ต่างไปจากข้ออื่น ๆ

การวิเคราะห์ความตรงและความเที่ยง: การวิเคราะห์ความเที่ยงของ THLA-S ใช้ Cronbach's alpha การหาความตรงใช้วิธีการต่อไปนี้ 1) การเปรียบเทียบคะแนน THLA-S ระหว่างผู้ที่ประเมินตนเองว่าสามารถทำกิจกรรมที่สำคัญต่อ HL ได้ง่าย/ง่ายที่สุด กับผู้ที่ตอบว่ายาก/อยากที่สุด (ตัวอย่างคำถาม คือ การจดจำข้อมูลที่ข้าพเจ้าเคยอ่านมาเป็นเรื่องยากหรือง่ายเพียงไร การจดจำข้อมูลที่มีคนเคยบอกข้าพเจ้าเป็นเรื่องยากหรือง่ายเพียงไร ฯลฯ) การเปรียบเทียบใช้ independent samples t test หากผู้ที่ตอบว่าง่าย ได้คะแนน THLA-S มากกว่าผู้ที่ตอบว่ายาก แสดงว่า THLA-S มีความตรง 2) การเปรียบเทียบคะแนน THLA-

S ระหว่างผู้ที่มีระดับการศึกษาต่างกัน ใช้การวิเคราะห์ความแปรปรวนทางเดียว หากผู้ที่มีระดับการศึกษาสูงกว่าได้คะแนน THLA-S มากกว่า แสดงว่า THLA-S มีความตรงการทดสอบความเท่าเทียมกันของความแปรปรวนในแต่ละกลุ่มระดับการศึกษาใช้ Levene test หากพบว่า ความแปรปรวนไม่แตกต่างกัน ใช้การทดสอบความแปรปรวนทางเดียวตามปกติ และการเปรียบเทียบรายคู่ใช้ Tukey test และ 3) การหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจาก THLA-S กับตัวชี้วัด HL คือ คะแนนรวมของคำถามคัดกรอง HL จำนวน 3 ข้อ คะแนนจากแบบวัดความสามารถในการอ่าน และความสามารถในการใช้ฉลากโภชนาการ หากค่าสัมประสิทธิ์สหสัมพันธ์เป็นบวกและมีค่าสูง แสดงว่า THLA-S มีความตรง

การหาเกณฑ์คะแนนตัดสินระดับ HL

ในปัจจุบัน ยังไม่มีวิธีวัดใดที่ถือว่าเป็น gold standard (GS) หรือแบบวัดมาตรฐานที่ได้รับการยอมรับโดยทั่วไปว่าวัดระดับ HL ได้ดี งานวิจัยนี้จึงเลือกใช้แบบวัด THLA-N (16) เป็น GS เพราะมีเกณฑ์สำหรับตัดสินคะแนนที่ชัดเจน คือ ผู้ที่มี THLA-N >4.83 ถือว่ามี HL เพียงพอ การวิเคราะห์โค้ง ROC สำหรับ THLA-N พบพื้นที่ใต้โค้ง ROC 0.76 และ 0.81 ขึ้นอยู่กับ GS ที่ใช้ แบบวัดมีความไว 0.71 และ 0.75 และความจำเพาะ 0.69 และ 0.75 ขึ้นกับ GS ที่ใช้ (16)

ในการวิเคราะห์โค้ง ROC จุดตัดคะแนน คือ คะแนนที่ทำให้ค่าความไวบวกค่าความจำเพาะมีค่าสูงสุด (21) การวิเคราะห์คำนวณค่า accuracy (ความถูกต้องในการทำนาย), positive predictive value (PPV+), negative predictive value (NPV-), LR+ (positive likelihood ratio) และ LR- (negative likelihood ratio) (22-24)

ผลการวิจัย

ข้อมูลทั่วไปของตัวอย่าง

ผู้เข้าร่วมการศึกษา 552 คน เป็นเพศหญิงจำนวน 296 คน (ร้อยละ 53.6) อายุอยู่ในช่วง 20-29 ปี (ร้อยละ 49.8) รองลงมา คือ 30-39 ปี (ร้อยละ 33.5) อายุเฉลี่ย 31.24 ± 9.22 ปี ตัวอย่างร้อยละ 68.6 มีความสามารถในการอ่านในระดับดีหรือดีมาก ตัวอย่างร้อยละ 1.3 เห็นว่า ตนมีความสามารถในการอ่านไม่ดีหรือไม่ดีอย่างยิ่ง เมื่อถามด้วยคำถามคัดกรองผู้ที่มี HL ไม่เพียงพอพบว่า ตัวอย่างร้อยละ

2.9 รายงานว่าไม่เข้าใจบ่อย ๆ หรือทุกครั้งเกี่ยวกับอาการตนเองจากการอ่านเอกสารที่ได้รับ ร้อยละ 10.1 ไม่มั่นใจหรือไม่มั่นใจเลยในการกรอกแบบฟอร์มทางการแพทย์ และร้อยละ 2.1 ขอให้คนอื่นช่วยอ่านเอกสารต่าง ๆ ให้ฟังบ่อย ๆ หรือทุกครั้ง

แบบสอบถามมีคำถามจำนวน 8 ข้อสำหรับให้ตัวอย่างประเมินความสามารถของตนเอง ตัวอย่างร้อยละ 42.2 ไม่เห็นด้วย/ไม่เห็นด้วยอย่างยิ่งกับคำถามว่า ตนเป็นคนที่ยึดเลขเก่ง ตัวอย่างร้อยละ 22.4, 20.6, 13.0, 18.8, 15.2, 13.9 และ 18.5 รายงานว่าเป็นเรื่องยากหรือยากมากในการทำกิจกรรมต่อไปนี้ การจดจำข้อมูลที่เคยอ่านมา การจดจำข้อมูลที่มีคนเคยบอกหรือได้ยิน การเข้าใจข้อมูลที่ได้อ่าน การเข้าใจข้อมูลที่คนอื่นบอก การอธิบายเรื่องเกี่ยวกับสุขภาพให้แพทย์ฟัง การค้นหาข้อมูลทางสุขภาพบนอินเทอร์เน็ต และการค้นหาข้อมูลสุขภาพที่ตนเองต้องการ

การวิเคราะห์คำถาม

แบบวัด THLA-S เริ่มต้นมีความยาว 26 ข้อ หลังจากปรับแก้ไขตามความเห็นของผู้เชี่ยวชาญและการทดสอบในตัวอย่างแล้ว ได้ตัดคำถามออก 2 ข้อเพราะมีตัวอย่างที่ตอบถูกต้องจำนวนน้อยมาก (น้อยกว่าร้อยละ 6) แสดงว่า ทั้งสองข้อไม่อาจจำแนกผู้ที่มี HL สูงและต่ำออกจากกันได้ และตัดคำถามเพิ่มเติม 1 ข้อออกเพราะมีค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนรวมของ THLA-S น้อย ($r=0.20$) แบบวัดฉบับที่ใช้ในการวิเคราะห์ข้อมูลมี 23 ข้อ ผู้สนใจสามารถขอแบบวัดได้จากผู้แต่งบทความคนที่สอง

ความเที่ยงและความตรงของ THLA-S

แบบวัด THLA-S (23 ข้อ) มี Cronbach's alpha เท่ากับ 0.81 ซึ่งมากกว่า 0.70 จึงถือว่ามีความเที่ยงสูงเป็นที่น่าพอใจ ตารางที่ 1 แสดงผลการเปรียบเทียบคะแนน THLA-S ระหว่างผู้ที่ประเมินตนเองว่า สามารถทำกิจกรรมที่สำคัญต่อการมี HL 8 อย่างได้ง่าย/ง่ายมาก (การคิดเลข การจดจำข้อมูล ฯลฯ) กับผู้ที่รายงานว่าทำได้ยาก/ยากมาก ผู้ที่สามารถทำ 6 กิจกรรม (จาก 8 กิจกรรม) ได้ง่ายมีคะแนน THLA-S สูงกว่าอย่างมีนัยสำคัญทางสถิติ ($P<0.05$) ซึ่งสอดคล้องกับสมมติฐานที่วางไว้ ความสามารถในการ 2 คำถามที่ไม่พบความแตกต่าง คือ การจดจำข้อมูลที่เคยอ่าน และการเล่าเรื่องสุขภาพของตนให้แพทย์ฟัง ผลการวิจัยโดยรวมบ่งชี้ว่า THLA-S มีความตรง

ตารางที่ 1. การเปรียบเทียบคะแนน THLA-S (คะแนนมีพิสัย 0-23) ระหว่างผู้ที่มีระดับความสามารถต่างกันในการทำกิจกรรมที่สำคัญต่อการมี HL (N=552)

ประเด็นที่คำถามวัด	คะแนน THLA-S				p ²
	ผู้มีความสามารถสูง ¹		ผู้มีความสามารถน้อย ¹		
	จำนวน	ค่าเฉลี่ย±SD	จำนวน	ค่าเฉลี่ย±SD	
ทักษะการคำนวณ	319	13.81±4.72	233	12.97±4.01	0.025
การจดจำข้อมูลที่เคยอ่าน	428	13.60±4.50	124	12.96±4.25	0.161
การจดจำข้อมูลที่เคยได้ยิน	437	13.70±4.44	114	12.58±4.38	0.017
การเข้าใจข้อมูลที่ได้อ่าน	480	13.69±4.42	72	11.85±4.33	0.001
การเข้าใจข้อมูลที่คนอื่นบอก	448	13.80±4.46	104	11.96±4.10	<0.001
การเล่าเรื่องสุขภาพให้แพทย์ฟัง	468	13.38±4.38	84	13.87±4.82	0.352
การค้นหาข้อมูลบนอินเทอร์เน็ต	475	13.79±4.48	77	11.38±3.67	<0.001
การค้นหาข้อมูลสุขภาพที่ตัวแทนต้องการ	450	13.78±4.46	102	12.03±4.12	<0.001

1: ผู้มีความสามารถสูง คือ ผู้ที่ประเมินตนเองว่าทักษะต่าง ๆ ในคำถามนั้นง่ายหรือง่ายมาก ผู้มีความสามารถน้อย คือ ผู้ที่ประเมินตนเองว่าทักษะในคำถามนั้นยากหรือยากมาก 2: independent samples t-test

ตารางที่ 2 แสดงผลการเปรียบเทียบคะแนน THLA-S ระหว่างผู้ที่มีระดับการศึกษาที่ต่างกัน โดยแบ่งระดับการศึกษาเป็น 4 กลุ่ม คือ น้อยกว่าเท่ากับ ประถมศึกษาปีที่ 6 (ป.6) มัธยมศึกษาปีที่ 1-3 (ม.1-ม.3) มัธยมศึกษาปีที่ 4-6 (ม.4-ม.6)/ปวช. และปวส./ปริญญาตรี หรือสูงกว่า ผลการวิเคราะห์ความแปรปรวนทางเดียวสรุปได้ว่า แต่ละกลุ่มการศึกษามีคะแนนไม่เท่ากัน (F=49.23, df=3, 548, P<0.001) หรือมีอย่างน้อย 1 คู่การศึกษาที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ การเปรียบเทียบรายคู่ด้วย Tukey's test พบว่า ทุกกลุ่มการศึกษามีคะแนน

แตกต่างกันอย่างมีนัยสำคัญทางสถิติ (P<0.05) โดยผู้ที่มีระดับการศึกษาที่สูงกว่ามีคะแนน THLA-S มากกว่า ซึ่งสอดคล้องกับสมมติฐานการวิจัยที่กำหนดไว้ และบ่งบอกความตรงของ THLA-S

ตารางที่ 3 แสดงค่าสัมประสิทธิ์สหสัมพันธ์ (r) ระหว่างคะแนน THLA-S กับตัวชี้วัด HL ต่าง ๆ ได้แก่ การรับรู้ความสามารถของตนเอง ความสามารถในการอ่านจากแบบทดสอบ cloze test คะแนนจากแบบคัดกรองผู้มี HL ไม่เพียงพอ ความสามารถในการใช้ฉลากโภชนาการ และ THLA-N ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง THLA-S และ

ตารางที่ 2. ผลการวิเคราะห์ความแปรปรวนทางเดียวเพื่อเปรียบเทียบความแตกต่างของคะแนน THLA-S ระหว่างผู้ที่มีระดับการศึกษาที่ต่างกัน (n= 552)

ระดับการศึกษา	จำนวน	ค่าเฉลี่ยคะแนน THLA-S±SD ^{1,2}
ประถมศึกษาหรือต่ำกว่า	125	10.76±3.82 ^a
มัธยมศึกษาปีที่ 1-3	130	12.18±4.00 ^b
มัธยมศึกษาปีที่ 4-6/ปวช.	125	13.84±4.15 ^c
ปวส./ปริญญาตรีและสูงกว่า	172	16.08±3.89 ^d
รวม	552	13.45±4.45

1: พิสัยของคะแนน 0-23; Levene test: F=0.93, df=3, 548, P=0.425; ANOVA: F=49.23, df=3, 548, P<0.001

2: ผลการทดสอบด้วย Tukey's test ตัวเลขที่มีตัวอักษรภาษาอังกฤษต่างกัน แสดงว่ามีระดับคะแนน THLA-S ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 3. ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน THLA-S กับตัวชี้วัด HL ต่าง ๆ

ตัวชี้วัด HL	ค่า r ของ THLA-S กับตัวชี้วัด HL	P
การรับรู้ความสามารถของตนเอง	0.336	<0.001
ความสามารถในการอ่านจาก cloze test	0.398	<0.001
คะแนนจากแบบคัดกรองผู้มี HL ไม่เพียงพอ	0.156	<0.001
ความสามารถในการใช้ฉลากโภชนาการ	0.634	<0.001
THLA-N	0.606	<0.001

ตัวชี้วัด HL ทุกตัวมีค่าเป็นบวกและมีนัยสำคัญทางสถิติ ($P < 0.05$) ซึ่งสอดคล้องกับสมมติฐานที่วางไว้ ผลการศึกษาส่วนนี้จึงบ่งชี้ความตรงของ THLA-S เป็นที่น่าสังเกตว่าคะแนน THLA-S และความสามารถในการใช้ฉลากโภชนาการ รวมทั้งแบบวัด THLA-N มีความสัมพันธ์กันค่อนข้างสูง ($r > 0.60$) นอกจากนี้ความสัมพันธ์ระหว่างคะแนน THLA-S และคะแนนจากแบบคัดกรองผู้มี HL ไม่เพียงพอแม้มีความสัมพันธ์ทางบวก แต่ความสัมพันธ์ค่อนข้างต่ำ จึงไม่สามารถใช้คำถามจากแบบคัดกรองผู้มี HL ไม่เพียงพอมาวัด HL แทน THLA-S ได้

การหาเกณฑ์แปลผลคะแนน

การวิเคราะห์โค้ง ROC ได้ค่า AUC 0.84 ค่า AUC 0.5-0.7, 0.7-0.9 และ > 0.9 แสดงว่า แบบวัดมีความถูกต้องในการจำแนกระดับ HL ได้น้อย ปานกลาง และมากตามลำดับ (23) แสดงว่า THLA-S มีความถูกต้องในการจำแนกระดับ HL ได้ดีปานกลางถึงมาก คะแนนบนโค้ง ROC ที่ทำให้ความไวบวกความจำเพาะมีค่าสูงสุด คือ 13 ดังนั้น ผู้ที่ได้คะแนน ≤ 13 ถือว่ามี HL ไม่เพียงพอ ส่วนผู้ที่ได้คะแนนมากกว่า 13 ถือว่ามี HL เพียงพอ

ความไว คือ สัดส่วนของผู้ที่มี HL ไม่เพียงพอที่ได้คะแนน ≤ 13 (true positive) ความจำเพาะ คือ สัดส่วนของผู้ที่มี HL เพียงพอที่ได้คะแนน > 13 (true negative) แบบวัด THLA-S มีความไวร้อยละ 75.0 นั่นคือ สามารถทำนายผู้ที่มี HL ต่ำได้อย่างถูกต้องร้อยละ 75.0 ส่วนความจำเพาะคือ ร้อยละ 78.4 นั่นคือ สามารถทำนายผู้ที่มี HL สูงได้อย่างถูกต้องร้อยละ 78.4 ค่า accuracy หรือความถูกต้องในการทำนายโดยรวม คือ ร้อยละ 76.60

ส่วนค่า PPV+ คือ ความถูกต้องในการทำนายระดับ HL สำหรับกลุ่มที่วัดด้วย THLA-S แล้วให้ผลเป็นบวก (≤ 13 คะแนน) ค่า PPV+ ในที่นี้เท่ากับ ร้อยละ 79.12 ส่วน

ค่า NPV คือ ความถูกต้องในการทำนายสำหรับกลุ่มที่วัดด้วย THLA-S แล้วให้ผลเป็นลบ (> 13 คะแนน) ค่า NPV ในที่นี้เท่ากับ ร้อยละ 74.19

LR+ คือ โอกาสที่เกิด true positive หากด้วยโอกาสเกิด false positive นั่นคือ เมื่อผลประเมินด้วย THLA-S แล้วให้ผลเป็นบวก (ประเมินว่า HL ไม่เพียงพอ) LR+ คือ โอกาสที่ตัวอย่างจะมี HL ไม่เพียงพอหากด้วยโอกาสที่ตัวอย่างจะมี HL เพียงพอ ส่วน LR- คือ โอกาสที่เกิด false negative หากด้วยโอกาสเกิด true negative เมื่อผลประเมินด้วย THLA-S เป็นลบ (ประเมินว่า HL เพียงพอ) แบบวัดที่มี LR (likelihood ratio) มากกว่า 10.00/น้อยกว่า 0.10, 5 – 10/0.10 – 0.20 และ 2.00 – 5.00/0.20 – 0.50 สามารถให้ข้อมูลที่เพียงพอต่อการตัดสินใจทางคลินิกมาก ปานกลาง และน้อย (แต่อาจมีความสำคัญทางคลินิก) ตามลำดับ ส่วน LR ที่มีค่าระหว่าง 0.50 – 2.00 ให้ข้อมูลที่ไม่น่าเชื่อถือ (24) แบบวัด THLA-S ให้ค่า LR ระหว่าง 0.32 และ 3.47 จึงให้ข้อมูลที่เพียงพอต่อการตัดสินใจน้อยแต่อาจมีความสำคัญทางคลินิก

ความหมายของเกณฑ์คะแนน

ตารางที่ 4 เปรียบเทียบจำนวนผู้ที่ตอบคำถาม 23 ข้อใน THLA-S ได้อย่างถูกต้องระหว่างผู้ที่มี HL เพียงพอ (THLA-S > 13) และไม่เพียงพอ (THLA-S ≤ 13) ผู้ที่มี HL เพียงพอสามารถตอบคำถามทั้ง 23 ข้อได้ถูกต้องมากกว่าผู้ที่มี HL ไม่เพียงพออย่างมีนัยสำคัญทางสถิติ ($P < 0.001$) โดยมีค่า OR 3.34-15.50 นอกจากนี้มีผลการวิจัยที่น่าสนใจ เช่น ผู้ที่มี HL เพียงพอร้อยละ 68.10 สามารถค้นหาข้อมูลอาหารข้างเคียงบนเอกสารกำกับยา (คำถามข้อ 4) แต่ผู้ที่มี HL ไม่เพียงพอร้อยละ 24.54 สามารถค้นหาข้อมูลดังกล่าวได้ ผู้ที่มี HL เพียงพอร้อยละ 82.08 เข้าใจตารางแสดงขนาดยาบนเอกสารกำกับยาและทราบขนาดยาที่ต้องใช้ (คำถาม

ตารางที่ 4. จำนวนและร้อยละของผู้ที่ตอบคำถามใน THLA-S ถูกต้อง จำแนกตามระดับ HL ที่ประเมินโดย THLA-S (N=552)

ประเด็นที่คำถามวัด ¹	ผู้ที่มี HL ไม่เพียงพอ ² (N=273)		ผู้ที่มี HL เพียงพอ ² (N=279)		OR (95%CI) ³
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	
ข้อ 1 การค้นหาข้อมูลบนเอกสารในรูปแบบความเรียง (คำถามเกี่ยวกับคอเลสเตอรอล)	86	31.50	192	68.82	4.80 (3.35-6.88)
ข้อ 2 การแปลความหมายจากข้อมูลบนเอกสารในรูปแบบความเรียง (คำถามเกี่ยวกับคอเลสเตอรอล)	24	8.79	68	24.37	3.34 (2.03-5.51)
ข้อ 4 การค้นหาข้อมูลอาการข้างเคียงบนเอกสารกำกับยา (คำถามเกี่ยวกับยามานาซอล)	67	24.54	190	68.10	6.56 (4.52-9.53)
ข้อ 5 การทราบขนาดยาที่ควรใช้จากตารางขนาดยาบนเอกสารกำกับยา (คำถามเกี่ยวกับยามานาซอล)	93	34.07	229	82.08	8.86 (5.97-13.16)
ข้อ 6 การประยุกต์ใช้ข้อมูลความถี่ในการใช้ยาบนเอกสารกำกับยา (คำถามเกี่ยวกับยามานาซอล)	155	56.78	250	89.61	6.56 (4.17-10.32)
ข้อ 7 การค้นหาและเข้าใจข้อมูลบนเอกสารในรูปแบบความเรียง (คำถามเกี่ยวกับโรคต่อ)	128	46.89	243	87.10	7.65 (5.01-11.67)
ข้อ 8 การแปลความหมายจากข้อมูลบนเอกสารในรูปแบบความเรียง (คำถามเกี่ยวกับโรคต่อ)	186	68.13	269	96.42	12.58 (6.37-24.85)
ข้อ 9 การแปลความหมายจากข้อมูลบนเอกสารในรูปแบบความเรียง (คำถามเกี่ยวกับโรคต่อ)	128	46.89	260	93.19	15.50 (9.19-26.15)
ข้อ 10 การเข้าใจข้อมูลในรูปแบบข้อความเสียง (ใช้คลิปเสียงประกอบคำถาม)	211	77.29	264	94.62	5.17 (2.86-9.35)
ข้อ 11 ความสามารถแปลความหมายข้อมูลในรูปแบบข้อความเสียง (ใช้คลิปเสียงประกอบคำถาม)	202	73.99	262	93.91	5.42 (3.09-9.48)
ข้อ 12 ความสามารถค้นหาและสรุปความหมายข้อมูลในรูปแบบแผ่นผัง (การอ่านแผ่นผังโรงพยาบาล)	132	48.35	253	90.68	10.39 (6.51-16.60)
ข้อ 13 ความสามารถค้นหาและสรุปความหมายข้อมูลในรูปแบบแผ่นผัง (การอ่านแผ่นผังโรงพยาบาล)	61	22.34	164	58.78	4.96 (3.42-7.19)
ข้อ 14 การแปลความหมายจากข้อมูลบนเอกสารในรูปแบบความเรียง (คำถามเกี่ยวกับคำแนะนำ อย.)	45	16.48	148	53.05	5.72 (3.85-8.51)
ข้อ 15 การประยุกต์ใช้ข้อมูลบนเอกสารในรูปแบบความเรียง (คำถามเกี่ยวกับคำแนะนำ อย.)	24	8.79	93	33.33	5.19 (3.19-8.44)
ข้อ 16 การใช้งานเว็บไซต์เพื่อคำนวณค่า BMR	85	31.14	195	69.89	5.13 (3.57-7.37)
ข้อ 17 การใช้งานเว็บไซต์เพื่อคำนวณค่า BMI และแปลความหมายที่ได้	151	55.31	250	89.61	6.97 (4.43-10.95)

ตารางที่ 4. จำนวนและร้อยละของผู้ที่ตอบคำถามใน THLA-S ถูกต้อง จำแนกตามระดับ HL ที่ประเมินโดย THLA-S (N=552) (ต่อ)

ประเด็นที่คำถามวัด ¹	ผู้ที่มี HL ไม่เพียงพอ ² (N=273)		ผู้ที่มี HL เพียงพอ ² (N=279)		OR (95%CI) ³
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	
ข้อ 18 การค้นหาข้อมูลบนเอกสารในรูปแบบความเรียง (สัญญาณโรคหลอดเลือดสมอง)	208	76.19	268	96.06	7.61 (3.92-14.79)
ข้อ 19 การค้นหาและเข้าใจข้อมูลในตาราง (การนับพลังงานอาหาร)	28	10.26	92	32.97	4.30 (2.71-6.85)
ข้อ 20 การคำนวณพลังงานอาหารจากตารางข้อมูล (การนับพลังงานอาหาร)	81	29.67	197	70.61	5.69 (3.95-8.21)
ข้อ 21 การสรุปข้อมูลจากสื่อวีดิทัศน์ (โรคอ้วนลงพุง)	240	87.91	276	98.92	12.65 (3.83-41.77)
ข้อ 22 การแปลความหมายข้อมูลจากสื่อวีดิทัศน์ (โรคอ้วนลงพุง)	234	85.71	268	96.06	4.06 (2.03-8.11)
ข้อ 23 การแปลความหมายของคำว่าหนึ่งหน่วยบริโภคบนฉลากโภชนาการ	34	12.45	160	57.35	9.45 (6.14-14.54)
ข้อ 25 ความสามารถในการอ่านกราฟแท่ง	48	17.58	184	65.95	9.08 (6.10-13.52)

1: ไม่แสดงข้อ 3, 24 และ 26 เพราะเป็นข้อที่ถูกตัดออกในกระบวนการคัดเลือกคำถาม

2: ผู้มี HL เพียงพอ คือ ผู้ที่ได้คะแนน THLA-S >13 ส่วนผู้ที่มี HL ไม่เพียงพอ คือ ผู้ที่ได้คะแนน THLA-S ≤13

3: Chi-square test พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ (P<0.001) ในคำถามทุกข้อ

ข้อ 5) แต่ผู้ที่มี HL ไม่เพียงพอร้อยละ 34.07 เท่านั้นที่เข้าใจตารางดังกล่าว ผู้ที่มี HL เพียงพอร้อยละ 90.68 เข้าใจแผนที่แสดงตำแหน่งของหน่วยบริการต่าง ๆ ในโรงพยาบาล (คำถามข้อที่ 12) แต่ผู้ที่มี HL ไม่เพียงพอร้อยละ 48 35 เข้าใจแผนดังกล่าว

การอภิปรายผลและสรุป

การอภิปรายผลการวิจัยมุ่งเปรียบเทียบแบบวัด THLA-S กับแบบวัด HL ชนิดทั่วไปฉบับภาษาไทยอื่น ๆ ได้แก่ แบบวัด REALM ฉบับภาษาไทย (11) แบบวัด THLA-W (12) แบบวัด THLA-W+ (14) แบบวัด THLA-N (16) และเครื่องมือวัดความรู้ด้านสุขภาพคนไทย (Thailand Health Literacy Scales: THLS) (10)

เนื้อหาของแบบวัด THLA-S

REALM ฉบับภาษาไทย (11) แบบวัด THLA-W (12) แบบวัด THLA-W+ (14) มีข้อจำกัด คือ เนื้อหาของการ

ทดสอบจำกัดเฉพาะการอ่านค่าที่พบในทางสุขภาพได้ถูกต้องหรือความเข้าใจในค่าเหล่านั้น แบบวัด THLA-S 23 ข้อมีจุดแข็งตรงเนื้อหาการวัดที่ครอบคลุมแง่มุมต่าง ๆ ของ HL โดยวัดทักษะเกี่ยวกับข้อความ (9 ข้อ) ทักษะเกี่ยวกับเอกสารที่ไม่ใช่ข้อความ (เช่น ตาราง) (8 ข้อ) ทักษะด้านการพูด-ฟัง (4 ข้อ) และทักษะการใช้อินเทอร์เน็ตเพื่อแสวงหาข้อมูล (2 ข้อ) คำถามในแบบวัดประเมินการค้นหาและการเข้าใจข้อมูลเกี่ยวกับสุขภาพมี 5 ข้อ ประเมินการแปลความจากข้อมูลเกี่ยวกับสุขภาพ (10 ข้อ) การสรุปข้อมูลหรือประยุกต์ใช้ข้อมูลกับสถานการณ์หนึ่ง ๆ (5 ข้อ) การคำนวณ (1 ข้อ) และ การใช้อินเทอร์เน็ตในการแสวงหาข้อมูล (2 ข้อ) แม้ว่าแบบวัดมีการประเมินความสามารถในการคำนวณด้วยคำถามเพียงข้อเดียว แต่ในความเป็นจริงแล้ว ทักษะการคำนวณแฝงอยู่ในคำถามหลายข้อ แต่ถือเป็นทักษะเด่นที่สำคัญที่สุดในคำถามเพียงข้อเดียว เนื้อหาของคำถามที่ใช้ทดสอบประกอบด้วย 3 ประเด็นเหมือน HLSI คือ การสร้างเสริมสุขภาพ-การป้องกันโรค (14 ข้อ)

การรักษาและใช้บริการสาธารณสุข (5 ข้อ) และการเข้าถึงบริการในระบบสุขภาพ (4 ข้อ)

แบบวัด THLA-N (16) ประกอบด้วยคำถาม 8 ข้อ สามข้อแรกเป็นคำถามคัดกรอง HL โดยสอบถามความยากง่ายในการทำกิจกรรมสุขภาพ เช่น การอ่านเอกสารสุขภาพ การกรอกแบบฟอร์ม การไม่เข้าใจเอกสารสุขภาพ อีก 5 ข้อเป็นคำถามวัดความสามารถในการใช้ฉลากโภชนาการ จึงถือว่ามีความหมายทางสุขภาพที่กว้างมากขึ้นกว่า REALM, THLA-W และ THLA-W+ แต่แบบวัด THLA-N มีเนื้อหาจำกัดเมื่อเปรียบเทียบกับแบบวัด THLA-S ในการวิจัยนี้

สำหรับแบบวัด THLS (10) มีเนื้อหาที่ครอบคลุมมากถึง 5 ด้าน คือ การเข้าถึงข้อมูลและบริการสุขภาพ การเข้าใจข้อมูลและบริการสุขภาพที่เพียงพอต่อการปฏิบัติ การตรวจสอบข้อมูลและบริการสุขภาพ การสื่อสารและการสนับสนุนทางสังคม และการจัดการสุขภาพตนเอง THLS มีเนื้อหาที่ครอบคลุม HL มาก อีกทั้งผู้พัฒนาแบบวัดออกแบบให้ตัวอย่างรายงานความสามารถของตนเองบน Likert scale ทำให้การทดสอบเป็นไปได้โดยง่าย และสามารถตอบแบบสอบถามได้เสร็จในเวลาอันรวดเร็วแม้ว่าจะมีคำถามมากถึง 47 ข้อก็ตาม แต่แบบสอบถามชนิดประเมินตนเองมีข้อจำกัดประการหนึ่ง คือ มีลักษณะอัตวิสัย แบบสอบถามที่มีลักษณะเหมือนข้อสอบดังเช่น THLA-S ในงานวิจัยนี้มีข้อดี คือ คะแนนที่วัดได้มีลักษณะวัตถุวิสัยมากกว่า แต่จะนำมาซึ่งข้อจำกัด คือ การมีข้อสอบจำนวนมากเพื่อวัดให้ครอบคลุมเนื้อหาของ HL เป็นการสร้างภาระในการตอบแก้ตัวอย่าง การสร้าง THLA-S เป็นการพยายามถ่วงดุลความครอบคลุมเนื้อหาของ HL และความยากในการใช้แบบวัด **ความเที่ยง**

แบบวัด THLA-S มีค่า Cronbach's alpha เท่ากับ 0.81 ซึ่งถือว่าน่าพอใจ แบบวัด REALM ฉบับภาษาไทย (11) แบบวัด THLS (10) แบบวัด THLA-W (12) แบบวัด THLA-W+ (14) และแบบวัด THLA-N (16) มีความเที่ยง 0.95, 0.81-0.94 (ขึ้นกับมิติ), 0.90, 0.90 และ 0.71 ตามลำดับ โดยรวมแล้วแบบวัด HL ชนิดทั่วไปทุกชนิดมีความเที่ยงอยู่ในระดับที่น่าพอใจ ทั้งนี้เป็นที่น่าสังเกตว่าความเที่ยงขึ้นกับความยาวของจำนวนคำถามในแบบวัดแบบวัดที่มีความยาวมากจะมีความเที่ยงที่สูงกว่า **ความตรง**

แบบวัด THLA-S มีความตรงเพราะ 1) ผู้ที่ประเมินตนเองว่าสามารถทำ 6 กิจกรรมที่จำเป็นต่อการมี HL (จาก

ทั้งหมด 8 กิจกรรม) ได้อย่างง่าย/ง่ายมาก (เช่น การคิดเลข การจดจำข้อมูล การแสวงหาข้อมูล ฯลฯ) มีคะแนน THLA-S สูงกว่าผู้ที่รายงานว่าทำได้ยากหรือยากมาก 2) ผู้ที่มีระดับการศึกษาสูงกว่าได้คะแนน THLA-S มากกว่า 3) ความสัมพันธ์ของ THLA-S กับตัวชี้วัด HL เป็นบวกและมีนัยสำคัญทางสถิติ ความสัมพันธ์ระหว่าง THLA-S กับการรับรู้ความสามารถของตนเอง ความสามารถในการอ่านจาก cloze test ความสามารถในการใช้ฉลากโภชนาการ และ THLA-N มีค่า 0.34, 0.40, 0.63 และ 0.61 ตามลำดับ ซึ่งถือว่า มีขนาดความสัมพันธ์ในระดับปานกลางถึงดีมากเมื่อใช้เกณฑ์ที่ว่า ความสัมพันธ์ในขนาด 0-0.20, 0.21-0.40, 0.41-0.60, 0.61-0.80 และ 0.81-1.00 ถือเป็นระดับที่น้อย ปานกลาง ดี ดีมาก และ ดีเยี่ยม ตามลำดับ ยกเว้น ความสัมพันธ์ระหว่าง THLA-S กับคะแนนจากแบบคัดกรองผู้มี HL ไม่เพียงพอ ซึ่งพบ $r = 0.16$ ที่เป็นบวกและมีนัยสำคัญทางสถิติ ($P < 0.05$) แต่มีขนาดต่ำ ดังนั้น จึงไม่สามารถใช้คำถามจากแบบคัดกรองผู้มี HL ไม่เพียงพอ มาวัด HL แทนการวัดด้วยชุด THLA-S ได้

แบบวัด THLA-S มีข้อมูลเกี่ยวกับความตรงที่มากกว่าหรือมากพอ ๆ กับแบบวัดอื่น ๆ ที่นำมาอภิปรายผลเปรียบเทียบ แบบวัด REALM ฉบับภาษาไทยมีข้อมูลความตรงจำกัด คือ มี $r = 0.47$ กับความสามารถในการอ่าน ผู้ที่มีการศึกษาในระดับประถมศึกษาได้คะแนนน้อยกว่าผู้ที่มีการศึกษาสูงกว่า แต่ในกลุ่มการศึกษาต่าง ๆ ที่สูงกว่าระดับประถมศึกษา (เช่น มัธยมศึกษา ประกาศนียบัตรวิชาชีพ ปริญญาตรี) มีคะแนนไม่แตกต่างกัน (11) แสดงถึง ceiling effect (ตัวอย่างส่วนใหญ่มีแนวโน้มจะได้คะแนนสูงในเกือบทุกระดับการศึกษา) และแบบวัดยังไม่มีการทดสอบความตรงโดยหาความสัมพันธ์กับตัวชี้วัด HL ที่หลากหลาย เช่น การคำนวณขนาดยา การอ่านฉลากยา ส่วน THLS มีข้อมูลความตรงเชิงโครงสร้างจากผลการวิเคราะห์องค์ประกอบเชิงยืนยันที่มี $\text{Chi-Square} = 1470.99$, $\text{df} = 784$, $\text{RMSEA} = 0.047$, $\text{NFI} = 0.98$, $\text{CFI} = 0.99$ และ $\text{SRMR} = 0.084$ ซึ่งบ่งชี้ว่า แบบจำลอง 5 องค์ประกอบที่กำหนดไว้มีความสอดคล้องกับข้อมูลที่สังเกตพบในงานวิจัย (10) แต่ THLS ยังขาดข้อมูลความตรงเชิงโครงสร้างมากพอนั้นคือ ยังไม่มีข้อมูลความสัมพันธ์ระหว่างคะแนนจากแบบวัดกับตัวแปรอื่น ๆ ที่น่าจะมีความสัมพันธ์กับ HL ในทางทฤษฎี ดังนั้นจึงถือว่า REALM และ THLS มีข้อมูลความตรงจำกัดกว่าแบบวัดอื่น ๆ

สำหรับ THLA-W, THLA-W+ (คิดคะแนนจากการทดสอบความเข้าใจเพียงอย่างเดียว), THLA-N และ THLA-S มีข้อมูลความตรงในระดับเดียวกัน และใช้ระเบียบวิธีวิจัยที่คล้ายกันในการพิสูจน์ความตรง THLA-W, THLA-W+, THLA-N และ THLA-S มีค่า r กับตัวบ่งชี้ HL เท่ากับ 0.40-0.57, 0.20-0.48, 0.22-0.47 และ 0.34-0.63 ตามลำดับ (12, 14, 16) โดยค่า r ของแบบวัด THLA-S มีค่าสูงกว่าเล็กน้อย **เกณฑ์แปลผลคะแนน**

การวิเคราะห์โค้ง ROC โดยใช้ THLA-N เป็น GS พบว่า THLA-S มี AUC=0.84 แสดงว่า แบบวัดสามารถแยกผู้ที่มี HL พอและไม่เพียงพอออกจากกันได้ดีปานกลางถึงดีมาก ผู้ที่ได้คะแนน THLA-S ≤ 13 ถือว่ามี HL ไม่เพียงพอ แบบวัดมีความไวร้อยละ 75.0 ความจำเพาะคือร้อยละ 78.4 ค่า PPV และ NPV เท่ากับร้อยละ 79.12 และ 74.19 ตามลำดับ ส่วน LR+ และ LR- คือ 3.47 และ 0.32 ตามลำดับ การกำหนดเกณฑ์คะแนนของ THLA-W, THLA-W+, THLA-N และ THLA-S ใช้การวิเคราะห์โค้ง ROC เช่นกัน แต่เกณฑ์คะแนนของ REALM ใช้เกณฑ์ของแบบวัดฉบับภาษาอังกฤษซึ่งกำหนดขึ้นจากกลุ่มตัวอย่างชาวอเมริกัน ทั้งนี้ยังไม่มีการวิจัยเพื่อกำหนดเกณฑ์ของแบบวัดนี้ในคนไทย (11)

ในแบบวัด THLAS ผู้พัฒนาแบบวัดกำหนดเกณฑ์ว่า คะแนน <141 (<ร้อยละ 60 ของคะแนนเต็ม), 141-187.99 (ร้อยละ 60-79.99 ของคะแนนเต็ม) และ 188 คะแนน (\geq ร้อยละ 80 ของคะแนนเต็ม) คือ การมี HL ต่ำพอใช้ และดีมาก ตามลำดับ แต่ทั้งนี้ ผู้พัฒนาแบบวัดไม่ได้บอกถึงวิธีการได้มาซึ่งเกณฑ์ดังกล่าว อีกทั้งยังไม่มีข้อมูลความไวและความจำเพาะของเกณฑ์คะแนน อย่างไรก็ตาม ผู้พัฒนาแบบวัดอธิบายว่า ตัวอย่างมักประเมินตนเองโดยให้คะแนนสูงกว่าความจริง ดังนั้นเกณฑ์คะแนนสำหรับ HL ที่ต่ำจึงไม่กำหนดที่ร้อยละ 50 แต่กำหนดที่ร้อยละ 60 แทน

THLA-W, THLA-W+, THLA-N และ THLA-S ใช้การวิเคราะห์โค้ง ROC ในการกำหนดเกณฑ์ตัดสินคะแนน จึงมีความเป็นวัตถุวิสัย สำหรับค่า AUC ของ THLA-N และ THLA-S มีค่าใกล้เคียงกัน คือ 0.76-0.81 (ขึ้นกับ GS) และ 0.84 ตามลำดับ แสดงว่า แบบวัดทั้งสองมีความถูกต้องในการจำแนกระดับ HL ได้ดีปานกลางและอยู่ในระดับที่ใกล้เคียงกัน อย่างไรก็ตาม ข้อควรระวังในการแปลผลการวิจัย คือ แบบวัดทั้ง 2 ใช้ GS ที่แตกต่างกัน จึงอาจเปรียบเทียบผลการวิจัยโดยตรงไม่ได้ ส่วนค่า AUC ของ

แบบวัด THLA-W และ THLA-W+ นั้นเปรียบเทียบกับ การวิจัยครั้งนี้ได้ยาก เพราะการวิจัยดังกล่าวใช้ GS ในลักษณะที่ต่างจากการวิจัยนี้มาก

ความไวของ THLA-N และ THLA-S มีค่าใกล้เคียงกัน คือ 0.71-0.75 (ขึ้นกับ GS) และ 0.75 ตามลำดับ ความจำเพาะของ THLA-N และ THLA-S มีค่าใกล้เคียงกัน คือ 0.69-0.75 (ขึ้นกับ GS) และ 0.78 ตามลำดับ ส่วนค่าความไวและความจำเพาะของ THLA-W และ THLA-W+ นั้นเปรียบเทียบกับ การวิจัยครั้งนี้ได้ยาก เพราะการวิจัยดังกล่าวใช้ GS ในลักษณะที่ต่างจากการวิจัยนี้มาก

สรุปการเปรียบเทียบแบบวัดต่าง ๆ

REALM มีจุดแข็ง คือ ใช้สะดวกเพราะเป็นเพียงรายการคำเพื่อทดสอบการอ่าน แบบวัดมีความเที่ยงที่ดี แต่มีข้อเสีย คือ แบบวัดประเมิน HL จากความสามารถทางการอ่านเพียงอย่างเดียว การทดสอบต้องมีบุคคลากรคอยประเมินความถูกต้องในการอ่าน ซึ่งบางครั้งมีความเป็นอัตวิสัยมาก แบบวัดมีข้อมูลความตรงจำกัด โดยยังไม่มีการหาความสัมพันธ์กับตัวชี้วัด HL เช่น การคำนวณขนาดยา การอ่านฉลากยา นอกจากนี้แบบวัดยังมี ceiling effect เกณฑ์คะแนนแบบวัดไทยใช้เกณฑ์ของแบบวัดฉบับภาษาอังกฤษซึ่งกำหนดขึ้นจากกลุ่มตัวอย่างชาวอเมริกัน ทั้งนี้ยังไม่มีการวิเคราะห์ ROC เพื่อกำหนดเกณฑ์ของแบบวัดนี้ในคนไทย จึงยังไม่มีข้อมูลความไว-ความจำเพาะของแบบวัด ตลอดจนความถูกต้องในการทำนาย

THLS มีข้อดี คือ คำถามมีความครอบคลุมเนื้อหา ของ HL แบบวัดมีความเที่ยงที่ดี แต่แบบวัดมีความยาว 47 ข้อและเป็นการวัดโดยให้ประเมินตัวอย่างประเมินความสามารถของตนเอง จึงมีลักษณะอัตวิสัย แบบวัดมีข้อมูลความตรงจำกัดเพียงการทดสอบองค์ประกอบเชิงยืนยัน โดยยังไม่มีการหาความสัมพันธ์กับตัวชี้วัด HL เกณฑ์คะแนนถูกกำหนดขึ้นโดยไม่มีการวิเคราะห์ ROC จึงยังไม่มีข้อมูลความไว-ความจำเพาะของแบบวัด ตลอดจนความถูกต้องในการทำนาย

THLA-W มีจุดแข็งเหมือน RELAM คือ ใช้สะดวกเพียง 1-2 นาที และผู้ทำแบบวัดไปใช้ไม่ต้องรับการอบรม เพราะเป็นเพียงรายการคำเพื่อทดสอบการอ่าน แบบวัดมีความเที่ยงที่ดีและมีความตรงที่ดี ข้อเสีย คือ ประเมิน HL จากความสามารถทางการอ่านเพียงอย่างเดียว การทดสอบต้องมีบุคคลากรคอยประเมินความถูกต้องในการอ่าน

THLA-W+ ถูกพัฒนาเพื่อแก้ไขข้อเสียของ THLA-W โดยประเมินทั้งความสามารถในการอ่านค่าและความรู้เกี่ยวกับความหมายของค่า ๆ นั้นผ่านแบบทดสอบชนิดตัวเลือก ทำให้มีความครอบคลุมของเนื้อหาเหนือกว่า REALM และ THLA-W อย่างไรก็ตาม ผู้พัฒนาแบบวัดแนะนำว่า ควรคิดคะแนนของแบบวัดนี้จากจำนวนข้อที่ผู้ตอบเลือกตัวเลือกถูก โดยไม่จำเป็นต้องทดสอบการอ่านค่า เพราะมีคุณสมบัติการวัดทางจิตวิทยาเทียบเท่าคะแนนที่คำนวณจากการอ่านร่วมกับการทดสอบความเข้าใจ จุดแข็งของแบบวัด คือ ใช้สะดวกเพียง 6-7 นาที และผู้ให้นำไปใช้ไม่ต้องรับการอบรม แบบวัดมีความเที่ยงที่ดีและมีความตรงที่ดี ข้อเสีย คือ ประเมิน HL จากความเข้าใจคำศัพท์เท่านั้น

THLA-N เป็นแบบวัดที่สร้างขึ้นเพื่อหลีกเลี่ยงข้อเสียของ REALM, THLA-W, และ THLA-W+ แบบวัดนี้ประเมินความสามารถในการคำนวณ การค้นหาข้อมูล และการแปลความหมายของข้อมูลโดยใช้ฉลากโภชนาการเป็นสื่อในการทดสอบ นอกจากนี้ยังให้ตัวอย่างประเมินตนเองด้วยแบบคัดกรอง HL 3 ข้อ แบบวัดนี้จึงมีลักษณะผสมระหว่างแบบทดสอบที่เป็นอัตวิสัยและวัตถุวิสัยรวมอยู่ด้วยกัน ซึ่งมีข้อดีคือทำให้วัดเนื้อหาของ HL ได้กว้างขึ้น ขณะเดียวกันมีความยาวเพียง 8 ข้อ ทำให้แบบวัดนี้มีความน่าสนใจในการใช้ในทางคลินิก แบบวัดมีความเที่ยงและความตรงเป็นที่น่าพอใจ นอกจากนี้ยังมีความถูกต้องในการทำนาย ความไวและความจำเพาะเป็นที่น่าพอใจอีกด้วย อย่างไรก็ตามข้อเสียของแบบวัดนี้คือ ทดสอบความสามารถในการใช้ฉลากโภชนาการเพียงอย่างเดียว โดยไม่ได้ใช้คำถามที่มีเนื้อหาทางสุขภาพอื่น ๆ เลย

THLA-S ที่พัฒนาขึ้นในงานวิจัยนี้มีวัตถุประสงค์เพื่อให้เป็นแบบวัดที่ประเมินมิติความสามารถของ HL ครอบคลุมโดยใช้สื่อที่หลากหลายและเนื้อหาของ การทดสอบเป็นสิ่งที่พบได้ตลอดช่วงชีวิต แบบวัดมีความตรง ความเที่ยง และมีความถูกต้องในการทำนายเป็นที่น่าพึงพอใจ อย่างไรก็ตามความครอบคลุมของแบบวัดนำมาซึ่งข้อเสียคือแบบวัดมีความยาวถึง 23 ข้อและต้องใช้เวลาทดสอบค่อนข้างมาก

ข้อเสนอแนะในการปรับปรุง THLA-S

แบบวัด THLA-S มีความยาวถึง 23 ข้อและต้องใช้เวลามากกว่า 30 นาที ทำให้การใช้งานไม่สะดวก ดังนั้นในอนาคตควรพัฒนาแบบวัดฉบับย่อ โดยเลือกเฉพาะข้อที่

สามารถแยกแยะผู้ที่มี HL ในระดับสูงและต่ำออกจากกันได้ คำถาม 4 ข้อในแบบวัดมีการใช้สื่อประกอบเป็นเสียงและภาพเคลื่อนไหว อย่างไรก็ตาม คำถามดังกล่าวจำแนกผู้ที่มีระดับ HL ต่างกันได้ไม่ดีนัก หากการใช้สื่อที่เป็นเสียงหรือภาพเคลื่อนไหวมีความจำเป็น การศึกษาในอนาคตควรปรับปรุงคำถาม 4 ข้อนี้ในแบบวัดให้มีความยากมากขึ้น เพื่อเพิ่มความจำแนกในการวัด แบบวัดมีคำถาม 2 ข้อที่เกี่ยวกับการใช้งานเว็บไซต์ ในการศึกษาพบตัวอย่างที่ไม่เคยใช้คอมพิวเตอร์และไม่สามารถตอบคำถามนี้ได้ แต่ในปัจจุบันการติดต่อสื่อสารและการค้นข้อมูลผ่านสมาร์ตโฟนเป็นที่นิยมอย่างมาก ตัวอย่างที่ไม่เคยใช้คอมพิวเตอร์ อาจใช้สมาร์ตโฟนได้ ดังนั้นในอนาคตอาจต้องเพิ่มคำถามที่เกี่ยวกับการใช้สมาร์ตโฟนในการค้นหาข้อมูล

กิตติกรรมประกาศ

การวิจัยครั้งนี้สำเร็จได้ด้วยความอนุเคราะห์จากผู้ให้ข้อมูลทุกท่าน ตลอดจนผู้เกี่ยวข้องทุกรายที่ผู้วิจัยขอข้อมูลเพิ่มเติม ผู้วิจัยขอขอบคุณผู้เข้าร่วมวิจัยที่ให้ความร่วมมือเป็นอย่างดี จนสำเร็จลุล่วงไปได้ด้วยดี

เอกสารอ้างอิง

1. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, Brand H. Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health* 2012; 12:80. doi: 10.1186/1471-2458-12-80.
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 reports/ technology assessments* No.199 [online]. 2011 [cited Feb 12, 2016]. Available from: www.ncbi.nlm.nih.gov/pubmed/23126607.
3. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med* 2011; 155: 97-107.
4. Cho YI, Lee SY, Arozullah AM, Crittenden KS. Effects of health literacy on health status and health service utilization amongst the elderly. *Soc Sci Med* 2008; 66: 1809-16.

5. Institute of Medicine. Health literacy: a prescription to end confusion. Washington DC: National Academies ; 2004.
6. Neilsen-Bohlman L, Panzer AM, Hamlin B, Kindig DA. Health literacy: a prescription to end confusion. Washington DC: The National Academies Press; 2004.
7. Baker DW. The meaning and the measure of health literacy. *J Gen Intern Med.* 2006; 21: 878-83.
8. Parker RM, Kindig DA. Beyond the Institute of medicine health literacy report: are the recommendations being taken seriously. *J Gen Intern Med* 2006; 21: 891-2.
9. McCormack L, Bann C, Squiers L, Berkman ND, Squire C, Schillinger D, et al. Measuring health literacy: a pilot study of a new skills-based instrument. *J Health Commun* 2010; 5 Suppl 2: S51-71.
10. Ungsinun I. Creating and developing of Thai health literacy scales [online]. 2017 [cited Jun 16, 2018]. Available from: bsris.swu.ac.th/upload/268335.pdf.
11. Bangonsri J. Validity and reliability of the Thai version of health literacy screening tools for patients at Srinagarind Hospital, Khon Kaen, Thailand [master thesis]. Khon Kaen University; 2013.
12. Phantong W. Development of the Thai Health Literacy Assessment using Word List (THLA-W) [minor thesis]. Songkhla: Prince of Songkla University; 2016.
13. Haun JN, Valerio MA, McCormack LA, Sørensen K, Paasche-Orlow MK. Health literacy measurement: an inventory and descriptive summary of 51 instruments. *J Health Commun* 2014; 19: 302-33.
14. Janchotoo P. Development of the Thai Health Literacy Assessment Using Word List with Extended Questions to Test Comprehension (THLA -W+): Testing in Sadao Hospital [minor Thesis]. Songkhla: Prince of Songkla University; 2017.
15. 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.
16. Na Pattalung P. Development of a Health Literacy Screener for Thais Based on the Approach of the Newest Vital Sign [master thesis]. Songkhla: Prince of Songkla University; 2018.
17. 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.
18. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. *Fam Med* 2004; 36: 588-94.
19. Willis GB, Royston P, Bercini D. The use of verbal report methods in the development and testing of survey questionnaires. *Appl Cogn Psychol* 1991; 5: 251-67
20. Faul GF, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009; 41: 1149-60.
21. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32-5.
22. Zhou XH, Obuchowski NA, Obuschowski DM. Statistical methods in diagnostic medicine. New York: Wiley & Sons; 2002.
23. Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988; 240: 1285-93.
24. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence- based medicine working group. *JAMA* 1994; 271: 389-91.
25. Feise RJ, Michael MJ. Functional rating index: a new valid and reliable instrument to measure the magnitude of clinical change in spinal conditions. *Spine* 2001; 26: 78-86.