

บทที่ 2

วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาครั้งนี้ ผู้วิจัยได้ทบทวนเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้องเพื่อเป็นแนวทางในการศึกษา แล้วนำมาสรุปเป็นประเด็น ๆ หลังจากนั้นผู้วิจัยนำเนื้อหาสาระที่ได้ทั้งหมดมากำหนดเป็นกรอบแนวคิดในการศึกษาครั้งนี้ รายละเอียดต่าง ๆ มีดังต่อไปนี้

1. คุณภาพชีวิตด้านสุขภาพ

1.1 แนวคิดเกี่ยวกับคุณภาพชีวิต (Quality of Life)

1.1.1 ความหมายของคุณภาพชีวิต

คำว่าคุณภาพชีวิตมีการนำมาใช้อย่างกว้างขวางในหลายสาขา เป็นเป้าหมายในการพัฒนาคุณภาพชีวิตและความเป็นอยู่ของประชาชน โดยเชื่อว่าถ้าบุคคลมีคุณภาพชีวิตที่ดีก็จะนำไปสู่การพัฒนาในด้านต่างๆ ได้รวดเร็วและมีประสิทธิภาพ (Bergner, 1989) ในทางการแพทย์ได้กำหนดให้คุณภาพชีวิตเป็นเป้าหมายหลักในการให้บริการนอกเหนือไปจากช่วยให้มีชีวิตรอด รักษาไว้ซึ่งสุขภาพและลดอัตราการเจ็บป่วยลง (จิระภา รัตนกิจ, 2540) โดยที่องค์การอนามัยโลก (WHO) ให้ความหมายคุณภาพชีวิตหมายถึง ระดับของการมีชีวิตที่ดี ความสุข ความพึงพอใจในชีวิต เป็นเรื่องเกี่ยวข้องกับสภาพความเป็นอยู่ การดำเนินชีวิตของบุคคลในสังคม (WHO, 1993)

อาจสรุปได้ว่า คุณภาพชีวิตหมายถึง ระดับความพึงพอใจ ความเป็นสุข ความเป็นอยู่ที่ดีของบุคคลแต่ละคนในสังคมที่ตนดำรงอยู่ ซึ่งขึ้นอยู่กับ การรับรู้และประสบการณ์ชีวิตที่ผ่านมาของบุคคล

1.1.2 องค์ประกอบของคุณภาพชีวิต

เนื่องจากแนวคิดคุณภาพชีวิตมีผู้ศึกษาไว้มากมายและมีมิติที่หลากหลายตามแนวคิดของนักวิชาการที่นำเสนอเอาไว้ ดังนั้นผู้วิจัยจะยกตัวอย่างโดยสังเขป ดังนี้

Zhan (1992) ได้รวบรวมแนวคิดจากนักวิจัยหลายท่านและสรุปว่าองค์ประกอบของคุณภาพชีวิตมี 4 ด้าน ได้แก่ ด้านความพึงพอใจในชีวิตเกี่ยวกับสภาพความเป็นอยู่และการดำเนินชีวิต ด้านอ้อมโนทัศน์เป็นความรู้สึกนึกคิดที่มีต่อตนเอง ความรู้สึกมีคุณค่าในตนเอง ด้านสุขภาพและการทำงานของร่างกายเกี่ยวกับความสามารถในการปฏิบัติกิจกรรมต่าง ๆ ด้านสังคมและเศรษฐกิจ เป็นความพึงพอใจเกี่ยวกับอาชีพ การศึกษา รายได้และการมีปฏิสัมพันธ์กับผู้อื่น และ

องค์การอนามัยโลก (WHO, 1993) ได้สรุปองค์ประกอบคุณภาพชีวิตว่ามี 4 ด้าน ได้แก่ ด้านร่างกาย ด้านจิตใจ ด้านความสัมพันธ์ทางสังคม ด้านสิ่งแวดล้อม

1.1.3 การประเมินคุณภาพชีวิต

การประเมินคุณภาพชีวิตมีวิธีการและแนวคิดที่ต่างกัน ขึ้นอยู่กับวัตถุประสงค์ การวิจัย นักวิชาการ ได้ระบุแนวทางการประเมินคุณภาพชีวิตไว้ต่างกัน เช่น Stromborg (1984) กล่าวไว้ 3 ลักษณะคือ 1) การประเมินเชิงวัตถุวิสัย โดยวัดเป็นข้อมูลเชิงปริมาณ ซึ่งประเมินโดยบุคคลอื่น เช่น แพทย์และบุคลากรสาธารณสุข 2) การประเมินเชิงจิตวิสัย โดยวัดเป็นข้อมูลเชิงปริมาณ ประเมินด้วยตนเอง ซึ่งขึ้นอยู่กับประสบการณ์ของบุคคลนั้น 3) การประเมินเชิงจิตวิสัยโดยวัดเป็นข้อมูลเชิงคุณภาพ ประเมินด้วยตนเอง ซึ่งเป็นการบรรยายและบ่งบอกสภาพที่บุคคลนั้น เป็นอยู่

นอกจากนี้ Meberg (1993) ได้จำแนกตัวชี้วัดคุณภาพชีวิตเป็น 2 ประเภทคือ วัตถุประสงค์ โดยเป็นข้อมูลเชิงรูปธรรม ได้แก่ รายได้ อาชีพ การศึกษา และเชิงจิตวิสัย ได้แก่ ความพึงพอใจในชีวิต ความสุข ความรู้สึกมีคุณค่าในชีวิต เป็นต้น

1.2 แนวคิดเกี่ยวกับคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพ (Health-Related Quality of Life)

สุขภาพเป็นมิติหนึ่งที่มีความสำคัญต่อคุณภาพชีวิต ดังนั้นการวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพจึงมีความจำเป็นและเป็นผลลัพธ์ทางด้านสุขภาพ นอกเหนือจากอายุขัย (Life expectancy) หรือการมีชีวิตรอด (survival) โดยคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพเป็นผลลัพธ์ที่ไม่สามารถมองเห็นได้จากผลทางคลินิกหรือการตรวจทางห้องปฏิบัติการเพียงอย่างเดียว การประเมินคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพจึงมีความสำคัญในการทำความเข้าใจถึงโรคและการรักษาว่ามีผลต่อผู้ป่วยอย่างไร จากมุมมองของผู้ป่วยเอง

คุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพเป็นลักษณะที่เป็นนามธรรมหรือเรียกว่าเป็นตัวแปรที่เป็นโครงสร้าง (construct) ที่เกิดขึ้นด้วยหลายแนวคิดทางด้านสุขภาพ ได้แก่ สุขภาพด้านกายภาพ (physical health) สุขภาพด้านจิตใจ (mental health) สุขภาพด้านสังคม (social health) และภาวะสุขภาพโดยทั่วไป (general health) ซึ่งสอดคล้องกับนิยามของ “สุขภาพ” โดยองค์การอนามัยโลก ที่ว่า สุขภาพไม่ใช่แค่เพียงการปราศจากโรค แต่หมายถึงการมีความสมบูรณ์ทั้งทางด้านร่างกาย จิตใจและสังคม

ดังนั้น คุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพ (Health related Quality of Life) จึงหมายถึง ความเป็นอยู่ที่ดีของบุคคลที่เชื่อมโยงกับสุขภาพในระดับที่เจาะจงมากขึ้น สะท้อนภาวะสุขภาพที่เกี่ยวข้องกับโรคใดโรคหนึ่ง หรือสะท้อนการรับรู้ การประเมินที่เชื่อมโยงกับผลการรักษา การให้

คุณค่าและปฏิกิริยาตอบสนองของบุคคลต่อภาวะสุขภาพของตนเองและต่อมิตีอื่น ๆ ที่ไม่ใช่ด้านการแพทย์แต่เกี่ยวข้องกับชีวิตของบุคคลนั้น

1.3 ความสำคัญของการวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพ

คุณภาพชีวิตเป็นผลลัพธ์ที่ไม่สามารถมองเห็นได้จากผลทางคลินิกหรือการตรวจทางห้องปฏิบัติการ การประเมินคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพมีความสำคัญอย่างมากในการทำ ความเข้าใจว่าโรคและการรักษามีผลอย่างไรต่อชีวิตของผู้ป่วย จากมุมมองของผู้ป่วยเอง ความสำคัญของการประเมินคุณภาพชีวิตมีหลายประการ (Fayers and Machin, 2000) ดังต่อไปนี้

1.3.1 เพื่อใช้ในการประเมินผลการรักษาหรือการดูแลทางสุขภาพ เช่น การเปรียบเทียบผลของยา 2 ชนิด โดยดูผลลัพธ์ทางด้านคลินิก และคุณภาพชีวิต

1.3.2 เพื่อติดตามประเมินภาวะสุขภาพของประชากร เช่น การประเมินคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพของประชากรผู้สูงอายุในชุมชนหนึ่ง หรือในกลุ่มเด็กนักเรียนแห่งหนึ่ง หรือผู้ป่วยเบาหวานในโรงพยาบาลแห่งหนึ่ง

1.3.3 เพื่อใช้ในการติดตามดูแลผู้ป่วยทางคลินิก ซึ่งในทางปฏิบัติการประเมินคุณภาพชีวิตเป็นวิธีการหนึ่งที่ใช้ในการประเมินภาวะสุขภาพของผู้ป่วย นอกจากผลทางคลินิกหรือผลการตรวจทางห้องปฏิบัติการ

1.3.4 จำแนกปัญหาและจัดลำดับความสำคัญของปัญหา จากการบันทึกของแพทย์ บุคลากรทางการแพทย์และผู้ป่วย

1.3.5 มีความสะดวกในการสื่อสารระหว่างแพทย์และผู้ป่วยเนื่องจากการจัดลำดับความสำคัญของปัญหา แพทย์จึงให้การรักษาได้ตรงตามความต้องการของคนไข้จริง ๆ และเกิดการตัดสินใจร่วมกันเลือกแนวทางในการให้การรักษา

1.4 ประวัติการพัฒนาเครื่องมือวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพ (Fayers and Machin, 2000)

เครื่องมือวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพชนิดแรกที่ใช้ประเมินภาวะสุขภาพของคนไข้และใช้อย่างแพร่หลายตั้งแต่ปี ค.ศ. 1947 คือ Karnofsky Performance Scale ซึ่งกำหนดค่าคะแนนตั้งแต่ 0 ถึง 100 โดย 0 หมายถึง เสียชีวิต และ 100 หมายถึง สุขภาพปกติ บุคลากรทางการแพทย์ใช้เครื่องมือดังกล่าวเพื่อประเมินภาวะสุขภาพของคนไข้ช่วงระยะเวลาหนึ่ง ต่อมา มีผู้พัฒนาเครื่องมือชนิดอื่น ๆ เพิ่มขึ้น ได้แก่ Functional ability, Physical functioning และ Activity of daily living (ADL) บางครั้งมีการนำมาใช้วัดคุณภาพชีวิตของคนไข้แต่เนื่องจากเครื่องมือ

ดังกล่าววัดได้เพียงมิติเดียว และเครื่องมือก็ไม่มี ความเหมาะสมในการนำมาประเมินคุณภาพชีวิต (Quality of Life) และภาวะความอยู่ดีมีสุข (Well-being)

ต่อมาได้มีการพัฒนาเครื่องมือวัดคุณภาพชีวิตด้านสุขภาพในช่วงปลาย ค.ศ.1970 ทำให้มีเครื่องมือวัดภาวะสุขภาพทั่วไปเพิ่มมากขึ้น ซึ่งเครื่องมือที่สร้างในช่วงเวลานี้มีวัตถุประสงค์เพื่อวัด ภาวะทางกายภาพ (Physical functioning) อาการทางกายภาพและทางจิต (Physical and psychological symptoms) ผลจากความเจ็บป่วย (impact of illness) การรับรู้ความทุกข์ (perceived distress) และความพึงพอใจในชีวิต (life satisfaction) ตัวอย่างเครื่องมือ เช่น Sickness Impact Profile, Nottingham Health Profile ส่วนใหญ่ผู้ใช้เครื่องมือจะนำมาใช้วัดคุณภาพชีวิต ซึ่งในความเป็นจริงผู้สร้างเครื่องมือดังกล่าวไม่ได้มีวัตถุประสงค์ของการสร้างเพื่อวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพ

ในช่วงเวลาเดียวกัน Priestman และ Baum (1976 อ้างถึงใน Fayers และ Machin, 2000) ปรับปรุงวิธีการที่เรียกว่า Linear analogue self-assessment เพื่อประเมินคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพในคนไข้มะเร็งเต้านม บางครั้งเรียกอีกชื่อหนึ่งว่า Visual analogue scale (VAS) มีลักษณะเป็นเส้นตรงยาว 10 เซนติเมตร ปลายทั้งสองด้านของเส้นตรงมีคำอธิบายประกอบ โดยคนไข้จะเป็นผู้ทำเครื่องหมายลงบนเส้นตรงด้วยตนเอง ตัวอย่างเครื่องมือที่ใช้วิธีการนี้ ได้แก่ EuroQol ซึ่ง Priestman and Baum (1976 อ้างถึงใน Fayers และ Machin, 2000) ใช้วัดลักษณะที่เป็นนามธรรมหลายลักษณะด้วยกัน

ต่อมา มีการสร้างเครื่องมือวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพจำนวนมาก เพื่อการวัดลักษณะที่เป็นนามธรรม เช่น อารมณ์ (emotion) ภูมิกษณะ (role) กระบวนการเรียนรู้ทางสังคม (social and cognitive functioning) ที่พบโดยทั่วไปส่วนใหญ่ เป็นการให้ความสำคัญกับการวัดคุณภาพชีวิตโดยรวม โดยมีความเชื่อมโยงกับลักษณะจิตวิทยาและสังคม (psychological and social) รวมถึงความสมบูรณ์ทางกาย (functional capacity) ดังนั้นเมื่อคนไข้มีความบกพร่องทางด้านกายภาพ จิตวิทยาและสังคม อาจสรุปได้ว่า คนไข้มีคุณภาพชีวิตที่ไม่ดี ซึ่งคุณภาพชีวิตเป็นลักษณะเฉพาะแต่ละบุคคลอาจมีการเปลี่ยนแปลงเมื่อเวลาผ่านไป สำหรับการนำเครื่องมือวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพไปใช้ทางด้านคลินิก เมื่อต้องการเปรียบเทียบผลการรักษา แพทย์ควรตระหนักถึงความแตกต่างของอาการทางคลินิกและผลข้างเคียงที่เกิดจากการให้การรักษาและความแตกต่างที่อาจเกิดขึ้นหลังจากนำเครื่องมือวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพไปใช้งานจริง

1.5 เครื่องมือที่ใช้วัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพ

เครื่องมือที่ใช้วัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพ แบ่งออกเป็น 2 ประเภทดังนี้

1.5.1 เครื่องมือที่ใช้วัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพแบบทั่วไป พัฒนาขึ้นเพื่อวัดคุณภาพชีวิตในทุกกลุ่มประชากรทั้งผู้มีสุขภาพดี ผู้ป่วยโรคต่าง ๆ เช่น SF-36 WHOQoL (มีฉบับภาษาไทย) EUOQoL (มีฉบับภาษาไทย) เป็นต้น

1.5.2 เครื่องมือที่ใช้วัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพแบบจำเพาะ เป็นเครื่องมือวัดผลจากโรคนั้น ๆ โดยตรง ผลที่ได้จะแสดงความเปลี่ยนแปลงของโรคหรือปัญหาสุขภาพในกลุ่มโรคเดียวกัน เช่น Quality of Life Questionnaires for Asthma (ใช้วัดในผู้ป่วยโรคหอบหืด) The Functional Living Index (ใช้วัดในผู้ป่วยโรคมะเร็ง) The Index of Activity of Daily Living (ใช้วัดในผู้ป่วยผู้สูงอายุ)

2. เครื่องมือวัดประเภทสเกล

ตัววัดด้านสุขภาพที่ใช้ในงานสาธารณสุข เพื่อบ่งบอกปัญหาสุขภาพของประชากร ใช้วางแผนงานด้านนโยบายสุขภาพและวางแผนการรักษาด้านการแพทย์ พบว่า หลายตัววัดไม่สามารถวัดได้โดยตรงจากคำถามเพียงข้อเดียว ตัววัดลักษณะนี้เรียกว่าตัววัดที่มีโครงสร้าง จำเป็นต้องใช้คำถามหลายข้อประกอบกัน แล้วนำคำตอบจากคำถามทุกข้อรวมกันจึงจะได้ค่าที่บอกลักษณะของบุคคลที่ถูกวัด (จิราพร เขียวอยู่, 2552)

การวัดลักษณะที่สนใจด้วยคำถามหลายข้อ จะเรียกชุดรายการคำถามว่า สเกล คำถามแต่ละข้อมีการกำหนดคะแนนให้คำตอบซึ่งมีหลายลักษณะ ลักษณะหนึ่งที่นิยมใช้ได้แก่ การกำหนดคะแนนของคำตอบเป็นมาตราเรียงลำดับ (rating scale) เช่น เมื่อวัดความพึงพอใจ กำหนดคะแนนคำตอบจาก 1 ถึง 5 คะแนน โดย 1 หมายถึง พอใจน้อยที่สุด และ 5 หมายถึง พอใจมากที่สุด

เนื่องจากเครื่องมือวัดประเภทสเกลมักใช้กับตัววัดที่มีลักษณะเป็นนามธรรม (subjective) มีหลายมิติและมีการเปลี่ยนแปลงตามเวลา (Wilkin และคณะ, 1992) และอยู่ในรูปของชุดข้อคำถาม (multi-item scales) ดังนั้น คำถามต่าง ๆ จึงควรมีความสัมพันธ์กัน (จิราพร เขียวอยู่, 2552) ซึ่ง Coste และคณะ (1995) เรียกตัวแปรดังกล่าวว่า composite variable และเรียกเครื่องมือวัดตัวแปรที่มีความซับซ้อนนี้ว่า composite measurement scales

ปัจจุบันการสร้างเครื่องมือประเภทสเกลเพื่อใช้พบได้ 2 ลักษณะคือ นำเครื่องมือที่มีอยู่แล้วที่มักสร้างโดยนักวิจัยต่างประเทศมาแปลเป็นภาษาไทยแล้วปรับปรุงให้เหมาะสมกับบริบทของคนไทยและการสร้างขึ้นใหม่โดยนักวิจัย ในกรณีการนำเครื่องมือจากต่างประเทศมาใช้ ก่อนใช้จริงจะต้องผ่านกระบวนการแปลภาษา ทดสอบคุณสมบัติความตรงเชิงเนื้อหา ทดสอบคุณสมบัติความ

เที่ยง (reliability) (Fayers and Machin, 2000) สำหรับกรณีพัฒนาเครื่องมือขึ้นใหม่ จะเป็นกรณีที่เครื่องมือชนิดเดิมที่มีอยู่แล้วมีประเด็นไม่ครอบคลุมในเรื่องที่ต้องการวัด (Spector, 1992; Streiner and Norman, 2003) ในกรณีนี้จะต้องดำเนินการสร้างด้วยวิธีการมาตรฐาน

ประเภทของสเกลที่สร้างขึ้นเพื่อวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพส่วนใหญ่จะเป็นประเภทลิเคิทสเกล (Likert scale หรือ rating scale หรือ summative scale) คือประกอบด้วยชุดรายการคำถามหลาย ๆ ข้อ แต่ละคำถามมีความสัมพันธ์กันและวัดในเรื่องเดียวกัน มีการให้คะแนนตัวเลือกที่เป็นตัวเลขเรียงลำดับจากน้อยไปมาก เช่น เรียงจาก 1 ถึง 5 หรือ 1 ถึง 7 (Fayers and Machin, 2000) ระเบียบวิธีการสร้างเครื่องมือประเภทสเกลจะนำเสนอรายละเอียดในหัวข้อถัดไป

3. ขั้นตอนการสร้างสเกล

ผู้วิจัยได้ศึกษาวิธีการสร้างสเกลจาก Spector (1992), Devellis (2003), Streiner and Norman (2003) และ Fayers and Machin (2000) ซึ่งดำராแต่ละเล่มได้เสนอขั้นตอนที่แตกต่างกันบ้างแต่เมื่อพิจารณาในภาพรวมพบว่า มีขั้นตอนที่สำคัญเหมือนกัน 6 ขั้นตอน รายละเอียดสรุปได้ดังต่อไปนี้

3.1 การกำหนดแนวคิด คำนิยามหรือโครงสร้างของตัวแปรที่ต้องการสร้างสเกลเพื่อวัด

(Define construct)

ก่อนที่จะเริ่มสร้างเครื่องมือ ในการกำหนดโครงสร้างเชิงเนื้อหาของสเกลประกอบด้วย

- 1) การให้ความหมายหรือกำหนดคำนิยามของสิ่งที่ต้องการวัด
- 2) กำหนดวัตถุประสงค์ประสงค์ของการวัด ซึ่งครอบคลุมกลุ่มเป้าหมายและวัตถุประสงค์เบื้องต้นของการใช้เครื่องมือ
- 3) การกำหนดและจำแนกแต่ละด้านของข้อคำถามเบื้องต้น เป็นการแบ่งมิติขององค์ประกอบของเครื่องมือโดยยึดถือตามแนวคิดหรือทฤษฎีที่เกี่ยวข้องกับสิ่งที่ต้องการวัด ขึ้นอยู่กับผู้สร้างและการให้คำนิยาม Fayers and Machin (2000)

สำหรับรายละเอียดการกำหนดแนวคิด คำนิยามหรือโครงสร้างของตัวแปรที่ต้องการสร้างสเกลเพื่อวัด ทำได้หลายวิธี ได้แก่ การทบทวนวรรณกรรม แนวคิดหรือเนื้อหาทางทฤษฎีที่เกี่ยวข้อง งานวิจัยที่ผ่านมา และสเกลที่ปรากฏอยู่ในปัจจุบันที่มีความเกี่ยวข้องกับสิ่งที่นักวิจัยต้องการสร้างสเกลเพื่อวัด (Spector, 1992)



3.2 การรวบรวมข้อคำถาม (Generate the potential items)

การสร้างข้อคำถามแต่ละข้อมาจากตัวชี้วัดแต่ละตัวหรือตัวชี้วัดหนึ่งอาจมีหลายข้อคำถาม เมื่อกำหนดด้านหรือประเด็นหลักของเนื้อหาแล้ว ต้องแจกแจงตัวชี้วัดให้ครบถ้วนเพื่อให้เขียนข้อคำถามได้ครอบคลุมทุกตัวชี้วัด โดยวิธีการรวบรวมข้อคำถามมีข้อแนะนำ 2 วิธี ได้แก่ การลดทอนเนื้อหา (deductive approach) และการเก็บรวบรวมแล้วจัดกลุ่มในภายหลัง (inductive approach) (Spector, 1992)

สำหรับการรวบรวมข้อมูลเพื่อนำมาสร้างเป็นข้อคำถามทำได้หลายวิธี ได้แก่ ปรับปรุงจากงานวิจัยที่มีผู้ศึกษาไว้ เขียนขึ้นใหม่โดยอาศัยแนวคิดหรือทฤษฎีที่เกี่ยวข้อง การสัมภาษณ์จากผู้เชี่ยวชาญและการระดมสมอง อาศัยความเห็นและข้อเสนอแนะที่ได้จากการสนทนากลุ่มจากบุคคลที่มีส่วนเกี่ยวข้องกับเรื่องที่จะสร้างสเกล (Streiner and Norman, 2003; Spector, 1992)

หลังจากที่ได้ร่างข้อคำถามมาแล้ว จะนำมาประเมินความตรงเชิงเนื้อหา (content validity) โดยอาศัยผู้เชี่ยวชาญตรวจสอบอย่างน้อย 3 คนแล้วนำมาแก้ไข และยังมีอีกวิธีการหนึ่งที่วัดได้เชิงปริมาณคือการพิจารณาเป็นดัชนีความตรงเชิงเนื้อหา (content validity index: CVI) ที่พิจารณาเป็น 4 ระดับ โดยที่ Lynn (1986) เสนอเกณฑ์พิจารณาว่า ควรมีสัดส่วนผู้เชี่ยวชาญที่เห็นด้วยกับข้อคำถาม โดยให้ระดับการประเมินระดับที่ 3 หรือ 4 อย่างน้อยร้อยละ 80 ของผู้เชี่ยวชาญทั้งหมด (Streiner and Norman, 2003)

ห้องสมุดงานวิจัย
วันที่..... 24 S.A. 2555.....
เลขทะเบียน..... 203316.....
เลขวิทยานิพนธ์.....

3.3 การกำหนดประเภทและรูปแบบของสเกล (Design the scale)

วิธีการสร้างสเกลมีหลายวิธี ได้แก่ ลิเคิร์ตสเกล (Likert scale) เซอร์สโตนสเกล (Thurstone scale) กัทแมนสเกล (Guttman scale) และซีแมนติกดิฟเฟอเรนเชียลสเกล (Semantic differential scale)

สำหรับตัวเลือกแบบ Likert scale เป็นสเกลวัดที่ประกอบด้วยรายการคำถามหลายข้อที่ระบุสิ่งที่ต้องการวัดเป็นค่าคะแนน แต่ละข้อคำถามไม่มีถูกผิดซึ่งมีลักษณะให้ผู้ตอบประเมินค่าได้ด้วยตัวเอง การกำหนดระดับของตัวเลือกที่พบตั้งแต่ 3-7 ระดับ ขึ้นอยู่กับผู้สร้าง ที่พบบ่อย 4-5 ระดับ (Fayers and Machin, 2000) แต่การใช้ตัวเลือกน้อยเกินไปทำให้ความแม่นยำ (precision) และความเที่ยง (reliability) ต่ำ ดังนั้น ควรกำหนดประมาณ 5-9 ระดับ (Spector, 1992)

3.4 การทดสอบสเกลเบื้องต้น (Pre-testing)

นำเครื่องมือไปทดสอบในกลุ่มตัวอย่างขนาดเล็กประมาณ 10 – 30 คน มีเป้าหมายเพื่อทดสอบความเข้าใจเกี่ยวกับความหมายของข้อคำถาม วิธีการตอบ และการเรียงลำดับของข้อคำถาม

อาจมีการจับเวลาที่ใช้ในการตอบคำถาม ผลที่ได้จะนำไปแก้ไขต่อไป เช่น การปรับปรุงภาษา เพิ่มรายการคำถามที่ขาดไป ตัดรายการคำถามที่สร้างความลำบากใจให้ผู้ตอบเพื่อลดปัญหาการไม่ตอบในข้อนั้น ๆ (missing data) (Fayers and Machin, 2000)

3.5 การวิเคราะห์รายการคำถาม (Item analysis)

ในการตรวจสอบโครงสร้างของเครื่องมือ (scale construction) หรือการวิเคราะห์รายการคำถาม (item analysis) ต้องผ่านขั้นตอนการตรวจสอบความตรงเชิงเนื้อหา (content validity) โดยผู้เชี่ยวชาญ ซึ่งเป้าหมายของการวิเคราะห์รายการคำถามคือ การลดข้อคำถาม (data reduction) ให้เหลือเฉพาะที่ต้องการและจำเป็นเท่านั้น โดยใช้ผลที่ได้จากการวิเคราะห์รายการคำถาม กระทำได้ 5 วิธีดังต่อไปนี้ (จิราพร เขียวอยู่, 2548)

3.5.1 การพิจารณาการกระจายคำตอบของแต่ละข้อคำถาม (distribution of response) พิจารณาการกระจายของคำตอบควรมีการกระจายในทุกระดับ ถ้าคำตอบอยู่ที่ระดับใดระดับหนึ่ง แสดงว่าข้อคำถามนั้นจำแนกไม่ได้ควรตัดออก เกณฑ์พิจารณาคือสัดส่วนการเลือกรายการคำตอบมีค่าระหว่าง 0.2 ถึง 0.8 (Streiner and Norman, 2003)

3.5.2 วิเคราะห์ความสัมพันธ์ระหว่างแต่ละรายการคำถาม (inter-item correlation) พิจารณาความสัมพันธ์แต่ละข้อคำถามภายในสเกลย่อยเดียวกัน ควรมีความสัมพันธ์กันในทิศทางบวกไม่น้อยกว่า 0.2 ในข้อคำถามที่มีความสัมพันธ์กันในทิศทางลบให้กลับค่าคะแนนตัวเลือกเป็นตรงข้ามกับค่าเดิม (recode)

3.5.3 วิเคราะห์ความสัมพันธ์ระหว่างคะแนนของแต่ละข้อคำถามกับคะแนนรวม (item-total correlation) เป็นความสัมพันธ์ของข้อคำถามแต่ละข้อกับคะแนนรวมของข้อคำถามอื่นๆ ที่ไม่รวมข้อคำถามนั้น ควรมีระดับความสัมพันธ์มากกว่า 0.2

3.5.4 วิเคราะห์ความสอดคล้องภายในระหว่างแต่ละรายการคำถาม (inter-item consistency) ข้อคำถามในแต่ละสเกลย่อยควรสะท้อนสิ่งที่ต้องการวัดในเรื่องเดียวกัน จะได้จากการหาค่าสัมประสิทธิ์ความสอดคล้องภายใน (internal consistency) ด้วยวิธีการสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's alpha coefficient) มีค่า 0 ถึง 1 และค่า Coefficient alpha ไม่ควรต่ำกว่า 0.7 แต่ไม่เกิน 0.9 (Nunnally, 1978)

3.5.5 วิเคราะห์ปัจจัย (factor analysis) มีการนำมาใช้ใน 3 กรณีคือ 1) การลดข้อคำถามในแต่ละสเกลย่อย 2) การวิเคราะห์หาโครงสร้างของข้อมูลว่ามีการรวมตัวกันหรือไม่ และเมื่อรวมตัวกันแล้วมีโครงสร้างร่วมกันอย่างไร จำแนกออกเป็นกี่มิติ เรียกวิธีการนี้ว่า การวิเคราะห์ปัจจัยเพื่อหาโครงสร้างของข้อมูล (exploratory factor analysis) 3) การยืนยันแนวคิดทางทฤษฎีเพื่อ



ตรวจสอบความตรงเชิงโครงสร้างของเครื่องมือชนิดใหม่ โดยผู้สร้างต้องการทราบว่าเครื่องมือที่สร้างขึ้นมีความสอดคล้องกับทฤษฎีหรือไม่ วิธีการวิเคราะห์ปัจจัยกรณีนี้เรียกว่า การวิเคราะห์ปัจจัยเชิงยืนยัน (confirmatory factor analysis)

เกณฑ์การคัดเลือกหรือตัดข้อคำถามพิจารณาค่าน้ำหนักองค์ประกอบ (factor loading) ต่ำกว่า 0.4 ในองค์ประกอบย่อยเดียวกันออกไป และการคัดเลือกจำนวนด้านหรือองค์ประกอบใช้ค่า eigenvalue มากกว่า 1

สำหรับวิธีการทางสถิติในการวิเคราะห์รายการคำถามแต่ละประเภทได้แก่ สถิติ Pearson correlation coefficient ใช้ประเมินค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม (item-total correlation) และวิเคราะห์ความสัมพันธ์ระหว่างแต่ละรายการคำถาม (inter-item correlation) สถิติ Cronbach's alpha coefficient ใช้ประเมินค่าความเที่ยงประเภทค่าความสอดคล้องภายใน (internal consistency) สถิติ Exploratory factor analysis ใช้ในการวิเคราะห์ปัจจัย (Spector, 1992; Devellis, 2003; Streiner and Norman, 2003)

3.6 การตรวจสอบคุณสมบัติด้านจิตเมตริกของสเกล (Psychometric property)

หลังจากผ่านขั้นตอนการวิเคราะห์รายการคำถามและได้ข้อคำถามที่เหมาะสมมาจำนวนหนึ่งแล้ว ต่อไปเป็นการตรวจสอบคุณสมบัติด้านจิตเมตริกของสเกล รายละเอียดดังนี้

3.6.1 คุณสมบัติด้านจิตเมตริกของสเกล

คุณสมบัติด้านจิตเมตริกของสเกลที่สำคัญมี 2 ประการ คือความตรง (validity) และความเที่ยง (reliability) แต่สำหรับสเกลที่นำมาใช้วัดตัววัดด้านสุขภาพ มีการให้ความสนใจกับคุณสมบัติที่สำคัญอีกหนึ่งประการ คือคุณสมบัติด้านความไวต่อการเปลี่ยนแปลง (sensitivity to change) (จิราพร เขียวอยู่, 2548)

กรณีความตรง แบ่งเป็น 3 ประเภท ได้แก่ ความตรงตามเนื้อหา (content validity) ความตรงตามโครงสร้าง (construct validity) และความตรงตามเกณฑ์สัมพัทธ์ (criterion-related validity) สำหรับความหมายของความตรงทั้ง 3 ประเภท มีดังนี้ ความตรงตามเนื้อหา หมายถึง เครื่องมือมีเนื้อหาที่ครอบคลุมทุกประเด็นของคุณสมบัติของตัวแปรที่ต้องการวัด โดยเป็นเนื้อหาที่สอดคล้องกับแนวคิดและเป็นไปตามค่านิยมเชิงปฏิบัติการของตัวแปรนั้น ความตรงตามโครงสร้าง หมายถึง ผลการวัดของลักษณะหรือตัวแปรที่สนใจตรงกับหรือสอดคล้องกับหรือเป็นไปตามแบบแผนของความสัมพันธ์ ที่คาดหวังไว้ตามแนวคิดหรือทฤษฎีของเนื้อหาสาระที่สนใจนั้น ๆ พิจารณาได้ 2 ลักษณะคือ เมื่อใช้เครื่องมือคนละชนิดวัดคุณลักษณะเดียวกัน จะมีความสอดคล้องกันเรียกความตรงนี้ว่าความตรงเชิงลู่เข้า (Convergent validity) และใช้เครื่องมือชนิดเดียวกันวัด

คุณลักษณะต่างกัน จะไม่มีความสอดคล้องกัน เรียกความตรงนี้ว่าความตรงเชิงจำแนก (Discriminant validity) และความตรงตามเกณฑ์สัมพัทธ์ หมายถึง เครื่องมือวัดได้ตรงกับเกณฑ์ หรือเครื่องมือวัดมาตรฐาน (gold standard) แบ่งออกเป็น 2 ประเภท คือความตรงร่วมสมัย (concurrent validity) หมายถึง เครื่องมือวัดได้ตรงกับเครื่องมือวัดมาตรฐาน (gold standard) ในช่วงเวลาปัจจุบัน และความตรงเชิงทำนาย (predictive validity) หมายถึง เครื่องมือวัดได้ตรงกับเครื่องมือวัดมาตรฐาน (gold standard) ในช่วงเวลาในอนาคต

ความเที่ยง คือ การวัดที่เมื่อทำการวัดหลาย ๆ ครั้ง แล้วให้ค่าตรงกันทุกครั้ง โดยการวัดแต่ละครั้งกระทำภายใต้สภาพการณ์เดียวกัน แบ่งเป็น 3 ประเภทได้แก่ ความคงที่ของการวัด (Test-retest reliability) ความสอดคล้องภายใน (Internal consistency) และความเที่ยงคู่ขนาน (Alternative form reliability) สำหรับความหมายของความเที่ยงทั้ง 3 ประเภทมีดังนี้ ความคงที่ของการวัด หมายถึง ความเที่ยงของการวัดเนื่องมาจากการทำการวัด ณ เวลาต่างกัน ความสอดคล้องภายใน หมายถึง คุณสมบัติของเครื่องมือประเภทสเกลที่สร้างเพื่อวัดลักษณะที่สนใจบางลักษณะ โดยลักษณะนั้นมีโครงสร้างตามแนวคิดหรือทฤษฎีบางประการ การประเมินค่าความสอดคล้องภายใน ได้แก่ วิธีการแบ่งครึ่ง (Split-half method) และวิธีสัมประสิทธิ์แอลฟา (Alpha coefficient) และความเที่ยงคู่ขนาน หมายถึง คุณสมบัติของเครื่องมือสองชุดใช้วัดลักษณะที่สนใจลักษณะเดียวกัน ทดลองใช้กับกลุ่มตัวอย่างกลุ่มเดียวกันในช่วงเวลาที่ต่างกัน หากคะแนนที่ได้จากแบบทดสอบสองชุดมีความสัมพันธ์กันระดับสูงแสดงว่าแบบทดสอบทั้งสองมีความเที่ยงคู่ขนาน

ความไวต่อการเปลี่ยนแปลง หมายถึง ความสามารถของเครื่องมือในการประเมินความเปลี่ยนแปลงของสุขภาพที่เกิดขึ้นในช่วงเวลาถัดไป (Maurice et al., 1998; Fayers and Machin, 2000)

3.6.2 การทดสอบคุณสมบัติด้านต่าง ๆ

การทดสอบคุณสมบัติด้านต่าง ๆ ดังกล่าวข้างต้นรายละเอียดดังนี้

ความตรงตามเนื้อหา ทำได้โดยให้ผู้เชี่ยวชาญตรวจสอบอย่างน้อย 3 คนแล้วนำมาแก้ไข และยังมีอีกวิธีการหนึ่งที่วัดได้เชิงปริมาณคือการพิจารณาเป็นดัชนีความตรงเชิงเนื้อหา (content validity index: CVI) ที่พิจารณาเป็น 4 ระดับ โดยที่ Lynn (1986) เสนอเกณฑ์พิจารณาว่าควรมีสัดส่วนผู้เชี่ยวชาญที่เห็นด้วยกับข้อคำถาม โดยให้ระดับการประเมินระดับที่ 3 หรือ 4 อย่างน้อยร้อยละ 80 ของผู้เชี่ยวชาญทั้งหมด (Streiner and Norman, 2003)

ความตรงตามโครงสร้าง ประเมินได้หลายวิธี เมื่อต้องการประเมินความตรงเชิงคู่เข้าและความตรงเชิงจำแนก จะใช้วิธีการทางสถิติค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) ระหว่างตัวแปรสองตัว อีกวิธีหนึ่งเป็นการเปรียบเทียบค่าของตัวแปรที่วัดระหว่างกลุ่ม

ที่ทราบความแตกต่าง (know group method) (Maurice et al., 1998) โดยการนำเครื่องมือวัดไปทดสอบในกลุ่ม 2 กลุ่ม คือกลุ่มที่มีลักษณะที่ต้องการวัดและกลุ่มที่ไม่มีลักษณะที่ต้องการวัด แล้วนำผลที่ได้มาเปรียบเทียบกันในกรณีที่มีข้อมูลมีการแจกแจงปกติใช้สถิติ Independent sample t-test หากไม่มีการแจกแจงแบบปกติใช้สถิติไม่อิงพารามิเตอร์ (non-parametric) ทดสอบโดยใช้สถิติ Mann-Whitney U test และการวิเคราะห์องค์ประกอบ (factor analysis) เป็นวิธีการตรวจสอบการรวมตัวของตัวแปรที่มีความสัมพันธ์กันไว้ในกลุ่มเดียวกัน ตัวแปรที่มีความสัมพันธ์กันมากจะอยู่ในปัจจัย (factor) เดียวกัน ทำได้โดยใช้วิธีการทางสถิติการวิเคราะห์ปัจจัยเชิงสำรวจ (exploratory factor analysis)

ความตรงตามเกณฑ์สัมพัทธ์ ประเมินได้หลายวิธี เมื่อต้องการประเมินความตรงร่วมสมัย จะประเมินด้วยวิธีการทางสถิติประเภทการหาความสัมพันธ์ (correlation) เช่น ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson Correlation Coefficient) หรือสถิติในกลุ่มการวิเคราะห์ความพ้อง (agreement) เช่น สถิติแคปป่า (Kappa) และสัมประสิทธิ์ความสัมพันธ์ภายในกลุ่ม (intra-class correlation coefficient: ICC) สัมประสิทธิ์ความสอดคล้อง (concordance correlation coefficient: CCC) สำหรับการประเมินความตรงเชิงทำนายใช้วิธีการทางสถิติการวิเคราะห์ถดถอย ทั้งนี้การเลือกใช้วิธีการทางสถิติแบบใดให้พิจารณาสเกลการวัดของตัวแปร

ความเที่ยง มีวิธีการประเมินตามประเภทความเที่ยงดังนี้ **ความคงที่ของการวัด** ทำได้โดยใช้เครื่องมือวัดซ้ำในช่วงเวลาต่างกันในกลุ่มคนเดียวกัน ผู้เชี่ยวชาญแนะนำช่วงเวลาที่เหมาะสมควรอยู่ระหว่าง 2-14 วัน (Streiner and Norman, 2003) แล้วนำค่าคะแนนทั้งสองครั้งหาค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) (Spector, 1992) ควรมีค่า 0.8 (Streiner and Norman, 2003) **ความสอดคล้องภายใน** มีหลายวิธีดังนี้ วิธีการแบ่งครึ่ง (Split-half method) ใช้สเกลในการทดสอบ 1 ชุด แล้วแบ่งสเกลออกเป็น 2 ส่วน จากนั้นคำนวณค่าสหสัมพันธ์ของทั้งสองสเกลย่อย โดยใช้วิธีการทางสถิติค่าสัมประสิทธิ์สหสัมพันธ์สเปียร์แมน บราวน์ (Spearman-Brown's formular) สำหรับการประเมินความสอดคล้องภายในด้วยวิธีสัมประสิทธิ์แอลฟา (Alpha coefficient) รายละเอียดกล่าวไว้ในหัวข้อ 3.5.4 และ**ความเที่ยงคู่ขนาน** ทำได้โดยใช้เครื่องมือ 2 ฉบับ มีลักษณะที่เท่าเทียมกัน (Equivalent form) ทดลองใช้กับกลุ่มตัวอย่างกลุ่มเดียวกันในช่วงเวลาที่ต่างกัน เวลาห่างกันประมาณ 1 ถึง 2 สัปดาห์ แล้วนำคะแนนจากทั้งสองฉบับคำนวณค่าความเที่ยงโดยใช้วิธีการทางสถิติค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) หากคะแนนที่ได้จากเครื่องมือทั้งสองฉบับมีความสัมพันธ์กันในทิศทางเดียวกัน แสดงว่าเครื่องมือทั้งสองฉบับมีความเที่ยงคู่ขนานกัน

ความไวต่อการเปลี่ยนแปลง มีวิธีประเมินหลายวิธี (Maurice et al., 1998; Fayers and Machin, 2000) ดังนี้

ตารางที่ 1 วิธีการวัดความไว (sensitivity to change)

วิธีการวัดความไว	สูตรการคำนวณ	ตัวหาร
Effect size (ES)	D/SD	SD ของ baseline
Standardized response mean (SRM)	D/SD diff	SD ของ Change
t-test comparison	D/SE diff	SE ของ Change
Responsiveness statistic	D/SD stable	SD ของ Change

โดยที่

D หมายถึง ความแตกต่างของคะแนนระหว่างการวัดสองครั้ง

SD หมายถึง ค่าส่วนเบี่ยงเบนมาตรฐาน

SD stable หมายถึงค่าส่วนเบี่ยงเบนมาตรฐานของความแตกต่าง(D)ในกลุ่มสถานะสุขภาพคงที่

SD diff หมายถึง ค่าส่วนเบี่ยงเบนมาตรฐานของความแตกต่างของคะแนนที่วัดสองครั้ง

SE diff หมายถึง ค่าความคลาดเคลื่อนมาตรฐานของความแตกต่างของค่าคะแนนที่วัดสองครั้ง

ค่าที่ได้จากการคำนวณหากมีค่าสูง แสดงว่าเครื่องมือวัดความเปลี่ยนแปลงได้ดี แต่ยังไม่ได้ว่าวิธีการคำนวณที่กล่าวมาวิธีใดดีที่สุด ดังนั้นจึงควรใช้หลายวิธีเพื่อยืนยันผล อีกทั้งควรพิจารณาความสัมพันธ์ของค่าความไวที่คำนวณได้กับมาตรฐานการรักษาอื่นๆ เช่น การเปลี่ยนแปลงทางคลินิก แล้วจึงพิจารณาความสามารถของเครื่องมือในการวัดความแตกต่างระหว่างกลุ่ม (Fayers and Machin, 2000; Streiner and Norman, 2003)

3.7 การรายงานการนำไปใช้ (พรรณทิพา ศักดิ์ทอง, 2550)

เครื่องมือที่สร้างใหม่ควรมีคู่มือการใช้ มีวิธีการคิดคะแนนและการแปลผล และมีรายละเอียดของหน่วยงานหรือบุคคลที่สามารถติดต่อเพื่อขอ นำเครื่องมือไปใช้

4. งานวิจัยเกี่ยวกับการสร้างสเกลวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพ

จากการศึกษาข้อมูลเบื้องต้นโดยผู้วิจัย ซึ่งทำการสำรวจบทความวิจัยที่พิมพ์เผยแพร่ในวารสารจดหมายเหตุทางการแพทย์ที่พิมพ์เผยแพร่ทั้งหมด 1,637 บทความ ระหว่างเดือนมกราคม พ.ศ.2543 ถึง เดือนธันวาคม พ.ศ.2552 พบว่าเป็นบทความวิจัยประเภทสร้างสเกลเพื่อวัดคุณภาพ



ชีวิตที่เกี่ยวข้องกับสุขภาพประเภทสร้างใหม่จำนวน 2 บทความ (Penpimol, 2007; Sucheera, 2005) รายละเอียดในบทความทั้ง 2 บทความ ปรากฏดังตารางที่ 3 ภาคผนวก ก

จาก 2 บทความ นักวิจัยสร้างเครื่องมือเพื่อวัดตัววัดด้านสุขภาพของโรคเรื้อรัง ได้แก่ โรคสายตาเลือนราง และปัญหาสุขภาพจิต

ประเด็นการรายงานระเบียบวิธีในการสร้างเครื่องมือประเภทสเกลในบทความวิจัยทั้ง 2 บทความ สรุปได้ดังนี้

การกำหนดแนวคิดหรือโครงสร้างของตัวแปรที่ต้องการสร้างสเกลเพื่อวัด (Define construct) ทบทวนบทความ 2 บทความ พบว่าทั้ง 2 ของบทความที่ทบทวน มีการกำหนดโครงสร้างเชิงเนื้อหาของสิ่งที่ต้องการวัด พบรายงานในรายละเอียดของการทบทวนแนวคิด/ทฤษฎี ร่วมกับการศึกษาจากงานวิจัยที่ผ่านมา รวมทั้งการสัมภาษณ์เชิงลึกจากผู้เชี่ยวชาญหรือทำ การสนทนากลุ่มจากกลุ่มเป้าหมาย

การรวบรวมรายการคำถาม (Generate the potential items) ทบทวนบทความ 2 บทความ พบว่าทั้ง 2 บทความ รายงานการรวบรวมรายการคำถาม ด้วยวิธีการทบทวนแนวคิด/ทฤษฎี ร่วมกับการศึกษาจากงานวิจัยที่ผ่านมา รวมทั้งการสัมภาษณ์เชิงลึกจากผู้เชี่ยวชาญหรือทำ การสนทนากลุ่มจากกลุ่มเป้าหมาย

การกำหนดประเภทและรูปแบบของสเกล (Design the scale) ทบทวนบทความ 2 บทความ พบว่าทั้ง 2 บทความที่ทบทวน ใช้สเกลของลิเคิร์ต (Likert scale) ที่กำหนดตัวเลือกของ ระดับคำตอบเป็น 5 ระดับเหมือนกัน

การทดสอบสเกลเบื้องต้น (Pretesting) ทบทวนบทความ 2 บทความ พบว่า 1/2 ของ บทความที่ทบทวนรายงานการทดสอบสเกลเบื้องต้น

การวิเคราะห์รายการคำถาม (Item analysis) ทบทวนบทความ 2 บทความ พบว่า 2/2 ของบทความที่ทบทวน รายงานการวิเคราะห์รายการคำถามด้วยวิธีการตรวจสอบสอดคล้องภายใน (internal consistency) ด้วยวิธีการทางสถิติ Cronbach's alpha coefficient และ 1/2 ของบทความที่ ทบทวนการวิเคราะห์ความสัมพันธ์ระหว่างคะแนนรายข้อและคะแนนรวม (item total correlation) ด้วยวิธีการทางสถิติ Pearson correlation และการวิเคราะห์ปัจจัยเพื่อหาโครงสร้างของข้อมูล โดย วิธีการ exploratory factor analysis

ความตรง (Validity) การรายงานการประเมินความตรงตามเนื้อหา (content validity) ทบทวนบทความ 2 บทความ พบว่า 1/2 ของบทความที่ทบทวน รายงานการทดสอบความตรงตาม เนื้อหาด้วยวิธีการสอบถามจากผู้เชี่ยวชาญ การรายงานการประเมินความตรงตามเกณฑ์สัมพันธ์ (criterion relate validity) ทบทวนบทความ 2 บทความ พบว่า 1/2 ของบทความที่ทบทวน รายงาน

การประเมินความตรงตามเกณฑ์สัมพัทธ์ประเภทความตรงร่วมสมัย (concurrent validity) โดยใช้วิธีการทางสถิติ Pearson correlation การรายงานการประเมินความตรงตามโครงสร้าง (construct validity) ทบทวนบทความ 2 บทความ พบว่า 2/2 ของบทความที่ทบทวน ในจำนวนนี้รายงานการประเมินความตรงตามโครงสร้างประเภทความตรงเชิงคู่เข้า 2/2 บทความ และความตรงเชิงจำแนก 2/2 บทความ โดยใช้วิธีการทางสถิติ Pearson correlation และ factor analysis

การรายงานการประเมินความเที่ยง (Reliability) ทบทวนบทความ 2 บทความ พบว่า 2/2 ของบทความที่ทบทวน รายงานการประเมินความเที่ยงประเภทความสอดคล้องภายใน (internal consistency) ด้วยวิธีสัมประสิทธิ์แอลฟา โดยใช้วิธีการทางสถิติ Cronbach's alpha coefficient

การทดสอบคุณสมบัติความไวต่อการเปลี่ยนแปลง (Sensitivity to change) ทบทวนบทความ 2 บทความ ไม่พบการรายงานการทดสอบคุณสมบัติความไวต่อการเปลี่ยนแปลง

5. งานวิจัยประเภทประเมินคุณภาพของงานวิจัยประเภทสร้างสเกล

จากการทบทวนบทความวิจัยประเภททบทวนระเบียบวิธีการสร้างเครื่องมือประเภทสเกลเพื่อวัดตัวแปรเชิงโครงสร้าง (construct) พบงานวิจัยจำนวน 4 บทความ (Hinkin, 1995; Hensley, 1999; Costes, 1995; Costes, 1997) ที่พิมพ์เผยแพร่ระหว่างปี ค.ศ. ถึง 1984 ถึง ปี ค.ศ. 1996 ที่มีวัตถุประสงค์เพื่อทบทวนระเบียบวิธีการสร้างสเกลที่พิมพ์เผยแพร่ในวารสารต่าง ๆ โดยนักวิจัยจะกำหนดประเด็นที่ศึกษาและเกณฑ์ในการพิจารณา ซึ่งมีรายละเอียดที่เกี่ยวข้องกับระเบียบวิธีวิจัยและวิธีการทางสถิติในการสร้างสเกลรวมถึงการตรวจสอบคุณสมบัติด้านจิตเมตริก (Psychometric properties) ได้แก่ ความเที่ยง (reliability) ความตรง (validity) และความไวต่อการเปลี่ยนแปลง (sensitivity to change)

บทความวิจัยทั้งงานวิจัยทั้ง 4 บทความนี้ทบทวน จะสร้างสเกลเพื่อวัดตัวแปรต่างๆ ดังนี้

Coste (1995) และ Coste (1997) ศึกษาบทความในบริบทด้านการแพทย์ เช่น ความเสี่ยง ความรุนแรงในการเกิดโรคและคุณภาพชีวิต Hinkin (1995) ศึกษาบทความในบริบทของพฤติกรรมองค์กร ทักษะคิดและการรับรู้ Henley (1999) ศึกษาบทความในบริบทของการจัดการองค์กร

ประเด็นการรายงานระเบียบวิธีวิจัยและสถิติในการสร้างเครื่องมือประเภทสเกลในบทความต่าง ๆ ที่ผู้วิจัยของ 4 บทความข้างต้นทำการศึกษา สรุปได้ดังนี้

5.1 การกำหนดแนวคิดหรือโครงสร้างของตัวแปรที่ต้องการสร้างสเกลเพื่อวัด (Define construct) การศึกษาของ Hinkin (1995) ซึ่งทบทวนบทความ 75 บทความ พบ ร้อยละ 83 ของบทความที่ทบทวน (62/75) มีการกำหนดนิยามของสิ่งที่ต้องการวัดเพื่อนำมากำหนดโครงสร้างเชิงเนื้อหาของตัวแปรที่ต้องการสร้างสเกลเพื่อวัด โดยใช้วิธีการทบทวนแนวคิด/ทฤษฎี และศึกษาจาก

รายงานวิจัยที่ผ่านมา ร้อยละ 11 ของบทความที่ทบทวน(8/75) ใช้วิธีการสนทนากลุ่มจากกลุ่มเป้าหมาย และ Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 6/6 ของบทความที่ทบทวน รายงานการกำหนดโครงสร้างเชิงเนื้อหาของตัวแปรที่ต้องการสร้างสเกลเพื่อวัด โดยใช้วิธีการทบทวนแนวคิด/ทฤษฎี

5.2 การรวบรวมรายการคำถาม (Generate the potential items) การศึกษาของ Hinkin (1995) ทบทวนบทความ 75 บทความ พบว่า 8/75 ของบทความที่ทบทวน รายงานการรวบรวมข้อคำถามด้วยวิธีการสอบถามจากคนไข้โดยตรง และ 9/75 ของบทความที่ทบทวน รายงานการรวบรวมข้อคำถามโดยขอคำแนะนำจากผู้เชี่ยวชาญ สำหรับการศึกษานี้ของ Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 6/6 ของบทความที่ทบทวน รายงานการรวบรวมข้อคำถามโดยไม่ระบุวิธีการที่ใช้

5.3 การกำหนดประเภทและรูปแบบของสเกล (Design the scale) จากการศึกษาของ Hinkin (1995) ทบทวนบทความ 75 บทความ พบว่า 31/75 ของบทความที่ทบทวน รายงานการกลับค่าคะแนนสำหรับข้อคำถามเชิงลบ และ 36/75 ของบทความที่ทบทวน รายงานการกำหนดจำนวนระดับคำตอบ 5 ตัวเลือกในประเด็นของการกำหนดรูปแบบสเกล สำหรับการศึกษานี้ของ Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 2/6 ของบทความที่ทบทวน รายงานการกลับค่าคะแนนสำหรับข้อคำถามเชิงลบ และ 5/6 ของบทความที่ทบทวน รายงานการกำหนดจำนวนระดับคำตอบ 5-7 ตัวเลือกในประเด็นของการกำหนดรูปแบบสเกล

5.4 การทดสอบสเกลเบื้องต้น (Pretesting) การศึกษาของ Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 6/6 ของบทความที่ทบทวน รายงานการทดสอบสเกลเบื้องต้น

5.5 การวิเคราะห์รายการคำถาม (Item analysis) การศึกษาของ Hinkin (1995) ทบทวนบทความ 75 บทความ พบว่า 73/75 ของบทความที่ทบทวน รายงานการตรวจสอบสอดคล้องภายในด้วยวิธีการ Cronbach's alpha และ 53/75 ของบทความที่ทบทวน รายงานการวิเคราะห์ปัจจัยโดยวิธีการ exploratory factor analysis ในประเด็นของการวิเคราะห์รายการคำถาม สำหรับการศึกษานี้ของ Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 6/6 ของบทความที่ทบทวน รายงานการตรวจสอบสอดคล้องภายในด้วยวิธีการ Cronbach's alpha และ 5/6 ของบทความที่ทบทวน รายงานการวิเคราะห์ปัจจัยโดยวิธีการ exploratory factor analysis ในประเด็นของการวิเคราะห์รายการคำถาม

5.6 การตรวจสอบตามคุณสมบัติด้านจิตเมตริกของสเกล (Psychometric property)

5.6.1 ความตรง (Validity)

การรายงานการประเมินความตรงตามเนื้อหา (content validity) จากการศึกษาของ Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 4/6 ของบทความที่ทบทวน รายงานการประเมินความตรงตามเนื้อหาด้วยวิธีการทบทวนวรรณกรรม ส่วนการศึกษาของ Coste (1995) ทบทวนบทความ 20 บทความ พบว่า 15/20 ของบทความที่ทบทวน รายงานการประเมินความตรงตามเนื้อหาด้วยวิธีทบทวนวรรณกรรมและสอบถามจากแพทย์ พยาบาลและคนไข้ และการศึกษาอีกเรื่องของ Coste (1997) ทบทวนบทความ 42 บทความ พบว่า 13/42 ของบทความที่ทบทวน รายงานการประเมินความตรงตามเนื้อหาด้วยวิธีพิจารณาเนื้อหาที่เกี่ยวข้อง 2/42 ของบทความที่ทบทวน รายงานการทดสอบความตรงตามเนื้อหาด้วยวิธีขอคำแนะนำจากผู้เชี่ยวชาญและ 1/42 ของบทความที่ทบทวน รายงานการประเมินความตรงตามเนื้อหาด้วยวิธีการสอบถามจากคนไข้

การรายงานการประเมินความตรงตามเกณฑ์สัมพันธ์ (criterion relate validity) Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 4/6 ของบทความที่ทบทวน รายงานการประเมินความตรงตามเกณฑ์สัมพันธ์โดยเปรียบเทียบสเกลกับเกณฑ์มาตรฐาน Coste (1995) ทบทวนบทความ 20 บทความ พบว่า 11/20 ของบทความที่ทบทวน รายงานการประเมินความตรงตามเกณฑ์สัมพันธ์โดยเปรียบเทียบสเกลกับเกณฑ์มาตรฐาน พบในการศึกษาประเภทไปข้างหน้า และ 9/20 ของบทความที่ทบทวน รายงานการประเมินความตรงตามเกณฑ์สัมพันธ์โดยเปรียบเทียบสเกลกับเกณฑ์มาตรฐาน พบในการศึกษาประเภทย้อนหลัง สำหรับ Coste (1997) ทบทวนบทความ 42 บทความ พบว่า 1/42 ของบทความที่ทบทวน รายงานการประเมินความตรงตามเกณฑ์สัมพันธ์โดยเปรียบเทียบสเกลกับเกณฑ์มาตรฐาน

การรายงานการประเมินความตรงตามโครงสร้าง (construct validity) Hinkin (1995) ทบทวนบทความ 75 บทความ พบว่า 7/75 ของบทความที่ทบทวน รายงานการประเมินความตรงตามโครงสร้างประเภทความตรงเชิงลู่เข้า และ 6/75 ของบทความที่ทบทวน รายงานการประเมินความตรงตามโครงสร้างประเภทความตรงเชิงจำแนก สำหรับ Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 1/6 ของบทความที่ทบทวน รายงานการประเมินความตรงตามโครงสร้างประเภทความตรงเชิงลู่เข้า และ 1/6 ของบทความที่ทบทวน รายงานการประเมินความตรงตามโครงสร้างประเภทความตรงเชิงจำแนก Coste (1995) ทบทวนบทความ 26 บทความ พบว่า 13/26 ของบทความที่ทบทวน รายงานการประเมินความตรงตามโครงสร้างประเภทความตรงลู่เข้า และความตรงเชิงจำแนก Coste (1997) ทบทวนบทความ 42 บทความ พบว่า 26/42 ของบทความที่

ทบทวน รายงานการประเมินความตรงตามโครงสร้างประเภทความตรงผู้เข้าและความตรงเชิง
จำแนก

5.6.2 การประเมินความเที่ยงของสเกล (Reliability)

ความสอดคล้องภายใน (internal consistency) Hinkin (1995) ทบทวนบทความ
75 บทความ พบว่า 73/75 ของบทความที่ทบทวน รายงานการประเมินความเที่ยงของสเกล ประเภท
ความสอดคล้องภายใน โดย Hensley (1999) ทบทวนบทความ 6 บทความ พบว่า 6/6 ของบทความ
ที่ทบทวนรายงานการประเมินความเที่ยง ประเภทความสอดคล้องภายใน ส่วน Coste (1995)
ทบทวนบทความ 20 บทความ พบว่า 5/26 ของบทความที่ทบทวน รายงานการประเมินความเที่ยง
ประเภทความสอดคล้องภายใน สำหรับ Coste (1997) ทบทวนบทความ 42 บทความ พบว่า 22/42
ของบทความที่ทบทวน รายงานการประเมินความเที่ยงประเภทความสอดคล้องภายใน

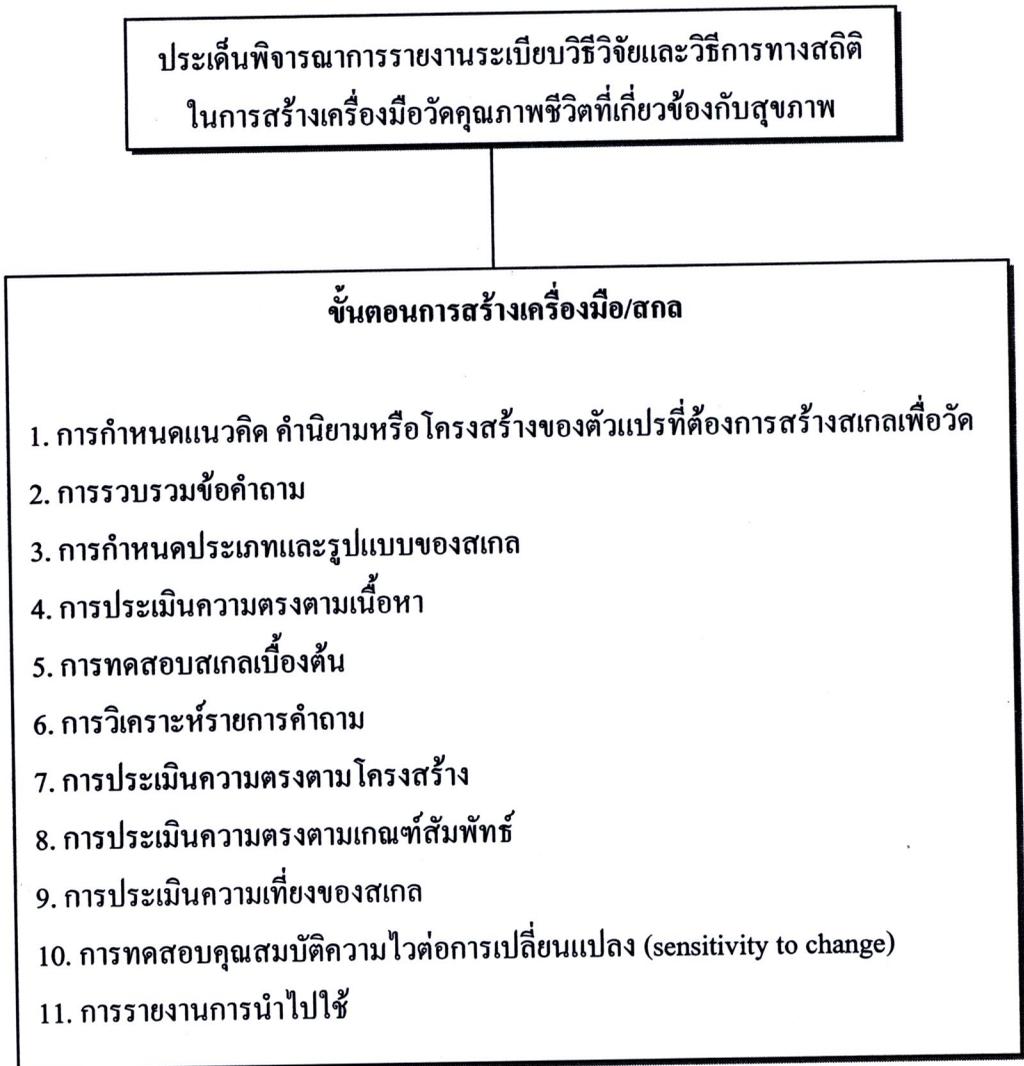
ความคงที่ของการวัด (test-retest reliability) จากการศึกษาของ Hinkin (1995)
ทบทวนบทความ 75 บทความ พบว่า 7/75 ของบทความที่ทบทวน รายงานการทดสอบความเที่ยง
ประเภทความคงที่ของการวัด ส่วนการศึกษาของ Coste (1995) ทบทวนบทความ 26 บทความ
พบว่า 6/26 ของบทความที่ทบทวน รายงานการทดสอบความเที่ยง ประเภทความคงที่ของการวัด
และ Coste (1997) ทบทวนบทความ 42 บทความ พบว่า 10/42 ของบทความที่ทบทวน รายงานการ
ทดสอบความเที่ยง ประเภทความคงที่ของการวัด

5.6.3 ความไวต่อการเปลี่ยนแปลง (Sensitivity to change) จากการศึกษาของ Coste (1995) ทบทวนบทความ 26 บทความ พบว่า 7/26 ของบทความที่ทบทวน รายงานการทดสอบความไว และการศึกษาอีกเรื่องหนึ่งของ Coste (1997) ทบทวนบทความ 42 บทความ พบว่า 1/42 ของบทความที่ทบทวน รายงานการทดสอบความไว

รายละเอียดของบทความวิจัยใน 4 รายการนี้ นำเสนอไว้ในตารางที่ 4 (ภาคผนวก ก)

6. กรอบแนวคิดในการวิจัย

ในการศึกษาระเบียบวิธีวิจัยและวิธีการทางสถิติในการสร้างเครื่องมือวัดคุณภาพชีวิตที่เกี่ยวข้องกับสุขภาพครั้งนี้ ผู้วิจัยพิจารณาระเบียบวิธีต่าง ๆ ตามขั้นตอนของการสร้างเครื่องมือประเภทสเกล ดังนำเสนอในภาพที่ 1 และในแต่ละขั้นตอนนี้จะพิจารณาถึงรายละเอียดของวิธีการต่างๆที่มีการรายงานไว้ในงานวิจัยที่ทบทวน



ภาพที่ 1 กรอบแนวคิดในการวิจัย