



การประเมินความน่าเชื่อถือของงานวิจัยรูปแบบ Randomized Controlled Trial ด้วย CASP Checklists version 2020

รหัสการศึกษาต่อเนื่อง 1010-1-000-002-06-2566

จำนวน 3.00 หน่วยกิต

วันที่รับรอง 21 มิถุนายน 2566

วันที่หมดอายุ 20 มิถุนายน 2567

โดย อ.ภก.ดุรงค์กร พลทอม

กลุ่มวิชาเภสัชกรรมปฏิบัติ คณะเภสัชศาสตร์ มหาวิทยาลัยอุบลราชธานี

*ติดต่อผู้นิพนธ์: คณะเภสัชศาสตร์ มหาวิทยาลัยอุบลราชธานี

85 ถ.สกลมารค์ ต.เมืองศรีโค อ.วารินชำราบ จ.อุบลราชธานี 34190 E-mail: durongkorn.p@ubu.ac.th

วัตถุประสงค์

เพื่อให้ทราบถึงการเปลี่ยนแปลงของแบบประเมินความน่าเชื่อถืองานวิจัยรูปแบบ randomized controlled trial ด้วย CASP checklists version 2020 ซึ่งเป็นฉบับที่ปรับปรุงขึ้นใหม่และการนำไปใช้

บทคัดย่อ

การอ่านและประเมินความน่าเชื่อถือของงานวิจัยเป็นหนึ่งในทักษะที่สำคัญสำหรับเภสัชกรเพื่อพิจารณาความน่าเชื่อถือของข้อมูลที่ได้มาใช้เป็นหลักในการตัดสินใจแก้ไขปัญหาทางคลินิก โดยเฉพาะอย่างยิ่งงานวิจัยในรูปแบบ randomized controlled trial ที่นับว่าเป็นแหล่งข้อมูลพบได้มากและมีความน่าเชื่อถือสูง ปัจจุบันได้มีแนวทางสำหรับการประเมินความน่าเชื่อถือของงานวิจัยรูปแบบ randomized controlled trial จากหลายสถาบัน หนึ่งในนั้นคือ CASP checklists ที่ถูกนำมาใช้กันอย่างแพร่หลาย และได้มีการปรับปรุงขึ้นใหม่ในปี 2020 ซึ่งมีโครงสร้างและรายละเอียดต่าง ๆ ที่แตกต่างไปจากฉบับเดิม เภสัชกรในฐานะผู้ใช้งานจึงควรทำความเข้าใจเพื่อให้สามารถใช้แบบประเมินนี้ได้อย่างถูกต้องเหมาะสมและนำไปสู่การตัดสินใจทางคลินิกที่มีแม่นยำ

คำสำคัญ: CASP checklists, randomized controlled trial, critical appraisal

บทนำ

ปัจจุบันโลกได้ก้าวเข้าสู่ยุคสมัยของการแพทย์เชิงประจักษ์ (evidence-based medicine, EBM)¹ การตัดสินใจเพื่อแก้ไขปัญหาทางคลินิกจำเป็นต้องตั้งอยู่บนพื้นฐานของงานวิจัยทางคลินิกที่มีความน่าเชื่อถือ ดังนั้นแล้ว นอกเหนือไปจากความรู้ความสามารถทางวิชาการ การอ่านและประเมินความน่าเชื่อถือของ



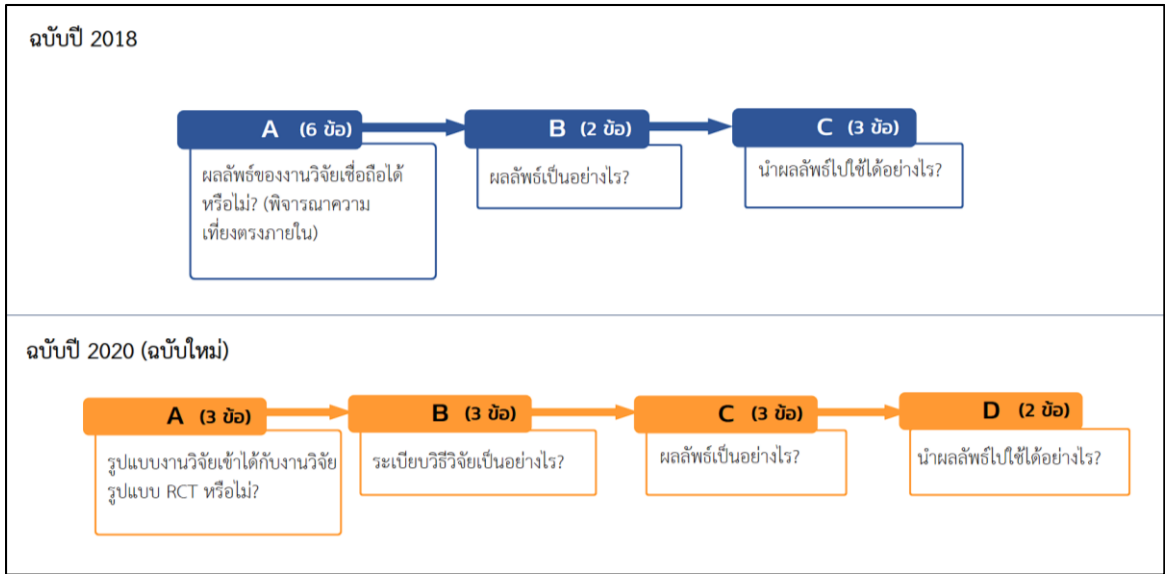
งานวิจัยเพื่อนำไปใช้ทางคลินิก (critical appraisal) จึงนับเป็นอีกหนึ่งทักษะที่สำคัญสำหรับบุคลากรทางการแพทย์รวมถึงเภสัชกร โดยเฉพาะอย่างยิ่งงานวิจัยในรูปแบบ randomized controlled trial (RCT) ที่พบได้มากในงานวิจัยการแพทย์ และนับว่าเป็นแหล่งข้อมูลที่มีความน่าเชื่อถือสูงสำหรับใช้สำหรับประกอบการตัดสินใจทางคลินิก เช่น การเลือกวิธีการรักษาที่มีประสิทธิภาพและความปลอดภัยให้แก่ผู้ป่วย เป็นต้น

ทั้งนี้ความน่าเชื่อถือของงานวิจัยเพื่อประกอบการตัดสินใจทางคลินิกนั้นสามารถพิจารณาได้จาก 2 ส่วน คือ ความเที่ยงตรงภายใน (internal validity) ซึ่งเป็นความถูกต้องเหมาะสมของระเบียบวิธีวิจัยที่ใช้ในงานวิจัยนั้น และความเที่ยงตรงภายนอก (external validity) ซึ่งเป็นความถูกต้องเหมาะสมของการนำวิธีที่ใช้ในการทดลองไปประยุกต์ใช้จริงกับผู้ป่วย ในปัจจุบันได้มีการพัฒนาแนวทางและแบบประเมินความน่าเชื่อถือของงานวิจัยจากหลากหลายสถาบัน และหนึ่งในแบบประเมินที่มีการใช้อย่างกว้างขวางคือแบบประเมินของ Critical Appraisal Skills Programme (CASP)²

CASP มีการพัฒนาแบบประเมินความน่าเชื่อถือของงานวิจัยรวมจำนวนทั้งสิ้น 8 รูปแบบ ได้แก่ randomized controlled trial, systematic review, qualitative study, cohort study, case control study, diagnostic study, economic evaluation และ clinical prediction rule สำหรับแบบประเมินความน่าเชื่อถือของงานวิจัยรูปแบบ RCT ฉบับเดิมในปี 2018 นั้น³ เป็นแบบประเมินที่พัฒนามาจากแนวทางการประเมินงานวิจัยของ JAMA 'Users' guides to the medical literature ปี ค.ศ. 1994 ซึ่งประกอบไปด้วยคำถามจำนวน 11 ข้อ แบ่งเป็น 3 ส่วน ได้แก่ Section A (ข้อที่ 1 ถึง 6) สำหรับประเมินความเที่ยงตรงภายใน Section B (ข้อที่ 7 และ 8) สำหรับประเมินผลลัพธ์ที่ได้ และ Section C (ข้อที่ 9 ถึง 11) สำหรับประเมินความเหมาะสมในการนำวิธีทดลองไปประยุกต์ใช้กับผู้ป่วย โดยในแต่ละข้อจะมีคำใบ้ (hint) ช่วยให้การคำแนะนำเพิ่มเติมว่าในแต่ละข้อคำถามนั้นมีประเด็นใดบ้างที่ต้องพิจารณาบ้าง

อย่างไรก็ตาม ในปี 2020 ได้มีการตีพิมพ์เครื่องมือประเมินงานวิจัยรูปแบบ RCT เป็นฉบับใหม่⁴ โดยในฉบับนี้ได้มีการปรับปรุงเพิ่มเติมตามแนวทางของ CONSORT 2010 guideline ตัวแบบประเมินยังคงประกอบไปด้วยคำถามจำนวนทั้งสิ้น 11 ข้อ แต่แบ่งออกใหม่เป็น 4 ส่วน ได้แก่ Section A (ข้อที่ 1 ถึง 3) สำหรับประเมินว่างานวิจัยชิ้นนั้นเข้าได้กับลักษณะงานวิจัยรูปแบบ RCT จริงหรือไม่ Section B (ข้อที่ 4 ถึง 6) สำหรับประเมินความเที่ยงตรงภายใน Sections C (ข้อที่ 7 ถึง 8) สำหรับประเมินผลลัพธ์ และ Sections D (ข้อที่ 8 ถึง 11) สำหรับประเมินความเหมาะสมของการนำวิธีทดลองไปประยุกต์ใช้จริง ทั้ง 4 ส่วนของแบบประเมินที่เพิ่มขึ้นนี้ทำให้ตัวข้อคำถามเองต้องเปลี่ยนแปลงไปด้วยเช่นกัน โดยยังคงมีหลักในการพิจารณาให้กับผู้ใช้ในแต่ละข้อเช่นเดิม รายละเอียดความแตกต่างระหว่างแบบประเมินทั้ง 2 ฉบับแสดงในแผนภาพที่ 1

คำถามแต่ละข้อสามารถเลือกตอบได้ 3 แบบคือ ใช่ ไม่ใช่ และไม่สามารถบอกได้ พร้อมอธิบายเหตุผลประกอบ โดยไม่มีการให้คะแนนเป็นตัวเลขและไม่มีการรวมคะแนนแต่อย่างใด เนื่องจากวัตถุประสงค์หลักของ CASP Checklists คือการช่วยเป็นแนวทางประกอบให้กับการตัดสินใจบนหลักฐานเชิงประจักษ์ ซึ่งต้องอาศัยองค์ความรู้ประกอบกับดุลยพินิจของผู้ประเมินในการวิเคราะห์ความน่าเชื่อถือของงานวิจัยด้วยตัวเอง



แผนภาพที่ 1 ความแตกต่างของแบบประเมิน CASP randomized controlled trial checklist

Section A: รูปแบบงานวิจัยเข้าได้กับงานวิจัยรูปแบบ RCT หรือไม่?

ในส่วนแรกนี้จะเป็นคำถามเพื่อประเมินการเป็นว่างานวิจัยชิ้นนี้มีการออกแบบที่เข้าได้กับลักษณะงานวิจัยรูปแบบ RCT หรือไม่ ประกอบไปด้วยคำถาม 3 ข้อ ได้แก่

1. Did the trial address a clearly focused issue?

Consider: Was the study designed to assess the outcomes of an intervention?

An issue can be ‘focused’ In terms of

- the population studied
- the intervention given
- the comparator chosen
- the outcomes measured

ข้อที่ 1 เป็นการถามถึงความชัดเจนของคำถามในงานวิจัย ซึ่งช่วยให้เห็นภาพรวมของวิจัยทั้งหมดได้ โดยพิจารณาจาก 4 องค์ประกอบหลักที่เรียกว่า PICO^{5,6} ซึ่งประกอบไปด้วย

- 1) Participant และ Population (P) คือลักษณะของกลุ่มตัวอย่างและประชากรที่เข้าร่วมงานวิจัย
- 2) Intervention (I) คือสิ่งทดลอง เป็นสิ่งที่สนใจศึกษาและต้องการทราบถึงผลลัพธ์ที่เกิดขึ้นจากสิ่งนั้น โดยกำหนดให้ใช้ในกลุ่มทดลอง ในงานวิจัยทางการแพทย์ intervention จะเกี่ยวข้องกับวิธีการวินิจฉัย หรือรักษาโรค เช่น ยาที่พัฒนาขึ้นใหม่ หรือแนวทางการรักษาใหม่ ๆ เป็นต้น
- 3) Comparator หรือ Control (C) คือสิ่งเปรียบเทียบ เป็นสิ่งที่ให้ในกลุ่มควบคุมเพื่อใช้เปรียบเทียบกับผลลัพธ์ที่ได้จากสิ่งทดลอง เช่น งานวิจัยที่ต้องการทราบถึงประสิทธิภาพของยา



ชนิดใหม่ ในกลุ่มควบคุมจึงอาจกำหนดให้ใช้ตามแนวทางการรักษาที่เป็นมาตรฐานซึ่งมีผลต่อการรักษาจริง (active control) หรืออาจกำหนดให้เป็นยาหลอก (placebo) ซึ่งไม่ได้มีผลในการรักษา

- 4) Outcomes (O) ผลลัพธ์ที่ได้จากสิ่งทดลอง ในทางการแพทย์จะเน้นไปที่อิทธิพลของการรักษา (treatment effect) เช่น การหายขาดจากโรค อาการเจ็บป่วยที่ดีขึ้น หรือผลทางห้องปฏิบัติการที่เกี่ยวข้องกับโรคนั้น เป็นต้น โดยวิเคราะห์เปรียบเทียบความแตกต่างของผลลัพธ์ที่ได้ระหว่างกลุ่มทดลองและกลุ่มควบคุม ทั้งระดับนัยสำคัญทางสถิติ (statistical significance) และนัยสำคัญทางคลินิก (clinical significance)

งานวิจัยที่ระบุองค์ประกอบทั้ง 4 ส่วนของ PICO ไว้อย่างชัดเจน แสดงถึงการมีคำถามของงานวิจัยที่ดี ช่วยให้ผู้อ่านสามารถรับรู้ได้ว่างานวิจัยชิ้นนั้นมีความสอดคล้องกับประเด็นปัญหาที่กำลังพบเจออยู่หรือไม่และง่ายต่อการนำไปประยุกต์ใช้เมื่อประเมินแล้วว่ามีแนวโน้มน่าเชื่อถือเพียงพอ

2. Was the assignment of participants to interventions randomized?

Consider:

- How was randomization carried out? Was the method appropriate?
- Was randomization sufficient to eliminate systematic bias?
- Was the allocation sequence concealed from investigators and participants?

ในข้อที่ 2 สิ่งที่ต้องพิจารณาคือวิธีการที่ผู้วิจัยใช้แบ่งกลุ่มตัวอย่างเข้ามาในงานวิจัยออกเป็น 2 กลุ่มคือกลุ่มทดลองและกลุ่มควบคุม ซึ่งงานวิจัยรูปแบบ RCT จำเป็นต้องใช้การสุ่ม (randomization)^{5,6} ในการแบ่งกลุ่มตัวอย่าง การสุ่มเป็นวิธีที่ทำให้กลุ่มตัวอย่างมีโอกาสได้รับสิ่งทดลองและสิ่งเปรียบเทียบเท่า ๆ กัน จึงมั่นใจได้ว่าการแบ่งของกลุ่มทดลองและกลุ่มควบคุมไม่ได้เกิดมาจากอคติ (bias) ของผู้วิจัย นอกจากนี้กระบวนการสุ่มที่ดีจะช่วยให้ทั้งกลุ่มทดลองและกลุ่มควบคุมมีคุณลักษณะต่าง ๆ ที่คล้ายคลึงกัน ณ จุดเริ่มต้นของงานวิจัย โดยปกติแล้วกลุ่มตัวอย่างแต่ละรายจะมีปัจจัยต่าง ๆ ทั้งที่รู้และไม่รู้ว่าจะส่งผลต่อผลลัพธ์ของงานวิจัยอย่างการเกิดโรคหรือผลการรักษาหรือไม่ การสุ่มที่ดีจะช่วยให้มีการกระจายตัวของคุณลักษณะเหล่านี้เท่า ๆ กันในกลุ่มทดลองและกลุ่มควบคุม ดังนั้นเมื่อการทดลองดำเนินไปจนได้ผลลัพธ์เกิดขึ้น จะทำให้มั่นใจได้มากขึ้นว่าผลลัพธ์ที่ได้เกิดจากสิ่งทดลองที่ให้ ไม่ใช่เพราะความแตกต่างของคุณลักษณะในกลุ่มตัวอย่างที่แตกต่างกัน งานวิจัยที่มีความชัดเจนจะระบุวิธีการสุ่มที่ใช้ซึ่งมีด้วยกันหลายวิธี เช่น การใช้ตารางสุ่ม การใช้รหัสสุ่มใส่ซองที่ปิดผนึก หรือการสุ่มด้วยระบบคอมพิวเตอร์ เป็นต้น และแสดงข้อมูลที่บอกถึงลักษณะของกลุ่มตัวอย่างที่ได้หลังจากสุ่มในตารางที่เรียกว่า ตารางคุณลักษณะกลุ่มตัวอย่าง (baseline characteristics) ซึ่งจะกล่าวถึงต่อไปในข้อที่ 5

การปกปิดการจัดสรร (allocation concealment) เป็นอีกประเด็นหนึ่งที่ต้องพิจารณาถึงเมื่อมีการสุ่มกลุ่มตัวอย่าง การปกปิดการจัดสรรคือการกำหนดให้ผู้ที่ทำหน้าที่จัดสรรกลุ่มตัวอย่างด้วยการสุ่มเป็นทีมงานที่ไม่เกี่ยวข้องกับการคัดเลือกกลุ่มตัวอย่างหรือการรักษาโดยตรง เช่น ในงานวิจัยที่แพทย์เป็นผู้คัดเลือกผู้ป่วย



เพื่อเข้ารับงานวิจัย อาจกำหนดให้เภสัชกรเป็นผู้ทำการสุ่ม รวมทั้งหลีกเลี่ยงการเผชิญหน้ากันตรง ๆ ระหว่างผู้วิจัยและผู้สุ่ม โดยใช้การติดต่อผ่านโทรศัพท์หรืออินเทอร์เน็ต

3. Were all participants who entered the study accounted for at its conclusion?

Consider:

- Were losses to follow-up and exclusions after randomization accounted for?
- Were participants analyzed in the study groups to which they were randomized (intention-to-treat analysis)?
- Was the study stopped early? If so, what was the reason?

ข้อที่ 3 เป็นการประเมินว่าในช่วงระหว่างการทดลองนั้น มีเหตุการณ์ที่ทำให้กลุ่มทดลองและกลุ่มควบคุมต้องขอลอกจากงานวิจัย ขาดหายจากการติดตามผล สลับหรือย้ายกลุ่มการรักษาจากกลุ่มหนึ่งไปยังอีกกลุ่มหนึ่งหรือไม่ ซึ่งเหตุการณ์เหล่านี้สามารถส่งผลถึงความน่าเชื่อถือของงานวิจัยเพราะความเท่าเทียมของคุณลักษณะภายในกลุ่มตัวอย่างที่ได้มาจากการสุ่มในขั้นตอนแรกมีการเปลี่ยนแปลง และส่งผลต่ออำนาจทางสถิติ (power) ผู้อ่านจึงจำเป็นต้องพิจารณาว่ามีการระบุถึงปัญหาเหล่านี้ไว้หรือไม่ เช่น งานวิจัยดำเนินไปตามระยะเวลาที่กำหนดจริง มีการระบุถึงสัดส่วนกลุ่มตัวอย่างที่ขาดการติดตามผลการรักษา รวมถึงการระบุถึงเหตุผลที่กลุ่มตัวอย่างนั้นขาดการติดตามไปหรือไม่ เช่น ขาดหายไปเพราะหายดีแล้วจึงไม่คิดว่าจำเป็นต้องใช้ยาอีก หรือเพราะอาการแย่ลง เกิดผลข้างเคียงในการรักษาจนทนไม่ไหวทำให้หยุดการรักษา และพิจารณาต่อว่าสัดส่วนที่ขาดหายไปนั้นสูงมากกว่าอีกกลุ่มเกินไปหรือไม่ โดยเฉพาะในกรณีที่มีผลลัพธ์ของงานวิจัยที่ได้มีจำนวนการปรากฏที่ค่อนข้างน้อย เช่น การเกิดโรคที่มีอุบัติการณ์ต่ำหรือการเสียชีวิตจากโรค การสูญเสียกลุ่มตัวอย่างไปเพียงย่อมส่งผลให้เกิดอคติในการแปลผลลัพธ์ของงานวิจัย กระบวนการที่ผู้วิจัยใช้ในการวิเคราะห์ผลลัพธ์จึงมีความสำคัญ หากผู้วิจัยเลือกวิเคราะห์แบบ per-protocol analysis โดยการตัดข้อมูลจากกลุ่มตัวอย่างที่ขาดการติดตามออก คัดเอาเฉพาะข้อมูลที่ได้จากกลุ่มตัวอย่างที่อยู่ครบตามกระบวนการงานวิจัย วิธีการนี้จะทำให้คุณลักษณะภายในกลุ่มตัวอย่างที่ตั้งใจให้เท่าเทียมกันด้วยการสุ่มเมื่อเริ่มต้นสูญเสียไป ในขณะที่การวิเคราะห์แบบ intention-to-treat analysis ซึ่งเป็นการวิเคราะห์ผลลัพธ์ของงานวิจัยตามกลุ่มที่ได้รับการสุ่มไว้ตั้งแต่แรกเริ่มวิจัย จึงช่วยรักษาอิทธิพลของการสุ่ม คงความเท่าเทียมของคุณลักษณะภายในกลุ่มตัวอย่างไว้ได้ intention-to-treat analysis จึงเป็นวิธีการวิเคราะห์ที่เหมาะสม นอกจากนี้หากงานวิจัยจำเป็นต้องหยุดก่อนกำหนดก็จำเป็นต้องมีการระบุสาเหตุไว้เพราะการหยุดการทดลองอาจเป็นสัญญาณของความปลอดภัยที่เกิดขึ้นกับผู้เข้าร่วมงานวิจัย

Section A ทั้งสามข้อแรกนั้นถูกออกแบบให้เป็นคำถามคัดกรอง (screening questions) ที่สามารถอ่านและหาคำตอบได้อย่างรวดเร็วจากชื่อเรื่อง (title) บทคัดย่อ (abstract) และระเบียบวิธีวิจัย (method) ว่างานวิจัยที่อ่านนั้นเข้าได้หรือไม่กับลักษณะสำคัญของงานวิจัยรูปแบบ RCT อันได้แก่ มีกลุ่มทดลองและกลุ่มควบคุมและมีการสุ่ม มีการติดตามและวิเคราะห์ผลลัพธ์ที่เหมาะสม และหากสามารถระบุคำถามของงานวิจัยได้ตรงกับประเด็นปัญหาที่ต้องการตัดสินใจ จึงค่อยประเมินต่อด้วยคำถามชุดต่อไป



Section B: ระเบียบวิธีวิจัยเป็นอย่างไร?

ในส่วนนี้เป็นการประเมินถึงประเมินความเที่ยงตรงภายในของงานวิจัย โดยพิจารณาจากความถูกต้องเหมาะสมของขั้นตอนตั้งแต่ระเบียบวิธีวิจัยไปจนถึงการได้มาซึ่งผลลัพธ์ ในส่วนนี้ประกอบไปด้วยคำถามจำนวนทั้งสิ้น 3 ข้อ ได้แก่

4. Were the participants ‘blind’ to intervention they were given?

Were the investigators ‘blind’ to the intervention they were giving to participants?

Were the people assessing/analyzing outcome/s ‘blinded’?

ข้อที่ 4 นี้จะสังเกตเห็นว่าประกอบไปด้วย 3 คำถามที่เกี่ยวข้องกับการปกปิด (blinding)^{5,6} ซึ่งเป็นกระบวนการป้องกันไม่ให้มีการระบุตัวตนของกลุ่มตัวอย่างได้ว่าอยู่ในกลุ่มทดลองหรือกลุ่มควบคุม เพื่อลดอคติของการเก็บข้อมูล เพราะข้อมูลบางชนิดสามารถเปลี่ยนแปลงได้ขึ้นกับตามรับรู้ เช่น การประเมินผลการรักษา การวิเคราะห์หาสาเหตุของการเสียชีวิตซึ่งต้องอาศัยดุลยพินิจของผู้เก็บข้อมูล หรือความรู้สึกเจ็บปวดซึ่งเป็นสิ่งที่ผู้ป่วยรายงานด้วยความรู้สึกของตนเอง กระบวนการปกปิดนี้สามารถทำได้ทั้งจากฝั่งของกลุ่มตัวอย่างและฝั่งของผู้เก็บข้อมูล

การปกปิดข้อมูลในฝั่งของกลุ่มตัวอย่าง มีขึ้นเพื่อป้องกันอคติที่อาจเกิดขึ้นจากการรับรู้ถึงสิ่งทดลองที่ตนเองหรือผู้อื่นได้รับ เช่น ในงานวิจัยทางการแพทย์ ถ้าผู้ป่วยรู้ว่าตนเองอยู่ในกลุ่มทดลองและได้รับยาตัวใหม่ ซึ่งมีราคาแพงและคาดว่าจะมีประสิทธิผลดีกว่ายาเดิม ความรู้สึกเชื่อมั่นในยาใหม่ที่ได้รับอาจทำให้ผู้ป่วยรายงานนั้นรายงานผลลัพธ์ประเภทที่เกี่ยวข้องกับความรู้สึก เช่น ความเจ็บปวดหรือความพึงพอใจในทางที่ดีกว่าที่ควรจะเป็น นอกจากนี้ยังอาจส่งผลไปถึงพฤติกรรมอื่น เช่น ความสม่ำเสมอในการใช้ยา การเลือกทานอาหาร การออกกำลังกายที่มากขึ้นหรือน้อยลงกว่าปกติ การเปลี่ยนแปลงเหล่านี้ย่อมส่งผลกระทบต่อการรักษา ทั้งหมดจึงเป็นเหตุผลที่ควรมีการปกปิดข้อมูลจากฝ่ายกลุ่มตัวอย่าง

การปกปิดข้อมูลในฝั่งของผู้วิจัยหรือผู้เก็บข้อมูลจะช่วยป้องกันอคติจากการเลือกเก็บข้อมูลได้ เช่น หากผู้เก็บข้อมูลเชื่อว่ายาที่ใช้ทดลองมีประสิทธิผลดีกว่ายาเก่าและรู้ว่าผู้ป่วยคนใดอยู่ในกลุ่มทดลอง เมื่อถึงเวลาเก็บข้อมูล ผู้เก็บข้อมูลอาจเลือกเก็บข้อมูลที่แสดงถึงประสิทธิผลที่ดีจากกลุ่มทดลองมากกว่าอีกกลุ่มได้ การปกปิดไม่ให้ผู้เก็บข้อมูลระบุได้ว่าผู้ป่วยแต่ละคนนั้นมาจากกลุ่มใดย่อมลดปัญหาในส่วนนี้

งานวิจัยที่มีการปกปิดข้อมูลทั้งในฝ่ายของกลุ่มตัวอย่างและผู้เก็บข้อมูล เรียกว่าการ double blind ซึ่งจะช่วยลดอคติจากทั้งสองฝ่ายได้ตามที่ได้อธิบายในข้างต้น นอกจากนี้ ถ้ามีการปกปิดในฝ่ายผู้วิเคราะห์ข้อมูลด้วยจะยิ่งทำให้ผลงานวิจัยน่าเชื่อถือขึ้น อย่างไรก็ตามงานวิจัยบางประเภทก็ไม่สามารถปกปิดข้อมูลได้ทั้งหมด เช่น งานวิจัยที่ศึกษาผลของการออกกำลังกายต่อการไม่ออกกำลังกาย กลุ่มตัวอย่างย่อมรู้ได้ด้วยตัวเองว่าอยู่ในกลุ่มออกกำลังกายหรือไม่ออกกำลังกาย หรืองานวิจัยที่ทำในโรคหรือยาที่มีความเสี่ยงต่อชีวิตสูง แพทย์จำเป็นต้องรู้ว่าผู้ป่วยได้รับยาใดอยู่เพื่อที่จะได้ให้การรักษาได้ทันเวลาที่หากเกิดอันตรายที่รุนแรง งานวิจัยในลักษณะนี้อาจปกปิดได้เฉพาะฝ่ายใดฝ่ายหนึ่ง เรียกว่า single blind ผู้อ่านจึงต้องวิเคราะห์ว่า ในงานวิจัยที่อ่านมีการปกปิดหรือไม่ ปกปิดด้วยวิธีใด และทำในฝ่ายใดบ้าง หากมีฝ่ายที่ไม่ถูกปกปิด การรับรู้ของคนในฝ่าย



นั้นได้ส่งผลอย่างไรต่อผลลัพธ์ของงานวิจัย ทั้งนี้ รายละเอียดของกระบวนการปกปิดข้อมูลสามารถหาได้จาก ส่วนของระเบียบวิธีวิจัย

5. Were the study groups similar at the start of the randomized controlled trial?

Consider:

- Were the baseline characteristics of each study group (e.g. age, sex, socio-economic group) clearly set out?
- Were there any differences between the study groups that could affect the outcome/s?

แม้ว่าการสุ่มจะเป็นวิธีการที่ช่วยลดอคติจากความแตกต่างของปัจจัยที่แตกต่างกันระหว่างกลุ่มทดลอง และกลุ่มควบคุม แต่ยังคงมีความเป็นไปได้ที่สุดท้ายแล้วคุณลักษณะของกลุ่มตัวอย่างจะมีความแตกต่างกันโดยบังเอิญ ซึ่งสามารถพิจารณาได้จากตารางคุณลักษณะกลุ่มตัวอย่าง (baseline characteristics) ที่แสดงข้อมูลเป็นจำนวนหรือสัดส่วนของคุณลักษณะต่าง ๆ ทั้งในกลุ่มตัวอย่างและกลุ่มควบคุมตามผู้วิจัยได้เก็บรวบรวมไว้ รวมทั้งแสดงผลการวิเคราะห์ความแตกต่างทางสถิติของคุณลักษณะระหว่างกลุ่มตัวอย่างเป็นค่า p-value อย่างไรก็ตามในปัจจุบันหลายงานวิจัยจะไม่แสดงค่า p-value ในตารางคุณลักษณะกลุ่มตัวอย่างให้เห็น ยกเว้นบางข้อมูลคุณลักษณะที่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างกลุ่มจะมีการแสดงไว้

สิ่งที่ผู้อ่านจะต้องทำการวิเคราะห์จากตารางคุณลักษณะกลุ่มตัวอย่างคือการค้นหาว่าปัจจัยใดที่มีจำนวนหรือสัดส่วนแตกต่างกันระหว่างสองกลุ่มอย่างมีนัยสำคัญทางสถิติ หากคุณลักษณะของกลุ่มตัวอย่างในงานวิจัยไม่แตกต่างกันจะสามารถบอกได้ว่างานวิจัยชิ้นนี้มีกระบวนการสุ่มที่ดี แต่สิ่งที่ต้องวิเคราะห์ต่อคือ ยังคงมีปัจจัยอื่นที่สำคัญต่อผลลัพธ์งานวิจัยที่ผู้วิจัยไม่ได้เก็บหรือแสดงไว้ในตารางคุณลักษณะกลุ่มตัวอย่างหรือไม่ เช่น งานวิจัยเกี่ยวกับความเสี่ยงในการเกิดโรคหลอดเลือดสมองอุดตัน (stroke) ในงานวิจัยนั้นควรแสดงข้อมูลอายุ เพศ เชื้อชาติ โรคความดันโลหิตสูง โรคไขมันในเลือดสูง โรคเบาหวาน โรคหัวใจและหลอดเลือด ประวัติโรคหลอดเลือดสมองและหัวใจในครอบครัว การสูบบุหรี่ และการไม่ออกกำลังกาย ซึ่งเป็นปัจจัยเสี่ยงที่สำคัญในโรคนี้⁷ เพราะปัจจัยที่ไม่ได้เก็บรวบรวมหรือแสดงไว้ในตารางอาจเป็นปัจจัยที่มีความแตกต่างและอาจส่งผลต่อผลลัพธ์ของงานวิจัย ส่วนในกรณีที่พบว่าไม่มีปัจจัยที่มีความแตกต่างระหว่างกลุ่ม ผู้อ่านจำเป็นต้องพิจารณาต่อด้วยว่าปัจจัยดังกล่าวนั้นสามารถส่งผลกระทบต่อผลลัพธ์ของงานวิจัยหรือไม่ และผู้วิจัยได้จัดการกับปัจจัยนั้นด้วยวิธีการใด

6. Apart from the experimental intervention, did each study group receive the same level of care (that is, were they treated equally)?

Consider:

- Was there a clearly defined study protocol?
- If any additional interventions were given (e.g. tests or treatments), were they similar between the study groups?



- Were the follow-up intervals the same for each study group?

ข้อที่ 6 เป็นสิ่งที่ต้องพิจารณาในบางกรณี เช่น การศึกษาเกี่ยวกับผู้ป่วยโรคเรื้อรังที่ต้องมีการใช้ยาหลายชนิดในการรักษา ผู้วิจัยจะต้องมีการควบคุมให้การรักษานั้น ๆ นอกเหนือจากยาที่เป็นสิ่งทดลองและสิ่งเปรียบเทียบให้ได้รับเหมือนกันด้วย รวมไปถึงกระบวนการอื่น ๆ ที่มีต่อกลุ่มตัวอย่างทั้งสองกลุ่มต้องเป็นไปด้วยมาตรฐานเดียวกัน เช่น วิธีการวัดผล ระยะเวลาในการนัดติดตามผล เป็นต้น รายละเอียดต่าง ๆ เหล่านี้ควรมีการระบุไว้ในระเบียบวิธีวิจัยไว้ให้ชัดเจน

Section C: ผลลัพธ์เป็นอย่างไร?

ในส่วนนี้จะเป็นการประเมินถึงผลลัพธ์ที่แสดงในงานวิจัย ประกอบไปด้วยคำถาม 3 ข้อ ได้แก่

7. Were the effects of intervention reported comprehensively?

Consider:

- Was a power calculation undertaken?
- What outcomes were measured, and were they clearly specified?
- How were the results expressed? For binary outcomes, were relative and absolute effects reported?
- Were the results reported for each outcome in each study group at each follow-up interval?
- Was there any missing or incomplete data?
- Was there differential drop-out between the study groups that could affect the results?
- Were potential sources of bias identified?
- Which statistical tests were used?
- Were p values reported?

ข้อที่ 7 เป็นการประเมินว่าผลลัพธ์ของงานวิจัยมีการรายงานไว้อย่างครอบคลุมครบถ้วน โดยเริ่มจากการประเมินผลลัพธ์หลัก^{8,9} (primary outcome) เนื่องจากเป็นตัวแปรสำคัญในการคำนวณขนาดตัวอย่าง (sample size) และอำนาจทางสถิติ ทั้งนี้ผลลัพธ์หลักอาจมีทั้งประเภทที่วัดผลเพียงสิ่งเดียว หรืออาจเป็นประเภทที่ประกอบไปด้วยผลลัพธ์ย่อย ๆ หลายชนิดรวมกันได้ (composite outcome) เช่น งานวิจัยที่เกี่ยวข้องกับโรคหลอดเลือดหัวใจจะมีการกำหนดผลลัพธ์หลักที่เรียกว่า major adverse cardiovascular events (MACE)⁹ ประกอบด้วยสามผลลัพธ์ย่อย ได้แก่ โรคกล้ามเนื้อหัวใจขาดเลือด โรคหลอดเลือดสมอง และการเสียชีวิตจากโรคหลอดเลือดหัวใจ สิ่งที่ต้องระวังเกี่ยวกับ composite outcome คือผลลัพธ์แต่ละอย่างที่ว่านั้นต้องสัมพันธ์และจัดกลุ่มเข้ากันได้จริง



สิ่งที่สำคัญคือผลลัพธ์หลักต้องสอดคล้องและตอบวัตถุประสงค์ของงานวิจัย มีเกณฑ์ที่ใช้ในการวัดผลที่ชัดเจน เป็นรูปธรรมเพื่อลดอคติในการแปลผล ขนาดกลุ่มตัวอย่างและระยะเวลาในการติดตามผลที่เพียงพอ มีการรายงานระดับความร่วมมือในการรักษา นอกจากนี้ผลลัพธ์ที่ดีควรมีความสำคัญในทางคลินิกกับผู้ป่วย (patient-oriented outcome) เพราะบางกรณีผลลัพธ์ตัวแทนที่วัดค่าได้ (surrogate outcome) เช่น ระดับของสารเคมีในเลือด อาจไม่สัมพันธ์กับอาการทางคลินิก

ในงานวิจัยยังมีผลลัพธ์รอง (secondary outcome) เพื่อความครบถ้วนเกี่ยวกับสิ่งที่เกิดขึ้นจากสิ่งทดลอง โดยเฉพาะในงานวิจัยทางการแพทย์ที่ต้องมีการประเมินทั้งประสิทธิภาพและความปลอดภัยในการรักษา ผลลัพธ์รองจึงมักเป็นข้อมูลเปรียบเทียบผลทางห้องปฏิบัติการ หรือการเกิดอาการข้างเคียงจากการใช้ยา

ผลลัพธ์งานวิจัยสามารถนำเสนอได้ในหลายรูปแบบ¹¹ ทั้งการพรรณนาเป็นจำนวน สัดส่วน ร้อยละ หรือค่าเฉลี่ย การแสดงเป็นผลลัพธ์ที่สะท้อนอิทธิพลของการรักษา เช่น relative risk (RR), relative risk reduction (RRR), absolute risk reduction (ARR), และ number needed to treat (NNT) เป็นต้น (รายละเอียดแสดงในตารางที่ 1) และการเปรียบเทียบผลลัพธ์ระหว่างกลุ่มตัวอย่างสองกลุ่มโดยกำหนดระดับนัยสำคัญทางสถิติ ด้วยค่า p-value หากค่า p-value ที่ได้มีค่าน้อยกว่า 0.05 จะถือว่าผลลัพธ์ที่ได้มีนัยสำคัญทางสถิติ ช่วยให้คาดการณ์ได้ว่าความแตกต่างระหว่างกลุ่มที่เกิดขึ้นในงานวิจัยไม่น่าจะเกิดขึ้นจากความบังเอิญ แต่สิ่งที่ต้องคำนึงถึงเสมอคือนัยสำคัญทางสถิติอาจเกิดขึ้นได้จากตัวแปรกวน (confounding) เพราะขาดการควบคุมที่ดี หรือการมีอคติในงานวิจัย เรียกปัญหาที่เกิดขึ้นนี้ว่า type 1 error หรือ false positive นอกจากนี้ค่า p-value ไม่ได้บ่งบอกถึงอิทธิพลการรักษาที่จะเกิดขึ้นกับผู้ป่วย จึงไม่สามารถตัดสินได้ว่ายาที่มีค่า p-value น้อยกว่าจะดีกว่ายาอีกชนิดเมื่อนำมาใช้ เพราะการที่ p-value สูงกว่าหรือเท่ากับ 0.05 อาจเพราะขาดอำนาจทางสถิติเนื่องจากกลุ่มตัวอย่างมีขนาดเล็ก ระยะเวลาที่ติดตามไม่นานพอ หรืออัตราการเกิดผลลัพธ์ในงานวิจัยต่ำกว่าที่คาดการณ์ไว้ รวมถึงกรณีของผลลัพธ์รองซึ่งไม่ได้คำนวณขนาดกลุ่มตัวอย่างเพื่อเก็บมาตั้งแต่ต้น การไม่พบระดับนัยสำคัญทางสถิติจึงเป็นสิ่งที่เกิดขึ้นได้ การไม่พบระดับนัยสำคัญทางสถิติจากปัญหาอำนาจทางสถิติไม่เพียงพอ เรียกว่า type 2 error หรือ false negative

นัยสำคัญทางคลินิกจึงเป็นสิ่งที่ผู้อ่านต้องประเมินควบคู่ไปกับนัยสำคัญทางสถิติ ซึ่งดูได้จากผลลัพธ์ว่า อัตราการเกิดผลลัพธ์จริงในงานวิจัย (true event rate) เป็นไปตามที่คาดการณ์ไว้ตอนคำนวณกลุ่มตัวอย่างหรือไม่ (expected event rate) การอ่านผลค่า RR ที่บอกถึงโอกาสหรือความเสี่ยงที่เกิดขึ้นในกลุ่มทดลองต่อกลุ่มควบคุม ค่า ARR ที่บอกถึงความแตกต่างสมบูรณ์ของอัตราการเกิดผลระหว่างสองกลุ่ม ค่า RRR ที่บอกถึงการลดลงของอัตราการเกิดผลลัพธ์ในกลุ่มทดลองต่อกลุ่มควบคุม ค่า NNT ที่บอกจำนวนของผู้ป่วยที่ต้องได้รับการรักษาตามวิธีการที่ใช้ในกลุ่มทดลองเพื่อให้เห็นผลการรักษาในผู้ป่วยหนึ่งคนระดับนัยสำคัญทางคลินิกจะเป็นสิ่งที่ช่วยบอกให้ทราบถึงโอกาสที่จะเกิดขึ้นเมื่อนำวิธีการรักษาในงานวิจัยไปใช้กับผู้ป่วยจริง

8. Was the precision of the estimate of the intervention or treatment effect reported?

Consider



- Were confidence intervals (CIs) reported?

ข้อที่ 8 คือการประเมินความแม่นยำของงานวิจัย สามารถพิจารณาได้จากช่วงความเชื่อมั่น (confidence interval, CI)^{8,9,11} ซึ่งเป็นช่วงตัวเลขค่าหนึ่งที่มีจะมีการรายงานร่วมกับตัวเลขผลลัพธ์หลักด้วย เช่น OR=0.89 (95%CI, 0.89-0.98) ช่วงความเชื่อมั่นนั้นเป็นการประมาณค่าแบบช่วงเพื่ออนุมานผลลัพธ์ที่เกิดขึ้นในกลุ่มตัวอย่างไปยังกลุ่มประชากร โดยทั่วไปมักรายงานช่วงความเชื่อมั่นไว้ที่ระดับร้อยละ 95 (95%CI) คือ มั่นใจได้ 95% ว่าช่วงของตัวเลขที่ได้นี้จะครอบคลุมค่าคงที่จริงในประชากร โดยผู้อ่านสามารถประเมินความแม่นยำของผลงานวิจัยได้จากช่วงความเชื่อมั่นที่แคบ เปรียบเสมือนการปลูกดอกที่แม่นยำจะทำให้ลูกดอกไปปักอยู่ใกล้ ๆ กันเสมอ อย่างไรก็ตามความกว้างแคบของช่วงความเชื่อมั่นไม่มีเกณฑ์กำหนดตัดสินที่ชัดเจน นอกจากนี้ยังสัมพันธ์กับขนาดของกลุ่มตัวอย่างด้วย ถ้ากลุ่มตัวอย่างมีขนาดใหญ่ มักจะส่งผลให้ได้ช่วงความเชื่อมั่นที่แคบลง เพราะกลุ่มตัวอย่างที่เยอะขึ้นทำให้มีโอกาสเห็นผลลัพธ์ที่แท้จริงสูงขึ้น ความแม่นยำจึงเพิ่มขึ้น

9. Do the benefits of the experimental intervention outweigh the harms and costs?

Consider:

- What was the size of the intervention or treatment effect?
- Were harms or unintended effects reported for each study group?
- Was a cost-effectiveness analysis undertaken? (Cost-effectiveness analysis allows a comparison to be made between different interventions used in the care of the same condition or problem.)

สำหรับข้อที่ 9 นั้น นอกจากความกว้างหรือแคบของช่วงความเชื่อมั่น การอ่านช่วงความเชื่อมั่นยังต้องพิจารณาในมุมของอิทธิพลทางคลินิกที่จะเกิดขึ้นด้วย ตัวอย่างเช่น ในงานวิจัยลดไขมันในเลือดตัวใหม่พบว่า มีประสิทธิภาพในการลดระดับไขมันได้ดีกว่ายาเก่า 2.25% (95%CI, 1.55-3.05) แปลว่ายาชนิดใหม่นี้จะสามารถลดระดับไขมันในเลือดได้ดีกว่ายาเก่าประมาณ 1.55 ถึง 3.05% ที่ระดับความเชื่อมั่น 95% สิ่งที่ต้องพิจารณาตามต่อก็คือ การที่ยาตัวใหม่สามารถลดระดับไขมันลงได้ดีกว่ายาเก่าประมาณ 2-3% จะส่งผลต่อระดับไขมันที่ลดลงของผู้ป่วยอย่างน้อยเพียงใด ส่งผลต่ออาการทางคลินิกของผู้ป่วยอย่างไร ประโยชน์ทางคลินิกเมื่อเทียบกับความเสี่ยงของผลข้างเคียงที่อาจเกิดขึ้นได้นั้นคุ้มค่าหรือไม่ และคุ้มค่าต่อค่าใช้จ่ายที่ต้องเพิ่มขึ้นจากการเปลี่ยนไปใช้ยาใหม่ที่แพงกว่าหรือไม่ เป็นต้น



ตารางที่ 1 ผลลัพธ์ที่แสดงอิทธิพลการรักษา^{1,1}

| ประเภทของผลลัพธ์ | ความหมาย | ตัวอย่างการแปลผล |
|---|---|--|
| Relative Risk (RR) | <p>กลุ่มทดลองมีความเสี่ยงในการเกิดผลลัพธ์เป็นเท่าเท่าเมื่อเทียบกับกลุ่มควบคุม</p> <ul style="list-style-type: none"> RR = 1 ทั้งสองกลุ่มมีความเสี่ยงไม่แตกต่างกัน RR < 1 กลุ่มทดลองมีความเสี่ยงในการเกิดผลลัพธ์น้อยกว่ากลุ่มควบคุม RR > 1 กลุ่มทดลองมีความเสี่ยงในการเกิดผลลัพธ์มากกว่ากลุ่มควบคุม | <p>RR = 0.32 กลุ่มทดลองมีความเสี่ยงในการเกิดผลลัพธ์เป็น 0.32 เท่าของกลุ่มควบคุม หรือการให้สิ่งทดลองช่วยลดความเสี่ยงการเกิดผลลัพธ์ 68% เมื่อเทียบกับสิ่งควบคุม (RRR)</p> <p>RR = 0.67 กลุ่มทดลองมีความเสี่ยงในการเกิดผลลัพธ์เป็น 0.67 เท่าของกลุ่มควบคุม หรือการให้สิ่งทดลองช่วยลดความเสี่ยงการเกิดผลลัพธ์ได้ 33% เมื่อเทียบกับสิ่งควบคุม (RRR)</p> |
| Relative Risk Reduction (RRR) Relative Risk Increase (RRI) | <p>การรายงานผลโดยปรับจากค่า RR ให้อยู่ในเปอร์เซ็นต์</p> <p>ใช้บอกว่าสิ่งทดลองสามารถลดหรือเพิ่มผลลัพธ์ได้กี่เปอร์เซ็นต์เมื่อเทียบกับสิ่งควบคุม (RRR = 1 - RR)</p> | <p>RR = 1.50 กลุ่มทดลองมีความเสี่ยงในการเกิดผลลัพธ์เป็น 1.5 เท่าของกลุ่มควบคุม หรือการให้สิ่งทดลองเพิ่มความเสี่ยงการเกิดผลลัพธ์ 55% เมื่อเทียบกับสิ่งควบคุม (RRI)</p> |
| Absolute Risk Reduction (ARR) Absolute Risk Increase (ARI) | <p>บอกค่าความแตกต่างสัมบูรณ์ของอัตราการเกิดผลลัพธ์ในกลุ่มควบคุมต่อกลุ่มทดลอง</p> <ul style="list-style-type: none"> ARR = 0 ้ไม่มีความแตกต่างกันระหว่างสองกลุ่ม ARR < 0 กลุ่มทดลองเกิดผลลัพธ์มากกว่ากลุ่มควบคุม (เรียก ARI) ARR > 0 กลุ่มทดลองการเกิดผลลัพธ์น้อยกว่ากลุ่มควบคุม | <p>ARR = 0.04 กลุ่มทดลองเกิดผลลัพธ์น้อยกว่ากลุ่มควบคุม 4%</p> <p>ARR = 0.04 NNT = 1 / 0.04 = 25 ต้องให้การรักษาด้วยสิ่งทดลองแก่ผู้ป่วยจำนวน 25 คนจึงจะเห็นผลการรักษาในผู้ป่วยหนึ่งคน</p> |
| Number Needed to Treat (NNT) | <p>จำนวนผู้ป่วยที่ต้องได้รับการรักษาด้วยสิ่งทดลองตามวิธีในงานวิจัย เพื่อให้เห็นผลการป้องกันรักษาผลลัพธ์ในผู้ป่วยหนึ่งคน (NNT = 1/ ARR)</p> | <p>ARR = 0.04 NNT = 1 / 0.04 = 25 ต้องให้การรักษาด้วยสิ่งทดลองแก่ผู้ป่วยจำนวน 25 คนจึงจะเห็นผลการรักษาในผู้ป่วยหนึ่งคน</p> |



Section D: นำผลลัพธ์ไปใช้ได้อย่างไร?

ส่วนสุดท้ายเป็นการประเมินความเที่ยงตรงภายนอก (external validity) คือการประเมินความเหมาะสมของนำวิธีในการทดลองก่อนที่จะนำไปประยุกต์ใช้จริง

10. Can the results be applied to your local population/in your context?

Consider:

- Are the study participants similar to the people in your care?
- Would any differences between your population and the study participants alter the outcomes reported in the study?
- Are the outcomes important to your population?
- Are there any outcomes you would have wanted information on that have not been studied or reported?
- Are there any limitations of the study that would affect your decision?

ข้อที่ 10 คือการประเมินว่าผู้ป่วยที่เราต้องการรักษาหรือแก้ปัญหา มีคุณลักษณะที่เหมือนหรือแตกต่างไปจากผู้ป่วยที่คัดเข้ามาในงานวิจัยอย่างไร ถ้าคุณลักษณะมีความแตกต่างกันมาก เราจะไม่สามารถบอกได้ว่าเมื่อให้การรักษาตามแนวทางที่ใช้ในงานวิจัยต่อผู้ป่วยของเราไปแล้วจะได้ผลลัพธ์ในการรักษาไปในแนวเดียวกัน ผู้ป่วยที่เหมาะสมจึงต้องผู้ป่วยมีคุณลักษณะตรงกับเกณฑ์คัดเข้า (inclusion criteria) และไม่มีข้อห้ามใดตามที่กำหนดไว้ในเกณฑ์คัดออก (exclusion criteria) และมีคุณลักษณะโดยรวมที่ใกล้เคียงกับข้อมูลในตารางคุณลักษณะกลุ่มตัวอย่าง

นอกจากนี้อาจยังต้องพิจารณาว่ายังมีข้อมูลทางคลินิกใดอื่นอีกหรือไม่ที่ไม่ปรากฏในงานวิจัย แต่อาจเป็นประเด็นปัญหาในผู้ป่วยรายนี้ได้ เช่น การที่ผู้ป่วยมีโรคร่วมบางอย่างที่ไม่ได้มีการระบุไว้ในงานวิจัย อาจกลายเป็นข้อจำกัด (limitation) ต่อการตัดสินใจทางคลินิกในผู้ป่วยรายนี้ได้ ในกรณีดังกล่าวผู้อ่านย่อมจำเป็นที่จะต้องค้นหาข้อมูลจากแหล่งอ้างอิงที่น่าเชื่อถืออื่น ๆ เพิ่มเติมเพื่อประเมินถึงความปลอดภัย หรือในงานวิจัยต้องมีการตรวจติดตามผลทางห้องปฏิบัติการบางอย่างซึ่งในสถานการณ์นั้นไม่สามารถทำได้ ผู้อ่านจำเป็นจะต้องตัดสินใจว่าจะรับมือกับความเสี่ยงนี้ได้อย่างไร

11. Would the experimental intervention provide greater value to the people in your care than any of the existing interventions?

Consider:

- What resources are needed to introduce this intervention taking into account time, finances, and skills development or training needs?
- Are you able to disinvest resources in one or more existing interventions in order to be able to re-invest in the new intervention?



ในข้อสุดท้ายคือการประเมินว่าในท้ายที่สุดแล้วเมื่อมั่นใจว่างานวิจัยที่กำลังประเมินนั้นตอบโจทย์และมีความน่าเชื่อถือ ผู้ประเมินจะสามารถนำเอาสิ่งทดลองหรือวิธีการรักษาที่ปรากฏในงานวิจัยมาประยุกต์ใช้ได้จริงบนพื้นฐานของทรัพยากรที่มีและที่ต้องนำมาเพิ่ม เช่น เวลา เงิน เครื่องจักร หรือบุคลากร เป็นต้น

หลังจากประเมินครบทั้ง 11 ข้อ ให้ผู้ประเมินทำการสรุปประเด็นหลักที่ได้จากงานวิจัยอีกครั้งโดยรวมว่างานวิจัยนั้นตอบปัญหาทางคลินิกที่พบ มีความน่าเชื่อถือเพียงพอที่จะนำมาใช้เป็นหลักฐานในการตัดสินใจทางคลินิก หรือนำไปสู่เปลี่ยนแนวทางการปฏิบัติในระดับองค์กรต่อไป

บทสรุป

CASP Randomized Controlled Trial checklists version 2020 เป็นแนวทางในการประเมินความน่าเชื่อถือของงานวิจัยรูปแบบ Randomized Controlled Trial ที่ได้รับการปรับปรุงให้มีคำถามที่ชัดเจนขึ้น และเหมาะสมต่อการนำไปใช้ยิ่งขึ้น อย่างไรก็ตาม การประเมินความน่าเชื่อถือของงานวิจัยนั้นเป็นทักษะที่ต้องอาศัยทั้งความรู้ความเข้าใจเกี่ยวกับสถิติ ระเบียบวิธีวิจัย และต้องอาศัยการฝึกฝนเพื่อให้การตัดสินใจทางคลินิกมีความเหมาะสมกับปัญหาที่เผชิญและนำไปสู่การรักษาและแก้ไขปัญหของผู้ป่วยอย่างมีประสิทธิภาพและความปลอดภัย

เอกสารอ้างอิง

1. Thoma A. A Brief History of Evidence-Based Medicine (EBM) and the Contributions of Dr David Sackett [Internet]. 2015. Cited 2019 Feb 2. Available from: <https://academic.oup.com/asj/article/35/8/NP261/251339>
2. Critical Appraisal Skills Programme. CASP Bibliography [Internet]. 2018. Cited: 2019 Feb 2. Available from: <https://casp-uk.net/bibliography/>
3. Critical Appraisal Skills Programme. CASP Randomised Controlled Trial Checklist 2018 [Internet]. 2018. Cited: 2022 Mar 3. Available from: <https://casp-uk.net/casp-checklist-archive/>
4. Critical Appraisal Skills Programme. CASP Randomised Controlled Trial Checklist (Newly Updated) [Internet]. 2020. Cited: 2022 Mar 3. Available from: <https://casp-uk.net/casp-tools-checklists/>
5. Al-Jundi A, Sakka S. Critical Appraisal of Clinical Research. J Clin Diagn Res. 2017; 11(5): JE01-JE05.
6. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem solving. Br Med J. 1995 Apr 29; 310(6987): 1122.



7. Furie KL, Rost NS. Overview of secondary prevention of ischemic stroke [Internet]. 2017. Cited: 2019 Feb 2. Available from: <https://www.uptodate.com/contents/overview-of-secondary-prevention-of-ischemic-stroke/>
8. CourcierDuplantier S, Falissard B, Fender P. Subjective Outcome Measures of Drug Efficacy. *Thérapie*. 2003 May 1; 58(3): 26773.
9. Pocock SJ, Stone GW. The Primary Outcome Is Positive — Is That Good Enough? [Internet]. 2016. Cited: 2019 Feb 2. Available from: https://www.nejm.org/doi/full/10.1056/NEJMra1601511?fbclid=IwAR2lWYmRlnpzOFUNRo07c3eQ4x0reYz1Y1uZ-AKnY-Wiw_ZV-oiLLBUYt44#t=article
10. Kip KE, Hollabaugh K, Marroquin OC, Williams DO. The Problem with Composite End Points in Cardiovascular Studies: The Story of Major Adverse Cardiac Events and Percutaneous Coronary Intervention. *J Am Coll Cardiol*. Vol. 51(7): 701-7.
11. Dawson BT, Robert G. *Basic & clinical biostatistics*, 4 ed. United States: New York: Lange Medical Books-McGraw-Hill, Medical Pub. Division, c2004.; 2004.