

ความลำเอียงในการวิจัยทางการแพทย์เชิงปริมาณ: การระบุและการลด ข้อผิดพลาดที่พบบ่อย

วงจันทร์ เพชรพิเชฐเชียร¹ พย.ด. (การพยาบาล) อพย. (การพยาบาลผู้ใหญ่และผู้สูงอายุ)

บทคัดย่อ: ความลำเอียงในการวิจัยเป็นปัญหาที่พบได้บ่อยและอาจลดความถูกต้องและความน่าเชื่อถือของผลการวิจัยในงานวิจัยทางการแพทย์เชิงปริมาณ ความลำเอียงสามารถเกิดขึ้นได้ในหลายขั้นตอนของกระบวนการวิจัย ตั้งแต่การกำหนดปัญหาจนถึงการรายงานผลการศึกษา บทความนี้มีวัตถุประสงค์เพื่อระบุความลำเอียงที่พบบ่อยในงานวิจัยทางการแพทย์เชิงปริมาณและนำเสนอแนวทางการลดความลำเอียงเพื่อยกระดับคุณภาพงานวิจัย ได้แก่ ความลำเอียงในการกำหนดปัญหาการวิจัย ความลำเอียงทางกรอบทฤษฎี/แนวคิด ไม่ถูกต้อง การจำแนกประเภทของแบบการวิจัยไม่ถูกต้อง/ไม่เหมาะสม ความลำเอียงจากการคัดเลือกกลุ่มตัวอย่าง ความลำเอียงจากการจัดกลุ่มตัวอย่างเข้ากลุ่ม การขาดการควบคุม/จัดปัจจัยกวน แผนการทดลองไม่ถูกต้อง/ไม่ชัดเจน ความลำเอียงจากการวัดตัวแปร อำนาจการทดสอบต่ำและความลำเอียงในการวิเคราะห์ข้อมูล และความลำเอียงในการแปลผลการรายงานผล และการเสนอแนะผลการศึกษาไม่เหมาะสม โดยมีการอธิบายและยกตัวอย่างที่พบบ่อย พร้อมกับเสนอแนวทางการลดความลำเอียงเหล่านี้ ซึ่งหากนักวิจัยมีการดำเนินการวิจัยที่ถูกต้องในทุก ๆ ขั้นตอนของการวิจัย จะช่วยยกระดับคุณภาพงานวิจัยทางการแพทย์ ทำให้ผลการวิจัยมีความน่าเชื่อถือและสามารถนำไปใช้ในทางปฏิบัติได้ตลอดจนมีแนวทางการศึกษาวิจัยในอนาคตได้เหมาะสมยิ่งขึ้น

วารสารการปฏิบัติการพยาบาลและการผดุงครรภ์ไทย 2567; 11(2): 05-19

คำสำคัญ: ความลำเอียง; ความคลาดเคลื่อน; การวิจัยทางการแพทย์; การวิจัยเชิงปริมาณ; ข้อผิดพลาด

¹ผู้ช่วยศาสตราจารย์ นักวิชาการอิสระ จังหวัดสงขลา

วันที่รับบทความ 4 กันยายน 2567 วันที่แก้ไขบทความ 29 กันยายน 2567 วันที่ตอบรับบทความ 30 กันยายน 2567

Bias in Quantitative Nursing Research: Identifying and Mitigating Common Pitfalls

Wongchan Petpichetchian¹ Ph.D. (Nursing) (Adult and Gerontological Nursing), Dip. APMSN.

Abstract: Bias in research is a common issue that can reduce the validity and reliability of study findings. In quantitative nursing research, bias can occur at various stages of the research process, from problem formulation to the reporting of study results. This paper aims to identify common biases in quantitative nursing research and provide strategies to reduce these biases to improve research quality. The types of bias discussed include problem framing bias, theoretical/ conceptual bias, misclassification of research design, sampling bias, selection bias, uncontrolled confounders, inadequate protocol descriptions, measurement bias, low power and analysis bias, and bias in interpretation and reporting and inappropriate implications and recommendations. The paper explains and provides common examples of these biases, along with strategies to mitigate them. If researchers correctly conduct their studies at every stage, it can enhance the quality of nursing research, making the findings more reliable, applicable in practice, and more suitable for future research.

Thai Journal of Nursing and Midwifery Practice 2024; 11(2): 05-19

Keywords: Bias; Errors; Nursing research; Quantitative research; Pitfalls

¹Assistant Professor, Independent scholar, Songkhla

Received September 4, 2024; Revised September 29, 2024; Accepted September 30, 2024

บทนำ

การรายงานผลการวิจัยที่ปรากฏในวารสารวิชาการและวิจัยทางการแพทย์แทบทั้งในประเทศและต่างประเทศ มักพบปรากฏการณ์ของความลำเอียงหรืออคติ (biases) ประเภทต่าง ๆ ในแทบทุกขั้นตอนของการวิจัยอยู่เสมอ แม้ว่ารายงานการวิจัยเหล่านั้นได้รับการตีพิมพ์เผยแพร่ในวารสารชั้นนำแล้วก็ตาม ซึ่งเมื่อมีการประเมินงานวิจัยอย่างมีวิจารณญาณ (critical appraisal) อาจพบข้อบกพร่องหรือความลำเอียงที่ส่งผลต่อคุณภาพงานวิจัย ดังข้อค้นพบจากการศึกษาการทบทวนขอบเขตงานวิจัยทางการแพทย์พยาบาลไทยที่เผยแพร่ในระหว่างปี พ.ศ. 2561 ถึง พ.ศ. 2564 ที่พบว่า ในภาพรวมคุณภาพของงานวิจัยโดยนักวิจัยชาวไทยที่ตีพิมพ์ในวารสารระดับ 1 ของฐานข้อมูลของศูนย์ดัชนีการอ้างอิงวารสารไทย (669 เรื่อง) และวารสารนานาชาติที่อยู่ในฐานข้อมูล CINAHL, PubMed และ Scopus (170 เรื่อง) ซึ่งเมื่อประเมินโดยใช้แบบประเมินมาตรฐานตามลักษณะการออกแบบการวิจัย ได้แก่ Joanna Briggs Institute (JBI) critical appraisal tools และ The Mixed Methods appraisal tools (MMAT) พบว่า มีคุณภาพในระดับสูงเพียงหนึ่งในสี่ (ร้อยละ 25.4) เท่านั้น¹ ทำให้เกิดคำถามตามมาว่าข้อค้นพบจากการศึกษาเหล่านั้น หากมีการนำมาใช้เพื่อขึ้นนำการปฏิบัติ จะได้ผลเช่นเดียวกับการศึกษานั้น ๆ ได้จริงหรือไม่ เพราะหากไม่จริง ผลเสียก็จะเกิดกับผู้รับบริการ กล่าวคือ อาจทำให้ไม่เกิดประโยชน์ต่อผู้รับบริการโดยเฉพาะผู้ป่วยกลุ่มโรคต่าง ๆ และประชาชนทั่วไป หรือทำให้เกิดความเสี่ยงได้อีกทั้งยังส่งผลต่อการพัฒนาศาสตร์ทางการแพทย์ ซึ่งมีความสำคัญต่อกระบวนการพัฒนาวิชาชีพ² ทำให้ภาพลักษณ์ของวิชาชีพการพยาบาลขาดความน่าเชื่อถือ เนื่องจากผลการวิจัย

ไม่สามารถตอบสนองความต้องการในการป้องกันโรค ส่งเสริมสุขภาพ ดูแลรักษาต่อเนื่อง พื้นฟูสุขภาพ รวมถึงการป้องกันภาวะแทรกซ้อนต่าง ๆ ให้กับผู้รับบริการสุขภาพได้อย่างที่สังคมคาดหวัง

ผู้เขียนในฐานะที่มีประสบการณ์และมีบทบาทสำคัญในด้านที่เกี่ยวกับ (1) การสอน-ถ่ายทอดความรู้ด้านกระบวนการวิจัยแก่นักศึกษาพยาบาลโดยเฉพาะอย่างยิ่งนักศึกษาระดับบัณฑิตศึกษา อาจารย์พยาบาลและพยาบาลนักวิจัย (2) การอ่านตรวจสอบบทความวิจัยก่อนการตอบรับตีพิมพ์เผยแพร่ในวารสารต่าง ๆ ทั้งระดับชาติและนานาชาติ และ (3) การทำหน้าที่ในกองบรรณาธิการของวารสารทางการแพทย์ของประเทศ ได้ระบุบทบาทของผู้ที่เกี่ยวข้องในการช่วยกันพัฒนาคุณภาพการวิจัยทางการแพทย์ ได้แก่ ผู้สอนเนื้อหาด้านการวิจัย พยาบาลนักวิจัย และผู้ทรงคุณวุฒิที่ทำหน้าที่ในการตรวจสอบบทความวิจัยก่อนการตีพิมพ์ว่าควรดำเนินการอย่างไรบ้าง เพื่อช่วยกันยกระดับคุณภาพการวิจัยทางการแพทย์ อย่างไรก็ตามยังคงพบข้อผิดพลาดในการดำเนินการและการนำเสนอบทความวิจัยอยู่เสมอมา บทความนี้จึงมีเป้าหมายสำคัญเพื่อนำเสนอข้อผิดพลาดที่พบบ่อยในกระบวนการวิจัยและการนำเสนอผลการวิจัย รวมทั้งให้ข้อเสนอแนะการป้องกันและลดความลำเอียงที่พบได้บ่อยในรูปแบบบทอรรถการเพื่อเป็นแนวทางสำหรับผู้ที่เกี่ยวข้อง โดยเฉพาะอย่างยิ่งพยาบาลนักวิจัย สามารถนำไปใช้ในการตรวจสอบการทำงานวิจัยในทุก ๆ ขั้นตอนเพื่อให้เกิดการพัฒนาศาสตร์ทางการแพทย์มีความก้าวหน้าและสามารถตอบสนองต่อความต้องการของสังคมได้อย่างมีประสิทธิภาพยิ่งขึ้น โดยเนื้อหาของบทความจะเน้นที่งานวิจัยเชิงปริมาณเป็นสำคัญ

ข้อผิดพลาดที่พบบ่อยและข้อเสนอแนะ

การดำเนินการวิจัยเพื่อให้ได้ผลการศึกษาน่าเชื่อถือจะต้องคำนึงถึงความถูกต้องและปราศจากความลำเอียงหรืออคติหรือป้องกันให้เกิดความลำเอียงน้อยที่สุดในทุก ๆ ขั้นตอน เริ่มตั้งแต่การคิดโจทย์และที่มาของปัญหาการวิจัย กรอบแนวคิด วัตถุประสงค์/คำถามการวิจัย (ต้นน้ำ) วิธีดำเนินการวิจัย: การเลือกแบบวิจัยและการออกแบบการวิจัย ในเรื่องประชากรและกลุ่มตัวอย่าง เครื่องมือที่ใช้ในการวิจัย จริยธรรมการวิจัย การดำเนินการศึกษาและเก็บข้อมูล การวิเคราะห์ข้อมูล (กลางน้ำ) และการแปลผลการวิจัย การรายงานผลและการให้ข้อเสนอแนะจากผลการวิจัย (ปลายน้ำ) ในทางปฏิบัติจริงของการวิจัยภาคสนามอาจมีข้อจำกัดบางประการที่ไม่สามารถควบคุมได้ หรือเกิดขึ้นด้วยข้อจำกัดของนักวิจัยเอง โดยเฉพาะอย่างยิ่งนักวิจัยหน้าใหม่ (เช่น นักศึกษาระดับบัณฑิตศึกษาทำวิจัยเพื่อเป็นส่วนหนึ่งของการศึกษา หรือพยาบาลวิชาชีพที่ทำวิจัยเพื่อใช้ประกอบการเลื่อนตำแหน่ง) ซึ่งมักเป็นข้อจำกัดด้านสมรรถนะ (research incompetence)⁴ เช่น การขาดความรู้ ขาดทักษะ และความรู้เท่าไม่ถึงการณ์ สมรรถนะด้านการวิจัยสามารถพัฒนาได้โดยการเรียนรู้ในระบบการฝึกอบรม การมีระบบพี่เลี้ยงให้ความช่วยเหลือ ตลอดจนการเรียนรู้ผ่านกิจกรรมวิชาการอื่น ๆ⁵ ซึ่งหากมีความจำกัดในการเรียนรู้จากแหล่งเรียนรู้ต่าง ๆ เหล่านี้อาจนำไปสู่ปัญหาการเกิดข้อผิดพลาดในกระบวนการวิจัย นอกจากนักวิจัยหน้าใหม่ นักวิจัยที่มีประสบการณ์ก็อาจมีการดำเนินการวิจัยผิดพลาดได้เช่นกัน หากมีการยึดติดกับความรู้และประสบการณ์เดิม ๆ ที่เคยเรียนรู้มาในอดีต โดยขาดการแสวงหาความรู้ เทคนิคและวิธีการดำเนินการวิจัยใหม่ ๆ อย่างสม่ำเสมอ

ข้อผิดพลาดที่พบบ่อยในงานวิจัยทางการแพทย์ ส่งผลให้เกิดข้อสงสัยในเรื่องความถูกต้องเชิงวิชาการ (scientific merit or scientific integrity) และความซื่อสัตย์ทางจริยธรรม (ethical integrity) ด้วยข้อจำกัดของพื้นที่การนำเสนอ บทความฉบับนี้จึงไม่ได้กล่าวถึงประเด็นที่เกี่ยวข้องกับความซื่อสัตย์ทางจริยธรรม และจริยธรรมการเผยแพร่ผลงาน (publication ethics) ซึ่งมีความสำคัญมากเช่นกัน แต่จะกล่าวเน้นถึงข้อผิดพลาดที่กระทบเรื่องความถูกต้องเชิงวิชาการเป็นสำคัญ โดยจำแนกข้อผิดพลาดที่พบบ่อยเป็น 10 ประเด็น ดังนี้ (1) ความลำเอียงในการกำหนดปัญหาการวิจัย (problem framing bias) (2) ความลำเอียงทางกรอบทฤษฎี/แนวคิดไม่ถูกต้อง (theoretical/conceptual bias) (3) การจำแนกประเภทของแบบการวิจัยไม่ถูกต้อง/ไม่เหมาะสม (misclassification of research design) (4) ความลำเอียงจากการคัดเลือกกลุ่มตัวอย่าง (sampling bias) (5) ความลำเอียงจากการจัดกลุ่มตัวอย่างเข้ากลุ่ม (selection bias) (6) การขาดการควบคุม/จัดปัจจัยกวน (uncontrolled confounders) (7) แผนการทดลองไม่ถูกต้อง/ไม่ชัดเจน (inadequate protocol descriptions) (8) ความลำเอียงจากการวัดตัวแปร (measurement/detection bias) (9) อำนาจการทดสอบต่ำและความลำเอียงในการวิเคราะห์ข้อมูล (low power and analysis bias) (10) ความลำเอียงในการแปลผลและการรายงานผลและการเสนอแนะผลการศึกษาไม่เหมาะสม (bias in interpretation and reporting and inappropriate implications and recommendations) ในแต่ละประเด็นมีรายละเอียดและข้อเสนอแนะสำหรับการป้องกันและลดข้อผิดพลาด (เคล็ดลับ) ดังนี้

1. ความลำเอียงในการกำหนดปัญหาการวิจัย ปัญหาเนี้ยมักเกิดจากการที่นักวิจัยมีความคิด/ความเชื่อส่วนตัวเกี่ยวกับปรากฏการณ์ที่สนใจศึกษา จึงนำเสนอข้อมูลไม่ครบถ้วนตามความเป็นจริง มีความจำกัดในการทบทวนวรรณกรรมในเรื่องที่เกี่ยวข้องนำไปสู่การกล่าวถึงความสำคัญของสิ่งที่ต้องการศึกษาเกินจริงหรือบ่อยครั้งที่พบว่าปัญหาการวิจัยได้มาจากการทบทวนวรรณกรรมของนักวิจัยที่ศึกษาก่อนหน้า แต่ขาดการวิเคราะห์ประเด็นสำคัญของการศึกษาที่ผ่านมาเพื่อบอกถึงช่องว่างขององค์ความรู้ทำให้เกิดปรากฏการณ์ของการศึกษาแบบซ้ำ ๆ ที่อาจไม่ได้ให้องค์ความรู้ใหม่แต่อย่างใด ดังนั้น ในขั้นตอนของการกำหนดปัญหาการวิจัย ซึ่งจะต้องปรากฏอยู่ในส่วนต้นของรายงานการวิจัยคือ บทนำหรือความเป็นมาและความสำคัญของการวิจัย ควรมีการดำเนินการดังนี้

“อ่านให้รอบตอบให้ชัดเจนให้เห็น

ช่องว่างเด่นเป็นจุดจุดค้นหา

ต้องวิจัยหาคำตอบเพื่อพัฒนา

ความเป็นมาบอกกล่าวให้ชัดเจน

ความรู้ใหม่จะสรรค์สร้างจากงานนี้

แม่นยำนิยตประโยชน์ทวิชี้ให้เห็น

ใครมาอ่านต่างก็รู้ว่าจำเป็น

เขียนชัดเจนทำ (วิจัยนี้) กันเถิดเกิดผลดี”

2. ความลำเอียงทางกรอบทฤษฎี/แนวคิด ไม่ถูกต้อง กรอบแนวคิดการวิจัยมีความสำคัญในงานวิจัยเชิงปริมาณ ที่นักวิจัยใช้เพื่อช่วยอธิบายความสัมพันธ์ระหว่างตัวแปรที่สนใจศึกษาอย่างมีตรรกะ (logical explanations) เพื่อคาดการณ์สิ่งที่คาดว่าจะพบในการศึกษานั้น ๆ กรอบแนวคิดการวิจัยอาจมีฐานความคิดมาจากทฤษฎีโมเดล หรือกรอบแนวคิดทางกายภาพและ/หรือศาสตร์ที่เกี่ยวข้องในสาขาอื่น ๆ เช่น ชีว-พยาธิวิทยา จิตวิทยา มนุษยวิทยา

สังคมวิทยา ที่มีผู้พัฒนาไว้แล้ว (extant theories, model, or frameworks) หรือนักวิจัยสังเคราะห์ขึ้นมาจากการทบทวนวรรณกรรมที่เกี่ยวข้องเพื่อทำการศึกษาปรากฏการณ์ที่สนใจ ช่วยให้การศึกษานั้น ๆ มีทิศทางที่ชัดเจน อย่างไรก็ตามในทางปฏิบัติ พบว่ามีการวิจัยจำนวนไม่น้อยที่มีการตีความทฤษฎีที่มีอยู่แล้วคลาดเคลื่อนไปจากต้นฉบับ ส่งผลให้เมื่อมีการคัดเลือกตัวแปรมาศึกษาหรือการอธิบายความสัมพันธ์ระหว่างตัวแปรมีความคลาดเคลื่อนไปด้วยนำไปสู่การอธิบายข้อค้นพบจากการวิจัยคลาดเคลื่อน จึงควรมีการดำเนินการดังนี้

“กรอบแนวคิดการวิจัยคือเสาหลัก

ช่วยชี้ชักพานำให้ค้นหา

ใช้ตรรกะนำทางคัดสรรมา

ตัวแปรหลักที่ศึกษาต้องไตร่ตรอง

อาจใช้กรอบทฤษฎีที่มีอยู่

หรือสังเคราะห์จากผู้รู้หมู่เพื่อนมอง

อย่าบิดเบือนหลักการต้องคล้อยจง

คิดถูกต้องทำให้ดีจึงมีชัย”

3. การจำแนกประเภทของแบบการวิจัย ไม่ถูกต้อง/ไม่เหมาะสม ในปัจจุบันแนวปฏิบัติที่ดีของการนำเสนอบทความวิจัยมักจะระบุให้มีการเขียนชื่อแบบวิจัยให้ชัดเจนตั้งแต่ชื่อเรื่องและ/หรือบทคัดย่อ เช่น CONSORT ให้ระบุว่าเป็นงานวิจัยทดลองแบบสุ่มเข้ากลุ่ม (randomized trials)⁶ หรือ STROBE ให้ระบุว่าเป็นงานวิจัยสังเกตปรากฏการณ์⁷ เป็นต้น เมื่อผู้อ่านเห็นชื่อเรื่องก็จะเข้าใจว่าวัตถุประสงค์หรือคำถามการวิจัยของการศึกษานั้น ๆ จะต้องมีการดำเนินการวิจัยอย่างไรเพื่อให้ได้คำตอบที่ต้องการ ตัวอย่างเช่น หากเป็นการวิจัยทดลองแบบสุ่มเข้ากลุ่มและมีกลุ่มเปรียบเทียบ (randomized controlled trial: RCT) หรือการวิจัยกึ่งทดลอง (quasi-

experimental study) ผู้อ่านย่อมคาดหวังที่จะเห็นกระบวนการของการทดลองหรือการจัดกระทำ (manipulation) ซึ่งเป็นตัวแปรต้น (independent variable) ที่จะส่งผลต่อการเปลี่ยนแปลงของตัวแปรผลลัพธ์ (dependent variable) หรือไม่ อย่างไรก็ตาม ตัวอย่างที่พบความผิดพลาด เช่น นักวิจัยนำแบบคัดกรองความเสี่ยงมากกว่าหนึ่งชนิดไปทดสอบเพื่อประเมินว่าแบบคัดกรองอันใด มีความสามารถในการคัดกรองความเสี่ยงได้ถูกต้องมากกว่า นักวิจัยระบุว่าการใช้การวิจัยแบบทดลองซึ่งไม่ถูกต้อง เพราะการศึกษาในลักษณะดังกล่าวการนำแบบคัดกรองความเสี่ยงมาทดสอบ ไม่ใช่การจัดกระทำที่จะไปส่งผลให้กลุ่มตัวอย่างที่ศึกษามีหรือไม่มีอาการเปลี่ยนแปลงไปจากเดิมที่เขาเป็นอยู่แล้ว ดังนั้นการศึกษาลักษณะนี้ จึงเป็นการศึกษาจากการสังเกตปรากฏการณ์ (observational study) โดยอาจใช้แบบการวิจัยเป็นการศึกษาแบบภาคตัดขวาง (cross-sectional study) หากมีการเก็บข้อมูลครั้งเดียว หรือระบุชัดเจนขึ้นว่าเป็นการศึกษาแบบเปรียบเทียบ (comparative study) เป็นต้น

นอกจากนี้ยังพบว่าการเรียกชื่อแบบวิจัยเป็น อย่างหนึ่งแต่ดำเนินการจริงเป็นอีกอย่างหนึ่ง ตัวอย่างเช่น ระบุว่าเป็นการวิจัยทดลองแบบข้ามกลุ่ม (crossover design) แต่ในขั้นตอนการดำเนินการวิจัย อธิบายว่ามีการศึกษากลุ่มเปรียบเทียบ/กลุ่มควบคุม จนแล้วเสร็จ จึงเริ่มศึกษากับกลุ่มทดลอง (quasi-experimental with a historic control group design) เป็นต้น ทั้งนี้การเลือกและเรียกชื่อแบบวิจัย จะต้องคำนึงถึงวัตถุประสงค์และคำถามการวิจัยโดยพยายามเลือกแบบการวิจัยที่จะช่วยให้ได้คำตอบที่น่าเชื่อถือมากที่สุดด้วยเสมอ เช่น หากต้องการทดสอบความเป็นเหตุและผล (causal relationship) เพื่อบอกถึง

ประสิทธิภาพหรือประสิทธิผล ก็ควรใช้การวิจัยแบบทดลอง RCT ยกเว้นในกรณีที่มีข้อจำกัดบางประการ ที่ทำไม่ได้ เช่น อาจเกิดประเด็นปัญหาในเชิงจริยธรรม หรืออาจมีปัจจัยคุกคามต่อความตรงภายใน (threats to internal validity) ดังนั้นเพื่อลดข้อผิดพลาดในลักษณะนี้ ควรมีการดำเนินการดังนี้

“การเลือกแบบการวิจัยที่เข้มข้น

ช่วยนำผลการศึกษานำเชื่อถือ

ไม่แน่ใจโปรดสอบถามจากคุณชื่อ

หมั่นเรียนรู้สม่ำเสมอ (ว่าเรา) ใช้งานเป็น

การเรียกชื่อแบบวิจัยอย่าสับสน

อย่าหลงกลเพียงเพราะใก้ยากลองเล่น

ต้องสอดคล้องโจทย์คำถามตรงประเด็น

ต้องเลือกพื้นที่ที่สอดคล้องตามมรรคา”

4. ความลำเอียงจากการคัดเลือกกลุ่มตัวอย่าง ในทางปฏิบัติงานวิจัยทางการแพทย์ส่วนใหญ่เป็นการศึกษาในกลุ่มคน (person) ณ สถานที่ (place or setting) และช่วงเวลา (time) เพียง “จำนวนหนึ่ง” เท่านั้น การคัดเลือกกลุ่มตัวอย่างจึงต้องมีความสามารถเป็นตัวแทนที่ดีของประชากร (representativeness) เพื่อให้ผลการศึกษาที่พบจากการวิจัยนั้น ๆ สามารถอ้างอิงกลับไปยังประชากร (generalizability) หรือเพื่อให้มีความตรงภายนอก (external validity)² นั่นเอง

กระบวนการและเทคนิคการคัดเลือกกลุ่มตัวอย่างเข้าศึกษาจึงมีความสำคัญโดยเริ่มต้นจากการออกแบบการเข้าถึงกลุ่มตัวอย่าง (การเชิญชวนเข้าร่วมศึกษา) การกำหนดเกณฑ์คุณสมบัติการคัดเลือกกลุ่มตัวอย่าง (เกณฑ์การคัดเข้า คัดออก ถอนตัวหรือยุติ) และเทคนิคที่ใช้ซึ่งมีทั้งแบบที่มีการสุ่มและไม่มีการสุ่ม (probability or random VS non-probability or non-random sampling) โดยนักวิจัยควรออกแบบ

การคัดเลือกกลุ่มตัวอย่าง (sampling plan) ให้เกิดความลำเอียงน้อยที่สุด ตัวอย่างที่พบเป็นข้อที่ทำให้ความสับสนเสี่ยงต่อความลำเอียงในเรื่องนี้ เช่น (1) การคัดเลือกสถานที่ศึกษาหรือกลุ่มคนที่เน้นความสะดวกของนักวิจัยเป็นหลัก (แคบเกินไป) ไม่มีการสุ่มกลุ่มตัวอย่าง แต่มักใช้การกำหนดเกณฑ์คุณสมบัติในการคัดเลือกกลุ่มตัวอย่างแบบเฉพาะเจาะจง (purposive sampling) การได้มาซึ่งกลุ่มตัวอย่างในลักษณะนี้จึงมีโอกาสเกิดปัจจัยคุกคามความตรงภายนอก เช่น การมีปฏิสัมพันธ์เกิดขึ้นระหว่างกระบวนการทดลองในหน่วยงานนั้น ๆ (interaction of treatments with settings) หรือผลลัพธ์การเปลี่ยนแปลงที่เกิดขึ้นมีความมีความจำเพาะกับบริบทนั้น ๆ เท่านั้น ที่ไปช่วยให้เกิดผลลัพธ์การศึกษาที่ดี ซึ่งหากนำไปใช้ในหน่วยงานอื่น อาจจะได้ผลเช่นนั้น⁸ เป็นต้น (2) นักวิจัยเป็นผู้ปฏิบัติงานอยู่ในพื้นที่ศึกษา (เช่น เป็นผู้ให้บริการโดยตรง) เป็นผู้เข้าไปเชิญชวนการเข้าร่วมโครงการวิจัย ดังนั้นกลุ่มตัวอย่างอาจเข้าสู่วิทยาศาสตร์ด้วยความเกรงใจ อย่างไรก็ตามแม้ว่าการสุ่มกลุ่มตัวอย่างจากประชากรเป้าหมายเป็นเทคนิคในอุดมคติในการปฏิบัติจริงอาจไม่สามารถทำได้ในทุกกรณี (เช่น หน่วยงานบางแห่งอาจไม่สมัครใจเข้าร่วม แม้จะถูกสุ่มเลือกเป็นสถานที่ศึกษา) บทบาทสำคัญของนักวิจัยคือการอธิบายเหตุผลไว้ให้ชัดเจน และเมื่อนำไปสู่การวิเคราะห์ข้อมูลและอภิปรายผลการศึกษา โดยเฉพาะเมื่อมีการใช้สถิติอนุมาน (inferential statistics) จะต้องกระทำด้วยความระมัดระวัง ไม่แปลผลเกินเลยมากเกินไป เพราะหนึ่งในข้อตกลงเบื้องต้นที่สำคัญของสถิติอนุมานคือกลุ่มตัวอย่างต้องได้มาจากการสุ่มจากประชากร² เคลล์คลับสำคัญในการลดความลำเอียงจากการคัดเลือกกลุ่มตัวอย่าง คือ

“หากทำได้ขอให้สุ่มกลุ่มตัวอย่าง
เพื่อช่วยเพิ่มความเป็นกลางความโปร่งใส
ช่วยขจัดอคติมีทั่วไป (กำหนดเกณฑ์ชัดเจน)
หากไม่ได้(สุ่ม) ไม่ว่ากันเลยตามเลย
การเข้าถึงกลุ่มตัวอย่างในบางกลุ่ม
อาจเสี่ยงอคติแม้เจียบเจย
ใช้คนนอกเข้าช่วยคัด (เลือก) นำชมเชย
เมื่อวิเคราะห์ (ทางสถิติ) อย่า (สรุป) เกินเลย
ต้องระวัง”

5. ความลำเอียงจากการจัดกลุ่มตัวอย่าง
เข้ากลุ่ม ความลำเอียงในเรื่องนี้ มักเกิดขึ้นในการศึกษาที่นักวิจัยต้องการเปรียบเทียบตัวแปรระหว่างกลุ่มตัวอย่างที่มีตั้งแต่สองกลุ่มขึ้นไปเพื่อศึกษาความสัมพันธ์เชิงสาเหตุ (causal relationship) ซึ่งอาจเป็นการศึกษาแบบศึกษาปรากฏการณ์ที่ไม่มีการทดลอง หรือการศึกษาแบบทดลองก็ได้⁹ สำหรับการศึกษาปรากฏการณ์ที่มีลักษณะดังกล่าว คือ การเกิดของกลุ่มที่นำมาเปรียบเทียบเป็นกลุ่มที่เกิดขึ้นโดยธรรมชาติไม่ได้เกิดจากการจัดกระทำของนักวิจัยที่ใช้บ่อย ได้แก่ (1) การศึกษาแบบ case-control study ตัวอย่างเช่น การศึกษาเปรียบเทียบระหว่างกลุ่มที่เกิดการติดเชื้อจากการใส่สายสวนปัสสาวะ-กลุ่ม case และกลุ่มไม่มีการติดเชื้อทางเดินปัสสาวะ-กลุ่ม control เพื่อค้นหาและอธิบายปัจจัยเสี่ยงที่ส่งผลต่อการติดเชื้อดังกล่าว (2) การศึกษาแบบ cohort study ตัวอย่างเช่น การศึกษาปัจจัยที่เกี่ยวข้องกับการทำงานของพยาบาล เช่น การอยู่เวรบ่อย-ดึก (กลุ่มที่อยู่เวรและกลุ่มที่ไม่อยู่เวร ซึ่งเป็นกลุ่มที่มีอยู่แล้วในธรรมชาติ) เพื่ออธิบายว่าไปส่งผลให้เกิดปัญหาสุขภาพหรือไม่และ (3) การศึกษาเปรียบเทียบแบบ comparative study ตัวอย่างเช่น การศึกษาเปรียบเทียบความสามารถของเครื่องมือวัดสองชนิด (เครื่องมือ ก. และเครื่องมือ

ข.) ในการคัดกรองความเสี่ยงต่อการพลัดตกหกล้ม เพื่อประเมินว่าเครื่องมือใดสามารถตรวจจับความเสี่ยงที่นำไปสู่การพลัดตกหกล้ม (เกิด หรือ ไม่เกิด) ได้ดีกว่ากัน จะเห็นได้ว่าการศึกษาปรากฏการณ์ที่ไม่ใช่การทดลองเพื่ออธิบายความเป็นเหตุ (หาปัจจัยที่นำมาก่อน) เป็นผล (หาผลลัพธ์ที่ตามมา) นั้น กลุ่มที่เกิดขึ้นตามธรรมชาติเหล่านั้น อาจมีความไม่เท่าเทียมกันอยู่แล้วตามธรรมชาติ (เช่น มีจำนวนเพศชายและเพศหญิงไม่เท่ากัน) จึงควรมีการออกแบบการได้มาของกลุ่มตัวอย่างเหล่านั้นให้มีความเท่าเทียมกันมากที่สุด เช่น อาจใช้การจับคู่ในปัจจุบันบางประการที่อาจส่งผลต่อคำตอบที่นักวิจัยสนใจ และควรมีการแสดงให้เห็นด้วยว่ากลุ่มเหล่านั้นมีความเหมือน-ความต่างกันอย่างไรหรือไม่

สำหรับการศึกษาแบบทดลอง ที่มีการจัดกลุ่มตัวอย่างเข้ากลุ่มตั้งแต่สองกลุ่มขึ้นไป (เช่น กลุ่มทดลอง และกลุ่มควบคุม) การจัดกลุ่มตัวอย่างเข้ากลุ่มเป็นการดำเนินการโดยนักวิจัย จึงมีโอกาสเกิดความลำเอียงในการจัดเข้ากลุ่มสูง โดยเฉพาะอย่างยิ่งการวิจัยแบบกึ่งทดลอง (quasi-experimental research) ที่ไม่ใช้การจัดเข้ากลุ่มโดยการสุ่ม (non-randomized controlled trial: non-RCT) ซึ่งแตกต่างไปจากการวิจัยแบบสุ่มที่มีกลุ่มควบคุม (randomized controlled trial: RCT) ที่ใช้การจัดกลุ่มเข้ากลุ่มโดยการสุ่ม (randomization) จึงมีความลำเอียงในลักษณะนี้น้อยกว่า อย่างไรก็ตามการวิจัย RCT นักวิจัยจะต้องระบุให้ชัดเจนว่ามีการจัดลำดับการเข้ากลุ่ม (sequence generation) อย่างไร มีกระบวนการอำพรางการเข้ากลุ่ม (allocation concealment) ด้วยหรือไม่ เพื่อให้มั่นใจว่าไม่มีความลำเอียงจากการจัดกลุ่มตัวอย่างเข้ากลุ่ม ดังปรากฏเป็นหนึ่งในข้อกำหนดของการรายงานบทความวิจัย RCT ตามแนวทางของ CONSORT

(Consolidated Standards of Reporting Trials)⁶ ตลอดจนการคำนึงถึงกรณีที่มีกลุ่มตัวอย่างสูญหาย (attrition) ภายหลังจากการจัดเข้ากลุ่มไปแล้ว เพราะจะเป็นแหล่งของความลำเอียงสำคัญเช่นกัน การดำเนินการที่เหมาะสมจะช่วยเพิ่มความตรงภายใน (internal validity) ของการวิจัยนั้น ๆ

“ความลำเอียงจัดเข้ากลุ่มพบเสมอ
อย่าได้ผลพลาดพลั้งรังแกมัน
ใช้การสุ่มเข้ากลุ่มคุณอนันต์
อย่างน้อยนั้นจับคู่ ก็ยังดี
จะดีขึ้นหากอำพรางไม่รู้ก่อน
ไม่แย่งย่อนหากศึกษาให้ถ้วนถี่
เพิ่มความตรงภายในของเราดี
เพราะเหตุนี้ (ตัวแปรต้น) จึงส่งผล (ตัวแปรตาม)
ดั่งตั้งใจ”

6. การขาดการควบคุม/ขจัดปัจจัยกวน
ปัจจัยกวน (confounding variables or confounders) เป็นตัวแปรที่อยู่นอกเหนือคำถามการวิจัย แต่ส่งผลกระทบต่อความสัมพันธ์ระหว่างตัวแปรที่นักวิจัยสนใจศึกษา² ซึ่งพบได้ทั้งในงานวิจัยสังเกตปรากฏการณ์และการวิจัยแบบทดลอง งานวิจัยที่มีการระบุปัจจัยกวนชัดเจน บอกให้ทราบว่ามีปัจจัยมีการดำเนินการเพื่อควบคุมและ/หรือขจัดปัจจัยกวนช่วยให้ผู้อ่าน/ผู้ใช้ประโยชน์จากงานวิจัยนั้น ๆ เกิดความมั่นใจว่าผลการศึกษามีความตรงภายใน การควบคุมและ/หรือขจัดปัจจัยกวนสามารถทำได้โดยการออกแบบการวิจัยหรือการใช้สถิติขั้นสูง¹⁰ ตัวอย่างการควบคุมโดยการออกแบบการวิจัยโดยกำหนดให้การคัดเลือกกลุ่มตัวอย่างมีความเท่าเทียมกันในปัจจัยกวนนั้น ๆ เช่น หากตัวแปรเพศเป็นปัจจัยกวน ต้องออกแบบการได้มาของกลุ่มตัวอย่างให้มีเพศชายและเพศหญิงในสัดส่วนเท่า ๆ กัน หรือใช้สถิติในการพิสูจน์ให้

เห็นว่าข้อค้นพบการศึกษาไม่แตกต่างกันไม่ว่าจะเป็นเพศหญิงหรือเพศชาย โดยการวิเคราะห์แยกกลุ่ม (subgroup analysis) หรืออาจใช้สถิติในการควบคุมตัวแปรกวน โดยจัดปัจจัยกวนเป็นตัวแปรร่วม (covariate) แล้วจึงใช้สถิติในการขจัดอิทธิพลของปัจจัยกวน เช่น สถิติ analysis of covariance (ANCOVA) รวมถึงการนำมาสู่การอธิบายหากผลการศึกษาไม่เป็นตามที่คาดการณ์ไว้ การทบทวนวรรณกรรมที่ช่วยให้การวิจัยสามารถค้นหาวาอาจมีปัจจัยใดเป็นปัจจัยกวนได้บ้าง ทั้งนี้อาจยังมีปัจจัยกวนอื่น ๆ ที่ยังไม่ทราบหรือไม่สามารถคาดการณ์ล่วงหน้าได้ นักวิจัยจึงต้องใช้การคิดวิเคราะห์เมื่อเกิดข้อสงสัยและนำไปสู่การอภิปรายผลต่อไป

“เจ้าตัวร้ายปัจจัยกวนป่วนขอบอก
คิดให้ออกบอกให้ได้ตัวไหนบ้าง
การทบทวนวรรณกรรมช่วยชี้ทาง
นำแฉ้วถางทางเตียนโล่งโปร่งสบาย
เก็บข้อมูลมาไว้ก่อนเป็นดีหนา
จากนั้นมาวิเคราะห์หาสื่อความหมาย
ขจัดออกให้พ้นทางห้ามกริดกราย
จึงบอกผลที่ได้อย่างมั่นใจ”

7. แผนการทดลองไม่ถูกต้อง/ไม่ชัดเจนในการวิจัยแบบทดลอง แผนหรือขั้นตอนการทดลอง (intervention protocol) มีความสำคัญเพราะเป็นตัวบ่งชี้ถึงปัจจัยที่เป็นเหตุ (X) ที่ส่งผลต่อการเปลี่ยนแปลงในตัวแปรผลลัพธ์ (Y) โดยข้อผิดพลาดที่พบบ่อย ได้แก่ (1) การกำหนดแผนการทดลองไม่สอดคล้องกับกรอบแนวคิดการวิจัยที่มีการนำทฤษฎี/กรอบแนวคิด/โมเดลที่มีอยู่แล้วมาใช้ซึ่งอาจเกิดจากความเข้าใจคลาดเคลื่อนของนักวิจัย หรือในบางครั้งมีเพียงการ “กล่าวนาม” ถึงทฤษฎี/กรอบแนวคิด/โมเดล โดยไม่ได้นำมาใช้ออกแบบการ

ทดลองแต่อย่างใด (2) การรายงานในบทความขาดรายละเอียดของแผนการทดลอง ทำให้ผู้อ่านขาดความมั่นใจว่าตอนดำเนินการทดลองจริง จะมีการดำเนินการทดลองได้ถูกต้องและสม่ำเสมอ (fidelity) ดังที่นักวิจัยต้องการได้มากนักน้อยเพียงใด และเกิดข้อจำกัดหากมีผู้สนใจต้องการนำไปใช้หรือไปศึกษาซ้ำ จึงควรมีการจัดทำแผนการทดลองและรายงานไว้ให้ชัดเจนโดยระบุรายละเอียดให้ทราบว่า ใครทำ (actor) ทำอะไร (what) ทำกับใคร (target) ทำเมื่อไร (temporality) ทำแค่ไหน มาก-น้อย ถี่-ห่าง (dose) หวังผลอะไร (outcome) และด้วยเหตุผลอะไร (justification)¹¹ และ (3) ขาดการใช้เทคนิคปกปิด (no blinding) เช่น ไม่ปกปิดกลุ่มตัวอย่างว่าตนเองอยู่ในกลุ่มใด (ทดลองหรือควบคุม) หรือไม่ปกปิดผู้กระทำการทดลองว่ากลุ่มตัวอย่างรายใดอยู่ในกลุ่มใด นำไปสู่อคติหรือความลำเอียงด้านการปฏิบัติ (performance bias)² ซึ่งในทางปฏิบัติอาจไม่สามารถกระทำการปกปิดได้ทั้งสองฝ่าย (ผู้ให้การทดลองและผู้ได้รับการทดลอง) เนื่องจากลักษณะธรรมชาติของการศึกษา (เช่น หากการทดลองคือการนวดกดจุด การปกปิดผู้ให้การทดลองไม่สามารถทำได้) แต่อาจปกปิดผู้ได้รับการทดลองได้ (เช่น ปกปิดโดยการให้การนวดกับทั้งสองกลุ่ม โดยกลุ่มทดลองทำการนวดกดจุดตามแผนการทดลอง ส่วนกลุ่มควบคุมทำการนวดหลอก) โดยสรุป นักวิจัยสามารถป้องกันข้อผิดพลาดเหล่านี้ได้โดย

“การออกแบบการทดลองต้องมีแผน
เขียนให้แน่นเพื่อนำ (ต่อ) ได้ไม่สับสน
ต้องถูกหลักวิชาการ (ทฤษฎี) อย่างแยบยล
ใครทำอะไร ทำทำไม ที่ไหนและเมื่อไร
อีกทั้งยังต้องควบคุมความถูกต้อง
ทั้งผู้รับ-ผู้ให้ (การทดลอง) ไม่สงสัย (blind)

หากมีเหตุชวนชั่งกวนบันทึกไว้
อธิบายได้ราบรื่นเมื่อถึงครา”

8. ความลำเอียงจากการวัดตัวแปร เป็นความลำเอียงที่เกิดขึ้นในกระบวนการวัด (measurement) หรือตรวจจับการเปลี่ยนแปลง (detection) มักเกิดขึ้นกับตัวแปรที่มีการวัดแบบอัตนัย (subjective measures) เช่น แบบสอบถาม แบบสังเกต แบบวัดคุณสมบัติบางประการ ๑ ที่ผู้ตอบเป็นผู้รายงานด้วยตนเอง (self-report) ในขณะที่ตัวแปรที่มีการวัดแบบปรนัย (objective measures) เช่น ส่วนสูง-น้ำหนัก ความดันโลหิต ระดับน้ำตาลในเลือด ๑ มีความลำเอียงประเภทนี้ต่ำกว่า² ตัวแปรที่มีการวัดแบบอัตนัย มักเป็นการวัดคุณลักษณะ (attributes) ที่เป็นนามธรรมเพื่อให้สามารถจับต้องได้ เช่น วัด ‘คุณภาพชีวิต’ ออกมาเป็นค่าตัวเลข ดังนั้นความสำคัญจึงอยู่ที่เครื่องมือที่นำมาใช้วัดตัวแปรเหล่านี้จะต้องมี “กฎ” บางประการในการวัดและแปลงคุณลักษณะเหล่านั้นเป็นค่าตัวเลขที่น่าเชื่อถือเพื่อนำไปสู่การวิเคราะห์ข้อมูลต่อไป กฎในความหมายของนันทลีและเบิร์นสไตน์¹² ครอบคลุมคุณสมบัติของเครื่องมือวัดที่ดีในเรื่อง (1) ความตรง (validity) คือ ความสามารถในการวัดคุณลักษณะที่ต้องการวัด (2) ความเที่ยง (reliability) คือ ความสม่ำเสมอของเครื่องมือวัด (3) และความคลาดเคลื่อนจากการวัด (measurement error)

ในการออกแบบการวัด ไม่ว่าจะเป็นการสร้างเครื่องมือขึ้นมาใหม่หรือการนำเครื่องมือที่มีผู้พัฒนาไว้ก่อนหน้า จะต้องมีการกำหนดนิยามตัวแปรที่ต้องการวัดที่ชัดเจน รายงานและแสดงผลให้ทราบเกี่ยวกับการตรวจสอบความตรงและความเที่ยงที่ถูกต้องและเหมาะสมกับลักษณะของเครื่องมือวัดนั้น ๆ ข้อผิดพลาดที่พบบ่อยคือ (1) การใช้เครื่องมือวัดที่อาจไม่สามารถวัดสิ่งที่ต้องการวัดได้อย่างแท้จริงจาก

การขาดการนิยามตัวแปรเชิงปฏิบัติการที่ชัดเจน หรือมีการนำเครื่องมือวัดที่มีอยู่แล้วที่ไม่ตรงกับนิยามมาใช้ หรือนักวิจัยพัฒนาเครื่องมือขึ้นมาใช้ โดยมีเพียงการตรวจสอบความตรงตามเนื้อหา (content validity) เท่านั้น ซึ่งอาจยังไม่เพียงพอ โดยเฉพาะเครื่องมือวัดตัวแปรที่มีคุณลักษณะหลายมิติ (multidimensional construct) จึงควรนำไปสู่การพัฒนาให้ได้เครื่องมือที่มาตรฐานมากขึ้นต่อไป (2) การตรวจสอบความเที่ยง พบความลำเอียงที่เกิดจากการใช้วิธีการทดสอบที่ไม่เหมาะสมกับลักษณะเครื่องมือ ทำให้ผลการทดสอบที่ได้ไม่น่าเชื่อถือ เช่น เครื่องมือวัดเป็นแบบสังเกต (observational checklist) ผู้วิจัยใช้การตรวจสอบความเที่ยงแบบความสอดคล้องภายใน (internal consistency reliability) และรายงานด้วยสถิติ Kuder-Richardson-20 (KR-20) หรือ Cronbach’s Alpha Coefficient¹² ที่เหมาะสมควรตรวจสอบความเที่ยงโดยการประเมินความเท่าเทียมระหว่างผู้ประเมิน (interrater reliability) และใช้สถิติ Fleiss Kappa หรือ Intraclass Correlation Coefficient (ICC)¹³ ขึ้นอยู่กับระดับของการวัดตัวแปรนั้น ๆ เป็นต้น (3) การใช้เครื่องมือที่พัฒนาขึ้นจากสังคมวัฒนธรรมและการใช้ภาษาที่แตกต่างไปจากพื้นที่ที่ศึกษา (เช่น เครื่องมือที่พัฒนาในต่างประเทศเป็นภาษาอังกฤษ) ซึ่งต้องเข้าสู่กระบวนการแปลที่ถูกต้องตามมาตรฐานสากล และต้องได้รับอนุญาตให้มีการแปลและการใช้เครื่องมือ นั้น ๆ อย่างถูกต้องด้วยเพื่อป้องกันปัญหาข้อโต้แย้งเรื่องลิขสิทธิ์ทางปัญญา หากมีการนำเครื่องมือที่นักวิจัยคนไทยแปลไว้ก่อนหน้านี้ ควรตรวจสอบความถูกต้องของการแปลเพื่อลดความเสี่ยงที่เกิดขึ้นจากความคลาดเคลื่อนในกระบวนการแปล และ (4) การได้รับอิทธิพลจากผู้เก็บข้อมูลทำให้ข้อมูลที่ได้มี

ความคลาดเคลื่อน เช่น นักวิจัยเป็นผู้ให้การทดลอง (interventionist) และเป็นผู้เก็บข้อมูลด้วยตนเอง (data collector) กลุ่มตัวอย่างอาจตอบคำถามหรือแบบสอบถามด้วยความเกรงใจ ดังนั้นในการวิจัยแบบทดลองใด ๆ หากสามารถทำได้ควรให้ผู้อื่นเป็นผู้เก็บข้อมูลและใช้กระบวนการปกปิดไม่ให้ผู้เก็บข้อมูลทราบว่ากลุ่มตัวอย่างรายใดอยู่ในกลุ่มใด (blinding data collector)

“เครื่องมือวัดมีความตรงและความเที่ยง
นิยมไว้มิพลาดเอียงไม่ตกหล่น
มาตรฐานการวัดเครื่องจักรกล
ตรวจสอบไว้มิสับสนไม่ติดใจ
หากมีการแปลมาจากสังคมอื่น
ขออย่าฝืนทนใช้หากสงสัย
โปรดอย่าลืมลิสทริ่งขอใช้
หากทำได้ให้ผู้อื่นเก็บข้อมูล”

9. อำนาจการทดสอบต่ำและความลำเอียง
ในการวิเคราะห์ข้อมูล การสรุปผลการศึกษานงานวิจัยเชิงปริมาณ มักมีการใช้สถิติเพื่อประมาณการค่าของประชากร (population parameters) โดยใช้ข้อมูลจากค่าสถิติที่ได้จากกลุ่มตัวอย่างที่ศึกษา (sample statistics) เพื่อให้ผลการวิเคราะห์ข้อมูลหรือการทดสอบทางสถิติมีความถูกต้อง ใกล้เคียงกับความเป็นจริงมากที่สุด (statistical conclusion validity) หรือที่เรียกกันว่าการวิจัยนั้น มีอำนาจการทดสอบ (power, $1 - \beta$)^{2,14} คือสามารถสรุปผลได้อย่างถูกต้องแม่นยำนั่นเอง

ปัจจัยสำคัญที่ส่งผลให้มีความคลาดเคลื่อนที่พบบ่อย ได้แก่ การใช้ขนาดกลุ่มตัวอย่างไม่เหมาะสมทั้งในขนาดที่มากเกินไปนำไปสู่การสรุปผลที่สูงเกินจริง (overpower) หรือต่ำเกินจริง (underpower) สำหรับข้อผิดพลาดในข้อนี้ขอเน้นที่การมีขนาดกลุ่ม

ตัวอย่างน้อยเกินไป (small sample size) นำไปสู่ข้อจำกัดเรื่องความสามารถในการเป็นตัวแทนของประชากร และความคลาดเคลื่อนของการสรุปทางสถิติประเภทที่ 2 (type 2 error, β) ซึ่งหมายถึงโอกาสของการสรุปผิดพลาดโดยการยอมรับสมมติฐานนัล (null hypothesis: H_0) ทั้ง ๆ ที่สมมติฐานนัลไม่เป็นจริง (false negative) เช่น สรุปว่าโปรแกรมการทดลองไม่มีประสิทธิผล ทั้ง ๆ ที่ความเป็นจริงมีประสิทธิผลวิธีการที่ช่วยให้มั่นใจมากขึ้นว่าขนาดกลุ่มตัวอย่างมีจำนวนเพียงพอที่แสดงถึงอำนาจการทดสอบของวิจัยนั้น ๆ คือ การกำหนดขนาดกลุ่มตัวอย่างด้วยการวิเคราะห์อำนาจการทดสอบ (power analysis)¹⁴ ซึ่งต้องใช้ข้อมูลกำหนดในการวิเคราะห์อำนาจ ๆ ได้แก่ (1) ขนาดอิทธิพล (effect size: ES) ซึ่งอาจได้มาจากการศึกษานำร่องของนักวิจัย หรือจากการศึกษาก่อนหน้านี้ที่มีลักษณะใกล้เคียงกับการศึกษาในครั้งนี้ (2) ค่าความคลาดเคลื่อนประเภทที่ 1 (type 1 error, α) ซึ่งหมายถึงโอกาสของการสรุปผิดพลาดโดยการปฏิเสธสมมติฐานนัล (null hypothesis: H_0) ทั้ง ๆ ที่สมมติฐานนัลเป็นจริง (false positive) โดยทั่วไปมักกำหนดค่า α ที่ .05 และ (3) ค่าอำนาจการทดสอบ (power) ที่ .80 โดยมีสูตรในการคำนวณขนาดกลุ่มตัวอย่างที่หลากหลายขึ้นอยู่กับสถิติที่ใช้ในการทดสอบสมมติฐานการวิจัยนั้น ๆ ซึ่งในปัจจุบันมีโปรแกรมสำเร็จรูปเพื่อช่วยในการกำหนดขนาดกลุ่มตัวอย่างโดยใช้หลักการ power analysis นี้โปรแกรมที่เป็นที่นิยมในงานวิจัยทางการแพทย์ในประเทศไทยคือ G*Power ซึ่งมีความครอบคลุมการกำหนดขนาดกลุ่มตัวอย่างสำหรับงานวิจัยที่มีการวิเคราะห์ข้อมูลที่หลากหลาย¹⁵⁻¹⁶ อย่างไรก็ตามแม้ว่านักวิจัยจะระบุว่ามีการใช้โปรแกรมดังกล่าวในการกำหนดขนาดกลุ่มตัวอย่าง พบข้อผิดพลาดที่เกิดขึ้นบ่อย คือ การเลือก

ES เข้าสู่การคำนวณไม่เหมาะสม เช่น กำหนด ES แบบ สะดวกตามความคิดเห็นส่วนตัวของนักวิจัยโดยไม่ ระบุเหตุผล หรือใช้งานวิจัยที่มีความใกล้เคียงน้อยกับ สิ่งที่นักวิจัยสนใจศึกษา หรือเกิดจากความเข้าใจ ผิดพลาด เช่น นักวิจัยต้องการศึกษาอิทธิพลของ ตัวแปร X ต่อตัวแปร Y แต่ไปนำผลการศึกษาที่ศึกษา อิทธิพลของตัวแปร Y ต่อตัวแปร Z มาใช้ ดังนั้นค่าตัว เลขจำนวนกลุ่มตัวอย่าง (N) ที่ได้จึงขาดความน่าเชื่อ ถือ นอกจากนี้ยังพบการเลือก ES มาแทนค่าใน โปรแกรมที่ไม่สอดคล้องกับการวิเคราะห์ข้อมูลที่จะ ใช้ในการศึกษานั้น ๆ ตัวอย่างเช่น ในโปรแกรม G*Power นักวิจัยเลือกสถิติ t-tests เพื่อทดสอบความ แตกต่างของค่าเฉลี่ยระหว่างกลุ่มตัวอย่างสองกลุ่ม [Means: Difference between two independent means (two groups)] แต่คำถาม/สมมติฐานการวิจัยต้องการ ทดสอบความแตกต่างทำนองเดียวกัน แต่มีการวัดซ้ำ 3 ครั้ง ดังนั้น ที่ถูกต้องจะต้องเลือกสถิติเป็น F tests ด้วย ANOVA-Repeated measures within-between interaction ความผิดพลาดในตัวอย่างดังกล่าวข้างต้น ส่งผลให้การวิจัยนั้น ๆ มีอำนาจการทดสอบต่ำ

นอกจากขนาดกลุ่มตัวอย่างที่อาจส่งผลให้เกิด การสรุปข้อค้นพบที่ผิดพลาด ปัญหาที่พบบ่อยอีก ประการคือ การเลือกใช้สถิติในการวิเคราะห์ข้อมูลไม่ เหมาะสมกับลักษณะข้อมูลที่ได้มา เช่น ใช้สถิติอ้างอิง แบบพารามетริกกับชุดข้อมูลที่ไม่เป็นไปตามข้อ ตกลงเบื้องต้นของสถิตินั้น ๆ (เช่น ชุดข้อมูลไม่มีการ แจกแจงแบบปกติ ฯ) หรือการไม่ให้ความสำคัญและ ไม่จัดอิทธิพลของตัวแปรกวนทั้งที่ทราบว่าอาจมี อิทธิพลจากตัวแปรกวนที่ทำให้ผลการวิจัยคลาด เคลื่อนไปจากความเป็นจริง และเมื่อได้ผลการ วิเคราะห์ข้อมูลแล้ว ในการรายงานผลจะต้องมีการ

รายงานข้อค้นพบให้ครบถ้วน ไม่ว่าข้อค้นพบนั้นจะ เป็นไปตามสมมติฐานที่คาดการณ์ไว้แต่ต้นหรือไม่ ก็ตาม

“การวิเคราะห์ข้อมูลสำคัญยิ่ง
ตอบทุกสิ่งทุกคำถามตามมาตรฐาน
ผลสรุปมีอำนาจการทำนาย (power)
ใช้คือใช่ (significant) ไม่ใช่ไม่ (non-
significant) จงบอกมา
กลุ่มตัวอย่างต้องเพียงพอให้ตรวจจับ
(อิทธิพล)
เลือกสถิติขานรับเหมาะสมหนา
ตรวจสอบการใช้งาน(ข้อตกลง)ตามตำรา
เพิ่มคุณค่าผลที่ได้ใกล้ความจริง”

10. ความลำเอียงในการแปลผล รายงานผล และเสนอแนะผลการศึกษาไม่เหมาะสม การแปล ผลการศึกษาหรือการอภิปรายผลการวิจัย เป็นเนื้อหา ส่วนสุดท้ายที่ปรากฏในการรายงานผลการวิจัยที่มี ความสำคัญ เพราะผู้อ่านมักใช้ในการตัดสินใจว่าการ ศึกษาที่น่าเชื่อถือมากน้อยเพียงใด ผู้วิจัยมีความ ลำเอียงในการแปลผลและรายงานผลบ้างหรือไม่ (reporting bias) ปัจจัยสำคัญที่ทำให้นักวิจัยมีความ ลำเอียงในการรายงานผลการศึกษา ได้แก่¹⁷ (1) แรงจูงใจ (motivation) เป็นเหตุผลพื้นฐานของ นักวิจัยที่จะนำเสนอผลการศึกษามาที่ตนสนใจหรือ ต้องการนำเสนอเท่านั้น (2) วิธีการ (means) เกี่ยวข้อง กับการเลือกใช้วิธีการจัดการกับข้อมูลและนำเสนอ ข้อมูล หรือปรับการวิเคราะห์ทางสถิติเพื่อให้ได้ ผลลัพธ์ตามที่ต้องการ (3) ความขัดแย้งและการถ่วง ดุลผลประโยชน์ (conflicts and balancing of interest) นักวิจัยเผชิญกับความขัดแย้งระหว่างมาตรฐาน ทางจริยธรรมและความถูกต้องเชิงวิชาการกับผล ประโยชน์ของตนเองในการรักษาความสัมพันธ์กับ

ผู้มีส่วนได้ส่วนเสีย (เช่น ผู้สนับสนุนทุนวิจัย) และ (4) แรงกดดันจากแวดวงวิชาการและสังคม (pressures from science and society) ให้รายงานผลการศึกษาที่ดูเป็นบวกหรือมีนัยสำคัญเกินกว่าความเป็นจริง ปัจจัยเหล่านี้ผลักดันให้นักวิจัยจำนวนไม่น้อยรายงานผลการวิจัยคลาดเคลื่อน ซึ่งถือว่าคุกคามต่อการพัฒนาศาสตร์และองค์ความรู้ในเรื่อง/สาขานั้น ๆ

ดังนั้น ในขั้นตอนการแปลผลและรายงานผลการวิจัย นักวิจัยควรนำเสนอและอธิบายข้อค้นพบที่ได้ อย่างสมเหตุสมผลตามความเป็นจริง โดยใช้ทักษะทางปัญญาในการคิดวิเคราะห์และประมวลผลเชิงตรรกะที่มีพื้นฐานมาจากทฤษฎี โมเดล หรือกรอบแนวคิดต่าง ๆ ที่นำมาใช้ตั้งแต่ขั้นตอนแรกของการวางแผนและออกแบบการวิจัย หลักฐานสนับสนุนจากการศึกษาอื่น ๆ ก่อนหน้านี้ ตลอดจนการวิเคราะห์และระบุข้อจำกัดหรือข้อผิดพลาดบางประการที่อาจพบในแต่ละขั้นตอนของการดำเนินการวิจัยที่อาจส่งผลต่อคำตอบการวิจัย โดยนักวิจัยระมัดระวังเรื่องการแปลผลที่เกินจริง (overclaim) ซึ่งมักเกิดจากขาดการวิเคราะห์ข้อจำกัดของการวิจัยนั้น ๆ ที่อาจเกิดขึ้นในข้อใดข้อหนึ่งหรือหลาย ๆ ข้อ ดังกล่าวมาแล้วข้างต้น (ข้อ 1-9) และถึงแม้ผลการศึกษาก่อมาจะเป็นไปตามสมมติฐาน ก็ควรพิจารณาด้วยว่าการได้ผลการวิเคราะห์ว่ามีนัยสำคัญทางสถิติ (statistical significance) แล้วนั้น มีนัยสำคัญทางคลินิกหรือการปฏิบัติ (clinical or practical significance) ด้วยหรือไม่ เนื่องจากการมีนัยสำคัญทางสถิติอาจเป็นได้จากการมีกลุ่มตัวอย่างขนาดใหญ่มาก¹⁸ เช่น การศึกษาความสัมพันธ์ระหว่างตัวแปร X และ Y กับกลุ่มตัวอย่างขนาดใหญ่ เช่น 500 คน พบค่าสัมประสิทธิ์สหสัมพันธ์ต่ำมากและมีนัยสำคัญทางสถิติ (เช่น $r_{xy} = .103, p < .05$) ในทางกลับกันหากข้อค้นพบไม่เป็น

ไปตามสมมติฐาน เช่น ไม่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ ไม่ควรรีบด่วนสรุปว่าการทดลองนั้นไม่มีผล แต่อาจพิจารณาค่าตัวเลขที่บ่งชี้ความแม่นยำ (precision) โดยดูจากค่าขนาดอิทธิพล ES และค่าช่วงเชื่อมั่น (confidence interval) ประกอบด้วยว่าเป็นอย่างไร เพราะอาจช่วยบอกถึงความสำคัญในทางคลินิกได้¹⁹ ในขณะเดียวกันต้องตรวจสอบว่ามีข้อผิดพลาดเกิดขึ้นในขั้นตอนใดของการวิจัย ควบคู่กันไปเสมอ

“ขออย่าใช้อคตินำมาก่อน

อย่าถอดถอนข้อไม่ซิก (not significant)

ทำตกหาย

ความจริงใจต้องแสดงทั้งใจกาย

ไม่วุ่นวายหาทางแก้ให้มัน (นัยสำคัญ)

ช่วยตรวจสอบเมื่อพบเห็นความแตกต่าง

มีนัยยะ (ทางคลินิก) อยู่บ้างใช้หรือไม่

ของที่พบเหมือนหรือต่างเพื่อนอย่างไร

ทฤษฎีนั้นไซ้ใช้ชี้ทาง”

การอภิปรายผลการวิจัยอย่างโปร่งใส ตรงไปตรงมา จะนำไปสู่การสรุปผลการวิจัยที่มีน้ำหนักน่าเชื่อถือ ไม่ว่าผลการศึกษาก่อมาอย่างไร เนื่องจากไม่มีงานวิจัยชิ้นใดสมบูรณ์แบบ ทุกครั้งที่นักวิจัยพยายามขจัดความลำเอียงอย่างหนึ่ง ก็อาจส่งผลกระทบเกิดความลำเอียงในรูปแบบอื่นตามมา ตัวอย่างเช่น นักวิจัยออกแบบการจัดกลุ่มตัวอย่างเข้ากลุ่มโดยการศึกษาในกลุ่มควบคุมให้เสร็จก่อนแล้วจึงทำการศึกษาในกลุ่มทดลองเพื่อป้องกันการปนเปื้อนจากการแลกเปลี่ยนข้อมูลที่เป็นส่วนหนึ่งของการทดลองไปยังอีกกลุ่มหนึ่ง (ลดปัญหา performance bias) แทนการศึกษาโดยการสุ่มกลุ่มตัวอย่างเข้ากลุ่มและศึกษากลุ่มทดลองและกลุ่มควบคุมคู่ขนานกันไป (parallel, RCT) ซึ่งสามารถป้องกัน selection bias

ได้แต่อาจหลีกเลี่ยง performance bias จากการแลกเปลี่ยนข้อมูลระหว่างกลุ่มไม่ได้ เป็นต้น ดังนั้น นักวิจัยจึงจำเป็นต้องรายงานเหตุผลความจำเป็น และวิเคราะห์ข้อจำกัดต่าง ๆ และนำไปสู่การให้ข้อเสนอแนะการนำผลการวิจัยไปใช้ตลอดจนการศึกษาวิจัยในอนาคตที่สอดคล้องกับวิธีการดำเนินการวิจัยที่ใช้และข้อค้นพบจากการศึกษานั้น ๆ

“บทสรุปส่งท้ายมีความหมาย
หมวดความรู้ไว้เพื่อสานต่อ
ให้เดินตามผลวิจัยที่ดีพอ
อย่าปล่อยเกินจริงจนเกินไป
ไม่มีงานชิ้นไหนสุดประเสริฐ
พบบกพร่องก่อเกิดเรียนรู้ใหม่
จึงต้องบอกข้อจำกัดอย่างจริงจัง
แนะนำวิจัยหน้าพากันเดิน (ต่อไป)”

สรุปและข้อเสนอแนะ

การส่งเสริมให้การวิจัยทางการแพทย์เชิงปริมาณมีคุณภาพสูงขึ้น นักวิจัยควรตระหนักถึงข้อผิดพลาดหรือความลำเอียงที่อาจเกิดขึ้นในทุกขั้นตอนของการวิจัย และใช้แนวทางที่นำเสนอในบทความนี้ในการตรวจสอบและลดข้อผิดพลาดต่าง ๆ ให้ได้มากที่สุด เพื่อให้ผลการวิจัยมีความถูกต้องเชื่อถือได้ ผู้อ่านสามารถนำข้อค้นพบจากการวิจัยไปใช้ประโยชน์ได้จริง อันจะส่งผลต่อการพัฒนาวิชาชีพการพยาบาล และการดูแลผู้ป่วยอย่างแท้จริง

เอกสารอ้างอิง

1. Wannasuntad S, Petpichetchian W, Vachprasit R, et al. A scoping review of Thai nursing research over the past 3 years. JTNMC. 2022; 37(2): 143–66. Thai.
2. Polit DF, Beck CT. Nursing research: Generating and assessing evidence for nursing practice 11th ed. Philadelphia, PA: Wolters Kluwer; 2021.
3. Petpichetchian W, Hanucharuenkul S, Turale S. Optimizing quality nursing research: A call for action. PRIJNR. 2022; 26(4): 549–54. Thai.
4. Hines S, Ramsbotham J, Coyer F. Registered nurses' experiences of reading and using research for work and education: a qualitative research study. BMC Nursing. 2022; 21(1): 114. doi: 10.1186/s12912-022-00877-3.
5. Chen Q, Sun M, Tang S, et al. Research capacity in nursing: a concept analysis based on a scoping review. BMJ Open. 2019; 9(11): e032356. doi: 10.1136/bmjopen-2019-032356.
6. Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. Ann Intern Med. 2010; 152(11): 726–32. doi: 10.1186/1741-7015-8-18.
7. Von Elm E, Altman DG, Egger M, et al. The Strengthening and Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for reporting observational studies. Ann Intern Med. 2007; 147(8): 573–7. doi: 10.7326/0003-4819-147-8-200710160-00010.

8. Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental designs for generalized causal inference. NY: Houghton Mifflin; 2002.
9. Ellenberg JH. Selection bias in observational and experimental studies. *Stat Med*. 1994; 13(557-67): 557-67.
10. D'Onofrio BM, Sjölander A, Lahey BB, et al. Accounting for confounding in observational studies. *Annu Rev Clin Psychol*. 2020; 16(1): 25-48. doi: 10.1146/annurev-clinpsy-032816045030.
11. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: Recommendations for specifying and reporting. *Implement Sci*. 2013; 8: 1-11. doi: 10.1186/1748-5908-8-139.
12. Nunnally JC, Bernstein I. Psychometric theory. 3rd ed. NY: McGraw-Hill; 1994.
13. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2): 155-63. doi: 10.1016/j.jcm.2016.02.012.
14. Cohen J. Statistical power analysis for the behavioral science 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
15. Faul F, Erdfelder E, Buchner A, et al. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods*. 2009; 41(4): 1149-60. doi:10.3758/BRM.41.4.1149.
16. Kang H. Sample size determination and power analysis using the G* Power software. *JEEHP*. 2021; 18. doi: <https://doi.org/10.3352/jeehp.2021.18.17>.
17. van der Steen J, ter Riet G, van den Bogert C, et al. Causes of reporting bias: A theoretical framework. *F1000Res*. 2019; 8: 280. doi: 10.12688/f1000research.18310.2.
18. Polit DF. Clinical significance in nursing research: a discussion and descriptive analysis. *Int J Nurs Stud*. 2017; 73(2017): 17-23. doi: 10.1016/j.ijnurstu.2017.05.002.
19. Hemming K, Javid I, Taljaard M. A review of high-impact journals found that misinterpretation of non-statistically significant results from randomized trials was common. *J Clin Epidemiol*. 2022; 145: 112-20. doi: 10.1016/j.jclinepi.2022.01.014.