

บทนำ

1

- 1.1 ความนำ
- 1.2 แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม
- 1.3 แนวทางการประเมินผลกระทบจากสิ่งแทรกแซง
- 1.4 การตัดสินใจเลือกใช้วิธีการทางสถิติในการวิเคราะห์ข้อมูล
- 1.5 บทสรุป
- 1.6 เอกสารอ้างอิง

1.1 ความนำ

ในงานวิจัยทางวิทยาศาสตร์สุขภาพ ตัวแปรผลลัพธ์ที่มีลักษณะแบบต่อเนื่อง เช่น คะแนนพฤติกรรม ระดับความดันเลือด systolic หรือ ปริมาณกรดยูริกในร่างกาย เป็นต้น ถูกนำมาใช้ค่อนข้างแพร่หลาย โดยเฉพาะการตอบคำถามวิจัยเชิงเปรียบเทียบ ภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม ครอบคลุมทั้งงานวิจัยที่มีการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่ม (randomized trial) และไม่สุ่ม (non-randomized trial) เพื่อศึกษาขนาดความแตกต่างที่เกิดขึ้นระหว่างกลุ่มในรูปแบบ “ผลกระทบของวิธีการ (treatment effect)” และสรุปผลเพื่อตอบคำถามวิจัยที่มีความเป็นไปได้ ทั้งในประเด็น “ประสิทธิภาพ (efficacy)” ซึ่งเป็นผลกระทบทางชีวภาพที่เกิดขึ้นจริงจากสิ่งแทรกแซง (intervention) ที่นักวิจัยจัดกระทำขึ้น หรือ ในประเด็น “ประสิทธิผล (effectiveness)” ซึ่งเป็นผลกระทบจากการปฏิบัติอย่างเป็นประจำด้วยสิ่งแทรกแซง (*Estrada, Arancibia et al., 2020*) เพื่อนำไปใช้เป็นหลักฐานเชิงประจักษ์ (evidence-based practice, EBP) ประกอบการตัดสินใจเชิงนโยบาย หรือ การเป็นทางเลือกที่เหมาะสมในการดูแลรักษาทางคลินิกและการส่งเสริมสุขภาพ (*Fox, Brathwaite et al., 2004*) อาทิ งานวิจัยเพื่อประเมินผลการเปลี่ยนแปลงคุณภาพชีวิตด้านสุขภาพ (HRQoL) ของผู้ป่วยโรคข้อเข่าเสื่อมก่อนและหลังการเข้าร่วมโครงการ “WellNet” ในประเทศออสเตรเลีย (*John, Tannous et al., 2020*) การศึกษาประสิทธิผลของโปรแกรมทางจิตวิทยา เพื่อลดความทุกข์ใจทางอารมณ์ของผู้ป่วยมะเร็งเต้านมหลังผ่าตัดในประเทศสหรัฐอเมริกา (*Andersen, Farrar et al., 2007*) หรือ การประเมินผลกระทบโปรแกรมการออกกำลังกายกลางแจ้งเพื่อสุขภาพที่ดีในกลุ่มผู้สูงอายุชาวจีนในประเทศฮ่องกง (*Lai, Ou et al., 2022*) เป็นต้น จากแนวทางการศึกษาดังกล่าว ผลกระทบของวิธีการ หรือ treatment effect ถือเป็นข้อค้นพบสำคัญ ซึ่งจำเป็นต้องอาศัยวิธีการทางสถิติที่เหมาะสมมาใช้ในการวิเคราะห์ข้อมูล เพื่อให้ได้ค่าประมาณที่มีความแม่นยำและไม่เอนเอียง รวมถึงการมีอำนาจการทดสอบทางสถิติที่มากพอ ดังนั้นในบทนี้ จึงได้กล่าวถึงองค์ความรู้พื้นฐานที่เกี่ยวข้องเพื่อให้นักวิจัยได้เข้าใจแนวคิดและหลักการประเมินผลกระทบจากสิ่งแทรกแซง ประกอบด้วย แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม แนวทางการประเมินผลกระทบจากสิ่งแทรกแซง การตัดสินใจเลือกใช้วิธีการทางสถิติที่เหมาะสมและบทสรุป

1.2 แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม

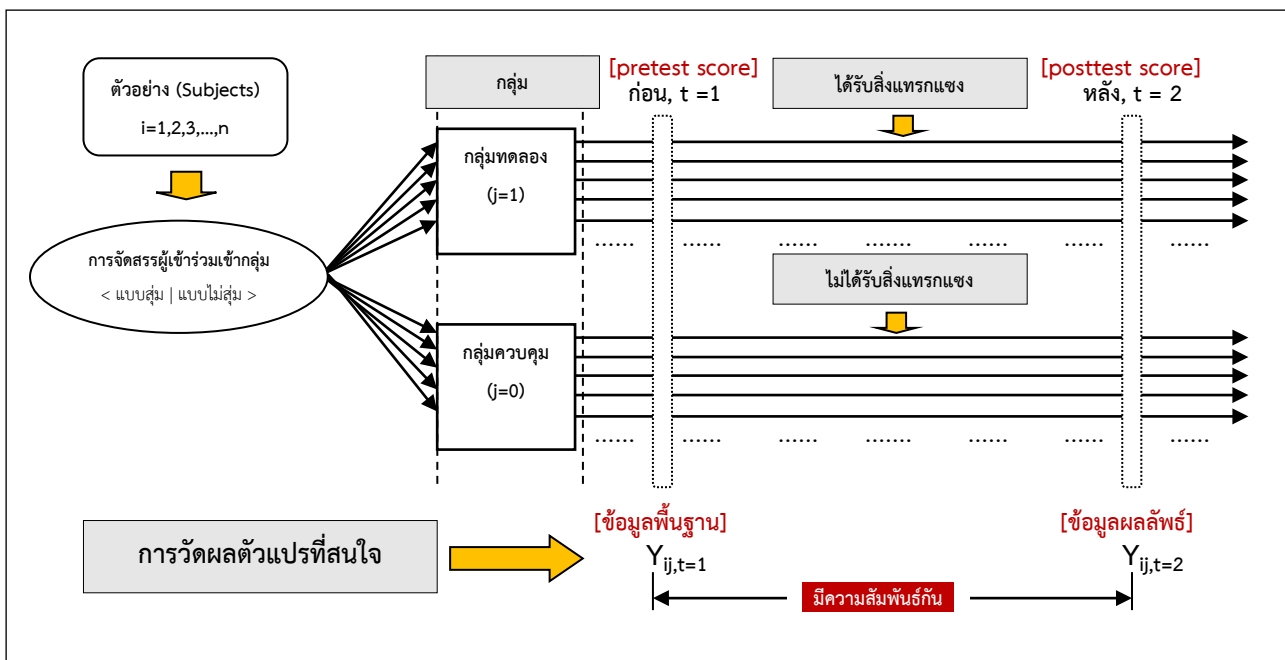
แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม ถูกนำมาใช้อย่างแพร่หลายในงานวิจัยทางวิทยาศาสตร์สุขภาพ (*Gliner, Morgan et al., 2003*) เพื่อตอบคำถามวิจัยเชิงเปรียบเทียบเกี่ยวกับการประเมินผลกระทบของวิธีการ โปรแกรม หลักสูตร หรือ สิ่งแทรกแซงอื่น ในกรณีตัวแปรผลลัพธ์ที่มีลักษณะแบบต่อเนื่อง ครอบคลุมทั้งงานวิจัยเชิงทดลองที่มีการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่มและไม่สุ่ม โดยเฉพาะงานวิจัยกึ่งทดลอง หรือ quasi-experimental study ขณะเดียวกันในการวิเคราะห์ข้อมูล เพื่อตอบคำถามวิจัยภายใต้แบบแผนดังกล่าว นักวิจัยไม่เพียงแต่ต้องทำความเข้าใจแนวทางและขั้นตอนของวิธีการทางสถิติที่นำมาใช้เท่านั้น หากแต่ยังต้องทำความเข้าใจแนวคิดและหลักการพื้นฐานที่สำคัญและเกี่ยวข้องภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูล

ก่อน-หลังแบบวัดซ้ำสองกลุ่ม ทั้งนี้เพื่อนำแนวคิดและหลักการดังกล่าว มาใช้อธิบายและยืนยันแนวปฏิบัติตามขั้นตอนของการวิเคราะห์ข้อมูลเกี่ยวกับการลดความลำเอียงและความผันแปรต่างๆ ที่อาจจะเกิดขึ้นและส่งผลกระทบต่อผลลัพธ์ของการวิเคราะห์ข้อมูลให้มีความถูกต้องและครอบคลุม โดยมีประเด็นที่เกี่ยวข้อง ดังนี้

- 1.2.1 ลักษณะทั่วไปของแบบแผนงานวิจัย
- 1.2.2 แนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่ม
- 1.2.3 ความตรง (validity) ของแบบแผนงานวิจัย
- 1.2.4 การพิจารณาข้อมูลก่อนและหลังทดลอง

1.2.1 ลักษณะทั่วไปของแบบแผนงานวิจัย

แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม เป็นรูปแบบงานวิจัยที่ประกอบด้วยกลุ่มศึกษา 2 กลุ่ม ได้แก่ **กลุ่มทดลอง** หรือ “treatment group” ซึ่งถือเป็นกลุ่มที่ผู้เข้าร่วมได้รับสิ่งแทรกแซงที่นักวิจัยต้องการศึกษาและ **กลุ่มควบคุม** หรือ “control group” ซึ่งเป็นกลุ่มที่ผู้เข้าร่วมไม่ได้รับสิ่งแทรกแซงที่ศึกษา แต่อาจได้รับสิ่งที่เป็แนวปฏิบัติเดิม หรือ สิ่งที่เทียบเคียงกับสิ่งแทรกแซงที่นำมาศึกษาแทนได้ เช่น ยาหลอก (placebo) เป็นต้น ขณะเดียวกัน การวัดผลตัวแปรแบบต่อเนื่องที่สนใจ จะถูกพิจารณาแบ่งการวัดออกเป็น 2 ครั้ง ตามลำดับเวลา ทั้งกลุ่มทดลองและกลุ่มควบคุม ประกอบด้วย **การวัดครั้งที่ 1** ซึ่งถือเป็นการวัด ณ จุดเริ่มต้นก่อนมีการจัดกระทำสิ่งแทรกแซง เรียกว่า ข้อมูลก่อนทดลอง หรือ คะแนนก่อนทดลอง (pretest score) และถือเป็นข้อมูลพื้นฐาน (baseline) และ **การวัดครั้งที่ 2** เป็นการวัด ณ จุดสุดท้ายหลังจากที่มีการจัดกระทำสิ่งแทรกแซงเสร็จสิ้น เรียกว่า ข้อมูลหลังทดลอง หรือ คะแนนหลังทดลอง (posttest score) ซึ่งในทางปฏิบัติถือเป็นข้อมูลผลลัพธ์ (outcome) ดังแผนภาพที่ 1.1



แผนภาพที่ 1.1 แสดงลักษณะทั่วไปของแบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม

จากแบบแผนงานวิจัยเชิงทดลองดังแผนภาพที่ 1.1 ข้างต้น พบว่า ในภาพรวมภายใต้แบบแผนดังกล่าว ประกอบด้วย กลุ่มศึกษา 2 กลุ่ม ได้แก่ กลุ่มทดลอง (ได้รับสิ่งแทรกแซง) และกลุ่มควบคุม (ไม่ได้รับสิ่งแทรกแซง) โดยมีการวัดผลตัวแปรแบบต่อเนื่องที่สนใจ จำนวนสองครั้งตามลำดับเวลา ได้แก่ ก่อนและหลังการทดลองเช่นเดียวกัน ทั้งสองกลุ่ม ซึ่งจากลักษณะดังกล่าว จึงทำให้มีประเด็นสำคัญที่นักวิจัยควรคำนึงถึง เนื่องจากส่งผลกระทบต่อ การกำหนดแนวทางการตัดสินใจเลือกใช้วิธีการทางสถิติและแนวทางการวิเคราะห์ข้อมูลที่แตกต่างกัน โดยเฉพาะใน ประเด็นแนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่ม ประเด็นความตรง (validity) ของแบบแผนงานวิจัยและประเด็น การพิจารณาข้อมูลก่อนและหลังทดลอง ซึ่งรายละเอียดแต่ละประเด็นได้มีการกล่าวถึงในหัวข้อถัดไป

1.2.2 แนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่ม

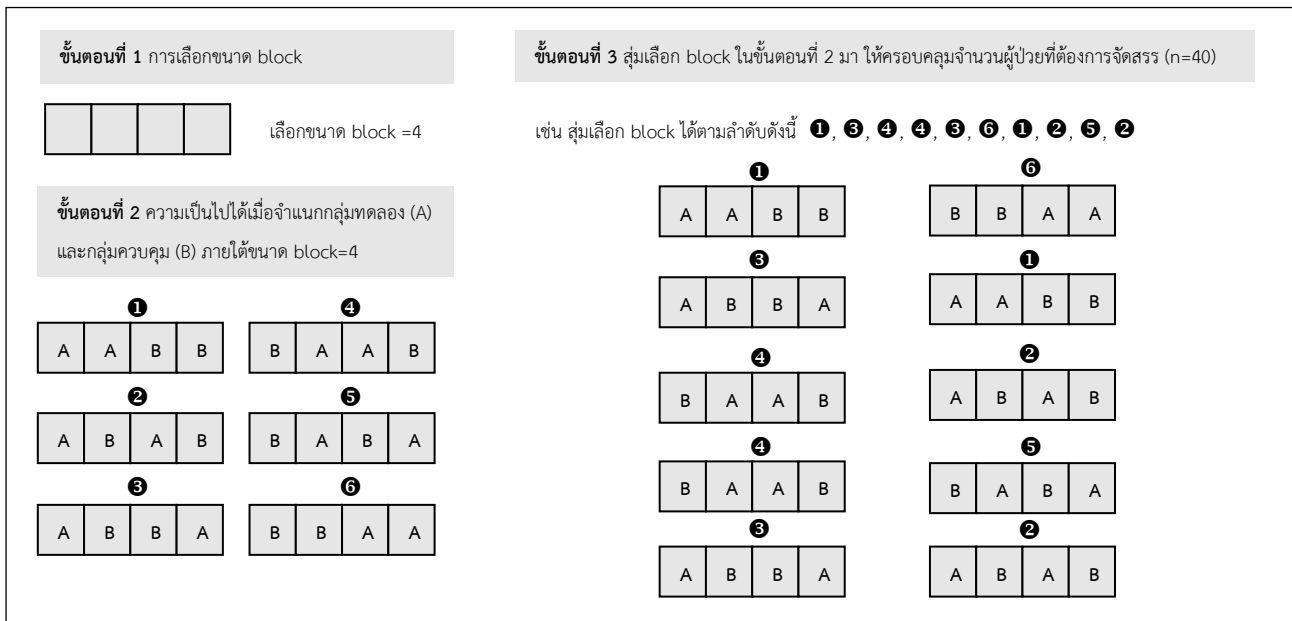
ในงานวิจัยทางวิทยาศาสตร์สุขภาพ โดยเฉพาะภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูล ก่อน-หลังแบบวัดซ้ำสองกลุ่ม แนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่ม ถือเป็นองค์ประกอบสำคัญที่ส่งผลโดยตรงต่อ ความลำเอียงที่เกิดขึ้นจากการเลือก (selection bias) และความกลมกลืนของตัวแปรที่สนใจและตัวแปรแทรกซ้อน ทั้งที่ทราบและไม่ทราบระหว่างสองกลุ่ม รวมถึงประเด็นการเลือกใช้วิธีการทางสถิติและแนวทางการวิเคราะห์ข้อมูล ให้มีความสอดคล้องและเหมาะสม (Berger, Bour et al., 2021) โดยจากแผนภาพที่ 1.1 พบว่า เมื่อผู้เข้าร่วม หรือ กลุ่มตัวอย่างที่ถูกพิจารณาโดยผ่านเกณฑ์คัดเข้า (inclusion criteria) และเกณฑ์คัดออก (exclusion criteria) ตามแนวทาง หรือ protocol ที่นักวิจัยได้กำหนดไว้เสร็จสิ้นแล้ว กลุ่มตัวอย่างดังกล่าว จะถูกนำมาพิจารณาเพื่อเข้าสู่ แนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่ม โดยในทางปฏิบัติ ประกอบด้วย 2 แนวทาง ดังนี้

- (1) แนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่ม (randomization)
- (2) แนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบไม่สุ่ม (non-randomization)

(1) **แนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่ม** หรือ randomization ถือเป็นวิธีการจัดสรร ผู้เข้าร่วมเข้ากลุ่ม ภายใต้แนวคิดการทำให้ปัจจัย หรือ ตัวแปรทั้งที่ทราบและไม่ทราบ ซึ่งมีอยู่และส่งผลกระทบต่อ ตัวแปรผลลัพธ์ที่สนใจ มีการกระจายอย่างสม่ำเสมอและใกล้เคียงกันทั้งกลุ่มทดลองและกลุ่มควบคุม นั่นคือ เมื่อนักวิจัย มีการนำแนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่มมาใช้ ตามกฎจำนวนมาก (law of large number) จะพบว่า ค่าเฉลี่ยของค่าต่างๆ ภายใต้คุณลักษณะของผู้เข้าร่วมทั้งสองกลุ่ม จะมีแนวโน้มเท่าเทียมกัน ดังนั้นเมื่อการศึกษา ได้มีการมุ่งเป้าไปที่การประเมินขนาดผลกระทบของกลุ่มที่แตกต่างกันจากสิ่งแทรกแซงที่สนใจ หรือ treatment effect จึงทำให้การศึกษานั้น สามารถสะท้อนผลลัพธ์ของความแตกต่างได้อย่างชัดเจนมากขึ้น ด้วยการลดโอกาส ผลกระทบที่มาจากคุณลักษณะของผู้เข้าร่วมที่แตกต่างกัน (Berger, Bour et al., 2021) ซึ่งโดยทั่วไป วิธีการ แบบ randomization มีได้หลายวิธีและถูกเลือกนำมาใช้บนพื้นฐานหลักคิดและวิธีการที่แตกต่างกัน เช่น วิธี simple randomization, วิธี block randomization หรือ วิธี stratified randomization เป็นต้น (Lim and In, 2019) ในทางปฏิบัติ นักวิจัยสามารถพบแนวทางดังกล่าวนี้ได้ในงานวิจัยทางคลินิก โดยเฉพาะงานวิจัยเชิงทดลองแบบสุ่มที่มี กลุ่มควบคุม หรือ randomized controlled trial (RCT) โดยสามารถแสดงตัวอย่างแนวทางการจัดสรรผู้เข้าร่วม เข้ากลุ่มแบบสุ่ม ด้วยวิธีการแบบ block randomization ได้ดังเช่น กรณีตัวอย่างที่ 1.1

กรณีตัวอย่างที่ 1.1

นักวิจัยท่านหนึ่ง ต้องการศึกษาระสิทธิภาพของยาลดอาการเจ็บปวดชนิดหนึ่ง ในกลุ่มผู้ป่วยโรคมะเร็งปอด จำนวน 40 ราย โดยกำหนดให้ผู้ป่วยได้รับยาลดอาการเจ็บปวดที่นำมาศึกษา หรือ “กลุ่มทดลอง” จำนวน 20 ราย และผู้ป่วยได้รับยาหลอก (placebo) หรือ “กลุ่มควบคุม” จำนวน 20 ราย และนักวิจัยได้เลือกใช้วิธีการสุ่มผู้ป่วยเข้ากลุ่มด้วยวิธีการ block randomization โดยมีแนวทางแสดงได้ ดังแผนภาพที่ 1.2



แผนภาพที่ 1.2 แสดงการสุ่มด้วยวิธี block randomization กรณีงานวิจัยเชิงทดลองแบบ RCT

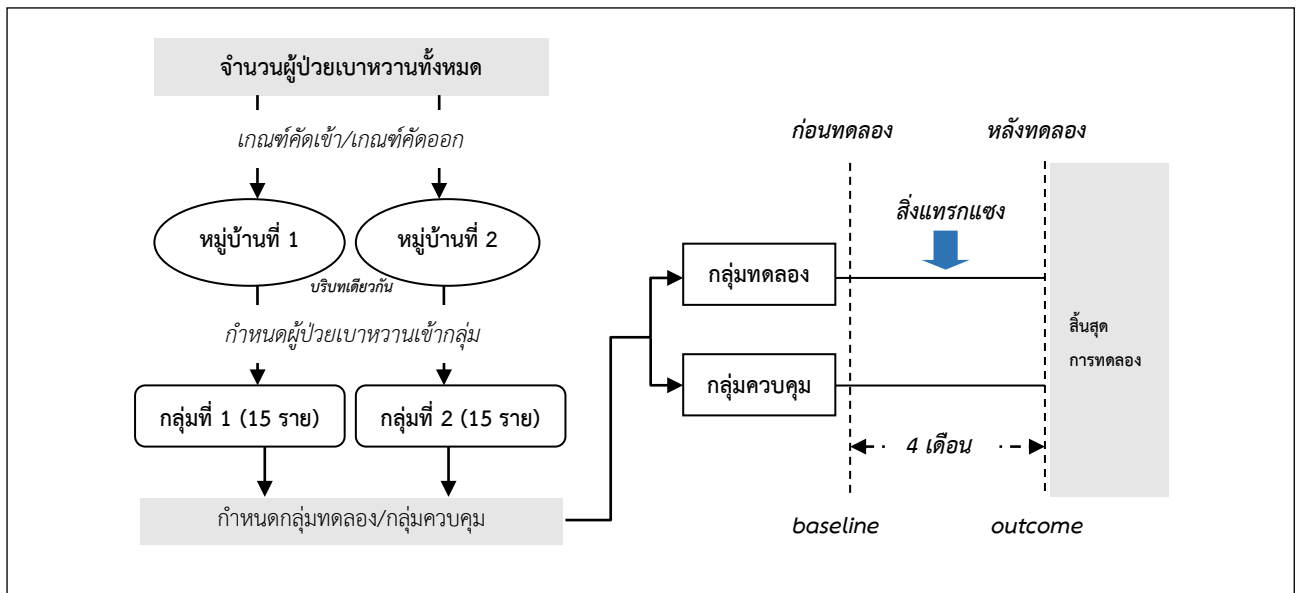
แผนภาพนี้ถูกปรับปรุงมาจากต้นฉบับตามแนวคิดของ Minsoo Kang (Kang, Raşan et al., 2008)

จากแผนภาพที่ 1.2 พบว่า นักวิจัยเลือกใช้วิธีการสุ่มแบบ block randomization โดยเริ่มต้นจาก **ขั้นตอนที่ 1** การเลือกขนาด block เท่ากับ 4 เพื่อรองรับกลุ่มทดลอง แทนด้วย “A” และกลุ่มควบคุม แทนด้วย “B” จากนั้นจึงเข้าสู่ **ขั้นตอนที่ 2** ซึ่งเป็นการพิจารณาความเป็นไปได้ทั้งหมดของการจัดเรียงสัญลักษณ์ “A” และ “B” ภายใต้อาณาเขตขนาด block = 4 โดยมีจำนวนที่เป็นไปได้ทั้งหมดเท่ากับ 6 วิธี ได้แก่ AABB, ABAB, ABBA, BAAB, BABA และ BBAA จากนั้นจึงเข้าสู่ **ขั้นตอนที่ 3** ซึ่งเป็นการจัดสรรผู้ป่วยจำนวน 40 ราย เข้ากลุ่มตามลำดับของ block ที่ถูกสุ่มเลือกมาได้ เช่น จากแผนภาพข้างต้น ผู้ป่วยลำดับที่ 1, 2, 3, 4 จะถูกจัดสรรเข้ากลุ่มตามลำดับของ block 1 นั่นคือ ผู้ป่วยลำดับที่ 1 และ 2 จะถูกจัดสรรให้เข้ากลุ่มทดลอง (“A”) ขณะที่ผู้ป่วยลำดับที่ 3 และ 4 จะถูกจัดสรรให้เข้ากลุ่มควบคุม (“B”) ต่อมา ผู้ป่วยลำดับที่ 5, 6, 7, 8 จะถูกจัดสรรเข้ากลุ่มตามลำดับของ block 6 นั่นคือ ผู้ป่วยลำดับที่ 5 และ 8 จะถูกจัดสรรให้เข้ากลุ่มทดลอง (“A”) ขณะที่ผู้ป่วยลำดับที่ 6 และ 7 จะถูกจัดสรรให้เข้ากลุ่มควบคุม (“B”) และผู้ป่วยรายอื่นๆ ต่อมาก็สามารถดำเนินการจัดสรรได้ด้วยวิธีการเดิม ตามลำดับ block ที่ถูกสุ่มมา เป็นต้น

(2) แนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบไม่สุ่ม หรือ non-randomization ถือเป็นวิธีการที่ถูกนำมาใช้ในกรณีงานวิจัยที่ไม่สามารถดำเนินการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่มได้ ทั้งนี้อาจเนื่องมาจากข้อจำกัดเกี่ยวกับความเป็นไปได้ในดำเนินการแบบสุ่ม หรือ เงื่อนไขทางจริยธรรม ดังเช่น งานวิจัยกึ่งทดลอง หรือ quasi-experimental study เป็นต้น ดังนั้นวิธีการที่ถูกนำมาใช้ภายใต้แนวทางนี้ จึงมีความเป็นไปได้ค่อนข้างหลากหลายรูปแบบ แต่ไม่เข้าข่ายและถือเป็นวิธีการสุ่มอย่างมีหลักเกณฑ์ดังเช่นวิธีการแบบ randomization เช่น การใช้วิธีการแบบจับคู่ (matching) ตามตัวแปร เช่น อาการทางคลินิก ลักษณะทางประชากร หรือ พื้นที่ของผู้เข้าร่วม ภายใต้บริบทที่นักวิจัยสนใจศึกษาและมุ่งเน้นการนำไปใช้ประโยชน์ในทางปฏิบัติ เป็นต้น (Miller, Smith et al., 2020) โดยสามารถแสดงตัวอย่างแนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบไม่สุ่มในงานวิจัยกึ่งทดลองได้ ดังเช่นกรณีตัวอย่างที่ 1.2

กรณีตัวอย่างที่ 1.2

นักวิจัยท่านหนึ่ง ต้องการศึกษาประสิทธิผลของโปรแกรมการปรับเปลี่ยนพฤติกรรมควบคุมระดับน้ำตาลในเลือดของผู้ป่วยเบาหวาน โดยมีขนาดตัวอย่างเท่ากับ 30 ราย (15 รายต่อกลุ่ม) ซึ่งกำหนดจากผู้ป่วยเบาหวานที่ผ่านเกณฑ์คัดเข้าและเกณฑ์คัดออก จากหมู่บ้าน 2 หมู่บ้านที่มีบริบทใกล้เคียงกัน หมู่บ้านละ 15 ราย จากนั้นจึงกำหนดให้ “กลุ่มทดลอง” เป็นกลุ่มที่ผู้เข้าร่วมได้รับโปรแกรมการปรับเปลี่ยนพฤติกรรมฯ และ “กลุ่มควบคุม” เป็นกลุ่มที่ไม่ได้รับโปรแกรมการปรับเปลี่ยนพฤติกรรมฯ แต่ได้รับแนวทางการปฏิบัติเช่นเดิมตามปกติ และกำหนดให้ระดับน้ำตาลในเลือดสะสม หรือ ค่า HbA1c เป็นตัวแปรผลลัพธ์ที่มีลักษณะแบบต่อเนื่อง ซึ่งผู้ป่วยเบาหวานทั้งสองกลุ่ม จะต้องถูกวัดจำนวน 2 ครั้งตามลำดับเวลา ได้แก่ ครั้งที่ 1 ก่อนเริ่มโปรแกรม และครั้งที่ 2 ในเดือนที่ 4 หลังจากเดือนที่เริ่มโปรแกรมการทดลอง ดังแผนภาพที่ 1.3



แผนภาพที่ 1.3 แสดงวิธีการจัดสรรคนเข้ากลุ่มแบบไม่สุ่ม กรณีงานวิจัยกึ่งทดลอง

จากแผนภาพที่ 1.3 ข้างต้น พบว่า นักวิจัยเลือกใช้วิธีการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบไม่สุ่มด้วยการพิจารณาเลือกตามพื้นที่หมู่บ้านที่มีบริบทใกล้เคียงกัน โดยมีการกำหนดเกณฑ์คัดเข้า/คัดออกและจัดสรรผู้เข้าร่วมเข้ากลุ่มโดยวิธีการจับคู่ตามลักษณะทางประชากร ได้แก่ เพศ อายุและระยะเวลาที่ภูมิลำเนาเดิมว่า เป็นโรคเบาหวาน จากนั้นจึงทำการกำหนดกลุ่ม โดยจำแนกเป็นกลุ่มทดลอง ได้แก่ กลุ่มที่ได้รับโปรแกรมและกลุ่มควบคุม ได้แก่ กลุ่มที่ไม่ได้รับโปรแกรม แต่มีการนำแนวปฏิบัติเดิมมาใช้แทน จากนั้นจึงเริ่มเข้าสู่กระบวนการทดลอง ด้วยการวัดข้อมูลก่อนทดลองทั้งสองกลุ่มและดำเนินการทดลองตามโปรแกรม จนถึงเดือนที่ 4 จึงทำการประเมินผลลัพธ์สุดท้ายของข้อมูลหลังการทดลองทั้งสองกลุ่มอีกครั้ง ซึ่งถือเป็นการเสร็จสิ้นการทดลองและมีการดำเนินการจัดเตรียมนำข้อมูลที่ได้ไปสู่ขั้นตอนตามแนวทางการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติที่เลือกนำมาใช้ต่อไป

1.2.3 ความตรง (validity) ของแบบแผนงานวิจัย

ในงานวิจัยทางวิทยาศาสตร์สุขภาพ เกี่ยวข้องโดยตรงกับการพัฒนาข้อมูล ซึ่งเป็นหลักฐานเชิงประจักษ์ที่ได้มาจากสิ่งแทรกแซง (evidence-based intervention) เพื่อนำมาใช้ปรับปรุงคุณภาพของระบบบริการในการดูแลรักษา การส่งเสริมและป้องกันสุขภาพ ครอบคลุมทั้งในระดับบุคคลและชุมชน โดยสิ่งแทรกแซงดังกล่าวส่วนใหญ่เป็นหนึ่งใน “7 Ps” ประกอบด้วย Programs (โครงการ) Practices (วิธีการปฏิบัติ) Principles (หลักปฏิบัติ) Procedures (ขั้นตอน) Products (ผลิตภัณฑ์) Pills (ยาเม็ด) และ Policies (นโยบาย) (Handley, Lyles et al., 2018) ดังนั้นผลลัพธ์จากงานวิจัยที่มีความถูกต้องและเชื่อถือได้ จึงเป็นประเด็นสำคัญที่นักวิจัยจำเป็นต้องคำนึงถึง เพราะส่งผลโดยตรงต่อการนำผลงานวิจัยที่ได้ไปใช้ประโยชน์ โดยทั่วไป ในการประเมินคุณภาพผลสรุปที่ได้จากงานวิจัยสามารถพิจารณาได้จาก **ความตรง หรือ validity** ซึ่งหมายถึง **“ผลลัพธ์ที่ได้จากการศึกษาของกลุ่มตัวอย่างนั้นเป็นตัวแทนข้อค้นพบจริง เช่นเดียวกับการศึกษาในกลุ่มตัวอย่างเดียวกันในที่อื่น ได้มากน้อยเพียงใด?”** (Patino and Ferreira, 2018) ซึ่งในทางปฏิบัติ สามารถแบ่งการพิจารณาออกได้เป็น 2 ลักษณะ ดังนี้

- (1) ความตรงภายใน (internal validity)
- (2) ความตรงภายนอก (external validity)

(1) **ความตรงภายใน** หรือ internal validity เป็นการบ่งชี้ข้อสรุปที่ได้จากงานวิจัยว่า “ผลลัพธ์จากงานวิจัยนั้น สามารถเป็นตัวแทนข้อค้นพบจริงของประชากรที่ศึกษาได้มากน้อยเพียงใด” หรือ อาจกล่าวอีกนัยหนึ่งได้ว่า “ผลการศึกษานั้น สามารถตอบคำถามวิจัย ได้หรือไม่?” โดยเฉพาะเมื่อความตรงภายในดังกล่าว ถูกนำมาใช้ภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม สามารถสื่อความหมายได้ว่า

“ความแตกต่างที่พบและเกิดขึ้นระหว่างกลุ่ม หรือ treatment effect ที่ได้ เป็นผลกระทบบางอย่างจากสิ่งแทรกแซงที่นักวิจัยนำมาศึกษาอย่างถูกต้องแล้ว หรือไม่ ?”

นั่นคือ ความตรงภายใน ถือเป็นประเด็นที่สะท้อนให้เห็นว่า แบบแผนงานวิจัยเชิงทดลองที่นักวิจัยนำมาใช้ รวมถึงแนวทางการดำเนินงานตามขั้นตอนต่างๆ ของการทดลองนั้น สามารถการันตีวิธีการในการควบคุมและการลดความคลาดเคลื่อนที่เกิดขึ้นได้อย่างถูกต้องและครบถ้วน เพียงพอแล้ว หรือยัง? (Rothwell, 2006)

ในทางปฏิบัติ ความตรงภายใน หรือ internal validity ภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม ส่วนใหญ่ถูกรบกวนจากความคลาดเคลื่อน 2 ลักษณะ ได้แก่

1.1) ความคลาดเคลื่อนอย่างเป็นระบบ (systematic error)

1.2) ความคลาดเคลื่อนตามโอกาส (chance error)

1.1) ความคลาดเคลื่อนอย่างเป็นระบบ หรือ เรียกว่า “**ความลำเอียง (bias)**” ถือเป็นความคลาดเคลื่อนที่เกิดขึ้นจากความบกพร่องในการดำเนินงานตามระเบียบวิธีของแบบแผนงานวิจัยที่กำหนด เช่น การสุ่มตัวอย่าง การเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูลและการแปลผลข้อมูล เป็นต้น โดยผลจากความคลาดเคลื่อนนี้ จะส่งผลโดยตรงต่อความถูกต้อง หรือ ความตรงของผลลัพธ์ที่ได้จากการวิจัย ทำให้ค่าประมาณของขนาดความแตกต่างที่เกิดขึ้นระหว่างกลุ่ม หรือ treatment effect อาจบิดเบือนไป ซึ่งมีความเป็นไปได้ทั้งในกรณีที่มีค่ามากเกินไป (overestimation) หรือ ต่ำกว่าความเป็นจริง (underestimation) ดังนั้นโดยสรุป จึงกล่าวได้ว่าความคลาดเคลื่อนอย่างเป็นระบบเป็น “**ความลำเอียงที่ส่งผลกระทบต่อค่าประมาณพารามิเตอร์ที่สนใจ ทำให้มีแนวโน้มมาก หรือ ต่ำกว่าความเป็นจริงอย่างเป็นระบบ อันเนื่องมาจากความบกพร่องในการดำเนินงานต่างๆ ภายใต้ระเบียบวิธีของแบบแผนงานวิจัยที่นำมาใช้**” (Barraza, Arancibia et al., 2019) ในทางปฏิบัติความลำเอียงที่เกิดขึ้นและพบเป็นส่วนใหญ่ โดยเฉพาะในงานวิจัยเชิงทดลองแบบ RCT ประกอบด้วย 4 ลักษณะ ได้แก่ ความลำเอียงจากการเลือก (selection bias), ความลำเอียงจากการจัดกระทำ (performance bias), ความลำเอียงจากการประเมินผล (detection bias) และความลำเอียงจากการลดจำนวน (attrition bias) (Cipriani, Purgato et al., 2011)

1.2) ความคลาดเคลื่อนตามโอกาส หรือ “chance error” ในทางปฏิบัติถือเป็นความคลาดเคลื่อนอย่างสุ่ม (random error) หรือ บางครั้งอาจถูกเรียกว่า ความคลาดเคลื่อนทางสถิติ (statistical error) นั่นคือเป็นความคลาดเคลื่อนที่เกิดขึ้นจากความผันแปรของผลลัพธ์ตามโอกาส ซึ่งความคลาดเคลื่อนนี้ แม้ไม่สามารถคาดการณ์การเกิดขึ้นได้อย่างแน่ชัด แต่นักวิจัยสามารถลดโอกาสของการเกิดความคลาดเคลื่อนดังกล่าวนี้ได้ โดยจากการศึกษาที่ผ่านมา พบว่า มี 3 ปัจจัยสำคัญที่ส่งผลต่อความคลาดเคลื่อนอย่างสุ่ม ได้แก่ (1) ระดับความผันแปรภายในบุคคลและระหว่างบุคคล, (2) ขนาดตัวอย่าง, และ (3) ขนาดผลกระทบของความแตกต่างที่เกิดขึ้นจากการเปรียบเทียบ (Barraza, Arancibia et al., 2019) ดังนั้นจากปัจจัยที่ส่งผลดังกล่าว นักวิจัยจึงควรให้ความสำคัญกับค่าประมาณของผลลัพธ์ที่ได้ เพราะถือเป็นผลสะท้อนถึงแนวโน้มการเกิดความคลาดเคลื่อนอย่างสุ่ม โดยเฉพาะเมื่อมีการใช้ขนาดตัวอย่างที่มากพอ รวมถึงการเลือกวิธีการทางสถิติที่มีประสิทธิภาพและสอดคล้องกับคำถามวิจัย เพราะขนาดตัวอย่างที่ใหญ่มากพอ นั้น จะทำให้ค่าประมาณที่ได้มีความแม่นยำและสามารถลดโอกาสของการเกิดผลกระทบที่มาจากความคลาดเคลื่อนอย่างสุ่มลงได้ แต่อย่างไรก็ตาม ไม่ได้หมายความว่า ขนาดตัวอย่างของการศึกษา จะถือเป็นตัวบ่งชี้ที่มีผลต่อความตรงภายใน หรือ internal validity ของการศึกษาเสมอไป เพราะแม้จะมีการเพิ่มขนาดตัวอย่างให้ใหญ่เพียงใด แต่หากขั้นตอนในการดำเนินงานต่างๆ ตามแนวทางของระเบียบวิธีวิจัยที่นำมาใช้ ยังคงมีความไม่เหมาะสมและก่อให้เกิดความลำเอียงอื่นดังที่กล่าวไปข้างต้นตามมา ขนาดตัวอย่างที่เพิ่มขึ้น ก็ไม่ได้ช่วยทำให้

ความตรงภายในดีขึ้นแต่อย่างไร ดังนั้นการควบคุม หรือ ป้องกันความคลาดเคลื่อนอย่างสุ่ม ในทางปฏิบัติ จึงสามารถพิจารณาผ่านแนวทางการวิเคราะห์ข้อมูลเชิงอนุมาน ประกอบด้วย การทดสอบสมมติฐานทางสถิติ ด้วยค่า p-value และการประมาณค่าพารามิเตอร์ ด้วยช่วงเชื่อมั่นที่ระดับ 95% เนื่องจากในการวิเคราะห์ข้อมูลภายใต้แนวทางทั้งสองวิธีดังกล่าว มีการเกี่ยวข้องกับความคลาดเคลื่อนในรูปความน่าจะเป็น 2 ชนิด ได้แก่ ความคลาดเคลื่อนชนิดที่ 1 หรือ แอลฟา (α) ซึ่งหมายถึง ความน่าจะเป็นที่จะปฏิเสธสมมติฐาน H_0 ทั้งที่สมมติฐาน H_0 เป็นจริง และความคลาดเคลื่อนชนิดที่ 2 หรือ เบต้า (β) หมายถึง ความน่าจะเป็นที่จะไม่ปฏิเสธสมมติฐาน H_0 ทั้งที่สมมติฐาน H_0 ไม่เป็นจริง ซึ่งผลจากการเกี่ยวข้องดังกล่าว จึงทำให้นักวิจัยจำเป็นต้องกำหนดค่าความน่าจะเป็นทั้งสองดังกล่าว ก่อนนำแนวทางการทดสอบสมมติฐานทางสถิติและการประมาณค่าพารามิเตอร์มาทำการวิเคราะห์ข้อมูล ดังเช่น การกำหนดระดับนัยสำคัญ (α) เท่ากับ 0.05 หรือ การกำหนดระดับความเชื่อมั่นที่ระดับ 95% ซึ่งถูกพิจารณาจาก $(1-\alpha) \times 100$ เมื่อกำหนด $\alpha = 0.05$, และการกำหนดอำนาจการทดสอบทางสถิติ $(1-\beta)$ เท่ากับ 0.8 หรือ 80% เมื่อกำหนดค่า $\beta=0.2$ เพื่อนำมาใช้ในการคำนวณขนาดตัวอย่างของการศึกษาให้มากเพียงพอ เป็นต้น

อย่างไรก็ตาม ในการทดสอบสมมติฐานทางสถิติด้วยค่า p-value นักวิจัยอาจต้องพึงระวังในการนำมาใช้อย่างเข้าใจผิด (misunderstanding) และการชี้้นำข้อสรุปไปในทางที่ผิด (misleading) เพราะ **“ค่า p-value สามารถนำมาใช้เพียงเพื่ออธิบายเกี่ยวกับข้อสรุปทางสถิติของข้อมูลที่สังเกตได้จากการศึกษาภายใต้วิธีการทางสถิติที่ถูกเลือกนำมาใช้ว่า ข้อมูลนั้นสอดคล้องกับสมมติฐานหลัก (H_0) ที่กำหนดไว้หรือไม่?” (Bonovas and Piovani, 2023)** และโดยเฉพาะในกรณีไม่ปฏิเสธสมมติฐานหลัก (H_0) หรือ $p\text{-value} \geq 0.05$ นั่นคือ ในทางปฏิบัติ นักวิจัยจะกำหนดสมมติฐานหลัก (H_0) เป็น **“ไม่มีผลกระทบ (no effect)”** หรือ **“ไม่มีความแตกต่าง (no difference)”** ดังนั้นเมื่อ $p\text{-value} \geq 0.05$ หรือ ตัดสินใจว่า **“ไม่ปฏิเสธสมมติฐานหลัก (H_0)”** หรือ **not significant** จะอธิบายได้เพียงว่า **“จากข้อมูลที่ศึกษาภายใต้วิธีการทางสถิติที่ถูกเลือกนำมาวิเคราะห์พบว่า ยังไม่มีหลักฐานเพียงพอที่จะบ่งชี้การปฏิเสธสมมติฐานหลัก (H_0) หรือ ยังไม่มีหลักฐานเพียงพอที่จะบ่งชี้ความแตกต่างได้”** เท่านั้น แต่ไม่ได้หมายความว่า เมื่อมีการตัดสินใจดังกล่าวแล้ว ข้อสรุปที่ได้จะต้องเป็นไปตามสมมติฐานหลัก (H_0) ที่ระบุไว้เสมอ เพราะผลของการไม่ปฏิเสธสมมติฐานหลัก (H_0) ดังกล่าว อาจเนื่องมาจากอำนาจการทดสอบทางสถิติของวิธีการทางสถิติที่นำมาใช้นั้น มีไม่มากพอในการบ่งชี้ผลกระทบ (Smith, 2020) หรือกล่าวโดยสรุปอีกนัยหนึ่งได้ว่า ในการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติแล้วพบว่า **“ไม่มีนัยสำคัญทางสถิติ”** หรือ **“ค่า p-value ≥ 0.05 ”** นั้น ไม่ได้หมายความว่า ประเด็นดังกล่าว **“ไม่มีผลกระทบ”** หรือ **“ไม่มีความแตกต่าง”** ดังนั้นนักวิจัยจึงควรพึงระวังและหลีกเลี่ยงการสรุปผลที่ผิดไปจากความเป็นจริงดังกล่าว โดยเฉพาะการนำมาใช้สรุปผลเพื่อบ่งชี้ความเท่าเทียมกันของข้อมูลก่อนทดลอง เมื่อมีการเปรียบเทียบระหว่างกลุ่มทดลองและกลุ่มควบคุม ภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม ซึ่งรายละเอียดจะได้มีการกล่าวถึงในหัวข้อที่เกี่ยวข้องต่อไป

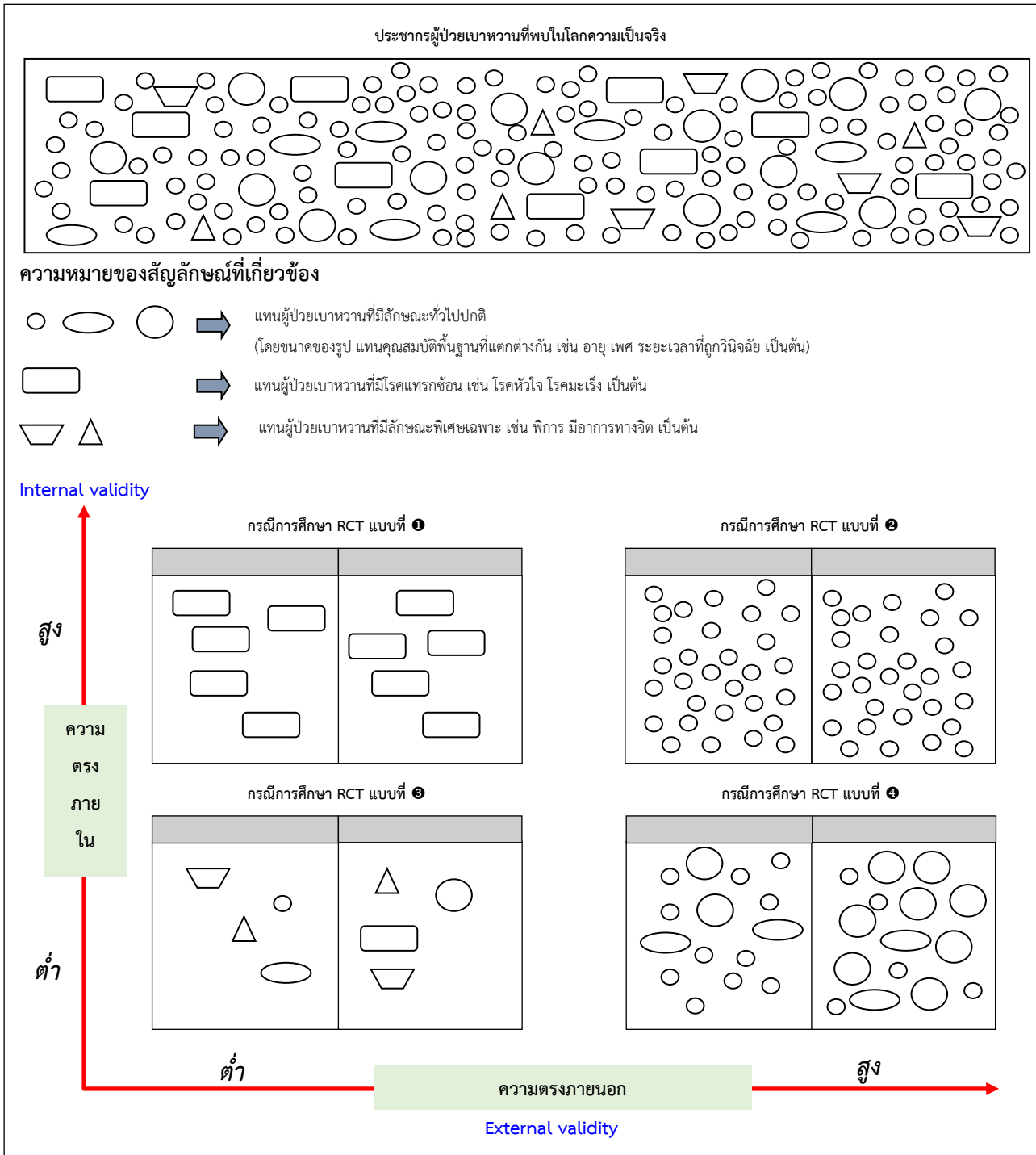
(2) ความตรงภายนอก หรือ external validity เมื่อถูกนำมาใช้ภายใต้แบบแผนงานวิจัยเชิงทดลอง สำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม สามารถสื่อความหมายได้ว่า

“ความแตกต่างที่พบและเกิดขึ้นระหว่างกลุ่ม หรือ treatment effect ที่ได้ สามารถถูกนำไปใช้อ้างอิงกับประชากรในบริบทอื่นได้ มากน้อยเพียงใด ?”

ซึ่งในการพิจารณาความตรงภายนอกดังกล่าว ในทางปฏิบัติจะขึ้นกับแบบแผนงานวิจัยและขั้นตอนในการดำเนินงานต่างๆ ตามกระบวนการทดลองที่ถูกเลือกนำมาใช้เป็นหลัก โดยพบว่า มีปัจจัยที่ส่งผลกระทบต่อประเด็นที่นักวิจัยควรให้ความสำคัญและคำนึงถึงเกี่ยวกับการพิจารณาความตรงภายนอก ประกอบด้วย (1) บริบท หรือ สภาพแวดล้อมของการศึกษา, (2) เกณฑ์คัดเข้าและเกณฑ์คัดออก, (3) คุณลักษณะของผู้เข้าร่วมที่ถูกสุ่ม, (4) สิ่งแทรกแซง, (5) วิธีการและการจัดการ, (6) การวัดผลลัพธ์และการติดตาม, และ (7) ผลที่ไม่พึงประสงค์ของวิธีการ (Rothwell, 2006, Cipriani, Purgato et al., 2011)

จากแนวทางการพิจารณาความตรง หรือ validity ของแบบแผนงานวิจัยดังกล่าวข้างต้น ในทางปฏิบัติ ความตรงภายใน หรือ internal validity อาจถือเป็นเงื่อนไขแรกที่ถูกนำมาพิจารณาและเกิดขึ้นก่อนความตรงภายนอก หรือ external validity เนื่องจากนักวิจัย จำเป็นต้องการผลลัพธ์ที่สามารถนำมาใช้เพื่อตอบโจทย์คำถามวิจัยที่กำหนดให้ได้ก่อนเป็นลำดับแรก (ซึ่งสอดคล้องกับความหมายของความตรงภายใน) จากนั้นจึงพิจารณาแนวทางการนำผลลัพธ์ที่ได้จากการศึกษานั้นไปขยายผล เพื่อคำนึงถึงการนำไปใช้ประโยชน์ในบริบทอื่นที่เกี่ยวข้อง (ซึ่งสอดคล้องกับความหมายของความตรงภายนอก) แต่อย่างไรก็ตาม ในความเป็นจริง ความสมดุลที่เกิดขึ้นระหว่างความตรงภายใน และความตรงภายนอก ควรต้องถูกคำนึงถึงและนำมาพิจารณาไปพร้อมกัน ภายใต้เงื่อนไข ข้อจำกัดและบริบทที่อาจมีความหลากหลายแตกต่างกันไป (Handley, Lyles et al., 2018) ดังเช่นกรณีงานวิจัยเชิงทดลองแบบ RCT แม้ได้รับการยอมรับและถือว่าเป็นงานวิจัยที่มีมาตรฐานดีที่สุด (gold standard) ภายใต้แบบแผนงานวิจัยเชิงทดลองที่อาศัยแนวคิดหลัก 3 ด้าน ได้แก่ การปกปิด (blinding or masking) การสุ่ม (randomization) และการมีกลุ่มควบคุม (control group) (Berger, Bour et al., 2021) ทำให้ผลลัพธ์ที่ได้ มีความตรงภายในค่อนข้างสูง แต่ขณะเดียวกัน หากไม่มีการคำนึงถึงข้อจำกัดทางคลินิกบางอย่าง ก็อาจนำไปสู่การทำให้ผลลัพธ์ที่ได้มีความตรงภายนอกต่ำได้เช่นกัน นั่นคือ ในบางการศึกษา อาจพบว่า ผลลัพธ์ที่ได้ มีค่าประมาณของผลกระทบจากสิ่งแทรกแซงค่อนข้างมาก หรือ ผลลัพธ์ที่ได้สามารถตอบโจทย์คำถามวิจัยที่กำหนดไว้ได้อย่างชัดเจน (ความตรงภายในค่อนข้างสูง) แต่การนำผลลัพธ์ หรือ ข้อค้นพบดังกล่าว ไปใช้ในทางปฏิบัติจริงกับกลุ่มเป้าหมายเดียวกันในบริบทอื่น อาจพบว่ามีข้อจำกัดทั้งความยุ่งยาก หรือ ซับซ้อน อันเนื่องมาจากเงื่อนไขของการทดลองที่มีความจำเพาะ หรือ แตกต่างไปจากปกติทั่วไป โดยเฉพาะความเป็นตัวแทนของผู้เข้าร่วม (representative) เพราะในโลกแห่งความเป็นจริง ประชากรผู้ป่วย ส่วนใหญ่มีความหลากหลาย แม้ผู้ป่วยทุกคน จะถูกการนำวิธีมาตรฐานซึ่งเป็นที่ยอมรับในระดับสากลมาใช้เป็นแนวทางในการวินิจฉัยโรคเช่นเดียวกัน มีการนำแนวทางการจัดสรรผู้เข้าร่วมแบบสุ่มที่เหมาะสมและมีการกำหนดเกณฑ์คัดเข้าและเกณฑ์คัดออกอย่างเที่ยงตรงแล้วก็ตาม แนวทางดังกล่าว อาจสามารถช่วยลดความแตกต่างระหว่าง

กลุ่มที่นำมาศึกษาอาจได้เป็นส่วนใหญ่ แต่อาจไม่ครอบคลุมทั้งหมด โดยเฉพาะในระดับกลุ่มย่อย ซึ่งสามารถพบความแตกต่างได้ เช่นในกรณีตัวอย่าง ดังแผนภาพที่ 1.4



แผนภาพที่ 1.4 แสดงการเปรียบเทียบความตรงภายในและความตรงภายนอกจากความเป็นตัวแทนของผู้เข้าร่วมในงานวิจัยเชิงทดลองแบบ RCT
 แผนภาพนี้ถูกปรับปรุงมาจากต้นฉบับตามแนวคิดของ Andrea Cipriani (Cipriani, Purgato et al., 2011)

จากแผนภาพที่ 1.4 ประชากรผู้ป่วยเบาหวานที่พบในโลกแห่งความเป็นจริง แม้จะถูกบ่งชี้การเป็นผู้ป่วยเบาหวานด้วยแนวทางการวินิจฉัยที่เป็นเกณฑ์มาตรฐานเดียวกันก็ตาม แต่เมื่อพิจารณาคุณสมบัติเฉพาะของผู้ป่วยกลับพบว่า มีความแตกต่างกัน (heterogenous) ดังนั้นในการพิจารณาเลือกผู้ป่วยมาทำการศึกษาภายใต้แนวทางการวิจัยเชิงทดลองแบบ RCT จึงจำเป็นต้องคำนึงถึงประเด็นความเป็นตัวแทนของประชากร โดยในทางปฏิบัติควรให้ครอบคลุมทั้งความตรงภายในและความตรงภายนอก ซึ่งสามารถสังเกตจากแผนภาพข้างต้นได้ดังนี้

กรณีการศึกษา RCT ที่ ❶ พบว่า ผู้ป่วยเบาหวานที่ถูกนำมาศึกษา ถูกคัดเลือกมาเฉพาะผู้ป่วยที่มีโรคแทรกซ้อนอื่น เช่น โรคหัวใจ หรือ โรคมะเร็ง เป็นต้น และถูกสุ่มเข้ากลุ่มทดลองและกลุ่มควบคุมในลักษณะที่ใกล้เคียง หรือ เท่าเทียมกัน ดังนั้นจึงส่งผลให้กรณีการศึกษาดังกล่าวนี้ มีแนวโน้มที่จะมีความตรงภายในค่อนข้างสูง ขณะเดียวกันหากพิจารณาคุณสมบัติของผู้ป่วยเบาหวานที่ถูกนำมาศึกษาเมื่อเทียบกับในระดับประชากรที่พบในโลกแห่งความเป็นจริง จะพบว่า เป็นเพียงส่วนหนึ่ง ดังนั้นจากเงื่อนไขคุณสมบัติของความเป็นตัวแทนผู้ป่วยเบาหวานที่มีความเฉพาะเจาะจงดังกล่าว จึงทำให้ผลการศึกษาที่ได้ แม้จะมีแนวโน้มความตรงภายในสูงก็ตาม แต่ความตรงภายนอกซึ่งมุ่งเน้นการนำผลลัพธ์ที่ได้ไปสู่การขยายผลในประชากรและบริบทอื่น อาจมีแนวโน้มต่ำ เนื่องจากประชากรที่เป็นผู้ป่วยเบาหวานที่มีโรคแทรกซ้อนอื่น มีจำนวนน้อย เมื่อเทียบกับประชากรผู้ป่วยเบาหวานในกลุ่มอื่น

กรณีการศึกษา RCT ที่ ❷ พบว่า ผู้ป่วยเบาหวานที่ถูกนำมาศึกษา ถูกคัดเลือกมาจากผู้ป่วยที่มีลักษณะทั่วไปปกติ ซึ่งเป็นประชากรส่วนใหญ่ในโลกแห่งความเป็นจริงและถูกสุ่มเข้ากลุ่มทดลองและกลุ่มควบคุมในลักษณะที่ใกล้เคียง หรือ เท่าเทียมกัน ดังนั้นจึงส่งผลให้กรณีการศึกษาดังกล่าว มีแนวโน้มที่จะมีความตรงภายในค่อนข้างสูง ขณะเดียวกันเมื่อพิจารณาคุณสมบัติของผู้ป่วยเบาหวานที่ถูกนำมาศึกษาเมื่อเทียบกับในระดับประชากรที่พบในโลกแห่งความเป็นจริง พบว่า เป็นส่วนใหญ่ของกลุ่มประชากร ดังนั้นจากเงื่อนไขคุณสมบัติของความเป็นตัวแทนผู้ป่วยเบาหวานที่มีลักษณะทั่วไป (generalization) ดังกล่าว จึงทำให้ผลการศึกษาที่ได้ นอกจากจะมีแนวโน้มความตรงภายในที่สูงแล้ว ยังมีแนวโน้มความตรงภายนอกสูงขึ้นไปด้วย เนื่องจากผู้ป่วยเบาหวานที่นำมาศึกษา มีลักษณะใกล้เคียงและพบมากเป็นส่วนใหญ่ในโลกแห่งความเป็นจริง นั่นคือ นอกจากผลลัพธ์ที่ได้จากงานวิจัยนี้ จะสามารถตอบคำถามวิจัยที่กำหนดไว้ได้อย่างถูกต้องอย่างชัดเจนแล้ว ผลลัพธ์จากงานวิจัยนี้ ยังสามารถถูกนำไปใช้ขยายผลในประชากรส่วนใหญ่ที่พบในโลกแห่งความเป็นจริงได้อีกด้วย

กรณีการศึกษา RCT ที่ ❸ พบว่า ผู้ป่วยเบาหวานที่ถูกนำมาศึกษา ถูกคัดเลือกมาจากผู้ป่วยที่มีลักษณะพิเศษเฉพาะ เช่น พิกการ มีอาการทางจิต เป็นต้น ซึ่งเป็นประชากรส่วนน้อยในโลกแห่งความเป็นจริงและจากการสุ่มเข้ากลุ่มทดลองและกลุ่มควบคุม พบว่า มีลักษณะที่แตกต่างกันค่อนข้างชัดเจน ดังนั้นจึงส่งผลให้กรณีการศึกษาดังกล่าว มีแนวโน้มที่จะมีความตรงภายในค่อนข้างต่ำ ขณะเดียวกันเมื่อพิจารณาคุณสมบัติของผู้ป่วยเบาหวานที่ถูกนำมาศึกษาเมื่อเทียบกับในระดับประชากรที่พบในโลกแห่งความเป็นจริง พบว่า เป็นส่วนน้อยของกลุ่มประชากร ดังนั้นจากเงื่อนไขคุณสมบัติของความเป็นตัวแทนผู้ป่วยเบาหวานที่มีลักษณะเฉพาะเจาะจง จึงทำให้ผลการศึกษาที่ได้ นอกจากจะมีแนวโน้มความตรงภายในที่ต่ำแล้ว ยังมีแนวโน้มความตรงภายนอกต่ำตามไปด้วย เนื่องจากผู้ป่วยเบาหวานที่นำมาศึกษานั้น มีลักษณะเฉพาะที่แตกต่างและพบได้น้อยในโลกแห่งความเป็นจริง นั่นคือ นอกจากผลลัพธ์ที่ได้จากงานวิจัยในกรณีดังกล่าวนี้ จะมีแนวโน้มไม่สามารถตอบโจทย์คำถามวิจัยที่กำหนดไว้ได้แล้ว ผลลัพธ์จากงานวิจัยนี้ ยังไม่สามารถถูกนำไปใช้ขยายผลในประชากรกลุ่มอื่นบนโลกแห่งความเป็นจริงได้

กรณีการศึกษา RCT ที่ 4 พบว่า ผู้ป่วยเบาหวานที่ถูกนำมาศึกษา ถูกคัดเลือกมาจากผู้ป่วยที่มีลักษณะปกติทั่วไป ซึ่งเป็นประชากรส่วนใหญ่ในโลกแห่งความเป็นจริง แต่ภายใต้กลุ่มประชากรดังกล่าว ยังมีคุณสมบัติพื้นฐาน เช่น อายุ เพศ ระยะเวลาที่ถูกวินิจฉัย เป็นต้น แตกต่างกันไปและจากการสุ่มเข้ากลุ่มทดลองและกลุ่มควบคุม ยังพบว่า มีลักษณะที่แตกต่างกันค่อนข้างชัดเจน ดังนั้นจึงส่งผลให้กรณีการศึกษาดังกล่าว มีแนวโน้มที่จะมีความตรงภายในค่อนข้างต่ำ ขณะเดียวกันเมื่อพิจารณาคุณสมบัติของผู้ป่วยเบาหวานที่ถูกนำมาศึกษาเมื่อเทียบกับในระดับประชากรที่พบในโลกแห่งความเป็นจริง พบว่า เป็นส่วนใหญ่ของกลุ่มประชากร ดังนั้นจากเงื่อนไขคุณสมบัติของความเป็นตัวแทนของผู้ป่วยเบาหวานที่มีลักษณะทั่วไป (generalization) ดังกล่าว จึงทำให้ผลการศึกษาที่ได้ มีแนวโน้มความตรงภายในที่ค่อนข้างต่ำ ขณะที่แนวโน้มความตรงภายนอกสูง เนื่องจากผู้ป่วยเบาหวานที่นำมาศึกษา พบมากเป็นส่วนใหญ่ในโลกแห่งความเป็นจริง นั่นคือ ผลลัพธ์ที่ได้จากงานวิจัยกรณีนี้ อาจมีแนวโน้มการตอบโจทยคำถามวิจัยที่กำหนดไว้อย่างไม่ชัดเจน เนื่องจากมีความผันแปรในกลุ่มย่อยของผู้เข้าร่วมที่นำมาศึกษา แต่อย่างไรก็ตามหากปัญหาข้อจำกัดเกี่ยวกับความผันแปรดังกล่าว ได้ถูกควบคุมและจัดการอย่างเหมาะสม ผลลัพธ์ที่ได้จากงานวิจัยนี้ จะมีแนวโน้มถูกนำไปใช้เพื่อขยายผลในประชากรส่วนใหญ่ที่พบในโลกแห่งความเป็นจริงได้ค่อนข้างดี

อย่างไรก็ตาม นอกเหนือไปจากการพิจารณาความตรงภายนอก โดยคำนึงถึงประเด็นผู้เข้าร่วมดังที่กล่าวไปข้างต้นแล้ว ในกรณีการวิจัยเชิงทดลอง โดยเฉพาะภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม ยังพบว่า การมีปฏิสัมพันธ์ร่วมกัน (interaction) ระหว่างการทดสอบก่อนทดลอง (pretest score) กับวิธีการ (treatment) ยังถือเป็นอีกประเด็นปัญหาสำคัญที่ส่งผลกระทบต่อความตรงภายนอก ซึ่งนักวิจัยจำเป็นต้องนำมาพิจารณา โดยการเกิดปฏิสัมพันธ์ดังกล่าว สามารถพบได้เมื่อ การทดสอบก่อนทดลองนั้น มีแนวโน้มทำให้ผู้เข้าร่วมเกิดความรู้สึกที่ไว หรือ ได้รับการกระตุ้นเพื่อตอบสนองต่อวิธีการที่แตกต่างกันมากกว่าการไม่มีการทดสอบก่อนทดลอง เช่น การศึกษาผลของแนวทางการรักษาภาวะระดับความดันในลูกตาสูง (ocular hypertension) ด้วยตนเอง โดยนำผู้ป่วยที่มีระดับความดันในลูกตาสูงเข้าร่วมและถูกทดสอบก่อนได้รับโปรแกรมที่เป็นแนวทางในการรักษาเกี่ยวกับการลดระดับความดันในลูกตาดด้วยตนเอง (อาทิ จำนวนครั้งของการนวดใต้ตาในรอบอาทิตย์ที่ผ่านมา หรือ จำนวนครั้งของการบริหารสายตาด้วยการกรอกลูกตาไป-มาต่อวัน เป็นต้น) ซึ่งในการตอบคำถามเกี่ยวกับการทดสอบก่อนทดลองดังกล่าว อาจมีแนวโน้มไปกระตุ้นให้ผู้เข้าร่วม อยากเพิ่มจำนวนครั้งของการทำกิจกรรมที่เกี่ยวข้องดังกล่าวด้วยตนเองมากขึ้น โดยไม่คำนึงถึงสิ่งแทรกแซง ทั้งที่ในความเป็นจริงแล้ว ผลที่เกิดจากการกระทำดังกล่าว มีผลโดยตรงต่อผลลัพธ์ที่ถูกประเมินและเป็นผลกระทบจากสิ่งแทรกแซง เป็นต้น ซึ่งจากปฏิสัมพันธ์ที่เกิดขึ้นดังกล่าว แม้จะทำให้ผลลัพธ์ที่ได้มีความตรงภายในสูง นั่นคือ สามารถตอบโจทยคำถามวิจัยได้ แต่เมื่อพิจารณาการนำผลลัพธ์ดังกล่าวไปสู่การขยายผลในบริบทอื่น อาจพบว่า มีข้อจำกัด เนื่องจากปฏิสัมพันธ์ที่เกิดขึ้นดังกล่าว ถือเป็นความจำเพาะในบริบทภายใต้การศึกษานี้เท่านั้น ซึ่งไม่ใช่ลักษณะทั่วไปที่สามารถนำไปขยายผลยังส่วนอื่นที่เกี่ยวข้องได้ ดังนั้นจึงทำให้ความตรงภายนอก มีแนวโน้มลดลง หากพบว่า มีผลกระทบดังกล่าวนี้เกิดขึ้น

1.2.4 การพิจารณาข้อมูลก่อนและหลังทดลอง

ในงานวิจัยทางวิทยาศาสตร์สุขภาพ ภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม โดยเฉพาะจากแผนภาพที่ 1.1 จะพบว่า ข้อมูลก่อนทดลอง หรือ pretest score (Y_{i1}) ถือเป็นข้อมูลสำคัญที่นักวิจัยจำเป็นต้องคำนึงถึงทั้งในขั้นตอนการวางแผนและการดำเนินงานเก็บรวบรวมข้อมูล ขณะเดียวกันข้อมูลดังกล่าว ควรต้องถูกนำมาพิจารณาทุกครั้ง โดยเฉพาะในขั้นตอนของการวิเคราะห์ข้อมูล เนื่องจากข้อมูลก่อนทดลอง จะถูกนำมาใช้ เพื่อลดความแปรปรวนของความคลาดเคลื่อน (error variance) ที่อาจจะเกิดขึ้นจากการวัดภายในตัวบุคคล (*Dimitrov and Rumrill, 2003*) ดังเช่นในกรณีการวัดระดับความดันเลือดขณะหัวใจบีบตัว (systolic blood pressure: SBP) ของผู้ป่วยเบาหวาน หากนักวิจัยวัดค่าดังกล่าวเพียงครั้งเดียว ผลลัพธ์ที่ได้ อาจมีค่าที่เป็นไปได้ทั้งค่าระดับความดันเลือดที่มาก หรือ น้อยกว่าระดับปกติที่ควรจะเป็นของบุคคลดังกล่าว ซึ่งขึ้นกับหลายปัจจัย โดยเฉพาะปัจจัยภายในตัวบุคคล ดังนั้นการนำข้อมูลก่อนทดลองมาใช้ จึงทำให้ค่าความดันเลือดขณะหัวใจบีบตัวที่วัดได้ในครั้งต่อมาถูกปรับความผันแปร โดยเฉพาะความผันแปรภายในตัวบุคคล ด้วยค่าความดันเลือดที่วัดก่อน และส่งผลโดยตรงให้ค่าผลลัพธ์ที่ต้องการ สามารถสะท้อนค่าจริงได้อย่างแม่นยำมากกว่าการวัดเพียงครั้งเดียว เป็นต้น

ขณะเดียวกัน หากเมื่อต้องการประเมินความแตกต่างระหว่างกลุ่ม ซึ่งเป็นผลกระทบจากสิ่งแทรกแซง หรือ treatment effect ที่นักวิจัยสนใจศึกษา การละเลยข้อมูลก่อนทดลอง อาจส่งผลทำให้ข้อสรุปที่ได้จากการนำข้อมูลหลังทดลอง หรือ posttest score (Y_{i2}) เพียงอย่างเดียวมาเปรียบเทียบระหว่างกลุ่ม มีความไม่ชัดเจน ดังเช่นหากข้อสรุปที่ได้ พบว่า มีความแตกต่างอย่างมีนัยสำคัญเกิดขึ้นระหว่างกลุ่ม ($p\text{-value} < 0.05$) นักวิจัยอาจไม่สามารถระบุความแตกต่างนั้น ได้อย่างชัดเจนว่า ผลกระทบนั้น เป็นผลกระทบที่เกิดขึ้นและมาจากสิ่งแทรกแซงที่ศึกษา หรือ เป็นผลกระทบแฝงที่เจือปนมาจากความไม่เท่าเทียมกันของข้อมูลก่อนทดลองที่ขาดการคำนึงถึงและถูกนำมาพิจารณา โดยนักวิจัยตั้งแต่เริ่มทำการทดลอง (*Vickers and Altman, 2001*) ในทางตรงกันข้าม ภายใต้การศึกษาเดียวกัน หากข้อสรุปที่ได้ พบว่า มีความแตกต่างอย่างไม่มีนัยสำคัญเกิดขึ้นระหว่างกลุ่ม ($p\text{-value} \geq 0.05$) นั้นอาจเป็นเพราะสิ่งแทรกแซงที่นักวิจัยสนใจศึกษานั้น “ไม่มีผล” หรือ “ไม่ส่งผลกระทบอย่างแท้จริง” หรือ อาจเป็นเพราะผลกระทบแฝงที่เจือปนมาจากความไม่เท่าเทียมกันของข้อมูลก่อนทดลองระหว่างกลุ่ม จึงทำให้ไม่สามารถบ่งชี้ความแตกต่างที่เกิดขึ้นได้อย่างชัดเจน เป็นต้น ดังนั้นการนำข้อมูลก่อนทดลอง หรือ pretest score มาพิจารณาประกอบในการวิเคราะห์ข้อมูล ภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม จึงไม่เพียงแต่การช่วยลดความแปรปรวนของความคลาดเคลื่อนภายในกลุ่มเท่านั้น แต่ยังสามารถเพิ่มอำนาจการทดสอบทางสถิติในการบ่งชี้ความแตกต่างระหว่างกลุ่มได้มากกว่ากรณีที่ข้อมูลก่อนทดลอง หรือ คะแนนก่อนทดลองไม่ถูกนำมาพิจารณา (*Dimitrov and Rumrill, 2003*)

อย่างไรก็ตาม นอกจากประเด็นการนำข้อมูลก่อนทดลอง หรือ pretest score มาพิจารณาดังกล่าวข้างต้นแล้ว ความสัมพันธ์ที่เกิดขึ้นระหว่างข้อมูลก่อนและหลังทดลองในแต่ละบุคคล หรือ ผู้เข้าร่วมในการศึกษา ยังถือเป็นอีกประเด็นที่มีความสำคัญและนักวิจัยจำเป็นต้องคำนึงถึง เพราะจากแนวปฏิบัติภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม ดังแผนภาพที่ 1.1 พบว่า ตัวแปรแบบต่อเนื่องที่สนใจของแต่ละผู้เข้าร่วม จะถูกวัดจำนวน 2 ครั้ง ตามลำดับเวลาที่แตกต่างกัน (ก่อนและหลังทดลอง) จึงส่งผลให้ข้อมูลการวัดตัวแปรที่สนใจทั้งสองครั้งที่ได้จากผู้เข้าร่วมคนเดียวกัน มีความสัมพันธ์กัน เช่น ผู้เข้าร่วมคนที่ 1 ข้อมูลก่อนทดลอง หรือ

pretest score (Y_{11}) จะมีความสัมพันธ์กับข้อมูลหลังทดลอง หรือ posttest score (Y_{12}) ของตนเอง เป็นต้น ดังนั้นในการพิจารณาแนวทางการประเมินผลกระทบจากสิ่งแทรกแซง ซึ่งจะได้กล่าวถึงรายละเอียดในหัวข้อที่เกี่ยวข้องถัดไป จึงเป็นประเด็นสำคัญว่า ความสัมพันธ์ระหว่างข้อมูลก่อนและหลังทดลองที่ถูกวัดซ้ำในผู้เข้าร่วมคนเดียวกันดังกล่าว ต้องถูกนำมาพิจารณาและประกอบการตัดสินใจเพื่อเลือกวิธีการทางสถิติและแนวทางการวิเคราะห์ข้อมูลภายใต้แบบแผนงานวิจัยดังกล่าว ให้มีความถูกต้องและครอบคลุมต่อไป

1.3 แนวทางการประเมินผลกระทบจากสิ่งแทรกแซง

การวิเคราะห์ข้อมูลภายใต้แบบแผนงานวิจัยสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่มในงานวิจัยทางวิทยาศาสตร์สุขภาพ จำเป็นอย่างยิ่งที่นักวิจัย ต้องทบทวนและทำความเข้าใจประเด็นที่เกี่ยวข้องกับแนวทางการประเมินผลกระทบจากสิ่งแทรกแซง เพื่อให้เกิดความชัดเจนในการตัดสินใจเลือกใช้วิธีการทางสถิติและกำหนดแนวทางการวิเคราะห์ข้อมูลได้อย่างถูกต้องและเหมาะสม โดยมีประเด็นที่นำมาพิจารณา ดังนี้

- 1.3.1 คำถามวิจัยที่เกี่ยวข้อง
- 1.3.2 ผลกระทบของวิธีการ
- 1.3.3 การมีกลุ่มเปรียบเทียบ
- 1.3.4 การนำข้อมูลก่อนทดลองมาพิจารณา

1.3.1 คำถามวิจัยที่เกี่ยวข้อง

ในทางปฏิบัติสำหรับการดำเนินงานวิจัย ผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูล ถือเป็นข้อค้นพบและคำตอบสำคัญที่มีความสอดคล้องและเกี่ยวพันโดยตรงกับคำถามวิจัย (Snyder, Foltz et al., 2019) หรือ อาจกล่าวได้ว่า คำถามวิจัยเป็นพื้นฐานสำคัญสำหรับการวิเคราะห์ข้อมูล นั่นคือ การวิเคราะห์ข้อมูลทางสถิติ ถือเป็นแนวทางหลักสำคัญที่ต้องถูกกระทำขึ้น เพื่อให้ได้มาซึ่งผลลัพธ์และเป็นคำตอบสำหรับคำถามวิจัยที่กำหนด (Twisk, 2013) ดังนั้นในการทำวิจัยภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่มก็เช่นเดียวกัน การทบทวนและทำความเข้าใจเกี่ยวกับคำถามวิจัยที่ถูกนำมาใช้ภายใต้แบบแผนงานวิจัยดังกล่าว จึงมีความสำคัญและจำเป็น ทั้งนี้เพื่อให้ นักวิจัยสามารถทำการวิเคราะห์ข้อมูล โดยมุ่งเน้นการหาผลลัพธ์ซึ่งเป็นคำตอบ หรือ ข้อค้นพบที่ได้จากงานวิจัยได้อย่างถูกต้องและตรงประเด็นมากขึ้น ซึ่งในทางปฏิบัติ สามารถจำแนกคำถามวิจัยออกเป็น 2 แนวคิด ภายใต้การพิจารณาข้อมูลก่อนและหลังทดลอง ดังนี้

- (1) คำถามวิจัยภายใต้แนวคิด “การนำข้อมูลก่อน-หลังทดลอง มาพิจารณาแบบวัดซ้ำ”
- (2) คำถามวิจัยภายใต้แนวคิด “การนำข้อมูลก่อนทดลอง มาพิจารณาแบบตัวแปรร่วม”

(1) คำถามวิจัยภายใต้แนวคิด “การนำข้อมูลก่อน-หลังทดลอง มาพิจารณาแบบวัดซ้ำ” นั้นคือ เป็นแนวคิดในการกำหนดคำถามวิจัยภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม โดยกำหนดให้ข้อมูลก่อน-หลังทดลองของทั้งกลุ่มทดลองและกลุ่มควบคุม เป็นตัวแปรผลลัพธ์ที่สนใจเช่นเดียวกัน เพียงแต่ถูกวัดซ้ำแตกต่างกันตามจุดเวลาที่นักวิจัยกำหนดไว้เท่านั้น ซึ่งนั่นก็หมายความว่า นอกจากนักวิจัย จะสนใจ ประเมินผลกระทบที่เกิดขึ้นจากกลุ่ม หรือ treatment effect ที่แตกต่างกันแล้ว นักวิจัยยังสนใจประเมินผลกระทบ ที่เกิดขึ้นจากจุดเวลาของการวัดทั้งสองครั้งที่แตกต่างกันอีกด้วย ดังนั้นในทางปฏิบัติสำหรับการวิเคราะห์ข้อมูล เพื่อตอบโจทย์คำถามวิจัยภายใต้แนวคิดดังกล่าวนี้ จึงพบว่า ผลกระทบที่นักวิจัยต้องการทราบ จะประกอบด้วย (1) ผลกระทบหลัก (main effect) จากตัวแปรกลุ่ม ได้แก่ กลุ่มทดลอง หรือ กลุ่มควบคุม, (2) ผลกระทบหลัก จากตัวแปรจุดเวลา ได้แก่ จุดเวลาที่ 1 (ก่อนทดลอง) หรือ จุดเวลาที่ 2 (หลังทดลอง), และ (3) ผลกระทบร่วม (interaction effect) ระหว่างตัวแปรกลุ่มและตัวแปรจุดเวลา ซึ่งสามารถเขียนรูปทั่วไปของคำถามวิจัยได้ ดังนี้

“ค่าเฉลี่ยปริมาณ (หรือ คะแนน) ของตัวแปรผลลัพธ์ที่สนใจก่อนและ หลังทดลองระหว่างกลุ่มทดลองและกลุ่มควบคุม แตกต่างกัน หรือไม่ ?”

จากคำถามวิจัยข้างต้น พบว่า ข้อมูลก่อนและหลังทดลองของทั้งกลุ่มทดลองและกลุ่มควบคุม ถูกกำหนดให้มีความสำคัญและนำมาพิจารณาในระดับเดียวกันแบบวัดซ้ำ นั่นคือ ถูกนำมาใช้เป็นตัวแปรผลลัพธ์ที่ถูกวัด ณ จุดเวลาเริ่มต้นก่อนทดลองและจุดเวลาหลังทดลอง จากนั้นจึงนำตัวแปรผลลัพธ์ทั้งสองจุดเวลาดังกล่าวมาทำการวิเคราะห์ข้อมูลเพื่อเปรียบเทียบความแตกต่างระหว่างกลุ่มทดลองและกลุ่มควบคุมไปพร้อมกัน เช่นตัวอย่าง คำถามวิจัย ดังแผนภาพที่ 1.5

คำถามวิจัย : “ค่าเฉลี่ยระดับ HbA1C ก่อนและหลังได้รับโปรแกรมการปรับพฤติกรรมควบคุมระดับน้ำตาลใน เลือดของผู้ป่วยเบาหวานในกลุ่มทดลอง แตกต่างจากกลุ่มควบคุม ซึ่งไม่ได้รับโปรแกรมฯ หรือไม่ ?”

คำอธิบาย : จากคำถามวิจัยดังกล่าว พบว่า ตัวแปรผลลัพธ์แบบต่อเนื่องที่สนใจ ได้แก่ ระดับ HbA1C ของผู้ป่วย เบาหวาน โดยมีการวัดผลสองจุดเวลา ได้แก่ ครั้งที่ 1 ก่อนทดลองและครั้งที่ 2 หลังทดลอง ประกอบด้วยกลุ่มศึกษาจำนวน 2 กลุ่ม ได้แก่ กลุ่มทดลอง (ได้รับโปรแกรมการปรับพฤติกรรม การควบคุมระดับน้ำตาลในเลือด) และกลุ่มควบคุม (ไม่ได้รับโปรแกรมการปรับพฤติกรรมฯ แต่จะ ได้รับการให้คำปรึกษาตามแนวปฏิบัติเดิมที่มีอยู่) ซึ่งภายใต้คำถามวิจัยดังกล่าว จะสังเกตได้ว่า ตัวแปรผลลัพธ์ทั้งสองจุดเวลา ถูกนำมาใช้เป็นตัวแปรในการพิจารณาผลกระทบที่เกิดขึ้น ครอบคลุมทั้ง 3 ลักษณะดังที่กล่าวไปข้างต้น ได้แก่ ผลกระทบหลักที่มาจากกลุ่ม ผลกระทบหลัก ที่มาจากจุดเวลาและผลกระทบร่วมที่มาจากกลุ่มและจุดเวลา ดังนั้นในการตัดสินใจเลือกใช้วิธีการ ทางสถิติสำหรับการวิเคราะห์ข้อมูล เพื่อตอบโจทย์คำถามวิจัยดังกล่าว นักวิจัยจึงจำเป็นต้องให้ ความสำคัญและคำนึงถึงวิธีการทางสถิติที่สามารถนำมาใช้ในการวิเคราะห์ข้อมูลและให้ผลลัพธ์ เพื่อตอบประเด็นผลกระทบภายใต้คำถามวิจัยได้อย่างครอบคลุมทั้ง 3 ลักษณะ

แผนภาพที่ 1.5 แสดงกรณีคำถามวิจัยตัวอย่างภายใต้แนวคิด “การนำข้อมูลก่อน-หลังทดลองมาพิจารณาแบบวัดซ้ำ”

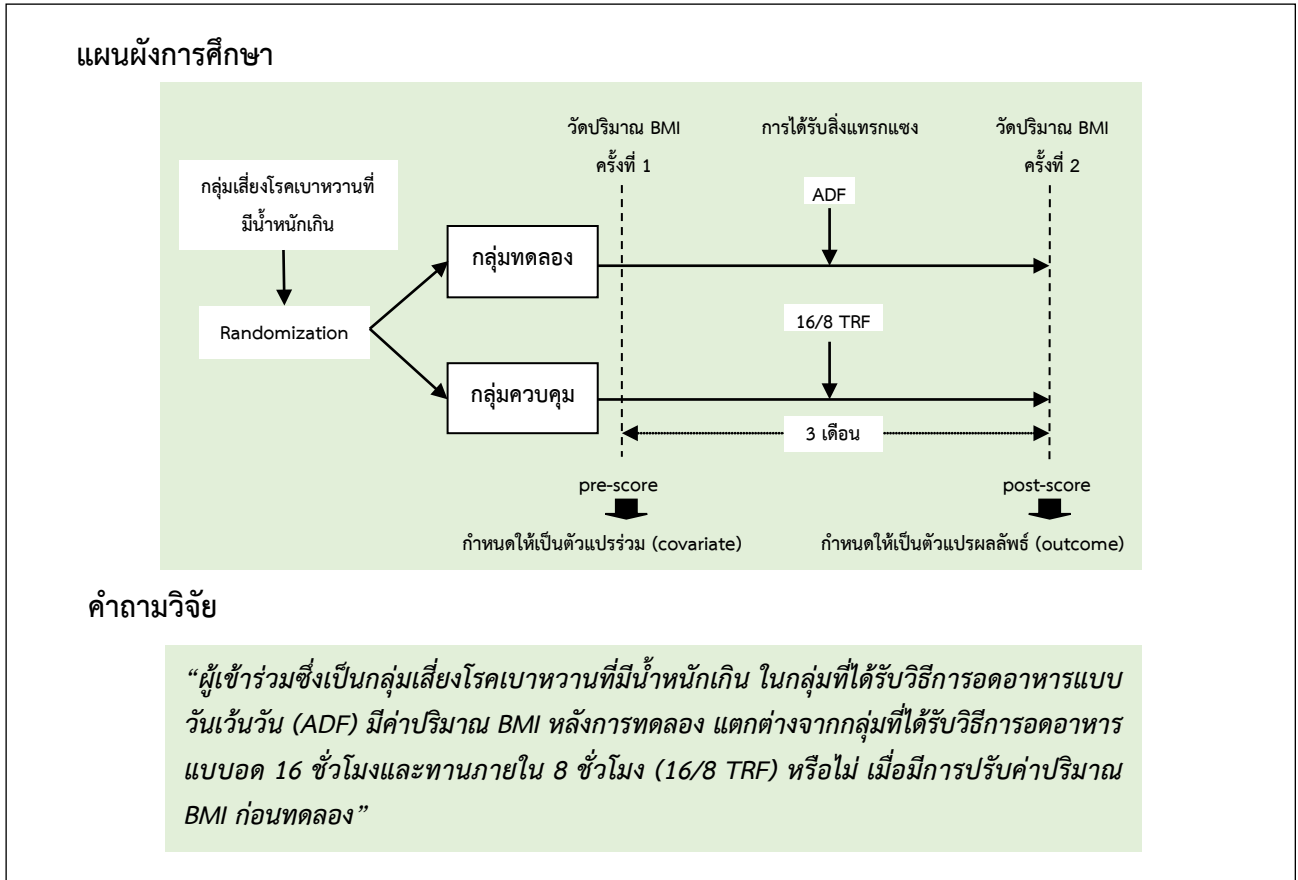
(2) แนวคิด “การนำข้อมูลก่อนทดลอง มาพิจารณาแบบตัวแปรร่วม” นั่นคือ เป็นการกำหนดคำถามวิจัยภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม โดยกำหนดให้ข้อมูลก่อนทดลอง เป็นตัวแปรร่วม หรือ covariate และกำหนดให้ข้อมูลหลังทดลอง เป็นตัวแปรผลลัพธ์ที่สนใจ ซึ่งนั้นก็หมายความว่า นักวิจัย มุ่งเน้นความสนใจไปที่การประเมินผลกระทบที่เกิดขึ้นจากกลุ่ม หรือ treatment effect เป็นหลักเพียงอย่างเดียว โดยไม่ได้คำนึงถึงและสนใจการประเมินผลกระทบที่เกิดขึ้นจากจุดเวลา ดังนั้นในทางปฏิบัติ จึงทำให้การกำหนดคำถามวิจัยภายใต้แนวคิดนี้ แตกต่างไปจากแนวคิดตั้งที่กล่าวไปข้างต้นอย่างชัดเจน เพราะการกำหนดคำถามวิจัยภายใต้แนวคิดนี้ จะมุ่งเน้นการวิเคราะห์ข้อมูล เพื่อให้ได้ผลลัพธ์ซึ่งเป็นคำตอบเกี่ยวกับผลกระทบหลักที่มาจากกลุ่มเท่านั้น ซึ่งสามารถเขียนรูปทั่วไปของคำถามวิจัยภายใต้แนวคิดดังกล่าวได้ ดังนี้

“ค่าเฉลี่ยปริมาณ (หรือ คะแนน) ของตัวแปรผลลัพธ์ที่สนใจหลังทดลอง ระหว่างกลุ่มทดลองและกลุ่มควบคุม แตกต่างกัน หรือไม่ ? เมื่อมีการนำปริมาณ (หรือ คะแนน) ก่อนทดลองมาปรับ”

จากคำถามวิจัยภายใต้แนวคิดตั้งรูปแบบข้างต้น สามารถถูกนำมาใช้ได้ทั้งงานวิจัยที่มีแนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่ม ดังเช่น งานวิจัยเชิงทดลองแบบ RCT เป็นต้น หรือ งานวิจัยที่มีแนวทางการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบไม่สุ่ม เช่น งานวิจัยกึ่งทดลอง หรือ quasi-experimental study เป็นต้น โดยตัวแปรผลลัพธ์แบบต่อเนื่องที่สนใจและถูกนำมาพิจารณา มีความเป็นไปได้ทั้งข้อมูลในลักษณะค่าปริมาณ หรือ ระดับสารต่างๆ ที่ถูกนำมาใช้ เพื่อประเมินความแตกต่าง ซึ่งอาจส่งผลทางตรง หรือ ทางอ้อม ไปสู่การเปลี่ยนแปลงลักษณะบางอย่างทางคลินิก หรือ การปรับเปลี่ยนพฤติกรรม ตามที่นักวิจัยสนใจศึกษา ดังเช่น **กรณีตัวอย่างที่ 1.3**

กรณีตัวอย่างที่ 1.3

จากการศึกษาวิธีการอดอาหารแบบเป็นช่วงๆ (intermittent fasting, IF) ในกลุ่มเสี่ยงโรคเบาหวานที่มีน้ำหนักเกิน ด้วยรูปแบบการทานแบบวันเว้นวัน (alternate-day fasting, ADF) กับการทานแบบอดอาหาร 16 ชั่วโมงและทานภายใน 8 ชั่วโมง (16/8 time-restricted fasting, 16/8 TRF) ภายใต้แนวทางการวิจัยเชิงทดลองแบบ RCT พบว่า ตัวแปรแบบต่อเนื่องที่สนใจและนำมาพิจารณาในการศึกษาครั้งนี้ ได้แก่ ปริมาณค่าดัชนีมวลกาย (BMI) โดยกำหนดให้ BMI ก่อนทดลอง เป็นตัวแปรร่วมและ BMI หลังทดลอง เป็นตัวแปรผลลัพธ์ที่สนใจในการนำมาเปรียบเทียบ เป็นต้น ดังนั้นจึงสามารถแสดงแผนผังการศึกษาและรูปแบบคำถามวิจัยในกรณีดังกล่าวนี้ได้ ดังแผนภาพที่ 1.6



แผนภาพที่ 1.6 แสดงกรณีตัวอย่างแผนผังการศึกษาและคำถามวิจัยภายใต้แบบแผนงานวิจัยเชิงทดลองแบบ RCT
 แผนภาพนี้ถูกปรับปรุงมาจากต้นฉบับตามแนวคิดของ Sek Ying CHAIR (CHAIR, CAI et al., 2022)

จากกรณีตัวอย่างคำถามวิจัยที่เกี่ยวข้องภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม ทั้งสองแนวคิดดังกล่าวข้างต้น ถือเป็นประเด็นสำคัญในการกำหนดแนวทางการประเมินผลกระทบจากสิ่งแทรกแซงที่นักวิจัยจำเป็นต้องทบทวนและตัดสินใจ เนื่องจากคำถามวิจัยภายใต้แนวคิดทั้งสองดังกล่าว มีแนวทางในการตัดสินใจเลือกวิธีการทางสถิติและกำหนดแนวทางในการวิเคราะห์ข้อมูลเพื่อประเมินผลกระทบที่แตกต่างกันอย่างชัดเจน นั่นคือ คำถามวิจัยภายใต้แนวคิดที่ 1 ซึ่งเป็นการนำข้อมูลก่อน-หลังทดลองมาพิจารณาแบบวัดซ้ำ จะมีผลกระทบที่สนใจ ประกอบด้วย (1) ผลกระทบหลักที่มาจากตัวแปรกลุ่ม, (2) ผลกระทบหลักที่มาจากตัวแปรจุดเวลา, และ (3) ผลกระทบร่วมระหว่างตัวแปรกลุ่มและตัวแปรจุดเวลา ซึ่งวิธีการทางสถิติที่ควรถูกเลือกนำมาใช้ที่เหมาะสม ได้แก่ การวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (repeated measures ANOVA, RM-ANOVA) (Winkens, van Breukelen et al., 2007) ขณะที่คำถามวิจัยภายใต้แนวคิดที่ 2 ซึ่งเป็นการนำข้อมูลก่อนทดลองมาพิจารณาเป็นตัวแปรร่วม หรือ covariate และให้ข้อมูลหลังทดลองเป็นตัวแปรผลลัพธ์ หรือ outcome จะให้ความสนใจผลกระทบที่เกิดขึ้นเพียงกรณีเดียว ได้แก่ ผลกระทบหลักที่มาจากตัวแปรกลุ่ม หรือ treatment effect เท่านั้น ซึ่งวิธีการทางสถิติที่ควรถูกเลือกนำมาใช้ที่เหมาะสม ได้แก่ การวิเคราะห์ความแปรปรวนร่วม (analysis of covariance, ANCOVA) (Winkens, van Breukelen et al., 2007) ดังนั้นประเด็นคำถามวิจัยที่เกี่ยวข้อง จึงถือเป็นอีกประเด็นหนึ่งที่มี

ความสำคัญสำหรับนักวิจัยในการนำมาพิจารณา เพื่อให้เกิดความชัดเจนและเป็นแนวทางในการประเมินผลกระทบจากสิ่งแทรกแซงที่ศึกษา รวมถึงการเป็นข้อมูลประกอบการตัดสินใจเพื่อเลือกวิธีการทางสถิติให้มีความถูกต้องและเหมาะสมกับผลกระทบที่ต้องการค้นหาคำตอบอีกด้วย

1.3.2 ผลกระทบของวิธีการ

ในงานวิจัยทางวิทยาศาสตร์สุขภาพ โดยเฉพาะทางด้านการแพทย์และสาธารณสุข “ผลกระทบของวิธีการ” หรือ “treatment effect” ถูกกล่าวถึงและนำมาใช้ในงานวิจัยเชิงทดลองก่อนข้างแพร่หลาย เพื่อบ่งชี้ข้อค้นพบสำคัญที่ได้จากการเปรียบเทียบความแตกต่างของตัวแปรผลลัพธ์ระหว่างกลุ่มทดลองและกลุ่มควบคุม ดังนั้น การทำความเข้าใจเกี่ยวกับผลกระทบของวิธีการ จึงมีผลโดยตรงต่อการกำหนดแนวทางการประเมินผลกระทบจากสิ่งแทรกแซง อีกทั้งยังส่งผลต่อเนื่องไปสู่ความชัดเจนในการตัดสินใจเลือกใช้วิธีการทางสถิติและการกำหนดแนวทางการวิเคราะห์ข้อมูลให้มีความถูกต้องและเหมาะสม โดยมีประเด็นที่เกี่ยวข้อง ดังนี้

- (1) ความหมาย “ผลกระทบของวิธีการ หรือ treatment effect”
- (2) ชนิดผลกระทบของวิธีการ
- (3) ตัวแบบผลกระทบของวิธีการ

(1) ความหมาย “ผลกระทบของวิธีการ หรือ treatment effect” หมายถึง การเปลี่ยนแปลง หรือ ความแตกต่างที่เกิดขึ้นของตัวแปรผลลัพธ์ที่ถูกพิจารณาจากผู้เข้าร่วม หรือ ค่าสังเกตของผู้เข้าร่วมในกลุ่มที่ได้รับสิ่งแทรกแซง (กลุ่มทดลอง) เทียบกับกลุ่มที่ไม่ได้รับสิ่งแทรกแซง (กลุ่มควบคุม) (Garrido, Dowd et al., 2016) ดังเช่น กรณีตัวอย่างในแผนภาพที่ 1.6 ข้างต้น พบว่า ผลกระทบของวิธีการ หรือ treatment effect หมายถึง การเปลี่ยนแปลง หรือ ความแตกต่างที่เกิดขึ้นของปริมาณค่าดัชนีมวลกาย (BMI) ที่วัดได้จากผู้เข้าร่วมในกลุ่มที่ได้รับรูปแบบการทานแบบวันเว้นวัน (alternate-day fasting, ADF) เมื่อเทียบกับรูปแบบการทานแบบอดอาหาร 16 ชั่วโมงและทานภายใน 8 ชั่วโมง (16/8 time-restricted fasting, 16/8 TRF) เป็นต้น

(2) ชนิดผลกระทบของวิธีการ จากความหมายผลกระทบของวิธีการ หรือ treatment effect ดังกล่าวข้างต้น ในทางปฏิบัติ ยังสามารถถูกแบ่งย่อยออกตามประเด็นคำถามที่นำมาพิจารณาเพื่อต้องการค้นหาคำตอบที่หลากหลาย ภายใต้เงื่อนไข หรือ ความจำเพาะที่แตกต่างกันได้ (Matsouaka, Liu et al., 2023) ดังนี้

- 2.1) ผลกระทบของวิธีการโดยเฉลี่ย (average treatment effect, ATE)
- 2.2) ผลกระทบของวิธีการโดยเฉลี่ยในกลุ่มได้รับวิธีการ (average treatment effect of the treated, ATT)
- 2.3) ผลกระทบของวิธีการโดยเฉลี่ยในกลุ่มไม่ได้รับวิธีการ (average treatment effect on the controls, ATC)
- 2.4) ผลกระทบของวิธีการโดยเฉลี่ยแบบมีเงื่อนไข (conditional average treatment effect, CATE)