KKU Journal for Public Health Research
Volume 19 Issue 2 May – August 2026

วารสารวิจัยสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น
ปีที่ 19 ฉบับที่ 2 เดือน พฤษภาคม – สิงหาคม 2569

นิพนธ์ต้นฉบับ

การออกแบบการวิจัยเชิงทดลองเพื่อเปรียบเทียบประสิทธิผลในรูปแบบความเหนือกว่า (Superiority), ความเท่าเทียม (Equivalence) และความไม่ด้อยกว่า (Non-inferiority)

เขมจิรา การเกตกลาง[(1)]*

วันที่ได้รับต้นฉบับ: 20 กุมภาพันธ์ 2569
วันที่ได้รับแก้ไข: 18 มีนาคม 2569
วันที่ตอบรับการตีพิมพ์: 20 มีนาคม 2569

** ผู้รับผิดชอบบทความ*
Email: oy.kemajira@hotmail.com
(1) ภาควิชาอายุรศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล

บทคัดย่อ

การวิจัยเชิงทดลอง โดยเฉพาะอย่างยิ่งในสาขาการแพทย์ การเปรียบเทียบวิธีการรักษาใหม่กับวิธีการรักษาแบบเดิมการดูแลรักษาที่มีอยู่เดิมเพื่อประเมินประสิทธิภาพเชิงเปรียบเทียบนั้นมีความสำคัญอย่างยิ่ง โดยทั่วไปแล้ว การออกแบบการทดลองจะแบ่งออกเป็น 4 ประเภทหลักตามคำถามวิจัยที่พบบ่อย ได้แก่ การทดลองแบบเปรียบเทียบความแตกต่าง การทดลองแบบเปรียบเทียบความเหนือกว่า การทดลองแบบเปรียบเทียบความเท่าเทียม และการทดลองแบบเปรียบเทียบความไม่ด้อยกว่า ความเข้าใจอย่างลึกซึ้งเกี่ยวกับการออกแบบเหล่านี้เป็นสิ่งจำเป็นสำหรับการตีความผลลัพธ์ทางคลินิกอย่างถูกต้อง ซึ่งจะช่วยให้สามารถระบุวิธีการรักษาที่มีประสิทธิภาพมากที่สุดและนำไปใช้กับผู้ป่วยได้อย่างเหมาะสม ส่งผลให้ได้รับประโยชน์ทางคลินิกสูงสุดและลดข้อผิดพลาดในการปฏิบัติทางการแพทย์ให้น้อยที่สุด

คำสำคัญ: *ความเหนือกว่า, ความเท่าเทียม, ไม่ด้อยกว่า, ช่วงเชื่อมั่นของความต่าง*

KKU Journal for Public Health Research
Volume 19 Issue 2 May – August 2026

วารสารวิจัยสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น
ปีที่ 19 ฉบับที่ 2 เดือน พฤษภาคม – สิงหาคม 2569

Original Article

A Comparative Study Evaluating the Therapeutic Efficacy through Superiority, Equivalence, and Non-Inferiority Trial Designs

Khemajira Karaketklang[(1)]*

Received Date: February 20, 2026
Revised Date: March 18, 2026
Accepted Date: March 20, 2026

** Corresponding author*
Email: oy.kemajira@hotmail.com
(1) Department of Medicine Faculty of Siriraj Hospital, Mahidol University

Abstract

In experimental research, particularly within the medical field, it is crucial to compare novel interventions with existing standards of care to evaluate their relative efficacy. Generally, experimental designs are categorized into four common types based on the specific research question: difference, superiority, equivalence, and non-inferiority trials. A profound understanding of these designs is essential for the accurate interpretation of clinical outcomes. This ensures that the most effective treatments are identified and appropriately administered to patients, thereby maximizing clinical benefits while minimizing errors in medical practice.

***Keywords*:** *Superiority, Equivalence, Non-Inferiority, Confidence Interval of the Difference*

บทนำ

การวิจัยเชิงทดลองเป็นกระบวนการที่มุ่งเน้นการศึกษาความสัมพันธ์เชิงเหตุและผลเพื่อทดสอบสมมติฐานอย่างเป็นระบบ โดยมีการเปรียบเทียบระหว่างวิธีการรักษาใหม่หรือยาชนิดใหม่กับวิธีการรักษาแบบเดิม การวิจัยในลักษณะนี้ต้องอาศัยการวางแผนการทดลองที่รัดกุมเพื่อควบคุมตัวแปรแทรกแซงที่อาจส่งผลกระทบต่อผลลัพธ์ ทั้งนี้มีวัตถุประสงค์เพื่อประเมินความปลอดภัยและประสิทธิภาพของสิ่งที่ศึกษาได้อย่างแม่นยำ ภายใต้เกณฑ์การยอมรับ (margin) ที่กำหนดไว้ว่าผลการทดสอบมีความเหนือกว่า (Superiority) เท่าเทียมกัน (Equivalence) หรือไม่ด้อยกว่า (Non-inferiority) ซึ่งเมื่อการควบคุมการวิจัยดำเนินไปอย่างเหมาะสม ผลลัพธ์ที่ได้จะสามารถนำมาอธิบายในรูปแบบต่าง ๆ ดังนี้ (อรุณ จิรวัฒน์กุล, 2550)

1. การประมาณค่าแบบจุด (Point estimate) คือ เป็นการหาค่าสถิติค่าเดียวของตัวแปรที่ได้จากกลุ่มตัวอย่าง ไปประมาณค่าพารามิเตอร์ของประชากร ซึ่งค่าที่ได้อาจมีความใกล้เคียงกับกลุ่มประชากร หรือมีความแตกต่างจากกลุ่มประชากร ขึ้นอยู่ผลการศึกษาที่เกิดขึ้นจริงหรือมีผลจากการสุ่มตัวอย่าง ถึงแม้จะมีกลุ่มตัวอย่างขนาดใหญ่ก็ตาม ดังนั้น หากเกิดความผิดพลาดจากการสุ่มตัวอย่าง การประมาณค่าแบบจุดก็จะไม่มีความน่าเชื่อ

2. การหาความแตกต่าง (Comparison) เป็นการทดสอบสมมติฐาน โดยการพิจารณาจากค่าระดับนัยสำคัญทางสถิติ (P-value) ซึ่งเมื่อพบความแตกต่างของการทดสอบทางสถิติ คือ ผลที่ได้จะต้องมีค่า P-value น้อยกว่าค่าระดับนัยสำคัญที่กำหนดไว้ เช่น ค่า P-value = 0.02 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญ 0.05 จะสามารถสรุปได้ว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

3. การประมาณค่าแบบช่วงเชื่อมั่น (Confidence Interval) เป็นการหาค่าสถิติที่ใช้ประมาณค่าพารามิเตอร์ของประชากรแบบช่วง โดยพิจารณาจากความน่าจะเป็นของค่าสถิติและความน่าจะเป็นที่ได้จากการสุ่มตัวอย่าง มีค่าความคลาดเคลื่อนมาตรฐาน (Standard Error: SE) มาใช้ในการประมาณค่าช่วงเชื่อมั่น การประมาณค่าแบบช่วงสามารถใช้พิจารณาการทดสอบสมมติฐาน เพื่อตัดสินใจสรุปผลการทดสอบ โดยพิจารณาจากช่วงเชื่อมั่นที่ต้องไม่ได้คร่อมศูนย์หรือมีศูนย์อยู่นอกช่วงเชื่อมั่น หากผลลัพธ์เป็นข้อมูลเชิงกลุ่ม จะพิจารณาจากช่วงเชื่อมั่นที่ต้องไม่ได้คร่อมหนึ่งหรือมีหนึ่งอยู่นอกช่วงเชื่อมั่น โดยการทดสอบสมมติฐานเป็นได้ทั้งแบบทางเดียว (One-tail alternative hypothesis) และสองทาง (Two-tail alternative hypothesis) แต่ส่วนใหญ่ถ้าต้องการวัดความเหนือกว่า มักทำการประมาณค่าแบบทางเดียว

บทความนี้ มีวัตถุประสงค์เพื่อให้ผู้อ่านได้มีความเข้าใจเบื้องต้นในงานวิจัยทางคลินิกที่มีรูปแบบการวิจัยเชิงทดลอง (Experimental) ที่มีการเปรียบเทียบผลการรักษาระหว่าง 2 วิธีการรักษา เพื่อพิจารณาความแตกต่างของผลการรักษา อย่างเช่น อัตราส่วน (odds ratio) สัดส่วน (proportion) หรือค่าเฉลี่ย (mean) เป็นต้น ซึ่งผู้วิจัยต้องมีการกำหนดขอบเขตการตัดสินใจของความแตกต่างระหว่างผลการรักษาของ 2 วิธีการ (margin: δ) ไว้ล่วงหน้า ว่าการรักษาใหม่นั้น "ดีกว่า หรือเหนือกว่า" “ความเท่าเทียม” หรือ “ไม่ด้อยกว่า” เทียบกับวิธีการแบบเดิม ซึ่งจะช่วยให้มีสามารถสรุปผลการทดลองได้ถูกต้องมากขึ้น โดยมีรูปแบบการทดลองเชิงการเปรียบเทียบ ผลการรักษาระหว่าง 2 วิธีการรักษา แบ่งออกเป็น 3 รูปแบบ ได้ดังนี้

Superiority trial

Superiority trial คือ การออกแบบการวิจัยเชิงทดลองต้องการดูว่า ผลการรักษาด้วยวิธีการรักษาใหม่เหนือกว่าการรักษาด้วยวิธีการแบบเดิมหรือไม่ โดยมีความคาดหวังว่า การรักษาด้วยวิธีใหม่ต้องดีกว่าการรักษาด้วยวิธีการแบบเดิม ค่าความต่างของข้อมูลจะเกี่ยวเนื่องกับอาการทางคลินิกที่ได้จากการพิจารณาจากขนาดที่มีความต่างจากการรักษาด้วยวิธีการแบบเดิม การออกแบบการทดลองนี้ แสดงถึงความเหนือกว่า แบ่งออกเป็น 2 ประเภท คือ การทดสอบที่แสดงความเหนือกว่าด้านสถิติ (Statistical Superiority Trial) และการทดสอบที่แสดงความเหนือกว่าทางด้านคลินิก (Clinical Superiority Trial) โดยมีการตั้งสมมติฐาน (Chow, Shao & Wang, 2003; Wang & Ji, 2020) ดังตารางที่ 1 และสามารถอธิบายผลการเปรียบเทียบแบบ Superiority trial ดังภาพที่ 1 และ 2

กรณีตัวอย่างงานวิจัยที่ทำการทดลองแบบ Superiority trial

การศึกษาการให้ยา Ganciclovir (GCV) และ Valganciclovir (VGCV) ในผู้ป่วยที่ได้รับการปลูกถ่ายอวัยวะ (Solid-organ transplant - SOT) ตามคำแนะนำมาตรฐานของผู้ผลิต มักส่งผลให้ระดับยาในร่างกายสูงเกินไปหรือต่ำเกินไป การใช้แบบจำลองทางเภสัชจลนศาสตร์ประชากร (Population Pharmacokinetics) ร่วมกับการคาดการณ์แบบ Bayesian อาจช่วยเพิ่มประสิทธิภาพการให้ยา เพื่อให้ระดับยาในเลือด (ค่า AUC) อยู่ในเกณฑ์ที่เหมาะสมตามเป้าหมายการรักษา คณะผู้วิจัยได้ทำการศึกษาแบบสุ่มและเปิดเผยข้อมูล (Randomized, Open-label) เพื่อทดสอบสมมติฐานความเหนือกว่าที่ร้อยละ 40 (40% superiority trial) ในผู้ป่วยผู้ใหญ่ที่ได้รับ SOT ซึ่งต้องใช้ยาเพื่อป้องกันหรือรักษาการติดเชื้อไซโตเมกะโลไวรัส (CMV) โดยแบ่งผู้ป่วยเป็น 2 กลุ่ม ผู้ป่วยในกลุ่ม B คือ ปรับขนาดยาตามค่าพยากรณ์จากแบบจำลอง Bayesian (NONMEM) เพื่อให้ได้ระดับยาตามเป้าหมาย ทำระดับเป้าหมายสูงถึง 88.6% ขณะที่กลุ่ม A คือ ได้รับยาตามคำแนะนำมาตรฐานของบริษัทผู้ผลิต ทำได้เพียง 22.2% ซึ่งบรรลุขอบเขตความเหนือกว่า 40% (superiority margin(δs)) ($P < 0.001$; 95% confidence interval [CI] difference, 47% to 86%) ซึ่ง 95% CI of difference (T–S) > δs (Padullés et al., 2016)

Equivalence trial

Equivalence trial คือ การออกแบบการวิจัยเชิงทดลองที่ไม่ได้ต้องการทราบว่า ผลการรักษาด้วยวิธีใหม่เหนือกว่าการรักษาด้วยวิธีแบบเดิม แต่คาดหวังว่าการรักษาด้วยวิธีใหม่ให้ผลเท่ากับหรือใกล้เคียงกับการรักษาด้วยวิธีแบบเดิม โดยมีการกำหนดขอบเขตความเท่ากันทางคลินิก (Δ) คือ ขอบเขตความต่างที่ยอมรับได้ทางคลินิก เช่น ถ้าการรักษา 2 วิธี พบว่า มีความเท่ากัน ช่วงเชื่อมั่นของความต่างควรจะอยู่ในช่วง $-\Delta$ ถึง $+\Delta$ และช่วงเชื่อมั่นคร่อมศูนย์ เช่น ผู้วิจัยทราบมาแล้วว่า ยาชนิดใหม่มีผลข้างเคียงของยา และราคาต่ำกว่ายาชนิดเก่า ถ้าผลลัพธ์ออกมาพบว่า ยาชนิดใหม่ให้ผลการรักษาเท่ากับหรือใกล้เคียงกับยาชนิดเก่า โดยที่มีผลข้างเคียงและราคาต่ำกว่ายาตัวเก่า ก็จะพิจารณาเลือกใช้ยาตัวใหม่แทน ในการทดสอบสมมติฐานจะทำการทดสอบแบบสองทาง (Chow, Shao, & Wang, 2003; Wang & Ji, 2020) ดังตารางที่ 1 และสามารถอธิบายผลการเปรียบเทียบแบบ Equivalence trial ดังภาพที่ 3

กรณีตัวอย่างงานวิจัยที่ทำการทดลองแบบ Equivalence trial

การศึกษาประสิทธิผลในการปรับแต่งสันเหงือกส่วนหน้าของเพดานเทียม 2 ชนิด ในการรักษาทารกปากแหว่งเพดานโหว่ด้านเดียวแบบสมบูรณ์ ผลการศึกษาพบว่า กลุ่มตัวอย่างเป็นทารกปากแหว่งเพดานโหว่ด้านเดียวแบบสมบูรณ์ (อายุตั้งแต่ 1 ถึง 26 วัน อายุเฉลี่ย 6.5 วัน) จำนวน 36 ราย กลุ่มตัวอย่างถูกแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มควบคุม 20 ราย และกลุ่มทดลอง 16 ราย กลุ่มควบคุมได้รับการใส่เพดานเทียมอะคริลิก (Acrylic obturator) ส่วนกลุ่มทดลองได้รับการใส่เพดานเทียมพอลิไวนิล (Polyvinyl obturator) กลุ่มตัวอย่างทั้งหมดได้รับการรักษาตามขั้นตอนมาตรฐานเดียวกันจนกระทั่งมีสุขภาพพร้อมสำหรับการผ่าตัดริมฝีปากเมื่ออายุประมาณ 3–4 เดือน ทำการวัดความกว้าง ความสูงของโค้งปลายเบ้าฟันและขนาดช่องโหว่จากแบบจำลองสันเหงือก จากการเปรียบเทียบค่าสัดส่วนของโค้งปลายเบ้าฟันที่เปลี่ยนแปลงไปหลังใส่เพดานเทียมด้วยสถิติทดสอบที เพื่อทดสอบความเท่าเทียมกัน โดยมีขอบเขตของความเท่าเทียมกันเท่ากับ 0.1 พบว่า สัดส่วนของโค้งปลายเบ้าฟันของทั้ง 2 กลุ่ม ค่าเฉลี่ยของสัดส่วนของโค้งปลายเบ้าฟัน ในกลุ่ม Acrylic obturator 0.76±0.09 และ ค่าเฉลี่ยของสัดส่วนของโค้งปลายเบ้าฟัน ในกลุ่ม Polyvinyl obturator 0.69±0.06 และรายงานค่า Mean difference เท่ากับ 0.002 (95% CI: -0.045, 0.049) ซึ่งช่วง 95% CI ของ Mean difference อยู่ในช่วง $-0.1 < (T-S) < 0.1$ มีความเท่าเทียมกันอย่างมีนัยสำคัญทางสถิติ ($p < 0.001$) แสดงว่าเพดานเทียมพอลิไวนิลให้ผลเรื่องการปรับแต่งสันเหงือกส่วนหน้าในทารกปากแหว่งเพดานโหว่ด้านเดียวแบบสมบูรณ์เท่าเทียมกันกับเพดานอะคริลิก (วรรณดี พลานุภาพ, ศิริชัย ธรรมชาติอารี, & พิมพ์นราพร พู่ทองคำ, 2556)

Non-inferiority trial

Non-inferiority trial คือ การออกแบบการวิจัยเชิงทดลองที่ต้องการดูว่า ผลการรักษาด้วยวิธีใหม่ให้ผลใกล้เคียงกับผลที่คาดหวัง ขอบเขตที่ไม่ด้อยกว่าจะถูกระบุในการทดลองที่ไม่ด้อยกว่า ขอบเขตที่ไม่ด้อยกว่าคือขอบเขตความแตกต่างที่กำหนดไว้ล่วงหน้าระหว่างการรักษาแบบใหม่และแบบมาตรฐาน ซึ่งมักใช้กับการศึกษาที่มีการสูญเสียประสิทธิผลสูงสุดของการรักษาแบบมาตรฐานเมื่อเทียบกับการรักษาแบบใหม่สามารถยอมรับได้เพื่อแลกกับผลประโยชน์ที่รับได้บางประการของการรักษาแบบใหม่ ขอบเขตที่ไม่ด้อยกว่าถูกกำหนดโดยสัมพันธ์กับผลประโยชน์ที่กำหนดไว้ของการรักษาแบบมาตรฐาน ยกตัวอย่างเช่น การทดลองที่ต้องการเปรียบเทียบว่า ผลการรักษาด้วยวิธีใหม่ให้ผลไม่ด้อยกว่าการรักษาด้วยวิธีแบบเดิมไม่เกิน 10 % ผู้วิจัยถึงจะยอมรับได้หรือเรียกอีกอย่างว่า ผลการรักษาด้วยวิธีเก่าให้ผลดีกว่าการรักษาด้วยวิธีใหม่มากกว่า 10% จะพิจารณารูปแบบการทดลองนี้คือ Non-inferiority trial ซึ่งจะแสดงให้เห็นว่า การรักษาด้วยวิธีใหม่มีผลกระทบไม่น้อยกว่าการรักษาวิธีแบบเดิม หรือมีผลกระทบมากกว่าหรืออาจมีผลกระทบใกล้เคียงกันทั้งสองวิธี ช่วงเชื่อมั่นของทั้งสองวิธีจะมีความใกล้เคียงกัน

ในบางครั้งอาจพิจารณาได้จากอีกทางหนึ่งคือ ถ้าการทดลองต้องการเปรียบเทียบว่า ผลการรักษาด้วยวิธีใหม่ให้ผลการรักษาได้ดีกว่าผลการรักษาวิธีแบบเดิม แต่เมื่อทำการทดสอบออกมาพบว่า ช่วงเชื่อมั่นของความต่างมีค่า lower limit อยู่ต่ำกว่าเกณฑ์ที่การยอมรับและค่า upper limit เกินเกณฑ์การยอมรับ แสดงว่าผลที่ได้จากการวิเคราะห์ยังไม่ถึงรูปแบบการทดลองแบบ Superiority trial และไม่ถูกจัดให้อยู่แบบ Equivalence trial ดังนั้นจะถูกจัดให้อยู่ในรูปการทดลองแบบ Non-inferiority trial

ผลจากการวิเคราะห์ข้อมูลของการทดลองแบบ Non-inferiority trial สามารถวัดได้จากช่วงเชื่อมั่นเช่นเดียวกับวิธีการของ Equivalence trial คือมีค่า lower limit ของช่วงเชื่อมั่นมีค่าเหนือกว่า $-\Delta$ โดยค่า upper limit ของช่วงเชื่อมั่นไม่ได้จำกัดขอบเขตไว้ และช่วงเชื่อมั่นไม่คร่อมเกณฑ์ที่การยอมรับ โดยการทดสอบสมมติฐานเป็นได้ทั้งแบบทางเดียวและสองทาง แต่การทดสอบสมมติฐานแบบสองทางควรพิจารณาเกณฑ์ของ Non-inferiority trial ที่ตั้งไว้และความถูกต้องในการให้ค่า $-\Delta$ เพราะอาจมีการคิดผิด ซึ่งทำให้เกิดความสับสนกับการทดลองแบบ Equivalence trial ซึ่งการตั้งสมมติฐานของ Non-inferiority trial (Chow, Shao & Wang, 2003; Wang & Ji, 2020) ดังตารางที่ 1 และสามารถอธิบายผลการเปรียบเทียบแบบ Non-inferiority trial ได้ดังภาพที่ 4 และ 5

กรณีตัวอย่างงานวิจัยที่ทำการทดลองแบบ Non-inferiority trial แบบกรณีผลลัพธ์เชิงบวก

งานวิจัยที่ทำการเปรียบเทียบประสิทธิผลการระงับปวดระหว่างยาที่ไม่ใช่กลุ่มฝิ่น (Non-opioid), โคเดอีน (Codeine) และออกซิโคโดน (Oxycodone analgesic regimens) ในผู้ใหญ่ที่มีอาการปวดปานกลางจากการบาดเจ็บที่แขนขา เป็นการทดลองแบบสุ่ม เป็นการเปรียบเทียบแบบ non-inferiority trial โดยมีกลุ่มที่ใช้ยา oxycodone analgesic regimens 59 ราย และกลุ่มเปรียบเทียบ คือ กลุ่มที่ใช้ยา non-opioid 61 ราย และ กลุ่ม codeine 62 ราย เพื่อศึกษาความแตกต่างของการเปลี่ยนแปลงของมาตราส่วนภาพเปรียบเทียบค่าเฉลี่ยของคะแนนระดับความเจ็บปวด (VAS) ระหว่างกลุ่ม ที่จุด 30 นาที โดยกำหนด inferiority margin 13 ผลการศึกษาพบว่า การเปลี่ยนแปลงของคะแนนระดับความเจ็บปวด (VAS) จาก Baseline จนถึงจุดเวลา 30 นาทีของกลุ่มที่ non-opioid เทียบกับกลุ่ม codeine -2.6 (-8.8 ถึง 3.6); กลุ่ม non-opioid เทียบกับกลุ่ม oxycodone analgesic regimens -2.7 (-9.3 ถึง 3.9) กลุ่ม codeine เทียบกับกลุ่ม oxycodone analgesic regimens 0.1 (-6.6 ถึง 6.4) ซึ่งขีดจำกัดบนของช่วงเชื่อมั่นของค่าความแตกต่าง ในการเปรียบผลลัพธ์ของแต่ละคู่ไม่เกิน inferiority margin 13 สรุปผลได้ว่า ผลการระงับปวดที่จุดเวลา 30 นาที ของกลุ่ม non-opioid กลุ่ม codeine และกลุ่ม oxycodone analgesic regimens ไม่ด้อยกว่ากัน (Graudins et al., 2016)

กรณีตัวอย่างงานวิจัยที่ทำการทดลองแบบ Non-inferiority trial แบบกรณีผลลัพธ์เชิงลบ

การศึกษาประเมินยา Rilpivirine (RPV) ซึ่งเป็นยาต้านไวรัสกลุ่ม NNRTI ตัวใหม่ มีประสิทธิภาพไม่ด้อยกว่ายามาตรฐานอย่าง Efavirenz (EFV) หรือไม่ เมื่อใช้

ร่วมกับยาพื้นฐานกลุ่ม N(t)RTI 2 ชนิด ในผู้ป่วยผู้ใหญ่ที่เพิ่งเริ่มรักษาเป็นครั้งแรก (Treatment-naive) เป็นการทดลองแบบสุ่มแบบ 1:1 ได้ทำการคัดเลือกผู้ป่วยที่ติดเชื้อ HIV-1 (อายุ ≥18 ปี) ที่ไม่เคยได้รับการบำบัดด้วยยาต้านไวรัสมาก่อน และมีปริมาณไวรัสในพลาสมาที่คัดกรองแล้ว 5,000 copies per mL หรือมากกว่า และมีความไวต่อไวรัสต่อ N(t)RTI พื้นฐาน ผลลัพธ์หลักคือไม่ด้อยกว่า (12% margin ในการวิเคราะห์การถดถอยแบบลอจิสติกส์) ที่ 48 สัปดาห์ในแง่ของการตอบสนองที่ได้รับการยืนยัน (viral load < 50 copies per mL, defined by the intent-to-treat time to loss of virologic response [TLOVR] algorithm) ในผู้ป่วยทุกรายจะได้รับยาที่ศึกษาอย่างน้อย 1 dose จากผลการศึกษาทำการคัดกรองผู้ป่วยจำนวน 947 ราย และรับสมัครผู้ป่วย 340 รายในแต่ละกลุ่ม ผู้ป่วยร้อยละ 86 (291 รายจาก 340 ราย) ที่ได้รับ Rilpivirine อย่างน้อย 1 dose ตอบสนองต่อการรักษา เมื่อเปรียบเทียบกับผู้ป่วยร้อยละ 82 (276 รายจาก 338 ราย) ที่ได้รับ efavirenz อย่างน้อย 1 dose(ความแตกต่าง 3.5% (95% CI: –1.7 ถึง 8.8); non-inferiority < 0.001) ซึ่งขีดจำกัดล่างของช่วงเชื่อมั่นของค่าความแตกต่างผลตอบรับที่ยืนยันใน 48 สัปดาห์ในแบบจำลองการถดถอยแบบลอจิสติกส์มีค่ามากกว่า -12% และ -10% ซึ่งยืนยันความไม่ด้อยกว่าที่ขอบ 12% margin และ 10% margin (p<0·001) (Cohen et al., 2011)

ผู้วิจัยได้พิจารณารูปแบบการทดลองทีมีความเป็นไปได้คือ Equivalence trial หรือ Non-inferiority trial เพราะผลที่ได้มีความใกล้เคียง แต่ยาชนิดใหม่ ใช้ง่ายกว่า ให้ผลข้างเคียงน้อยกว่า และประหยัดค่าใช้จ่ายกว่า ซึ่งผู้เชี่ยวชาญต้องให้หลักการและเหตุผลในความเท่ากันก่อนการทดลองจึงจะสามารถบอกได้ชัดเจนว่ามีรูปแบบการทดลองเป็นแบบใด ในงานวิจัยนี้ผู้วิจัยให้เป็นรูปแบบการทดลองแบบ Equivalence trial เพราะผลที่ได้มีความใกล้เคียง แต่ยาชนิดใหม่ ใช้ง่ายกว่า ให้ผลข้างเคียงน้อยกว่า และประหยัดค่าใช้จ่ายกว่า

ข้อควรพิจารณา

ก่อนการออกแบบการทดลองว่าจะเป็นแบบ Superiority trial, Equivalence trial หรือ Non-inferiority trial จะต้องมีเหตุผลในการพิจารณาก่อนออกแบบการทดลอง (Committee for proprietary medicinal products (CPMP), 2001; Walker & Nowacki, 2010) ดังนี้

1. ขนาดตัวอย่างในการศึกษาต้องเพียงพอ ถ้าพบว่า มีความไม่แตกต่างของผลการศึกษา ต้องคำนวณหาอำนาจการทดสอบว่าเพียงพอหรือไม่
2. ขนาดตัวอย่างของการทดลองแบบ Equivalence trial ใหญ่กว่าการทดลองแบบอื่น เพราะเป็นพิจารณาจากช่วงเชื่อมั่น ดังนั้นควรคำนวณขนาดตัวอย่างให้เพียงพอ เพื่อให้ผลการศึกษามีความชัดเจนและถูกต้องมากขึ้น
3. ต้องตั้งเกณฑ์การพิจารณาให้ชัดเจนก่อนว่าต้องการผลการศึกษาการทดลองแบบ Equivalence trial หรือ Non-inferiority trial เพื่อเป็นการไม่สับสนในการกำหนดรูปแบบการทดลองที่เหมาะสม
4. ควรมีการวางแผนการวิเคราะห์ข้อมูลอย่างเหมาะสม
5. การอธิบายผลการวิเคราะห์ควรอธิบายจากช่วงเชื่อมั่นจะสามารถอธิบายผลได้ดีกว่าการทดสอบสมมติฐาน
6. การออกแบบการทดสอบสมมติฐานทั้งทางเดียวและสองทางโดยใช้ช่วงเชื่อมั่นในการพิจารณาผลลัพธ์ที่ได้นั้น ในการทดสอบสมมติฐานแบบสองทางส่วนใหญ่จะใช้ในงานวิจัยทดลองทางคลินิกเพราะไม่ทราบแน่ชัดถึงผลที่จะได้ทางการรักษา แต่ท่ามกลางคุณประโยชน์ของยาชนิดใหม่ เมื่อเทียบกับยาชนิดเก่าจะสามารถบรรยายง่ายกว่าการทดลองแบบทางเดียว แต่ถ้าเป็นการทดสอบสมมติฐานแบบทางเดียว คุณประโยชน์ของยาชนิดใหม่มีความชัดเจนและถูกต้องมากขึ้น เพื่อเพิ่มความเชื่อมั่นว่า ยาชนิดใหม่ดีกว่ายาชนิดเก่าแล้วทำการทดลองอีกครั้งเพื่อดูผลลัพธ์ที่ได้จากการทดลอง และการลดโอกาสความผิดพลาดในการตัดสินใจให้น้อยลด เช่น จาก 95% CI เป็น 97.5% CI จะเป็นการเพิ่มความเชื่อมั่นของผลลัพธ์ที่ได้มากขึ้น
7. ควรควบคุมคุณภาพการทดลองให้ดีเพื่อป้องกันไม่ให้ผลของการวิเคราะห์ข้อมูลผิดไปจากคำถามการวิจัยที่ตั้งไว้ในตอนแรก
8. ส่วนวิเคราะห์ควรรายงานกลุ่มผู้ป่วยที่วิเคราะห์อย่างชัดเจน และรายงานผลการวิเคราะห์ทั้งแบบ intention to treat (ITT) และ per protocol (PP)

สรุป

บทความนี้ มีวัตถุประสงค์เพื่อให้ผู้อ่านสามารถเข้าใจการทดสอบและการแปลความหมายของผลการทดสอบความแตกต่างระหว่างกลุ่มในงานวิจัยทางคลินิกแบบ Superiority trial , Equivalence trial หรือ Non-inferiority trial โดยเน้นการพิจารณาการตั้งสมมติฐาน มีการวาดทิศทางของความต่างระหว่างกลุ่ม การพิจารณาผลลัพธ์ของผลการทดสอบ พร้อมยกตัวอย่างงานวิจัยทางคลินิกที่มีการทดสอบ ทั้ง 3 แบบ เพื่อให้มีความเข้าใจมากขึ้น แต่การพิจารณาค่า margin ขึ้นอยู่กับความเหมาะสมของงานวิจัย หากกำหนดน้อยไปก็อาจทำให้เสียโอกาสที่ได้ใช้การรักษาใหม่ที่มีประสิทธิภาพที่ใกล้เคียงกับการรักษาแบบเดิม ที่อาจจะมีค่าใช้จ่ายที่น้อยกว่า หรือหากกำหนดมากเกินไปอาจทำให้ได้รับการรักษาประสิทธิภาพไม่ดีพอเมื่อเทียบกับการรักษาแบบเดิม ซึ่งผู้วิจัยควรคำนึงถึงข้อควรพิจารณาในการออกแบบการวิจัยให้เหมาะสมกับกลุ่มตัวอย่าง วิธีการรักษาหรือเครื่องมือ การควบคุมคุณภาพการทดลองเพื่อให้ได้ผลลัพธ์ที่ถูกต้องมากที่สุด

เอกสารอ้างอิง

วรรณดี พลานุภาพ, ศิริชัย ธรรมชาติอารี, & พิมพ์นราพร พู่ทองคำ. (2556). ประสิทธิผลในการปรับแต่งสันเหงือกส่วนหน้าของเพดานเทียม 2 ชนิด ในการรักษาทารกปากแหว่งเพดานโหว่ด้านเดียวแบบสมบูรณ์. **วิทยาสารทันตแพทยศาสตร์, 63**(4), 156-164.

อรุณ จิรวัฒน์กุล. (2550). **ชีวสถิติ** (พิมพ์ครั้งที่ 4). ขอนแก่น: ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น.

Chow, S. C., Shao, J., & Wang, H. (2003). **Sample size calculations in clinical research.** New York: Marcel Dekker.

Cohen, C. J., Andrade-Villanueva J., Clotet, B., Fourie, J., Johnson, M. A., Ruxrungtham, K., et al. (2011). Rilpivirine versus efavirenz with two background nucleoside or nucleotide reverse transcriptase inhibitors in treatment-naive adults infected with HIV-1 (THRIVE): A phase 3, randomized, non-inferiority trial. **Lancet**, **378**(9787), 229-237.

Committee for proprietary medicinal products (CPMP). (2000). **Points to consider on switching between superiority and non-inferiority.** Retrieved April 20, 2025, https://www.tga.gov.au/sites/default/files/2024-12/enswitching_between_superiority_and_non-inferiority.pdf

Committee for Proprietary Medicinal Products. (2001). Points to consider on switching between superiority and non-inferiority. **British Journal of Clinical Pharmacology, 52**(3), 223-228.

Deere, K. C., Whitehouse, M. R., Porter, M., Blom, A. W., & Sayers, A. (2016). Assessing the non-inferiority of prosthesis constructs used in hip replacement using data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man: A benchmarking study. **BMJ Open, 9**(4), e026685.

Graudins, A., Meek, R., Parkinson, J., Egerton-Warburton, D., & Meyer, A. (2016). A randomized controlled trial of paracetamol and ibuprofen with or without codeine or oxycodone as initial analgesia for adults with moderate pain from limb injury. **Emergency Medicine Australasia, 28**(6), 666-672.

Padullés, A., Colom, H., Bestard, O., Melilli, E., Sabé, N., Rigo, R., et al. (2016). Contribution of population pharmacokinetics to dose optimization of ganciclovir-valganciclovir in solid-organ transplant patients. **Antimicrobial Agents and Chemotherapy, 60**(4), 1992-2002.

Piaggio, G., Elbourne, D. R., Pocock, S. J., Evans S. J. W., & Altman, D. G. (2012). Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. **JAMA, 308**(24), 2594–2604.

Sayers, A., Crowther, M. J., Judge, A., Whitehouse, M. R., & Blom, A. W. (2017). Determining the sample size required to establish whether a medical device is non-inferior to an external benchmark. **BMJ Open, 7**(8), e015397.

Schumi, J., & Wittes, J. T. (2011). Through the looking glass: understanding non-inferiority. **Trials, 12**, 106.

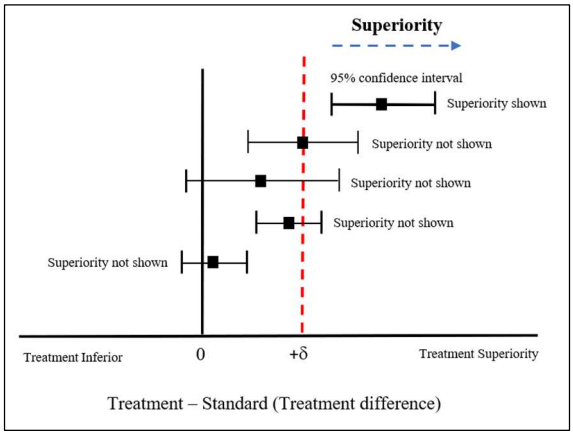
Walker, E., & Nowacki, A. S. (2011). Understanding equivalence and noninferiority testing. **Journal of General Internal Medicine, 26**(2), 192–196.

Wang, X., & Ji, X. (2020). Sample size estimation in clinical research: From randomized controlled trials to observational studies. **Chest, 158**(1S), S12–S20.

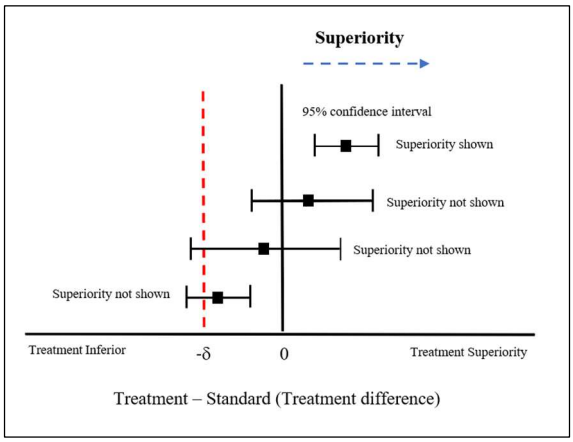
ตารางที่ 1 การกำหนดทดสอบสมมติฐานหลักและสมมติฐานรองในรูปแบบการทดลองเชิงเปรียบเทียบ

การทดลองเชิงเปรียบเทียบ	สมมติฐานหลัก (Null hypothesis : H_0)	สมมติฐานรอง (Alternative hypothesis : H_a)
ความเหนือกว่า (Superiority)	T – S ≤ superiority margin (δs)	T – S > superiority margin (δs)
เท่าเทียมกัน (Equivalence)	|T – S| ≥ equivalence margin (δe)	|T – S| < equivalence margin (δe) ซึ่งเท่ากับ -δe < (T-S) < δe
ไม่ด้อยกว่า (Non-inferiority)	T – S ≥ noninferiority margin (δni) T – S ≤ - noninferiority margin (δni)	T – S < noninferiority margin(δni) เมื่อ δni > 0 T – S > - noninferiority margin (δni) เมื่อ δni < 0

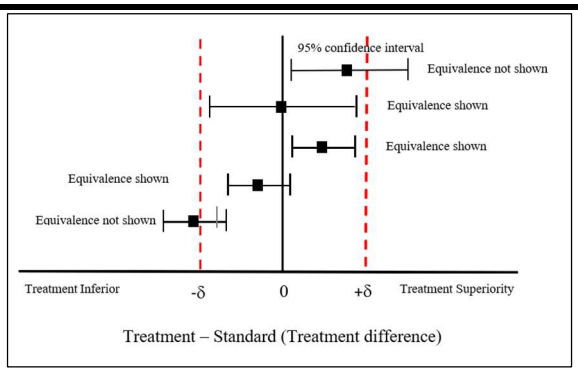
Treatment (T) and Standard (S).



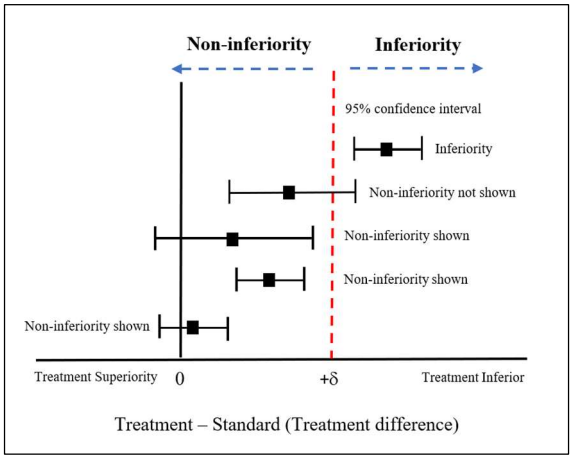
ภาพที่ 1 การพิจารณาความเหนือกว่า หรือ ดีกว่า ของวิธีการรักษาใหม่เทียบกับวิธีการรักษาแบบเดิม โดยผลการรักษาทางบวก เช่น อัตราความสำเร็จ (Success rate) (Committee for proprietary medicinal products (CPMP), 2000; Walker & Nowacki, 2010)



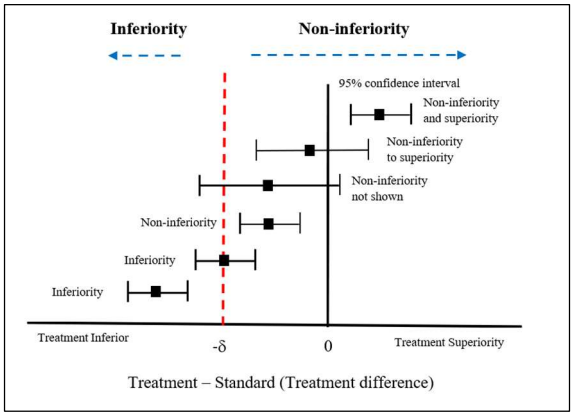
ภาพที่ 2 การพิจารณาความเหนือกว่า หรือ ดีกว่า ของวิธีการรักษาใหม่เทียบกับวิธีการรักษาแบบเดิม โดยผลการรักษาทางลบ เช่น อัตราความผิดพลาด (Failure rate) (Committee for Proprietary Medicinal Product, 2000; Walker & Nowacki, 2010; Sayers et al., 2017)



ภาพที่ 3 การพิจารณาความเท่ากันของวิธีการรักษาใหม่เทียบกับวิธีการรักษาแบบเดิม (Committee for Proprietary Medicinal Product, 2000; Walker & Nowacki., 2010; Sayers et al., 2017; Schumi & Wittes., 2011)



ภาพที่ 4 การพิจารณาไม่ด้อยกว่าของวิธีการรักษาใหม่เทียบกับวิธีการรักษาแบบเดิม โดยผลการรักษาทางบวก เช่น การลดลงของระดับความปวด (Walker & Nowacki., 2010; Deere et al., 2016; Piaggio et al., 2012)



ภาพที่ 5 การพิจารณาไม่ด้อยกว่าของวิธีการรักษาใหม่เทียบกับวิธีการรักษาแบบเดิม โดยผลการรักษาทางลบ เช่น อัตราความผิดพลาด (Failure rate) (Walker & Nowacki., 2010; Sayers et al., 2017; Deere et al., 2016; Piaggio et al., 2012)