

รายงานการฝึกอบรมเชิงปฏิบัติการเรื่อง “การวิเคราะห์ข้อมูลงานวิจัยทางการแพทย์ขั้นสูง”

ครั้งที่ 15 ระหว่างวันที่ 4-8 สิงหาคม 2568

เพื่อเผยแพร่ในเวปไซด์สาขาวิชาวิทยาศาสตร์สุขภาพ

.....

ในปัจจุบันการพัฒนางานวิจัย ถือเป็นสิ่งสำคัญอย่างยิ่งในการพัฒนาองค์ความรู้ต่างๆอย่างเป็นระบบ และต่อเนื่อง ซึ่งการวิจัยต้องทราบถึงระเบียบวิธีวิจัย แบบแผนการวิจัย วิธีดำเนินการวิจัย รวมถึงการวิเคราะห์ ประมวลผล เพื่อหาคำตอบของสมมติฐานการวิจัยนั้นๆ ในการวิเคราะห์และประมวลผลข้อมูล การวิจัยด้วย คอมพิวเตอร์โปรแกรมสำเร็จรูปทางสถิติ นับว่าเป็นกระบวนการวิจัยภาคปฏิบัติ ที่มีความสำคัญและจำเป็น อย่างยิ่ง ที่ผู้วิจัยจะต้องมีความรู้ในการเลือกใช้สถิติตามแบบแผนการวิจัยและชนิดของข้อมูล เพื่อนำไปสู่ผลการ สรุปลการวิจัย ได้อย่างถูกต้องและเป็นข้อมูลที่น่าเชื่อถือ ดังนั้นผู้ทำงานวิจัยต้องมีความรู้ ความสามารถในการ ประมวลผลด้วยคอมพิวเตอร์ได้ อีกทั้งยังสามารถใช้วิธีการทางสถิติทั้งขั้นพื้นฐานและขั้นสูงได้อย่างดี การทำ วิจัยถือเป็นสิ่งที่มีประโยชน์อย่างยิ่งในการพัฒนางาน พัฒนาวิชาชีพ การพัฒนาวิถีชีวิตของมนุษย์ และเป็นการ พัฒนาองค์ความรู้ต่างๆ อย่างเป็นระบบและต่อเนื่อง ซึ่งในปัจจุบันองค์กรจำนวนมากได้ สนับสนุนและส่งเสริม ให้บุคลากรในสังกัดมีความรู้ความสามารถด้านการวิจัย เพื่อใช้ในการยกระดับคุณภาพ งานหรือคุณภาพชีวิต ใน การทำงานวิจัยนั้นต้องทราบถึงหลักการและองค์ประกอบของการเขียนโครงร่าง งานวิจัย ระเบียบวิธีการวิจัย รูปแบบการวิจัย รวมถึงการบันทึก และการจัดการข้อมูลการวิจัยเพื่อนำไปใช้ใน การวิเคราะห์ข้อมูลได้อย่าง ถูกต้อง ดังนี้

1. การวัดการเกิดโรค (Measuring the occurrence of disease)

การวัด outcome งานวิจัย คือ กระบวนการประเมินผลลัพธ์ที่เกิดขึ้นจากการนำผลผลิต (Output) ของ งานวิจัยไปใช้ประโยชน์ โดยวัดการเปลี่ยนแปลงที่เกิดขึ้นกับกลุ่มเป้าหมายทั้งด้านพฤติกรรม ความรู้ ทักษะ และทักษะ

การวิจัยทางระบาดวิทยามีพื้นฐานอยู่บนความสามารถในการวัดปริมาณการเกิดโรค (หรือเหตุการณ์ที่ เกี่ยวข้องกับสุขภาพอื่นๆ) ในประชากร การวิจัยนี้ต้องระบุให้ชัดเจนดังนี้

- (1) Case หมายถึง บุคคลในกลุ่มประชากรที่มีโรค หรือประสบเหตุการณ์ที่สนใจ (เช่น เสียชีวิต)
- (2) ประชากรที่ผู้ป่วยเกิดขึ้น
- (3) ระยะเวลาที่รวบรวมข้อมูล

ในระบาดวิทยา การนิยามความหมายของคำว่า “case” นั้นไม่ใช่เรื่องง่าย แม้แต่กับโรคที่เป็นที่รู้จักกัน ดี นิยามของกรณีทางระบาดวิทยาไม่จำเป็นต้องเหมือนกับนิยามทางคลินิก และนักระบาดวิทยามักถูกบังคับให้ ใช้การทดสอบวินิจฉัยที่รุกรานน้อยกว่าและราคาถูกกว่าที่แพทย์ใช้กันโดยทั่วไป อย่างไรก็ตาม เพื่อวัตถุประสงค์ ในการศึกษา สิ่งสำคัญคือต้องทำให้นิยามของกรณีเป็นมาตรฐาน ตัวอย่างเช่น “ case มะเร็ง” ควรครอบคลุม เฉพาะกรณีที่ได้รับการยืนยันทางจุลพยาธิวิทยาเท่านั้นหรือไม่

ควรรวมรอยโรค ณ ตำแหน่งเดิมด้วยหรือไม่ สำหรับมะเร็งของอวัยวะคู่ (เช่น เต้านม อัณฑะ ไต) จำนวน case ที่นับได้ควรสะท้อนถึงจำนวนผู้ป่วยที่เป็นมะเร็งหรือจำนวนอวัยวะที่ได้รับผลกระทบหรือไม่ นัก ระบาดวิทยาโรคมะเร็งยังสนใจที่จะวัดความถี่ของเหตุการณ์ที่เกี่ยวข้องกับสุขภาพอื่นๆ เช่น ผู้ที่สูบบุหรี่ ไข้ยา คุมกำเนิดชนิดรับประทาน หรือใช้บริการสุขภาพบางอย่าง อาจถูกนับเป็น case ได้

ข้อควรพิจารณาที่สำคัญอีกประการหนึ่งเมื่อต้องรับมือกับภาวะที่ไม่ทำให้เสียชีวิตซ้ำ (เช่น ไข้หวัด ธรรมดา) คือการตัดสินใจว่าในแต่ละบุคคล ควรนับเหตุการณ์หรือเหตุการณ์แต่ละครั้งเป็น case หรือนับเฉพาะ การเริ่มเป็นครั้งแรก ในบทนี้ เราสมมติว่าบุคคลหนึ่งๆ สามารถมีอาการของโรคที่สนใจได้เพียงครั้งเดียว อย่างไรก็ตาม

ก็ตาม มาตรการการเกิดขึ้นที่นำเสนอสามารถปรับเปลี่ยนให้ครอบคลุมถึงเหตุการณ์ที่เกิดขึ้นซ้ำได้ case สามารถระบุได้ผ่านทะเบียนโรค ระบบแจ้งเตือน ใบมรณะบัตร บทคัดย่อบันทึกทางคลินิก การสำรวจประชากรทั่วไป ฯลฯ อย่างไรก็ตาม สิ่งสำคัญคือต้องแน่ใจว่าตัวเลือกรวม case ทั้งหมดที่เกิดขึ้นในประชากรที่ศึกษา และไม่รวม case จากที่อื่น ตัวอย่างเช่น เมื่อวัดการเกิดโรคในเมืองใดเมืองหนึ่ง ควรนับ case ศึกษาทั้งหมดที่เกิดขึ้นในหมู่ผู้อยู่อาศัยรวมอยู่ในตัวเลือก แม้กระทั่งผู้ป่วยที่ได้รับการวินิจฉัยที่อื่น ในทางตรงกันข้าม กรณีที่ได้รับการวินิจฉัยในผู้ที่ปกติอาศัยอยู่ในที่อื่น ควรได้รับการยกเว้น

2. ประเภทของการศึกษาทางระบาดวิทยา

การวิจัยทางระบาดวิทยาแบ่งออกเป็น 2 ประเภทหลัก ๆ คือ การวิจัยเชิงสังเกต (Observational studies) และการวิจัยเชิงทดลอง (Experimental studies) โดยการวิจัยเชิงสังเกตจะศึกษาปรากฏการณ์ที่เกิดขึ้นตามธรรมชาติ ในขณะที่การวิจัยเชิงทดลองจะมีการจัดกระทำหรือเปลี่ยนแปลงปัจจัยที่ศึกษา

1. การวิจัยเชิงสังเกต (Observational studies):

1.1 การศึกษาเชิงพรรณนา (Descriptive studies): มุ่งเน้นการบรรยายลักษณะของโรคหรือปัญหาทางสุขภาพ รวมถึงการกระจายของโรคในประชากร กลุ่มคน หรือพื้นที่ที่ศึกษา ตัวอย่างเช่น รายงานผู้ป่วย (Case report), การศึกษาภาคตัดขวางเชิงพรรณนา (Cross-sectional descriptive study).

1.2 การศึกษาเชิงวิเคราะห์ (Analytic studies): มุ่งเน้นการวิเคราะห์ความสัมพันธ์ระหว่างปัจจัยเสี่ยงกับผลลัพธ์ทางสุขภาพ โดยมักจะมีการเปรียบเทียบระหว่างกลุ่มที่มีและไม่มีปัจจัยเสี่ยง.

1.3 การศึกษาแบบตัดขวาง (Cross-sectional study): เป็นการศึกษาที่เก็บรวบรวมข้อมูลเกี่ยวกับปัจจัยเสี่ยงและผลลัพธ์ในเวลาเดียวกัน.

1.4 การศึกษาแบบไปข้างหน้า (Prospective study/Cohort study): เป็นการศึกษาที่ติดตามกลุ่มคนที่ไม่เป็นโรคในปัจจุบัน และติดตามไปข้างหน้าเพื่อดูว่าใครเป็นโรคร้าง โดยมีการแบ่งกลุ่มตามการสัมผัสปัจจัยเสี่ยง.

1.5 การศึกษาแบบย้อนหลัง (Retrospective study/Case-control study): เป็นการศึกษาที่เริ่มจากกลุ่มคนที่ป่วย (กรณี) และกลุ่มคนที่ไม่ได้ป่วย (กลุ่มควบคุม) แล้วย้อนกลับเพื่อดูว่าในอดีตกลุ่มไหนมีความเสี่ยงมากกว่า

2. การวิจัยเชิงทดลอง (Experimental studies):

2.1 การทดลองทางคลินิกแบบสุ่ม (Randomized Controlled Trials - RCTs): เป็นการทดลองที่มีการสุ่มให้ผู้เข้าร่วมการวิจัยได้รับหรือไม่ได้รับปัจจัยที่ศึกษา (เช่น ยา หรือ การรักษา).

2.2 การศึกษาการแทรกแซงแบบไม่สุ่ม (Non-randomized intervention studies): เป็นการศึกษาที่มีการจัดกระทำปัจจัยที่ศึกษา แต่ไม่ได้มีการสุ่มให้ผู้เข้าร่วมการวิจัย.

3. การเลือกชนิดของการวิจัย

การเลือกชนิดของการวิจัยขึ้นอยู่กับวัตถุประสงค์ของการศึกษา ปัญหาที่ต้องการศึกษา และทรัพยากรที่มีอยู่ การศึกษาเชิงพรรณนาเหมาะสำหรับการสร้างสมมติฐาน ส่วนการศึกษาเชิงวิเคราะห์เหมาะสำหรับการทดสอบสมมติฐานและหาความสัมพันธ์ระหว่างปัจจัยเสี่ยงและผลลัพธ์ทางสุขภาพ

4. Systematic Review & Meta analysis

4.1 Systematic Review เป็นการทบทวนเอกสารที่เกี่ยวข้องกับเรื่องที่เรากำลังศึกษา จากนั้นทำการประเมิน วิเคราะห์ และ สังเคราะห์ แล้วเขียนรายงานขึ้นมาใหม่

4.2 Meta Analysis เป็นการทบทวนเอกสารที่เกี่ยวข้องกับเรื่องที่เรากำลังศึกษา โดยเลือกเอกสารที่ออกแบบการศึกษา (Design study) ที่เหมือนกัน แล้วนำข้อมูลจากหลากหลายแหล่งเหล่านั้น มาทำการวิเคราะห์โดยใช้ชุดข้อมูลรวมเหล่านั้น (Combine analysis) เพื่อสรุปผลการศึกษาเป็นรายงานสรุปเดี่ยว (Single Summary Result)

5. การวิเคราะห์ความสัมพันธ์ทางระบาดวิทยาโดยใช้ 2x2 Table

5.1 นักระบาดวิทยา มักวิเคราะห์ความสัมพันธ์ ระหว่าง Exposure กับ Outcome ในรูปของ 2x2 Table เพื่อหา

5.1.1 RR (กรณีที่เป็น Cohort /Prospective หรือ Intervention Study)

5.1.2 OR (กรณีที่เป็น Cross sectional , Case Control / Retrospective study)

มากกว่า การใช้ chi square Test

5.2 การวิเคราะห์ความสัมพันธ์ ด้วย OR ,RR นั้น เป็นการวิเคราะห์แบบง่าย ที่ไม่ซับซ้อน แต่มีข้อจำกัด คือ สามารถวิเคราะห์ ตัวแปรที่เป็น Nominal หรือ Ordinal เท่านั้น แต่ในความเป็นจริง การศึกษานั้นมีตัวแปรประเภท Scale ด้วย ซึ่งจำเป็นต้อง ใช้โปรแกรมสถิติต่างๆ เช่น SPSS (เป็นโปรแกรมที่นักศึกษาส่วนใหญ่คุ้นเคย มากกว่า strata หรือ Epi Info) เช่น Compare Mean (T Test , Anova), Logistic Regression , Time Series Analysis , Survival Analysis เป็นต้น

6. Cohort study

ในการศึกษาแบบโคฮอร์ต (cohort study) จะมีการติดตามกลุ่มผู้เข้าร่วม (cohort) เป็นระยะเวลาหนึ่ง การประเมินจะดำเนินการตั้งแต่เริ่มต้น ระหว่างการติดตามผล และเมื่อสิ้นสุดการติดตามผล ดังนั้น การศึกษาแบบโคฮอร์ตจึงเป็นการศึกษาเชิงประจักษ์และเชิงยาว (longitudinal) โดยอาศัยข้อมูลจากกลุ่มตัวอย่าง นอกจากนี้ยังเป็นการสังเกตการณ์และ (โดยปกติ) เป็นไปตามธรรมชาติ การวิเคราะห์สามารถทำได้สำหรับกลุ่มตัวอย่างทั้งหมดหรือสำหรับกลุ่มย่อยซึ่งสามารถนำมาเปรียบเทียบกันได้ เนื่องจากไม่มีการสุ่มตัวอย่างไปยังกลุ่มย่อยที่สนใจ จึงไม่สามารถระบุความสัมพันธ์ระหว่างสาเหตุและผลได้ และความสัมพันธ์ระหว่างตัวแปรต้องระบุเป็นความสัมพันธ์ที่อาจได้รับอิทธิพลจากปัจจัยกวนหรือไม่ก็ได้ กลุ่มตัวอย่างที่ศึกษาสามารถกำหนดแบบไปข้างหน้าหรือแบบย้อนหลังได้ และแต่ละวิธีก็มีข้อดีและข้อเสียของตัวเอง ประเด็นเหล่านี้และประเด็นอื่นๆ ได้รับการอธิบายโดยใช้ตัวอย่างประกอบ ข้อสังเกตพิเศษเกี่ยวกับการศึกษาแบบโคฮอร์ตในจิตเวชศาสตร์อินเดีย

7. Case-control study

การศึกษาแบบ Case-control เป็นการศึกษาเชิงสังเกตแบบย้อนหลัง (retrospective) ที่เปรียบเทียบกลุ่มผู้ที่มีภาวะหรือผลลัพธ์ที่สนใจ (กลุ่ม case) กับกลุ่มที่ไม่มีภาวะนั้น (กลุ่ม control) เพื่อหาความสัมพันธ์ระหว่างปัจจัยเสี่ยง (exposure) กับผลลัพธ์ โดยการมอย้อนกลับไปในอดีตเพื่อประเมินปัจจัยเสี่ยงที่อาจนำไปสู่ภาวะดังกล่าว

8. Clinical trial study (การทดลองทางคลินิก) คือ การวิจัยทางการแพทย์ที่ทำในคนอาสาสมัคร เพื่อประเมินความปลอดภัยและประสิทธิผลของวิธีการรักษาใหม่ ๆ เช่น ยา อุปกรณ์ทาง

การแพทย์ หรือวิธีการอื่น ๆ ที่อาจช่วยในการรักษา ป้องกันโรค หรือส่งเสริมสุขภาพ โดยเปรียบเทียบกับวิธีการรักษามาตรฐาน หรือยาหลอก เพื่อให้ได้ข้อมูลที่น่าเชื่อถือในการตัดสินใจใช้รักษาผู้ป่วยในอนาคต

9. การวิเคราะห์หัตถ์แปร (Multivariate Analysis) คือ เทคนิคทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลที่มีตัวแปรตั้งแต่สองตัวขึ้นไปพร้อมกัน เพื่อค้นหารูปแบบ ความสัมพันธ์หรือความเชื่อมโยงระหว่างตัวแปรต่างๆ และทำความเข้าใจผลกระทบของตัวแปรเหล่านั้นต่อผลลัพธ์ที่สนใจ เทคนิคนี้ช่วยให้ได้ข้อมูลเชิงลึกที่มากกว่าการวิเคราะห์แบบตัวแปรเดียวหรือสองตัวแปร และถูกนำไปประยุกต์ใช้ในหลากหลายสาขาเช่น การตลาด การพยากรณ์อากาศ และการดูแลสุขภาพ

10. Survival Analysis คืออะไร

Survival Analysis (การวิเคราะห์การรอดชีพ) หรือเรียกอีกอย่างว่า การวิเคราะห์ระยะปลอดเหตุการณ์ เป็นสาขาย่อยหนึ่งในสถิติ มีเป้าหมายเพื่อวิเคราะห์และสร้างแบบจำลองที่ทำนายเวลาที่ใช้ในการเกิดเหตุการณ์บางอย่างที่เราสนใจ (Time to Event หรือ Survival Time) เช่น ระยะเวลาการเสียชีวิตของผู้ป่วยมะเร็ง ระยะเวลาที่เครื่องจักรจะเสีย หรือระยะเวลาที่ลูกค้าจะย้ายค่าย เป็นต้น Survival Analysis มีการประยุกต์ใช้มากในทางการแพทย์ ส่วนในสาขาอื่น ๆ ก็มีการนำวิธีนี้ไปใช้ในการวิเคราะห์ข้อมูลที่สนใจ ระยะเวลาการเกิดเหตุการณ์ในลักษณะนี้ด้วยเช่นเดียวกัน

ในทางการแพทย์มีการประยุกต์ใช้ Survival Analysis ค่อนข้างเยอะ โดยตัวอย่างเหตุการณ์ที่สนใจ อาจจะเป็นการเสียชีวิต การเข้ารับการรักษาซ้ำ การออกจากโรงพยาบาล หรือเหตุการณ์อื่น ๆ ตัวอย่าง เช่น การทดลองเปรียบเทียบประสิทธิภาพการรักษา โดยแบ่งเป็นสองกลุ่มทดสอบที่มีการให้ยาต่างชนิดกัน เพื่อเปรียบเทียบระยะเวลาในการรักษาหาย

11. Longitudinal data analysis

ความหมายและขอบเขตของข้อมูลระยะยาว ข้อมูลระยะยาว (longitudinal data) สามารถเกิดขึ้นได้จากหลายรูปแบบ ภายใต้การศึกษา ระยะยาว (longitudinal study) ซึ่งครอบคลุมทั้งการ ศึกษาแบบสังเกต และแบบทดลอง โดยในสาขาทาง เศรษฐศาสตร์ หรือสังคมศาสตร์ อาจเรียกข้อมูล ดังกล่าวนี้นว่า “panel data” ซึ่งถือเป็นข้อมูลชนิดเดียวกันที่สามารถนำมาใช้เรียกแทนระหว่างกันได้ (exchangeable) แต่อย่างไรก็ตาม เนื่องจากวิธีการ ทางสถิติที่ถูกนำมาใช้ในการวิเคราะห์ข้อมูลระยะยาว ส่วนใหญ่มุ่งเน้นและให้ความสำคัญกับการพิจารณา ข้อมูลตัวแปรผลลัพธ์ที่มีความสัมพันธ์กัน (correlated data) ซึ่งแตกต่างจากวิธีการทางสถิติแบบเดิมที่ ตัวแปรผลลัพธ์ถูกพิจารณาแบบอิสระ (independent outcome) ดังนั้น เพื่อให้ให้นักวิจัยได้เข้าใจ แนวคิดและ หลักการของข้อมูลระยะยาวสอดคล้องกับวิธีการทาง สถิติสำหรับการวิเคราะห์ข้อมูลระยะยาวไปพร้อมกัน

ข้อมูลระยะยาว หมายถึง ข้อมูลที่ตัวแปร ผลลัพธ์มีความสัมพันธ์กัน โดยความสัมพันธ์ของข้อมูล ดังกล่าวครอบคลุมทั้งความสัมพันธ์ที่เกิดขึ้นจากการวัดซ้ำ (repeatedly measured data) ความสัมพันธ์ภายในกลุ่ม (cluster data) และความสัมพันธ์ภายใน คู่ (paired data) ขอบเขตของข้อมูลระยะยาว มีความเป็นไปได้ทั้งการศึกษาแบบย้อนกลับ (retrospective studies) แบบภาคตัดขวาง (cross-sectional studies) และแบบไปข้างหน้า (prospective studies) โดยครอบคลุมฐานข้อมูลที่มีการเก็บรวบรวมเพื่อตอบ วัตถุประสงค์เดียวกัน หรือมีการเชื่อมโยงระหว่างกัน บนระบบออนไลน์ หรือ อิเล็กทรอนิกส์ ที่สะสมกันมา อย่างต่อเนื่อง ลักษณะข้อมูลตัวแปรผลลัพธ์ที่สัมพันธ์กัน ลักษณะข้อมูลตัวแปรผลลัพธ์ที่สัมพันธ์กัน ถือเป็นลักษณะจำเพาะของข้อมูลระยะยาวที่พบและ จำเป็นต้องอาศัยวิธีการทางสถิติที่เหมาะสมมาใช้ในการวิเคราะห์ ซึ่งในทางปฏิบัติลักษณะข้อมูลตัวแปร ผลลัพธ์ที่สัมพันธ์กัน ประกอบด้วย (1) ลักษณะข้อมูล ที่ตัวแปรผลลัพธ์ใน

บุคคลเดิมถูกวัดซ้ำและ (2) ลักษณะ ข้อมูลที่ตัวแปรผลลัพธ์แต่ละบุคคลถูกวัดภายใต้กลุ่มเดียวกัน ซึ่งลักษณะข้อมูลนี้ยังแบ่งย่อยได้ดังนี้

(1) ลักษณะข้อมูลที่ตัวแปรผลลัพธ์ใน บุคคลเดิมถูกวัดซ้ำ เป็นข้อมูลที่มีการวัดซ้ำ (repeated measurements) โดย “แต่ละบุคคล จะถูกวัดค่า ตัวแปรผลลัพธ์ซ้ำ ตั้งแต่สองครั้งขึ้นไป ภายใต้ จุดเวลาที่แตกต่างกันตามที่นักวิจัยกำหนด” จึงทำให้ ข้อมูลตัวแปรผลลัพธ์ที่วัดได้ในแต่ละจุดเวลาของ บุคคลเดียวกัน มีความสัมพันธ์กัน

(2) ลักษณะข้อมูลที่ตัวแปรผลลัพธ์แต่ละ บุคคลถูกวัดภายใต้กลุ่มเดียวกัน เป็นข้อมูลแบบกลุ่ม (clustered/grouped data) โดย “ตัวแปรผลลัพธ์ของแต่ละบุคคล จะไม่ ถูกวัดซ้ำ ด้วยเงื่อนไขของเวลา แต่จะถูกวัดภายใต้ กลุ่มเดียวกัน จึงทำให้ข้อมูลตัวแปรผลลัพธ์ที่ได้ มีความคล้ายคลึง หรือสัมพันธ์กัน” โดยแบ่งย่อยได้ 3 แบบ ดังนี้

2.1) ข้อมูลแบบกลุ่มระดับเดียว (one level clustered/grouped data) เป็นข้อมูลแบบ กลุ่มที่ “ตัวแปรผลลัพธ์แต่ละบุคคล ถูกวัดภายใต้ กลุ่มเดียวกันโดยพิจารณาเพียงระดับเดียว” จึงส่งผลให้ข้อมูลภายใต้กลุ่มเดียวกันเมื่อพิจารณา เพียงระดับเดียว มีความสัมพันธ์กัน

2.2) ข้อมูลแบบกลุ่มพหุระดับ (multilevel clustered/grouped data) เป็นข้อมูลแบบ กลุ่มที่ “ตัวแปรผลลัพธ์แต่ละบุคคล ถูกวัดภายใต้ กลุ่มเดียวกันตั้งแต่สองระดับขึ้นไป” จึงส่งผลให้ ข้อมูลภายใต้กลุ่มเดียวกันเมื่อพิจารณาตั้งแต่สองระดับ ขึ้นไป มีความสัมพันธ์กัน

2.2.1) ข้อมูลแบบกลุ่มระดับเดียว (one level) ในทางปฏิบัติที่นักวิจัยคุ้นเคย เมื่อมีการศึกษาและเก็บรวบรวมข้อมูลในโรงเรียนด้วยแบบแผนงานวิจัยภาคตัดขวาง นักวิจัยมักมุ่งเป้าไปที่กลุ่มตัวอย่าง ซึ่งเป็นนักเรียนแต่ละคนและมองว่า นักเรียนแต่ละคน ไม่ว่าจะมาจากห้องเรียนไหน ถือว่ามีความเป็นอิสระ หรือ มีข้อสมมติว่า ห้องเรียนที่แตกต่างกัน ไม่มีผลกระทบใดๆ ต่อตัวแปรผลลัพธ์ที่เก็บรวบรวมมาได้จากนักเรียน แต่ละคน ซึ่งในความจริงแล้ว แนวคิดและการปฏิบัติ ดังกล่าว ถือว่าไม่เหมาะสมและอาจก่อให้เกิดความ ลำเอียงต่อการประมาณค่าของผลลัพธ์และข้อสรุปที่ได้ เพราะแม้ในคำถามวิจัย จะมุ่งเน้นความสนใจไปที่ ตัวแปรผลลัพธ์ที่นักวิจัยเก็บรวบรวมข้อมูลได้โดยตรง จากนักเรียนระดับชั้นที่สนใจ แต่ค่าตัวแปรผลลัพธ์ที่ได้ อาจมีผลกระทบมาจากความแตกต่างของห้องเรียนที่ นักเรียนแต่ละคนสังกัดอยู่ เช่น สภาพบรรยากาศ สิ่งแวดล้อมภายในห้องเรียน รวมถึงครูผู้สอน เป็นต้น นั่นคือ นักเรียนที่อยู่ห้องเรียนเดียวกัน ย่อมได้รับสภาพ บรรยากาศและสิ่งแวดล้อมเช่นเดียวกัน รวมถึงวิธีการ จัดการเรียนการสอนของครูผู้สอนคนเดียวกัน ภายใต้ ห้องเรียนเดียวกัน ส่งผลให้นักเรียนแต่ละคน ที่มาจาก ห้องเรียนเดียวกัน มีลักษณะและโอกาสของการเกิดค่าของตัวแปรผลลัพธ์ที่คล้ายคลึง หรือ สัมพันธ์กัน เมื่อ เทียบกับนักเรียนแต่ละคนที่มาจากห้องเรียนอื่น ดังนั้น ในกรณีที่นักวิจัยพิจารณาเฉพาะความแตกต่างที่เกิดขึ้น ใน “ระดับห้องเรียน” ถือว่า ระดับข้อมูลแบบกลุ่มที่ ถูกนำมาพิจารณานี้ เป็นข้อมูลแบบกลุ่มระดับเดียว

2.2.2) ข้อมูลแบบกลุ่มพหุระดับ (multilevel) นอกจากนักวิจัยจะพิจารณาข้อมูลแบบกลุ่มระดับเดียว (ห้องเรียน) แล้ว นักวิจัยควรพิจารณาความ แตกต่างของระดับข้อมูลเพิ่มเติม เช่น โรงเรียนที่แตกต่างกัน อาจเป็นผลกระทบหนึ่งที่มีผลต่อข้อมูล ของนักเรียนแต่ละคนได้ เป็นต้น ดังนั้น เมื่อนักวิจัย นำความแตกต่างของโรงเรียนมาพิจารณา นั่นถือว่า นอกเหนือจากนักเรียนแต่ละคนที่อยู่ในห้องเรียน เดียวกัน จะมีความสัมพันธ์กันแบบกลุ่มแล้ว (ระดับห้องเรียน) การที่นักเรียนแต่ละคนอยู่ในโรงเรียน เดียวกัน ยังถือว่ามีความสัมพันธ์กันแบบกลุ่มอีก ระดับหนึ่งด้วย (ระดับโรงเรียน) เมื่อเทียบกับนักเรียน ที่มาจากรโรงเรียนอื่น ทำให้ในกรณีนี้ ถือเป็นข้อมูล แบบกลุ่มพหุระดับ ประกอบด้วย 2 ระดับ ได้แก่ ระดับโรงเรียนและระดับห้องเรียน

2.3) ข้อมูลแบบกลุ่มรายคู่ (clustered/ grouped pair data) เป็นข้อมูลแบบกลุ่มที่ “ตัวแปร ผลลัพธ์

แต่ละบุคคล ภู่วัดและพิจารณาภายใต้กลุ่มรายคู่” จึงส่งผลให้ข้อมูลภายใต้คู่เดียวกัน มีความสัมพันธ์กัน

วิธีการทางสถิติที่เหมาะสมสำหรับการวิเคราะห์ ข้อมูลระยะยาว จากข้อจำกัดและความไม่เหมาะสม รวมถึง ผลกระทบที่เกิดขึ้นเมื่อนำวิธีการทางสถิติเดิมที่มีอยู่ มาใช้กับข้อมูลระยะยาวข้างต้น ดังนั้น ผู้เขียนจึงได้นำเสนอทางเลือกที่เหมาะสมของวิธีการทางสถิติ สำหรับข้อมูลระยะยาว โดยเริ่มต้นจากวิธีการพื้นฐาน ไปสู่วิธีการที่ซับซ้อน ได้ดังนี้

(1) การวิเคราะห์คะแนนที่เปลี่ยนแปลง (Change score analysis) เป็นวิธีการทางสถิติพื้นฐานที่ถูกนำมาใช้กับ ข้อมูลระยะยาว กรณีตัวแปรผลลัพธ์แบบต่อเนื่อง เพื่อ ตอบคำถามวิจัยเกี่ยวกับความแตกต่างระหว่างค่าเฉลี่ยสองจุดเวลาที่วัดซ้ำในบุคคลเดิม เช่น paired t-test เป็นต้น จะมีความเหมาะสมเมื่อนักวิจัยต้องการบ่งชี้ ผลการเปลี่ยนแปลงเบื้องต้น แต่ยังไม่สามารถอธิบาย เชิงเหตุและผลได้ เนื่องจากยังมีข้อจำกัดหลายส่วน เช่น จำนวนสองจุดเวลาที่วัดซ้ำ อาจไม่เพียงพอในการ อธิบายการเปลี่ยนแปลงของบางผลลัพธ์ที่สนใจได้และ ไม่สามารถนำตัวแปรร่วมอื่นมาพิจารณาประกอบได้ เป็นต้น นอกจากนี้ ยังมีเงื่อนไขจำนวนข้อมูลในแต่ละ จุดเวลาที่วัดซ้ำต้องครบถ้วน หรือเท่าเทียมกัน รวมถึง ไม่สามารถนำข้อมูลสูญหายมาวิเคราะห์ร่วมได้

(2) การวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (Repeated Measures ANOVA, RM-ANOVA) เป็นวิธีการทางสถิติที่ถูกนำมาใช้กับตัวแปร ผลลัพธ์แบบต่อเนื่องเช่นเดียวกับ paired t-test เพื่อ หาความสัมพันธ์กับชุดตัวแปรร่วม (ตัวแปรกลุ่ม หรือ เวลา) ภายใต้คำถามวิจัยเกี่ยวกับการเปรียบเทียบค่า เฉลี่ยระหว่างจุดเวลาที่วัดซ้ำมากกว่าสองจุดเวลาขึ้นไป โดยสามารถนำมาใช้เปรียบเทียบได้ทั้งแบบกลุ่มเดียว หรือมากกว่า การวิเคราะห์ความแปรปรวนแบบวัดซ้ำนี้ ถือเป็นวิธีการที่ถูกนำมาใช้อย่างแพร่หลายในกรณี ข้อมูลระยะยาวแบบวัดซ้ำ แต่อย่างไรก็ตาม ยังพบว่า วิธีการดังกล่าวมีเงื่อนไขและจำกัดหลายอย่าง ซึ่ง นักวิจัยจำเป็นต้องคำนึงถึง

(3) การวิเคราะห์สมการประมาณค่าเฉลี่ยทั่วไป (Generalized Estimating Equation; GEE) เป็นวิธีการทางสถิติที่ถูกนำมาใช้กับตัวแปร ผลลัพธ์ที่ถูกวัดซ้ำ ทั้งกรณีข้อมูลแบบต่อเนื่องและไม่ ต่อเนื่อง หรือกรณีการแจกแจงทั้งแบบปกติและแบบ ไม่ปกติ ภายใต้คำถามวิจัยที่มุ่งเน้นการประมาณค่า ผลกระทบค่าเฉลี่ยประชากร (population-average effect) ภายในตัวแปรผลลัพธ์ที่วัดซ้ำกับตัวแปรร่วม ทั้งที่สามารถผันแปรได้ตามเวลาและแบบคงที่ หรือ บางครั้งถูกเรียกว่า “marginal models” ในทางปฏิบัติ วิธีการ GEE มีหลักการวิเคราะห์ข้อมูลอยู่ 2 ขั้นตอน ที่สำคัญ ประกอบด้วย

ขั้นตอนที่ 1 การสร้าง หรือ พัฒนาตัวแบบ ซึ่งดำเนินการภายใต้แนวคิดแบบอิสระ ของตัวแบบเชิงเส้นนัยทั่วไป (Generalized Linear Model, GLMs) ด้วยการเชื่อมโยงค่าเฉลี่ยของตัวแปร ผลลัพธ์เป็นฟังก์ชันเชิงเส้นของตัวแปรร่วมที่สนใจผ่าน ฟังก์ชันเชื่อมโยง (link function) เช่น ฟังก์ชันเชื่อมโยง identity จะถูกใช้กับตัวแปรผลลัพธ์แบบต่อเนื่อง ส่วนฟังก์ชันเชื่อมโยง logit หรือ log ถูกนำมาใช้กับ ตัวแปรผลลัพธ์แบบแจกที่มีสองค่า (binary data) หรือ แบบจำนวนนับ (count data) เป็นต้น

ขั้นตอนที่ 2 การประมาณค่าพารามิเตอร์ด้วยโครงสร้างความสัมพันธ์ ในรูปเมตริกซ์ หรือ working correlation structure (matrix) ซึ่งประกอบด้วยหลายรูปแบบ เช่น exchangeable, independent หรือ autoregressive เป็นต้น ซึ่งในขั้นตอนนี้ค่าความคลาดเคลื่อนมาตรฐาน (SE) ที่ได้จากขั้นตอนที่ 1 และถูกคำนวณภายใต้ แนวคิดอิสระ จะถูกปรับด้วยโครงสร้างความสัมพันธ์ ในรูปเมตริกซ์ดังกล่าว เพื่อให้มีความถูกต้องมากขึ้น โดยทั้งสองขั้นตอนจะถูกกระทำแบบวนซ้ำจนกระทั่ง ลู่เข้าสู่จุดอิมิต หรือเกณฑ์ที่เหมาะสมของตัวแบบ ดังนั้น จึงกล่าวได้ว่า GEE ไม่ถือเป็นตัวแบบ แต่เป็น วิธีการประมาณค่าที่ถูกนำมาใช้ภายใต้ตัวแบบที่นักวิจัย สนใจเท่านั้น¹⁴ เช่น การนำ GEE มาใช้ประมาณค่า ในกรณีตัวแปรผลลัพธ์ถูกวัดซ้ำกับตัวแบบสมการถดถอยลอจิสติก หรือ ตัวแบบสมการถดถอยแบบ ปัวซอง เป็นต้น อย่างไรก็ตาม แม้ GEE จะถูกนำมาใช้ ในการประมาณค่าของตัวแบบได้อย่างครอบคลุม ทั้งข้อมูลตัวแปรผลลัพธ์แบบต่อเนื่องและไม่ต่อเนื่อง แต่ยังมี

ข้อจำกัดบางอย่างที่นักวิจัยควรคำนึงถึง เช่น ความสามารถในการจัดการข้อมูลสูญหายได้ เฉพาะกรณี missing completely at random (MCAR) เท่านั้น การมีความไวต่อข้อมูลที่มีค่าผิดปกติ มาก (outlier) หรือไม่ สามารถให้ผลลัพธ์ในภาพ เฉพาะบุคคล (subject-specific result) ได้ เป็นต้น

(4) การวิเคราะห์ด้วยแบบผลกระทบผสม (Mixed Effects Models; MEM) เป็นวิธีการทางสถิติที่ก็นำมาใช้กับตัวแปร ผลลัพธ์ที่ถ่วงน้ำหนักได้ครอบคลุมทุกกรณีของตัวแปร ผลลัพธ์เช่นเดียวกับ GEE และสามารถตอบคำถามวิจัยได้อย่างครอบคลุมทั้งผลกระทบค่าเฉลี่ยประชากรและ ผลกระทบเฉพาะบุคคล ทั้งนี้เนื่องจากเป็นวิธีการที่สามารถอธิบายความสัมพันธ์ของการวัดซ้ำได้โดยตรง ด้วยการใช้ “ผลกระทบสุ่ม (random effect)” มา พิจารณาเพื่ออธิบายแนวโน้มเฉพาะบุคคล หรือกลุ่ม อย่างต่อเนื่อง ดังนั้น จึงถูกเรียกว่า “mixed effects” ซึ่งเป็นการพิจารณาทั้งสองผลกระทบพร้อมกัน ได้แก่ ผลกระทบคงที่ (fixed effect) และผลกระทบสุ่ม ทำให้วิธีการ MEM สามารถนำมาใช้กับข้อมูลระยะยาว ที่มีลักษณะสัมพันธ์กันได้ทั้งแบบวัดซ้ำ แบบกลุ่ม และ แบบรายคู่ และตอบคำถามวิจัยได้ครบทั้งสองประเด็น ข้างต้น แต่อย่างไรก็ตาม เนื่องจากวิธีการ MEM ค่อนข้างซับซ้อนทั้งการทำความเข้าใจและการนำไปใช้ จึงอาจเป็นข้อจำกัดสำคัญต่อการนำมาใช้ของนักวิจัยส่วนใหญ่ ดังนั้น การนำเสนอบทเรียนรู้และแนวปฏิบัติ ตรงไปตรงมาและง่ายต่อการทำความเข้าใจ ยังจำเป็นสำหรับวิธีการนี้

The image shows two side-by-side screenshots of a video conference. The top of each screen displays a row of participant video feeds with names like 'Dr. P. W. E. Harnong', 'Anaya Praditphai', 'Parichat Wonggorn', 'Abhorat Sukjeam-Bai', 'Pisit Charanrat', and 'Admin-Supak'. The left screen (slide 6) is titled 'Hierarchy of Epidemiological studies' and features a large red arrow pointing upwards. To the left of the arrow is a list of study types: 'Clinical Trial', 'Cohort', 'Cross-sectional / Case control', 'Cross-sectional', 'Case series', and 'Case report'. To the right of the arrow are three boxes labeled 'Experiment', 'Analytic', and 'Descriptive'. The right screen (slide 7) is titled 'Cohort Study' and lists three bullet points: 'The most powerful observational study for identifying an association between risk factors and a disease', 'The most time consuming', and 'The most expensive'.

Hierarchy of Epidemiological studies

- Clinical Trial
- Cohort
- Cross-sectional / Case control
- Cross-sectional
- Case series
- Case report

Experiment
Analytic
Descriptive

Cohort Study

- The most powerful observational study for identifying an association between risk factors and a disease
- The most time consuming
- The most expensive