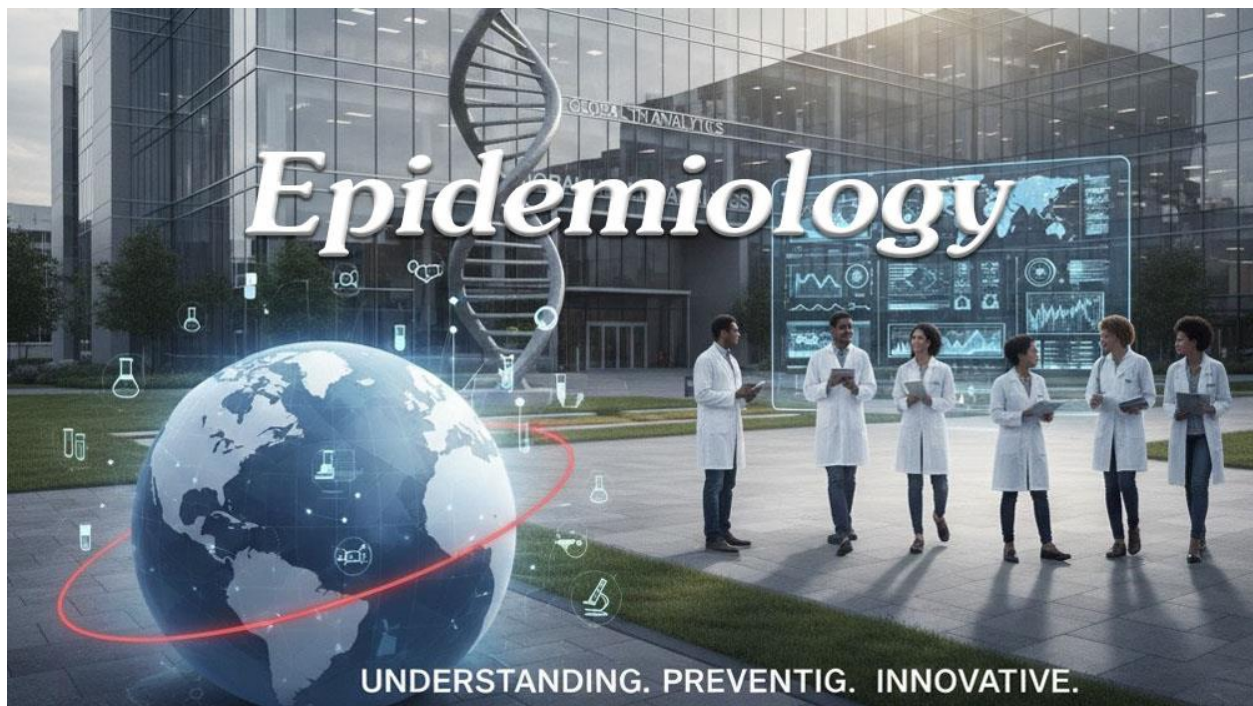


Module 2

Epidemiology



นพ.ชลทิศ อุไรฤกษ์กุล

M.D., M.P.H ,M.B.A

อว.เวชศาสตร์ป้องกัน (คลินิก)

สารบัญ

หน้า

Module 2 .Practical Epidemiology

1. Introduction to Epidemiology	1
2. Epidemiology Study	12
3. การทดสอบความสัมพันธ์ทางระบาดวิทยาและทางสถิติ	21
4. Bias in Clinical Research	39

Introduction to Epidemiology

Content

1. [เกริ่นนำ](#)
2. [ความหมายของระบาดวิทยา](#)
3. [ระบาดวิทยาของมะเร็งเต้านม](#)
4. Descriptive Epidemiology (ส่วนนี้คือส่วนที่ไปใช้ในอธิบาย distribution of disease ตาม Time place person)
 1. [อุบัติการณ์และความชุก](#)
 2. [Rate และ Ratio ต่างๆ](#)
 3. [Age Standardized Incidence Rate \(ASR\)](#)
5. ระบาดวิทยาส่วนของ Determinant (หาความสัมพันธ์หรือสาเหตุของโลก)
 1. [Epidemiology Study](#)
 2. [การทดสอบความสัมพันธ์ เพื่อหาสาเหตุหรือปัจจัยเสี่ยงที่ทำให้เกิดโรค](#)
 3. [Screening Test](#)
 4. [Bias in clinical Research](#)
6. เอกสารอ่านประกอบ
 1. [Introduction to Epidemiology](#)
 2. [Excel ช่วยคำนวณ Aged Adjust & ASR ¹](#)

เกริ่นนำ

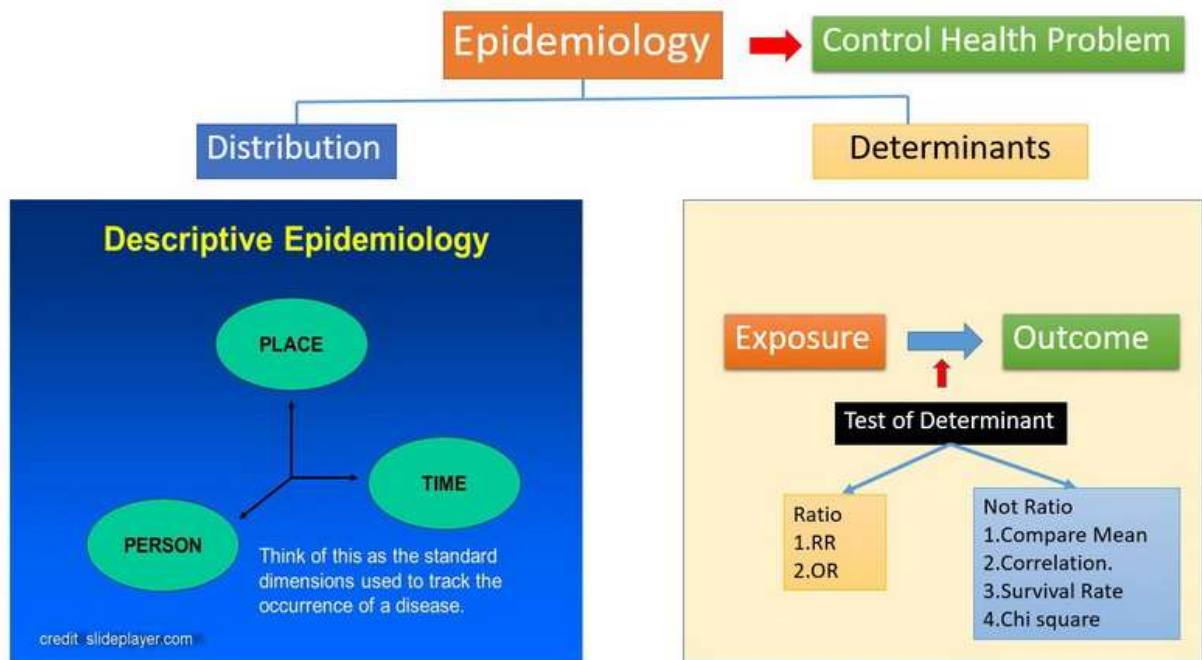
ระบาดวิทยาเป็นวิชาที่นักสาธารณสุขจำเป็นต้องใช้ในทุกขั้นตอนของการบริหารการสาธารณสุข ตั้งแต่การวางแผน และการ Implement แผน และแนวคิดของระบาดวิทยาสามารถนำมาใช้ได้ทั้งโรคติดต่อ และโรคไม่ติดต่อ เพราะถ้าเรารู้การกระจายของโรค ตาม Time place Person รู้ว่าอะไรเป็นสาเหตุ หรือ ปัจจัยเสี่ยงแล้ว อย่างรู้อย่างไรแล้ว จะทำให้การดำเนินการ จะเป็นแบบ Focus ไม่ได้แบบเหวี่ยงแห ทำให้สามารถใช้ทรัพยากรที่มีอยู่จำกัดให้เกิดประโยชน์สูงสุดได้

ความยากของระบาดวิทยาในตอนเรียน มีอยู่ 3 ประการคือ

1. ศัพท์และนิยามของศัพท์ต่างๆในวิชาระบาด ซึ่งจะนำมาใช้ในส่วนของ Descriptive Epidemiology มีมากมาย และมักจะก่อให้เกิดความสับสน ถ้าไม่แม่นในนิยาม เช่น อุบัติการณ์ (incidence) และ ความชุก (prevalence) รวมถึง Rate & Ratio ต่างๆที่มีมากมาย (ยังไม่นับรวมถึงวิธีการคำนวณเพื่อให้ได้ค่าต่างๆนั้นมา) ส่วนนี้จำเป็นที่ไม่ควรปล่อยผ่าน ต้องสามารถแยกแยะความแตกต่างของศัพท์ที่ใช้ในทางระบาดวิทยาเหล่านั้น
2. Concept ทางระบาดวิทยา ซึ่งต้องอาศัยความเข้าใจมากกว่าการท่องจำ เมื่อทำงานจริง มี Concept ที่มักจะถูกนำมาใช้บ่อยได้แก่
 1. การ Standardized และการ Adjust ต่างๆ แนวคิด คือ ถ้าสัดส่วนประชากรของแต่ละพื้นที่ต่างกัน แล้วจำเป็นต้อง Adjust เพื่อให้เกิดความยุติธรรม โดยมีองค์กรกลางเช่น WHO ได้กำหนด Standard Population เพื่อเทียบให้อัตราต่างๆด้วย Standard Population เดียวกัน
 2. การศึกษาทางระบาดวิทยา เช่น Prospective Study เริ่มจาก Exposure ไปหา Outcome คำนวณค่า Relative Risk (RR) ได้ Retrospective Study เริ่มจาก Outcome หรือ Disease แล้วถามย้อนไปว่าได้สัมผัส Exposure หรือ หา RR ไม่ได้ หาได้แต่ Odds ratio เป็นต้น
3. การคำนวณ เพื่อให้ได้ค่าต่างๆ ส่วนนี้เดิมเป็นปัญหาใหญ่ เพราะต้องใช้เครื่องคิดเลขหรือ Excel ช่วยในการคำนวณ แต่ปัจจุบันมีเครื่องมือทางสถิติช่วยมากมาย ทำให้ลดปัญหาที่ส่วนนี้ไปได้ยากมาก ถ้า คำศัพท์ นิยาม และ Concept เราแน่นแล้ว คิดว่าถาการคำนวณไม่น่าจะปัญหา เพราะจะแนะนำ tool ต่างๆที่นิยมใช้มาให้เราใช้กัน

Course นี้จึงเน้นที่การปู concept ให้เกิดความเข้าใจ และความสามารถในการ interpret Output ต่างๆ ที่ได้จากการคำนวณมากกว่า และเลือกเนื้อหาคิดว่าจำเป็นสำหรับ นักการสาธารณสุขเพื่อใช้ในการปฏิบัติงานจริง (ตามที่กำหนดไว้ใน Contents)

ความหมายของระบาดวิทยา



1. ระบาดวิทยา มาจากรากศัพท์ Epi = Upon ,demio=Population ,logos = Study
2. ระบาดวิทยา หมายถึง การศึกษา (logos) เกี่ยวกับการกระจาย (Distribution) และสาเหตุหรือปัจจัยเสี่ยง (Determinant) ของโรคหรือเหตุการณ์โดยในกลุ่มประชากร (demio = Population) เฉพาะ (epi = upon) เพื่อนำผลการศึกษาไปสู่การควบคุมปัญหาสุขภาพ

Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems

3. ระบาดวิทยา จึงเน้นหนักใน 2 เรื่องหลักคือ การกระจาย (Distribution) และ สาเหตุหรือปัจจัยเสี่ยงที่ทำให้เกิดโรคนั้น (Determinant)
4. การกระจาย ส่วนใหญ่จำแนกตาม Time ,Place ,Person
5. Determinants คือหาความสัมพันธ์ (Relationship) แบ่งเป็น
 1. Causal Relationship เป็นสาเหตุโดยตรงที่ทำให้เกิดโรค การศึกษาที่จะสรุปว่าปัจจัยเสี่ยง (Exposure) เป็นสาเหตุที่ทำให้เกิดโรค (Outcome) จึงต้องเป็นการศึกษาที่เริ่มจาก Exposure แล้วติดตามไปข้างหน้าว่าเกิด Outcome หรือไม่ ซึ่งก็คือ Prospective Study
 2. Non Causal Relationship มีความสัมพันธ์ แต่ไม่ได้เป็นสาเหตุโดยตรงที่ทำให้เกิดโรค โดยการศึกษาที่เริ่มจากโรค (Outcome) แล้วถอยย้อนกลับไปได้สัมผัส Exposure หรือไม่ จะสรุปได้แต่เพียงว่ามีความสัมพันธ์ แต่ไม่สามารถสรุปได้ว่า เป็นสาเหตุที่ทำให้เกิดโรคนั้น
6. เมื่อพิสูจน์ทางสถิติ พบว่ามีความสัมพันธ์ อย่างถ้วนสรุปว่าเป็น Causal Relationship อาจเกิดจาก Bias ,By Confounding ,By chance

1. Bias & By Cofounding Factors
2. By Chance (Type 1 & Type 2 Error and Screening Test)

ระบาดวิทยาของมะเร็งเต้านม

ระบาดวิทยาของมะเร็งเต้านม จะกล่าวถึง การกระจายของปัจจัยต่างๆ (Distribution) และปัจจัยเสี่ยงหรือสาเหตุของมะเร็งเต้านม (Determinant) นอกจากนั้นก็ควรพูดถึง Intervention ที่จะมาควบคุมมะเร็งเต้านม

1. ระบาดวิทยา
 1. การกระจาย (Distribution) ได้แก่
 1. อุบัติการณ์และความชุกของมะเร็งเต้านม
 2. อาการและอาการแสดงของมะเร็งเต้านม
 3. ประเภทของมะเร็งเต้านม (Breast Cancer Type)
 4. ระยะของมะเร็งเต้านม (Staging)
 2. ปัจจัยเสี่ยงและสาเหตุของมะเร็งเต้านม (Determinant)
 1. ปัจจัยเสี่ยงของมะเร็งเต้านม
 2. สาเหตุของมะเร็งเต้านม
2. Intervention ในการควบคุมมะเร็งเต้านม
 1. การคัดกรองมะเร็งเต้านม
 2. การรักษามะเร็งเต้านม

Incidence , Prevalence ,Age Standardized Incidence Rate (ASR)

1. อุบัติการณ์ (Incidence) เท่ากับ จำนวนรายใหม่ในช่วงเวลาใดเวลาหนึ่ง (มักใช้ 1 ปี)หาร ด้วย ประชากรกลุ่มเสี่ยง (Population at risk)
2. ความชุก เท่ากับ จำนวน รายเก่าที่สะสมมา (ที่มีชีวิต) + รายใหม่ปีนี้ – จำนวนรายที่เสียชีวิตในปีนั้น หาร ด้วย ประชากร (มักใช้ประชากรกลางปี เท่ากับ ประชากรต้นปี + ย้ายเข้า - ย้ายออก+จำนวนคนเกิด + จำนวนคนตาย)

โจทย์ข้อที่ 1 จังหวัด ก.มีประชากรกลางปี พ.ศ.2562 จำนวน 1 ล้าน คน เป็นชาย 5 แสนคน เป็นหญิง 5 แสนคน พบมะเร็งเต้านมในหญิงทั้งหมด โดย

1. cases สะสมจากสิ้นปี 2562 ผู้ป่วยมะเร็งเต้านมที่ยังมีชีวิตอยู่ จำนวน 1510 คน
2. พบรายใหม่ในปี 2562 จำนวน 200 คน
3. เสียชีวิตในปี 2562 จำนวน 10 คน

ในระหว่างปี มีผู้หญิงอพยพเข้า 2,000 คน อพยพออก 3,000 คน คนเกิดที่เป็นหญิง 400 คน คนตายที่เป็นหญิง 500

คำถาม

1. จังหวัด ก.มีอุบัติการณ์ของมะเร็งเต้านมเท่ากับเท่าไร
2. จังหวัด ก.มีความชุกของมะเร็งเต้านมเท่ากับเท่าไร
3. กรณีที่ความชุก มากกว่า อุบัติการณ์ แสดงว่าอะไร ถ้าน้อยกว่า แสดงว่าอะไร

Rate & Ratio อื่น ๆ

1. Mortality & Morbidity Rate ได้แก่
 1. Crude Morbidity Rate
 2. Crude Mortality Rate หรือ Crude Death Rate
 3. Case Fatality Rate
 4. Age Specific Morbidity Rate คือ อัตราการป่วยตามช่วงอายุต่างๆ ส่วนนี้จะนำไปคำนวณ Standardized morbidity rate ต่างๆ)
 5. Age Specific Mortality Rate คือ อัตราการตายตามช่วงอายุต่างๆ ส่วนนี้จะนำไปคำนวณ Standardized mortality rate ต่างๆ)
2. Fertility Rate
 1. General Fertility Rate (GFR)
 2. Total Fertility Rate (TFR)
 3. Net Reproductive Rate

Age standardized Incidence Rate (ASR)

1. บางโรคมีอัตราการเกิดโรคในแต่ละ Age Group ต่างกัน เช่น มะเร็งเต้านม จะพบในผู้สูงอายุ 50 ปีขึ้นไป มากกว่า กลุ่มที่มีอายุน้อยกว่า 50 ปี ทำให้ประเทศที่มีสัดส่วนของผู้สูงอายุที่มากกว่าจะพบจำนวนรายที่มากกว่า เมื่อคิดอุบัติการณ์อย่างหยาบ (Crude Incidence Rate) ก็จะได้พบอัตราที่สูงกว่า
2. เพื่อให้เกิดความเป็นธรรมในการเปรียบเทียบอุบัติการณ์ของมะเร็งเต้านม (รวมถึงโรคอื่นๆด้วย) จึงมีหน่วยงานกลาง ทำการกำหนด Standard Population ขึ้น เช่น WHO Standard Population เพื่อให้คิดอุบัติการณ์โดยปรับอายุให้เป็นมาตรฐานเดียวกัน เรียกว่า Age Standardized Incidence Rate (ASR) เพื่อให้อุบัติการณ์ที่ปรับแล้ว มาเปรียบเทียบอุบัติการณ์ระหว่างประเทศ
3. Standard Population จะทำการกำหนดประชากรแบ่งตาม age group ละ 5 ปี ว่าในแต่ละช่วงอายุมีกี่คน โดยผลรวมทั้งหมดจะเป็น 100,000 คน หรือ 100%
4. หา Age Specific morbidity rate ของแต่ละช่วงอายุ แล้วคูณด้วย Standard Population ของช่วงอายุนั้น จากนั้นหาผลรวมว่า มีจำนวนป่วยเท่าไร (Adjusted) จากนั้น หาคด้วย ประชากรกลุ่มเสี่ยง ก็จะได้ ASR

5. Area 1 และ Area 2 มีประชากรเท่ากันคือ 300,000 คน และมีผู้ป่วยรายใหม่เท่ากันคือ 400 คน แต่ทั้ง 2 Area มีสัดส่วนประชากรของแต่ละช่วงอายุ ต่างกัน ดังตารางด้านล่าง

Age Standardization incidence rate (ASR)

คำถาม Area ไหน มี ASR มากกว่ากัน

Area 1	B	C	Area 2	B	C
Age Group	Population	New Cases	Age Group	Population	New Cases
0-14	50,000	100	0-14	25,000	50
15-59	200,000	200	15-59	210,000	225
>=60 ปี	50,000	100	>=60 ปี	65,000	125
total	300,000	400	total	300,000	400
Crude incidence	133.3	/100000	Crude incidence	133.3	/100000

เมื่อ Download Excel ที่ให้มา แล้วกรอกตัวเลขเข้าไปใน Sheet Exercise1 จะได้ข้อมูลตามตารางข้างล่าง โดย 2 Area มี Crude Incidence rate เท่ากันคือ 133.3 / 100,000 ประชากร แต่มี ASR ต่างกัน โดย Area 1 มี ASR = 138.1 / 100,000 ประชากร ต่ำกว่า Area 2 ที่มี ASR = 141.6 / 100,000 ประชากร แสดงว่า Area 2 มีอุบัติการณ์เมื่อปรับแล้วสูงกว่า

Area1

WHO Standard Population 2020-2025

Exercise 1



Area 1	B	C	D=C/B	E	F	G=D*F
Age Group	Population	New Cases	Age specific rate	WHO Standard Population	Standard Pop to be used	Adjusted Cases
0-14	50,000	100	0.00200	26,139	26,139	52.28
15-59	200,000	200	0.00100	61,921	61,921	61.92
>=60 ปี	50,000	100	0.00200	11,940	11,940	23.88
total	300,000	400		100,000	100,000	138.1
Crude incidence	133.3	/100000		ASR	138.1	100000

Area 2	B	C	D=C/B	E	F	G=D*F
Age Group	Population	New Cases	Age specific rate	WHO Standard Population	Standard Pop to be used	Adjusted Cases
0-14	25,000	50	0.00200	26,139	26,139	52.28
15-59	210,000	225	0.00107	61,921	61,921	66.34
>=60 ปี	65,000	125	0.00192	11,940	11,940	22.96
total	300,000	400		100,000	100,000	141.6
Crude incidence	133.3	/100000		ASR	141.6	100000

6. ถ้าจะเปรียบเทียบมะเร็งเต้านมกับประเทศต่างๆ ก็มักจะปรับด้วย standard pop ถ้ากรอกข้อมูลเข้าไปใน Table Exercise2 จะได้ข้อมูลดังตารางด้านล่าง สรุป Crude incidence rate = 31.1 ต่อแสนประชากรต่อปี แต่ ASR = 30.1 ต่อแสนประชากรต่อปี

Age Group	WHO Standard Population
0	1,822
1-4	7,033
5-9	8,687
10-14	8,597
15-19	8,474
20-24	8,222
25-29	7,928
30-34	7,605
35-39	7,145
40-44	6,590
45-49	6,038
50-54	5,371
55-59	4,547
60-64	3,723
65-69	2,955
70-74	2,210
75-79	1,515
80-84	905
85+	632
total	100,000

WHO Standard Population 2020-2025

Exercise 2



ข้อมูลมะเร็งเต้านมโครงการสืบสานฯ 5 ปี (2556-2560)

	B	C	D=C/B	E	F	G=D*F
30-34	292,175	46	0.00016	7,605	7,605	1.20
35-39	282,208	173	0.00061	7,145	7,145	4.38
40-44	297,513	354	0.00119	6,590	6,590	7.84
45-49	290,650	554	0.00191	6,038	6,038	11.51
50-54	255,072	581	0.00228	5,371	5,371	12.23
55-59	209,463	534	0.00255	4,547	4,547	11.59
60-64	161,750	381	0.00236	3,723	3,723	8.77
65-69	113,240	333	0.00294	2,955	2,955	8.69
70-74		0	-	2,210	0	-
75-79		0	-	1,515	0	-
80-84		0	-	905	0	-
85+		0	-	632	0	-
total	1,902,071	2,956		100,000	43,975	66.2
Crude incidence rate /แสนปี		31.1			ASR	30.1

Epidemiology Study

การศึกษาทางระบาดวิทยา มีวัตถุประสงค์เพื่อศึกษาว่าสาเหตุที่ทำให้เกิดโรคคืออะไร (Causal Relationship) หรือถ้าไม่สามารถจะพิสูจน์หาสาเหตุได้ ก็หาว่าอะไรคือปัจจัยเสี่ยงที่สัมพันธ์กับโรคที่กำลังศึกษา การศึกษาทางระบาดวิทยานั้น แบ่งเป็น 2 กลุ่มใหญ่ และแต่ละกลุ่มก็สามารถแยกย่อยลงไปได้อีก

1. การศึกษาที่เริ่มจาก Outcome (Disease) แล้วถามย้อนไปว่า ได้รับ Exposure หรือไม่ เรียกว่า Retrospective Study กลุ่มนี้หาได้แต่ Odds ratio และไม่สามารถบอกได้ว่า Exposure เป็นสาเหตุที่ทำให้เกิด Outcome การศึกษากลุ่มนี้ได้แก่
 1. Case Control Study
 2. Cross Sectional Study (Prevalence Study)
2. การศึกษาที่เริ่มจาก Outcome (Disease) แล้วถามย้อนไปว่า ได้รับ Exposure หรือไม่ เรียกว่า Retrospective Study กลุ่มนี้หาได้แต่ Odds ratio และไม่สามารถบอกได้ว่า Exposure เป็นสาเหตุที่ทำให้เกิด Outcome การศึกษากลุ่มนี้ได้แก่
 1. Case Control Study
 2. Cross Sectional Study (Prevalence Study)
3. การศึกษาเริ่มจาก Exposure แล้วติดตามไปหา Outcome เรียกว่า Prospective Study การศึกษาประเภทนี้สามารถหา Relative Risk (RR) ได้ และอาจจะสรุปว่า Exposure เป็นต้นเหตุ (Causality) ได้ การศึกษากลุ่มนี้ได้แก่
 1. Cohort Study
 2. Intervention Study เช่น Randomized Case Control Study (RCT)

Type of Study and Epidemiological Test

Type of Study	Test	Causal Relation
1.Cross Sectional ¹	OR ³	No
2.Case Control (Retrospective Study) ¹	OR ³	No
3.Cohort (Prospective Study) ²	RR ⁴	Yes
4.Intervention Study (RCT) ²	RR ⁴	Yes

Remark : 1 = ศึกษาจาก Outcome ไปหา Exposure, 2=ศึกษาจาก Exposure แล้วติดตามการเกิด Outcome
3 = ย่อจาก Odds Ratio, 4 = ย่อจาก Relative Risk

การทดสอบความสัมพันธ์ เพื่อหาสาเหตุหรือปัจจัยเสี่ยงที่ทำให้เกิดโรค

ในการทดสอบความสัมพันธ์เพื่อหาสาเหตุหรือปัจจัยเสี่ยงที่ทำให้เกิดโรคนั้น จำเป็นต้องใช้การทดสอบทางสถิติ เนื่องจากวิชาระบาดวิทยากำเนิดเกิดมานานเป็น 100 ปีในช่วงที่ยังไม่มีเครื่องมือการวิเคราะห์ทางสถิติ เช่น Strata , SPSS จึงแบ่งการหาความสัมพันธ์เป็น 2 แบบ คือ

1. การวิเคราะห์แบบดั้งเดิม โดยใช้ 2x2 Table เพื่อหา OR หรือ RR และ Chi square ซึ่งเป็น Non parametric Statistic โดยใช้กับตัวแปรที่เป็น Nominal หรือ Ordinal
2. การวิเคราะห์โดยใช้เครื่องมือทางสถิติ เพื่อคำนวณเกี่ยวกับตัวแปรที่เป็น Scale โดยใช้ Parametric Statistic ได้แก่
 1. Compare Mean ด้วย t Test , Paired t Test , ANOVA
 2. Correlation
 3. Regression ทั้ง Linear & Logistic Regression ทั้ง Simple และ Multiple Regression
 4. Time Series Analysis
 5. Survival Analysis

ประเภทของตัวแปร

ประเภท	หา Mean & SD	Statistic
Norminal	ไม่ได้	Non parametric เช่น Chi Square ,OR,RR
Ordinal	ไม่ได้ แต่ หา Mean rank ได้	Non Parametric แบบ Ordinal
Scale	ได้	Parametric เช่น Compare mean (t Test,ANOVA),Correlation Regression, Time Series,survival rate

Screening Test

ความจำกัดทางทรัพยากรของการตรวจทางห้องปฏิบัติการ และความจำกัดของบุคลากรทางการแพทย์ที่มีความเชี่ยวชาญ ทำให้ต้องมี Screening Test ที่ง่าย หาได้โดยทั่วไป ไม่จำเป็นต้องใช้ผู้เชี่ยวชาญที่มีทักษะพิเศษ และที่สำคัญคือต้นทุนต่ำ เมื่อ Screening Test แล้ว +ve ค่อยตรวจเพื่อยืนยันอีกทีหนึ่ง เช่น ในเรื่องมะเร็งเต้านมนั้น Gold Standard ในการวินิจฉัย ต้องได้ผลชิ้นเนื้อ (Tissue Diagnosis) การจะเอาชิ้นเนื้อจากก้อนทุกก้อนของเต้านมที่ตรวจพบ เป็นเรื่องที่ยาก จึงใช้ Mammogram เป็น Screening Test ก่อน ถ้ารังสีแพทย์อ่านผลแล้วสงสัย (BIRADS ≥ 4) ค่อยส่งตรวจชิ้นเนื้อ จะทำให้ใช้ทรัพยากรที่มีอยู่จำกัด ให้เกิดประโยชน์สูงสุด ซึ่งในเรื่องของ Screening Test นั้น มีศัพท์ 6 ตัว ที่ควรทำความเข้าใจทั้งความหมายและวิธีการคำนวณค่าคือ

1. Sensitivity
2. Specificity
3. PPV (Positive Predictive Value)
4. NPV (Negative Predictive Value)
5. False Positive
6. False Negative

Screening Test						
		Confirmation Test		Total		
		+ve	-ve			
Screening Test	+ve	A	B	A+B	A=TP	B=FP
	-ve	C	D	C+D	C=FN	D=TN
Total		A+C	B+D	A+B+C+D		
Sensitivity =		$A/(A+C)$				
Specificity =		$D/(B+D)$				
Positive predictive value (PPV) =		$A/(A+B)$				
Negative Predictive Value (NPV) =		$D/(C+D)$				
Cancer Detection Rate (CDR/1000) =		$A/(A+B+C+D)$				
Cancer Rate (CR/1000) =		$(A+C)/(A+B+C+D)$				

กระบวนการก่อนหน้าคือการ Screen กระบวนการตามหลังคือการ Confirm



Bias

เมื่อทำการวิเคราะห์ความสัมพันธ์แล้ว พบว่ามีนัยสำคัญทางสถิติแล้ว ผู้ทำวิจัยต้องถามตัวเองว่า งานวิจัยของเรามี 3 B หรือไม่ คือ

1. By chance โดยเฉพาะในกรณีที่ค่า $p \text{ value} < .05$ เล็กน้อย เพราะมีโอกาส (Chance) ที่จะปฏิเสธสมมุติฐาน ทั้งที่เป็นจริง หรือ Type 1 Error
2. By Confounding การทดสอบแล้วมีความสัมพันธ์ อาจเกิดจาก Exposure ที่เราทำการศึกษา ไปสัมพันธ์กับ Confounding Variables ซึ่งเป็นสาเหตุของการทำให้เกิด Outcome นั้นโดยตรง ทำให้ Exposure ดังกล่าวมีความสัมพันธ์เทียมกับ outcome ที่ทำการศึกษา
3. Bias หรือ อคติในงานวิจัย หมายถึง สิ่งใดก็ตามที่ทำให้ข้อมูลเบี่ยงเบนไปจากข้อเท็จจริงทั้งโดยตั้งใจและไม่ตั้งใจ อคติที่พบบ่อยได้แก่
 1. Information process bias ตั้งแต่ Data Collection, Data Analysis, Data Interpretation bias
 2. อคติในการเผยแพร่ข้อมูล (Data publication) โดยมีแนวโน้มที่จะเผยแพร่ข้อมูลที่ good Result มากกว่า bad result
 3. อคติในการพิจารณาบทความ (Reviewing bias) ซึ่ง Reviewer ที่มีภูมิหลังต่อผู้ทำวิจัยทั้งทางบวกหรือทางลบ จะส่งผลต่อการมี Bias ที่จะรับหรือไม่รับงานวิจัยเพื่อตีพิมพ์
 4. อคติจากแหล่งทุนสนับสนุน (Funding or Sponsor bias) งานวิจัยที่ได้ทุน จากบริษัทที่เป็นเจ้าของผลิตภัณฑ์ที่เราทำการวิจัยนั้น ย่อมมี bias

Epidemiological Study (การศึกษาทางระบาดวิทยา)

Content

1. ความหมายและขั้นตอนการศึกษาทางระบาดวิทยา
2. ประเภทของการศึกษาทางระบาดวิทยา
 1. Cross sectional Study (Prevalence Study)
 2. Case Control (Retrospective Study)
 3. Cohort Study (Prospective Study)
 4. Intervention Study : Randomized Controlled Trial
3. Systematic Review & Meta analysis
4. การวิเคราะห์ความสัมพันธ์ทางระบาดวิทยาโดยใช้ 2x2 Table
5. เอกสารอ่านประกอบ
 1. Video : Epidemiology made Easy Dr.Ranil Appuhamy
 2. [Epidemiological Study](#)¹
 3. Excel ช่วยคำนวณ ASR,OR,OR,Chi square,Screening Test

ความหมายและขั้นตอนการศึกษาทางระบาดวิทยา

การศึกษาทางระบาดวิทยา คือกระบวนการทางวิทยาศาสตร์เพื่อที่จะตอบคำถาม โดยใช้ข้อมูลในระดับประชากร และดำเนินการศึกษาโดยคำนึงถึงเรื่องจริยธรรม (Ethic) โดยมีขั้นตอนดังต่อไปนี้

1. ตั้งคำถามการศึกษา
2. เลือกประเภทการศึกษา
3. เก็บรวบรวมและวิเคราะห์ข้อมูล
4. แปลผลผลลัพธ์ที่ได้
5. รายงานผลการศึกษา



Credit : Dr.Ranil Appuhamy :Epidemiology made Easy

Cross Sectional Study (Prevalence Study)

เป็นการศึกษา ข้อมูลทางสุขภาพของกลุ่มประชากรที่ถือเป็นตัวแทน (Representative) ในช่วงเวลาใดเวลาหนึ่ง (Point in time) โดยใช้แบบสอบถาม หรือจะทำ Health Survey เป็นวิธีการดำเนินการที่ค่าใช้จ่ายไม่สูง สามารถที่จะประเมินความต้องการทางสุขภาพ (health need) แม้จะสามารถประเมินสัมพันธ์ระหว่าง ปัจจัยเสี่ยง (Exposure) กับ ผลลัพธ์ (outcome) ได้ แต่ไม่สามารถบอกได้ว่า เป็นสาเหตุที่ทำให้เกิดโรคนั้น (Causal Relationship)

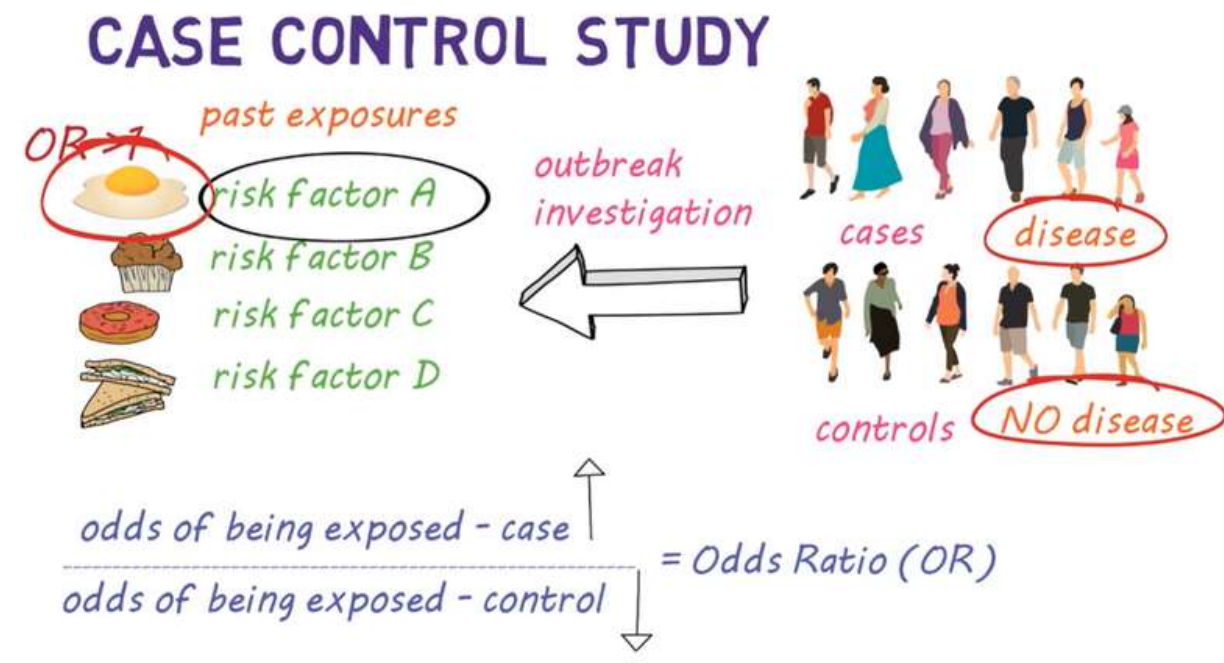
** ค่าสถิติที่ใช้ในการหาความสัมพันธ์ของการศึกษาแบบ Cross sectional study คือ Chi Square และ Odds ratio **



Credit : Dr.Ranil Appuhamy ;Epidemiology made Easy

Case Control Study (Retrospective Study)

เป็นการศึกษาที่เริ่มจาก Outcome โดยแยกเป็น 2 กลุ่มคือ กลุ่มที่เป็นโรค กับ กลุ่มที่ไม่เป็นโรค (outcome อาจจะไม่ใช่วโรคก็ได้ เช่น Early Cancer Staging เป็นต้น) จากนั้นถอยย้อนไปว่ามีการสัมผัสปัจจัยเสี่ยงหรือไม่ (Exposure) ทั้งกลุ่มที่เป็นโรคกับกลุ่มที่ไม่โรค



Credit : Dr.Ranil Appuhamy ;Epidemiology made Easy

การคำนวณความเสี่ยง (Odds ratio : OR)

Odds ratio (OR) in Retrospective or Cross sectional Study

		Outcome	
		Yes	No
Treatment	Yes	A	B
	No	C	D

$$OR = \frac{(A \times D)}{(B \times C)}$$

$$OR = \frac{a/b}{c/d}$$

$$= \frac{a \times d}{b \times c}$$

with the standard error of the log odds ratio being

$$SE[\ln(OR)] = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

and 95% confidence interval

$$95\% \text{ CI} = \exp \left(\ln(OR) - 1.96 \times SE[\ln(OR)] \right) \text{ to } \exp \left(\ln(OR) + 1.96 \times SE[\ln(OR)] \right)$$

Credit <https://www.medcalc.org>

Odds ratio คือ ตัวประมาณการค่าของความเสี่ยง เกิดจากค่า odds ของ Exposure หารด้วย Odds ของ กลุ่ม Non Exposure โดย

OR = ad/bc

การคำนวณ 95% Confidence Interval (CI) ของ OR โดยคำนวณจากสูตรในภาพ

1. OR แสดงถึง ว่า ปัจจัยเสี่ยงมีความสัมพันธ์กับการเกิดโรค (Associate)
2. $OR < 1$ แสดงถึง ว่า ปัจจัยเสี่ยงสามารถปกป้องไม่ให้เกิดโรคได้ (Protective Factors)
3. $OR = 1$ แสดงถึง ว่า ปัจจัยเสี่ยงไม่มีความสัมพันธ์กับการเกิดโรค (Not Associate)
4. 95% CI ของ OR ไม่มีค่าใดค่าหนึ่งผ่าน 1 แสดงว่า มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ
5. 95% CI ของ OR มีค่าใดค่าหนึ่งผ่าน 1 แสดงว่า ไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ



Credit : Dr.Ranil Appuhamy ;Epidemiology made Easy

จุดเด่นและจุดด้อยของการศึกษาแบบ Case Control

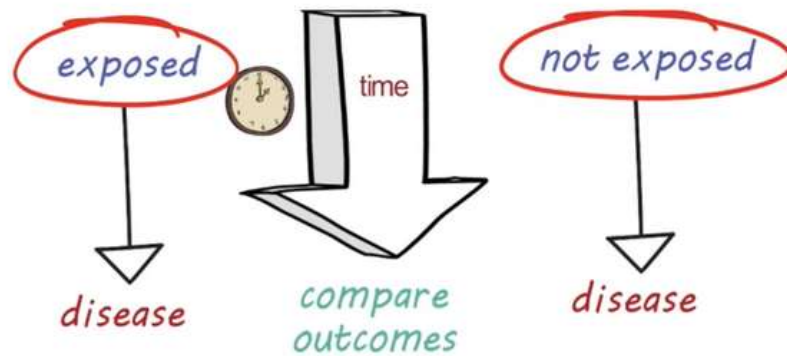
1. จุดเด่น เป็นการศึกษาที่ทำได้ง่าย ค่าใช้จ่ายไม่สูง
2. จุดด้อย คือ ใช้ไม่ดีกับ Exposure ที่สัมผัสบ่อย การเลือกกลุ่ม Control และการถามย้อนว่าได้สัมผัสปัจจัยเสี่ยงหรือไม่ อาจมีปัญหาเรื่อง Recall

Cohort Study (Prospective Study)

Cohort Study เป็นการศึกษาที่เริ่มจาก แบ่งกลุ่มเป็นกลุ่มสัมผัสปัจจัยเสี่ยง (Exposure) กับไม่สัมผัสปัจจัยเสี่ยง (Non Exposure) จากนั้นติดตามไปข้างหน้าว่า แต่ละกลุ่มป่วยเป็นโรคกี่คน ไม่ป่วยกี่คน

COHORT STUDY

risk factors



$$\text{Relative Risk (RR)} = \frac{\text{risk of disease (exposed)}}{\text{risk of disease (unexposed)}}$$

$RR > 1$ (increased risk)

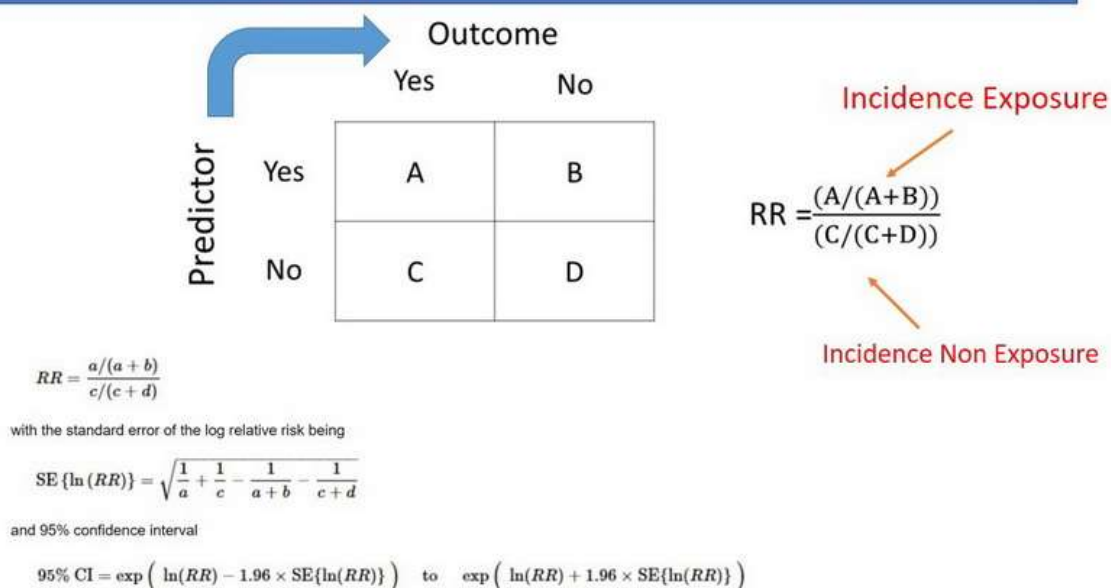
$RR = 1$ (same)

$RR < 1$ (lower risk)

Credit : Dr.Ranil Appuhamy ;Epidemiology made Easy

การคำนวณความเสี่ยง (Relative Risk :RR)

Relative Risk (RR) in prospective Study



Credit <https://www.medcalc.org>

Relative Risk คำนวณจาก Incidence ของการเกิดโรคในกลุ่ม Exposure หารด้วย Incidence ของการเกิดโรคในกลุ่ม Non Exposure โดย $RR = [a/(a+b)] / [c/(c+d)]$

การคำนวณ 95% Confidence Interval (CI) ของ OR โดยคำนวณจากสูตรในภาพ

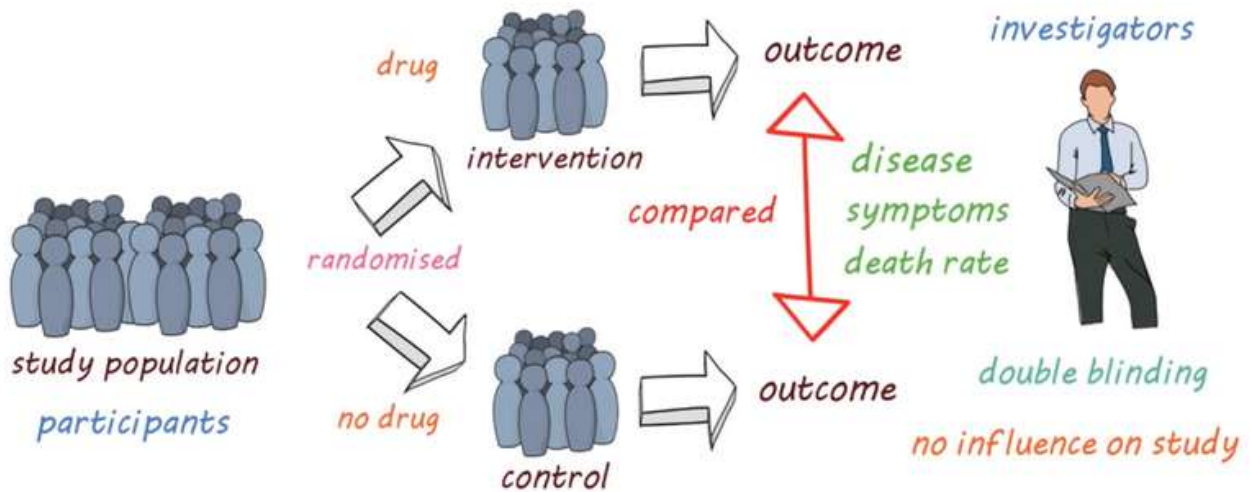
1. RR แสดงถึง ว่า ปัจจัยเสี่ยงมีความสัมพันธ์กับการเกิดโรค (Associate)
2. $RR < 1$ แสดงถึง ว่า ปัจจัยเสี่ยงสามารถปกป้องไม่ให้เกิดโรคได้ (Protective Factors)
3. $RR = 1$ แสดงถึง ว่า ปัจจัยเสี่ยงไม่มีความสัมพันธ์กับการเกิดโรค (Not Associate)
4. 95% CI ของ RR ไม่มีค่าใดค่าหนึ่งผ่าน 1 แสดงว่า มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ
5. 95% CI ของ RR มีค่าใดค่าหนึ่งผ่าน 1 แสดงว่า ไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ

จุดเด่นและจุดด้อยของการศึกษาแบบ Cohort Study

1. จุดเด่น เป็นการศึกษาไปข้างหน้าโดยติดตาม Exposure ไปข้างหน้า (Time Sequence) ถ้ามีความสัมพันธ์ จะเป็น ความสัมพันธ์ แบบ Casual Relationship หรือเป็นสาเหตุของการเกิดโรคนั้น
2. จุดด้อย คือ ค่าใช้จ่ายสูง ไม่เหมาะสำหรับการศึกษาในโรคที่พบบ่อย (rare disease) การติดตามกลุ่มเป้าหมายเป็น ระยะเวลาที่นาน ทำให้มีปัญหาในการติดตามกลุ่มที่ทำการศึกษา

Intervention Study : Randomized Controlled Trial

Randomised Controlled Trial best interventional study design



Credit : Dr.Ranil Appuhamy ;Epidemiology made Easy

Intervention Study ที่เป็น best study design คือ Randomized Control Trial (RCT) เนื่องจาก

1. ทำการสุ่ม (Randomised) กลุ่มศึกษา (study population) เป็นกลุ่มศึกษา (Intervention group) คือกลุ่มที่ได้ intervention หรือให้ยา หรือให้การรักษาด้วยวิธีการใดวิธีการหนึ่ง และกลุ่มควบคุม (Control group) คือกลุ่มที่ไม่ได้ Intervention หรือให้ยาหลอก (placebo)
2. วัดผลลัพธ์ (Outcome) ที่เกิดขึ้นจากทั้ง 2 กลุ่ม และทำการเปรียบเทียบว่ามีความแตกต่างกันอย่างมีนัยสำคัญหรือไม่
3. มีการ blind ผู้ที่ทำการศึกษา (Investigator) เพื่อไม่ให้ทราบว่าคุณุ่มไหนเป็นกลุ่มให้ intervention หรือกลุ่มไหนเป็นกลุ่มควบคุม

จุดเด่นและจุดด้อยของการศึกษา แบบ RCT

1. จุดเด่น เป็นการศึกษาไปข้างหน้าจากปัจจัยไปหาผลลัพธ์ ความสัมพันธ์ที่ได้ จึงเป็น Causal Relationship ,มีการสุ่มประชากรที่ศึกษา ทำให้ทั้ง 2 กลุ่มมีโอกาสถูกเลือกเท่าๆกัน หรือ ทั้ง 2 กลุ่มมีตัวแปรอื่นๆเหมือนกัน ต่างกันตรงที่ กลุ่มหนึ่งได้รับ Intervention อีกกลุ่มไม่ได้รับ Intervention
2. จุดด้อย คือ ต้นทุนสูง ต้องใช้กลุ่มตัวอย่างมาก และมีบางกรณีที่ไม่สามารถทำการศึกษาแบบ RCT ได้

Systematic Review & Meta analysis

SUMMARIES

systematic review



relevant studies

assess
synthesize
interprets

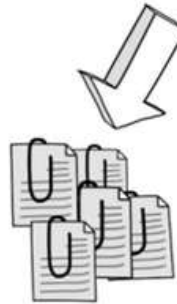


summary

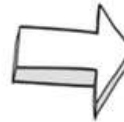
meta-analysis

studies of
similar design

single summary result



uses data



combined analysis

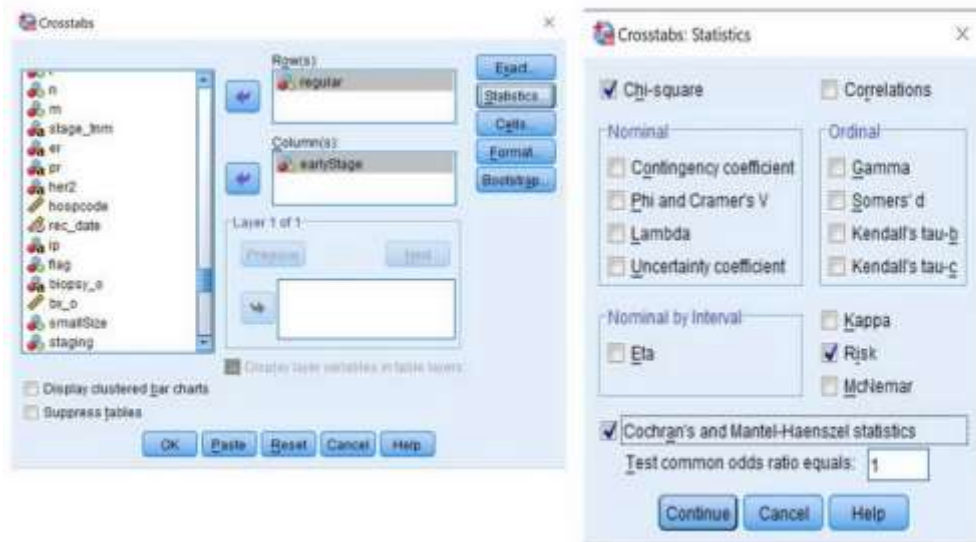
Credit : Dr.Ranil Appuhamy ;Epidemiology made Easy

1. Systematic Review เป็นการทบทวนเอกสารที่เกี่ยวข้องกับเรื่องที่เรากำลังศึกษา จากนั้นทำการประเมิน วิเคราะห์ และสังเคราะห์ แล้วเขียนรายงานขึ้นมาใหม่
2. Meta Analysisเป็นการทบทวนเอกสารที่เกี่ยวข้องกับเรื่องที่เราทำการศึกษา โดยเลือกเอกสารที่ออกแบบการศึกษา (Design study) ที่เหมือนกัน แล้วนำข้อมูลจากหลากหลายแหล่งเหล่านั้น มาทำการวิเคราะห์โดยใช้ชุดข้อมูลรวมเหล่านั้น (Combine analysis) เพื่อสรุปผลการศึกษาเป็นรายงานสรุปเดียว (Single Summary Result)

การวิเคราะห์ความสัมพันธ์ทางระบาดวิทยาโดยใช้ 2x2 Table

1. นักระบาดวิทยา มักวิเคราะห์ความสัมพันธ์ ระหว่าง Exposure กับ Outcome ในรูปของ 2x2 Table เพื่อหา
 1. RR (กรณีที่เป็น Cohort /Prospective หรือ Intervention Study)
 2. OR (กรณีที่เป็น Cross sectional ,Case Control / Retrospective study) มากกว่า การใช้ chi square Test
2. การวิเคราะห์ความสัมพันธ์ ด้วย OR ,RR นั้น เป็นการวิเคราะห์แบบง่าย ที่ไม่ซับซ้อน แต่มีข้อจำกัด คือ สามารถวิเคราะห์ตัวแปรที่เป็น Nominal หรือ Ordinal เท่านั้น แต่ในความเป็นจริง การศึกษานั้นมีตัวแปรประเภท Scale ด้วย ซึ่งจำเป็นต้องใช้โปรแกรมสถิติต่างๆ เช่น SPSS (เป็นโปรแกรมที่นักศึกษาส่วนใหญ่คุ้นเคย มากกว่า strata หรือ Epi Info) เช่น Compare Mean (T Test ,Anova) ,Logistic Regression , Time Series Analysis , Survival Analysis เป็นต้น

3. SPSS ก็มีโปรแกรมในการคำนวณ Risk เพื่อหา OR และ RR เช่นกัน ทำให้สะดวก เมื่อใช้ คำสั่ง Analysis --> Cross Tab ก็สามารถทำ 2x2 Table วิเคราะห์ได้ทั้ง Chi Square และ OR หรือ RR ได้



regular * earlyStage Crosstabulation

Count		earlyStage		Total
		Early stage	Late stage	
regular	Regular	1300	550	1850
	Non Regular	458	249	707
Total		1758	799	2557

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for regular (Regular / Non Regular)	1.285	1.069	1.544
For cohort earlyStage = Early stage	1.085	1.020	1.154
For cohort earlyStage = Late stage	.844	.747	.954
N of Valid Cases	2557		

Mantel-Haenszel Common Odds Ratio Estimate

Estimate	1.285
ln(Estimate)	.251
Std. Error of ln(Estimate)	.094
Asymp. Sig. (2-sided)	.007
Asymp. 95% Confidence Interval	
Common Odds Ratio	1.069
Lower Bound	1.069
Upper Bound	1.544
ln(Common Odds Ratio)	.067
Lower Bound	.067
Upper Bound	.435

The Mantel-Haenszel common odds ratio estimate is asymptotically normally distributed under the common odds ratio of 1.000 assumption. So is the natural log of the estimate.

4. เนื่องจากโปรแกรม SPSS มีลิขสิทธิ์ จึงใช้โปรแกรม Excel เขียนสูตรการคำนวณ OR, RR, 95% CI, Chi square และ รายงาน P value และ Significance เพื่อใส่ข้อมูลใน ตาราง 2x2 แล้ว จะคำนวณค่าทางสถิติให้ [Click เพื่อ Download โปรแกรม Excel](#)

		Early Stage		Total
		D+	D-	
Regular	E+	1,300	550	1,850
	E-	458	249	707
Total		1,758	799	2,557

Statistic Analysis				
		Lower	Upper	Sig
RR(MH)	1.085	1.020	1.154	Sig
OR (MH)	1.285	1.069	1.544	Sig

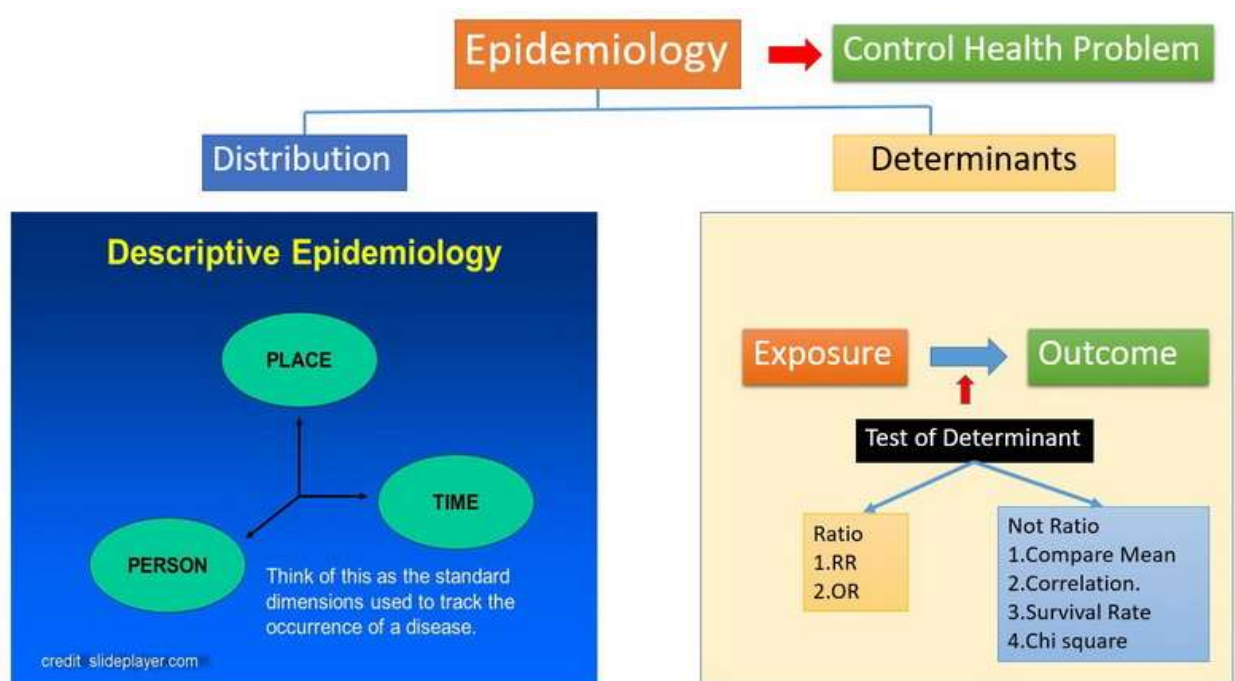
		Chi Square	P-value	Sig
Chi Square (Cochran)		7.18	0.02	Sig

การทดสอบความสัมพันธ์ทางระบาดวิทยาและทางสถิติ

Content

1. จุดมุ่งหมายของการทดสอบความสัมพันธ์
2. ประเภทของตัวแปร
3. การทดสอบสมมติฐาน
4. การทดสอบที่นักระบาดวิทยานิยมใช้ หา OR,RR,95% CI ,Chi square จาก 2x2 Table
 1. Exercise1 : Epidemiology study 1
 2. Exercise2 : Epidemiology study 2
5. การทดสอบด้วย Parametric Statistic
 1. การติดตั้ง Submenu **Data Analysis** ใน Menu Data ของ Excel เพื่อคำนวณค่าทางสถิติ
 2. Exercise 3 : การใช้ Excel ในการทดสอบ 2 samples t Test
6. เอกสารอ่านประกอบ
 1. [Video การ Install Analysis toolPack ใน Excel](#)
 2. [การทดสอบความสัมพันธ์ทางระบาดวิทยาและทางสถิติ¹](#)
 3. [Epidemiological Study²](#)
 4. [Excel ช่วย OR,OR,95%CI,Chi square³](#)

จุดมุ่งหมายของการทดสอบความสัมพันธ์



1. ระบาดวิทยา มาจากรากศัพท์ Epi = Upon ,demio=Population ,logos = Study หมายถึง การศึกษา (logos) เกี่ยวกับการกระจาย (Distribution) และหาสาเหตุหรือปัจจัยเสี่ยง (Determinant) ของโรคหรือเหตุการณ์โดยในกลุ่มประชากร (demio = Population) เฉพาะ (epi = upon) เพื่อนำผลการศึกษาไปสู่การควบคุมปัญหาสุขภาพ

Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems

2. การทดสอบความสัมพันธ์ คือ การใช้กระบวนการทางสถิติ ซึ่งเป็นกระบวนการทางวิทยาศาสตร์ เพื่อหาว่า อะไรเป็นสาเหตุ (Causal Relationship) หรือในกรณีที่ยังพิสูจน์หาสาเหตุไม่ได้ ก็จะหาว่าอะไรเป็นปัจจัยเสี่ยง (Non Causal Relationship) เพื่อจะนำไปสู่การหาสาเหตุที่แท้จริงต่อไป ซึ่งการทดสอบนั้นต้องใช้การทดสอบทางสถิติ ซึ่งสามารถแยกเป็น 2 ส่วน คือ การทดสอบทางสถิติเฉพาะ ที่นักระบาดวิทยาชอบใช้ (OR,RR,95% CI ของ OR,RR) และการทดสอบทางสถิติทั่วไปที่ใช้ในทุกวิชาซึ่งไม่เฉพาะนักระบาดวิทยา

ประเภทของตัวแปร

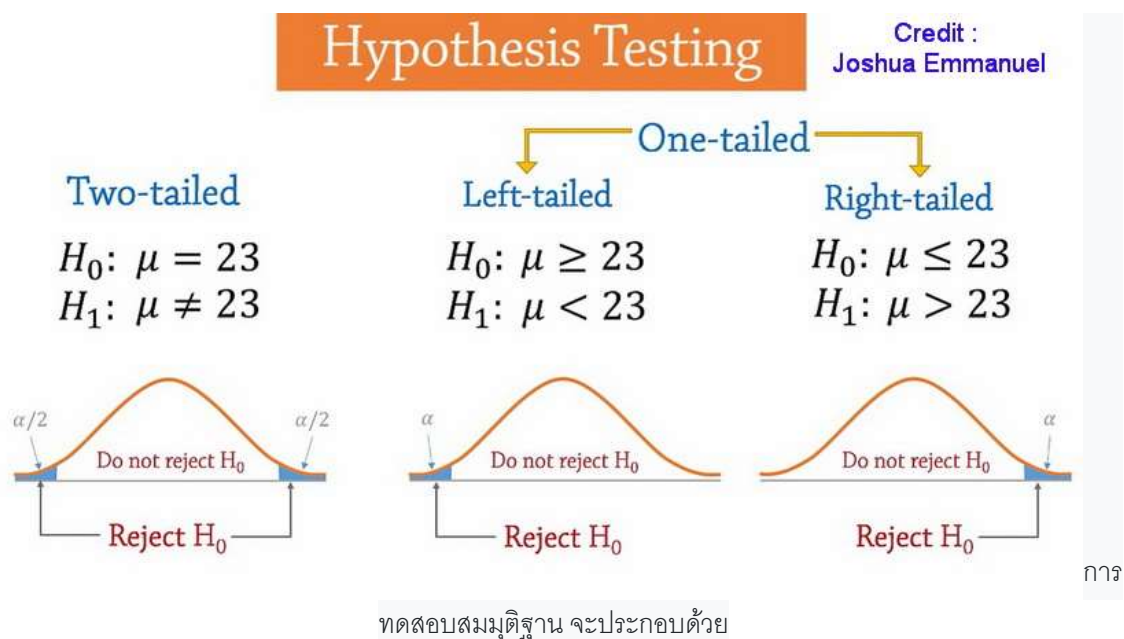
ตัวแปรแบ่งเป็น 3 ชนิด ได้แก่

1. ตัวแปรแบบ Nominal คือตัวแปรที่เป็นชื่อเรียกต่างๆ เช่น เพศ สถานภาพสมรส เป็นต้น แม้จะระบุค่าของตัวแปร เช่น 1=เพศชาย และ 2 เท่ากับเพศหญิง แต่ค่าตัวเลขดังกล่าวคือถือเป็น Character ไม่สามารถนำมาบวก ลบ คูณหาร กันได้
2. ตัวแปรแบบ ordinal คือตัวแปรประเภทลำดับ เช่น การวัดความพึงพอใจ เป็น 5 ลำดับ 1=พึงพอใจน้อยมาก ,2=พึงพอใจน้อย ,3=พึงพอใจปานกลาง ,4=พึงพอใจมาก 5=พึงพอใจอย่างมาก ตัวแปรประเภทนี้ ค่าของตัวเลข ไม่สามารถนำมาบวก ลบ คูณหาร กันได้ และ ความพึงพอใจ 4 ไม่ได้เป็น 2 เท่าของความพึงพอใจ 2 เพียงแต่สามารถบอกได้ว่า ความพึงพอใจ 5 มากกว่า 4 และ 4 มากกว่า 3 ฯลฯ ปัจจุบันมีเครื่องมือทางสถิติ ที่สามารถหาค่าสถิติ แบบ ordinal ได้ โดยทำการหา Mean Rank แทน Mean ที่เป็นค่าเฉลี่ย แทน ในทางปฏิบัติถ้าไม่เคร่งครัดในกฎเกณฑ์มากนัก อาจจะยอมให้นำมาบวก ลบ คูณหาร กันได้ เช่น การหาเกรดเฉลี่ยของการศึกษา ค่า GPAX เกิดจาก การนำเกรด x จำนวนหน่วยกิต แล้วนำมาหาเกรดเฉลี่ย เป็นต้น ซึ่งต้องระมัดระวังในการนำมาใช้ เพราะยังไเกรด A ,B ,C ,D ก็เป็นตัวแปรประเภท ordinal ไม่สามารถนำมาบวกลบกันได้
3. ตัวแปรแบบ Scale คือตัวแปรตัวเลขเป็นหน่วยวัดที่แท้จริง เช่น น้ำหนัก เป็น กิโลกรัม ส่วนสูงเป็นเมตร ค่า Cholesterol ในเลือดเป็น mg% เป็นต้น ตัวแปรประเภทนี้ สามารถหาค่า Mean & SD ได้ และส่วนสูง 2 เมตร จะเป็น 2 เท่าของส่วนสูง 1 เมตร
4. การทราบว่าตัวแปรเป็นประเภทใด จะทำให้เราเลือกได้ว่าจะใช้ Test อะไรในการหาความสัมพันธ์ ระหว่าง Exposures กับ Outcome (ดังตารางที่สรุป ด้านล่าง)

ประเภทของตัวแปร

ประเภท	หา Mean & SD	Statistic
Norminal	ไม่ได้	Non parametric เช่น Chi Square ,OR,RR
Ordinal	ไม่ได้ แต่ หา Mean rank ได้	Non Parametric แบบ Ordinal
Scale	ได้	Parametric เช่น Compare mean (t Test,ANOVA),Correlation Regression, Time Series,survival rate

การทดสอบสมมติฐาน



- H_0 หรือ Null Hypothesis หรือค่าเฉลี่ยของประชากรเท่ากัน ซึ่งสามารถกำหนดได้ 2 แบบ คือ
 - แบบ two tailed (กำหนดให้ค่าเฉลี่ยเท่ากัน)
 - แบบ one tailed โดย ถ้าเป็น left tailed (กำหนดให้ค่าเฉลี่ยมากกว่าหรือเท่ากับ) หรือ Right tailed (ค่าเฉลี่ยน้อยกว่าหรือเท่ากับ)
- H_1 หรือ Alternative Hypothesis หรือค่าเฉลี่ยของประชากร ไม่เท่ากัน (2 tailed) หรือ น้อยกว่า (กรณีที่เป็น Left tailed) หรือ มากกว่า (กรณีที่เป็น Right tailed) โดยถ้าทำการทดสอบทางสถิติแล้ว มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แสดงว่า ปฏิเสธ H_0 ซึ่งเท่ากับยอมรับ H_1

3. สำหรับการทดสอบด้วย OR หรือ RR นั้น จะทำการคำนวณหาค่า 95 % CI ของ OR หรือ RR ถ้าค่า 95% CI ไม่มีค่าใดผ่าน 1 แสดงว่ามีความแตกต่างอย่างมีนัยสำคัญทางสถิติ (ปฏิเสธ H_0 ยอมรับ H_1)

การทดสอบที่นักระบาดวิทยานิยมใช้ หา OR,RR,95% CI ,Chi square จาก 2x2 Table

1. ถ้าเป็นการศึกษาที่เริ่มจาก Outcome ก่อน แล้ววกไปหาว่ามี Exposure หรือไม่ ได้แก่ Case Control หรือ Cross sectional Study แล้ว การทดสอบที่นักระบาดวิทยานิยมใช้ คือ Odds Ratio (OR) ซึ่งทดสอบได้ว่าเป็นความสัมพันธ์กัน แต่ไม่สามารถสรุปว่าเป็นความสัมพันธ์แบบ Causal Relation ship
2. ถ้าเป็นการศึกษาที่เริ่มจาก Exposure ก่อน แล้วติดตามไปข้างหน้าว่าเกิด Outcome หรือไม่ ได้แก่ Cohort หรือ Intervention Study แล้ว การทดสอบที่นักระบาดวิทยานิยมใช้ คือ Relative Risk (RR) ซึ่งถ้าทดสอบได้ว่าเป็นความสัมพันธ์กัน สามารถสรุปว่าเป็นความสัมพันธ์แบบ Causal Relation ship หรือ Exposure เป็นสาเหตุให้เกิด Outcome สอดคล้องกับ Time Sequence ของการศึกษา
3. สามารถอ่านรายละเอียดของ Epidemiological Study ที่ Link [รายละเอียด](#)

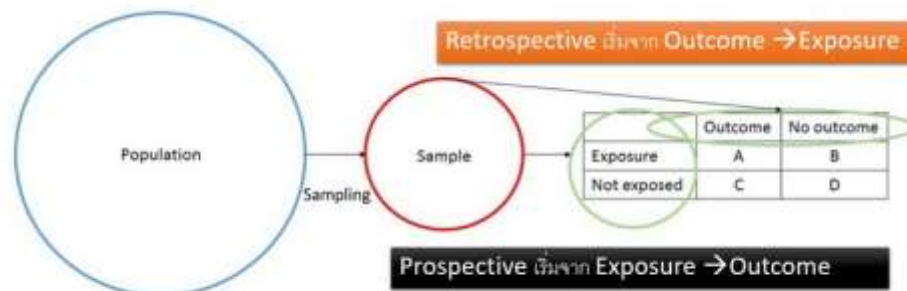
Type of Study and Epidemiological Test

Type of Study	Test	Causal Relation
1.Cross Sectional ¹	OR ³	No
2.Case Control (Retrospective Study) ¹	OR ³	No
3.Cohort (Prospective Study) ²	RR ⁴	Yes
4.Intervention Study (RCT) ²	RR ⁴	Yes

Remark : 1 = ศึกษาจาก Outcome ไปหา Exposure, 2=ศึกษาจาก Exposure แล้วติดตามการเกิด Outcome
3 = ย่อจาก Odds Ratio, 4 = ย่อจาก Relative Risk

แนวทางการศึกษา

Population ,Sample ,Exposure,Outcome(Event)

Credit : <https://www.scalefive.com>

สุ่มตัวอย่าง (Sampling) จากประชากร (Population) จากตัวอย่างที่ได้ จะทำเป็นตาราง 2x2 table โดยแบบแผนจะต้องตายตัวไม่ว่าจะเป็น Case Control(Retropective) หรือ Cohort (Prospective) เพื่อที่จะกำหนดค่า a,b,c,d คือ ส่วนแถว (row) จะเป็น Exposure (ได้รับ (+ve) หรือไม่ได้รับ (-ve) และส่วนของ Column คือส่วนของ Outcome เช่น Disease (ป่วย +ve ไม่ป่วย -ve) ต่างกันที่จุดเริ่มต้นหรือ Time Sequence โดย

1. ถ้าเริ่มจาก Outcome เช่น แบ่งเป็นกลุ่มป่วยกับไม่ป่วย แล้วถามย้อนไปหา Exposure ว่าสัมผัสหรือไม่ของทั้งกลุ่มป่วยและไม่ป่วย จะเป็นการศึกษาแบบ Case Control (Retrospective) หรือ Cross sectional จะคำนวณหาค่า OR และ 95% CI ของ Odds สูตรในการคำนวณค่า ดังภาพด้านล่าง โดยถ้าค่า 95% CI ของ OR ไม่มีค่าใดผ่าน 1 แสดงว่ามีความแตกต่างอย่างมีนัยสำคัญทางสถิติ หรือ มีความสัมพันธ์ระหว่าง Outcome กับ Exposure แต่ไม่สามารถบอกได้ว่าเป็น Causal Relationship

Odds ratio (OR) in Retrospective or Cross sectional Study

		Outcome	
		Yes	No
Treatment	Yes	A	B
	No	C	D

$$OR = \frac{(A \times D)}{(B \times C)}$$

$$OR = \frac{a/b}{c/d} = \frac{a \times d}{b \times c}$$

with the standard error of the log odds ratio being

$$SE\{\ln(OR)\} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

and 95% confidence interval

$$95\% \text{ CI} = \exp\left(\ln(OR) - 1.96 \times SE\{\ln(OR)\}\right) \text{ to } \exp\left(\ln(OR) + 1.96 \times SE\{\ln(OR)\}\right)$$

Credit <https://www.medcalc.org>

2. ถ้าเริ่มจาก Exposure เช่น แบ่งเป็นกลุ่มสัมผัสหรือไม่ได้สัมผัส Exposure แล้วติดตามไปข้างหน้า ว่าเป็นโรค ที่คนไม่เป็นโรคที่คน ของทั้ง 2 กลุ่ม เป็นการศึกษาแบบ Cohort (Prospective) หรือ Intervention study จะคำนวณหาค่า RR และ 95% CI ของ RR สูตรในการคำนวณหาค่า ดังภาพด้านล่าง โดยถ้าค่า 95% CI ของ RR ไม่มีค่าใดผ่าน 1 แสดงว่ามีความแตกต่างอย่างมีนัยสำคัญทางสถิติ หรือ มีความสัมพันธ์ระหว่าง Exposure กับ Outcome และอาจจะเป็นความสัมพันธ์แบบ Casual

Relationship

Relative Risk (RR) in prospective Study

		Outcome	
		Yes	No
Predictor	Yes	A	B
	No	C	D

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

with the standard error of the log relative risk being

$$SE\{\ln(RR)\} = \sqrt{\frac{1}{a} + \frac{1}{c} - \frac{1}{a+b} - \frac{1}{c+d}}$$

and 95% confidence interval

$$95\% \text{ CI} = \exp\left(\ln(RR) - 1.96 \times SE\{\ln(RR)\}\right) \text{ to } \exp\left(\ln(RR) + 1.96 \times SE\{\ln(RR)\}\right)$$

Incidence Exposure

Incidence Non Exposure

Credit <https://www.medcalc.org>

3. ในกรณีที่ต้องการคำนวณหาค่า Chi square ก็สามารถคำนวณโดยใช้สูตรด้านล่างนี้ โดยถ้าเป็น 2x2 Table ค่า Degree of Freedom (df)=(row-1)x(column-1) = (2-1)x(2-1) = 1 โดยค่า ค่า p=.05 นั้น ค่า Chi square ที่คำนวณได้เท่ากับ 3.841

Chi Square Test

		Outcome	
		Yes	No
Treatment	Yes	A	B
	No	C	D

$$\chi^2 = \frac{N(BC-AD)^2}{(A+B)(A+C)(B+D)(C+D)}$$

$$df = (row-1) \times (column-1) \\ = (2-1) \times (2-1) = 1$$

Critical values of the Chi-square distribution with d degrees of freedom

Probability of exceeding the critical value							
d	0.05	0.01	0.001	d	0.05	0.01	0.001
1	3.841	6.635	10.828	11	19.675	24.725	31.264

Epidemiological study 1

โจทย์ ทำการศึกษาความสัมพันธ์ระหว่างการได้รับฮอร์โมนทดแทนหลังหมดประจำเดือน (HRT) กับ การเกิดมะเร็งเต้านม โดยนำแฟ้มประวัติมะเร็งเต้านม จำนวน 800 ราย ได้รับ HRT 300 ราย ไม่ได้รับ 500 ราย และนำหญิงที่มาใช้บริการวัยทอง ที่ไม่เป็นมะเร็งเต้านม โดยเลือกช่วงอายุ และเศรษฐกิจที่ใกล้เคียงกับกลุ่มศึกษา จำนวน 1600 คน มาเป็นกลุ่มควบคุม โดยในกลุ่มควบคุมได้รับ HRT จำนวน 400 คน ไม่ได้รับจำนวน 1200 คน

คำถาม

1. การศึกษานี้เป็นการศึกษาประเภทใด
2. จะใช้ OR หรือ RR
3. ค่าที่ท่านตอบในข้อ 2 เท่ากับเท่าไร 95% CI เท่ากับเท่าไร
4. จะแปลผล 95% CI ในข้อ 3 อย่างไร มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ หรือไม่
5. Chi square เท่ากับเท่าไร มีนัยสำคัญทางสถิติหรือไม่
6. OR หรือ RR นั้นดีกว่า Chi Square อย่างไร

Epidemiological study 2

โจทย์ ทำการศึกษา การสูบบุหรี่กับมะเร็งปอด โดยติดตาม กลุ่มที่สูบบุหรี่ และไม่สูบบุหรี่ ไป 10 ปี พบว่า กลุ่มสูบบุหรี่ จำนวน 9,000 คน เป็นมะเร็งปอด 200 คน ส่วนกลุ่มที่ไม่สูบบุหรี่ จำนวน 18,000 คน เป็นมะเร็งปอด 100 คน

คำถาม

1. การศึกษานี้เป็นการศึกษาประเภทใด
2. จะใช้ OR หรือ RR
3. ค่าที่ท่านตอบในข้อ 2 เทียบเท่าไร 95% CI เทียบเท่าไร
4. จะแปลผล 95% CI ในข้อ 3 อย่างไร มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ หรือไม่
5. Chi square เทียบเท่าไร มีนัยสำคัญทางสถิติหรือไม่

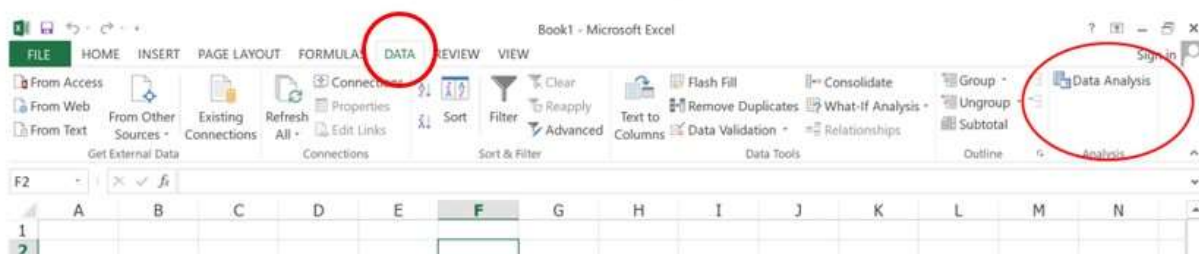
การทดสอบทางสถิติ แบบ Parametric Statistic

จะใช้ parametric ได้ในตัวแปรชนิด Scale ซึ่งสามารถหา Mean และ SD การทดสอบ แบบ Parametric Statistic เป็นการวิเคราะห์ที่ไม่สามารถใช้ 2x2 Table ได้ จำเป็นต้องใช้โปรแกรมทางสถิติ ซึ่งถ้าเป็นโปรแกรมที่เราคุ้นเคย เช่น SPSS ก็จะมีปัญหาเรื่องลิขสิทธิ์ ส่วนโปรแกรมที่ไม่มีลิขสิทธิ์มักจะเป็นโปรแกรมที่เราไม่คุ้นเคย โปรแกรมที่พวกเราคุ้นเคย และมีพร้อมในเครื่องคอมพิวเตอร์ของทุกคน คือ MS Excel ในส่วนของ Data Analysis ซึ่งสามารถคำนวณค่าสถิติได้ parametric Statistic ได้แก่

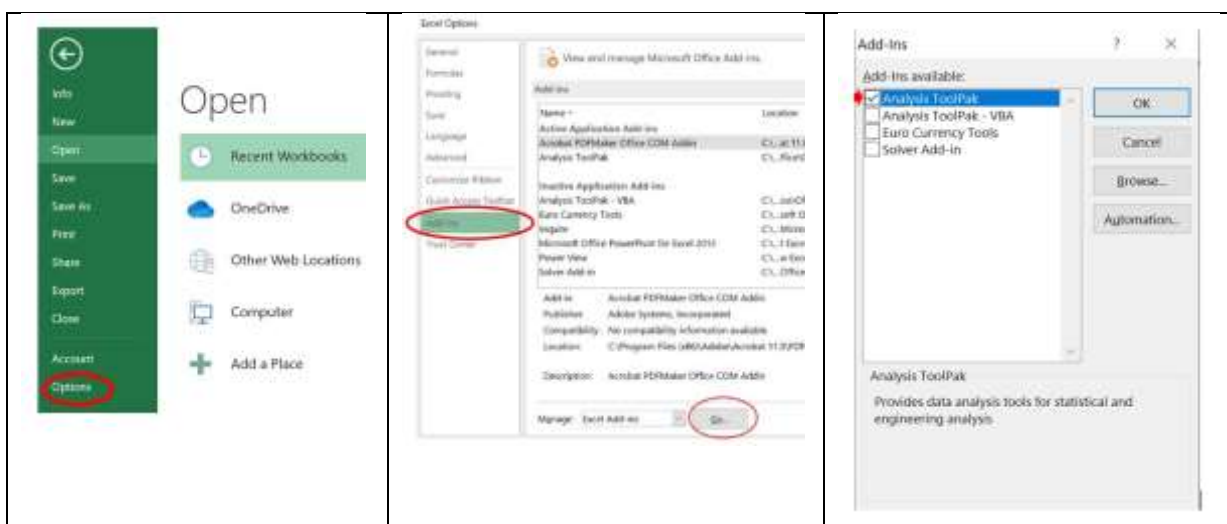
1. การเปรียบเทียบค่าเฉลี่ย (Compare mean) ด้วย t Test ,Paired t Test,ANOVA)
2. หาค่าสหสัมพันธ์ (Correlation)
3. Regression analysis (Linear หรือ Logistic regression)
4. Time series analysis
5. Survival Analysis

การติดตั้ง Submenu Data Analysis ใน Menu Data ของ Excel เพื่อคำนวณค่าทางสถิติ

1. ทำการตรวจสอบว่า Excel มี submenu : Data Analysis (ข้อมูลขวานบน) หรือไม่ โดย เปิดโปรแกรม Excel เลือก ที่ Menu Data แล้วดูว่า ที่มุมขวานบน มี Submenu ที่ชื่อ Data Analysis ซึ่งเป็นส่วนที่จะคำนวณทางสถิติหรือไม่ ถ้ามีแล้ว ไม่ต้องดำเนินการตามข้อ 2 ถ้าไม่มีให้ทำการ Add ตามข้อ 2



2. click ที่ menu File จากนั้น click ที่ Option (รูปที่ 1) จะเปิดมาที่หน้าใหม่ ให้ click ที่ Add-ins แล้ว Click ที่ปุ่ม Go ที่อยู่ด้านล่าง (รูปที่ 2) จะเปิดมาที่หน้าใหม่ click เพื่อเลือก check box --> Analysis ToolPak จากนั้น click OK



3. เมื่อกลับไป Excel เมื่อ Click ที่ Menu Data จะมี submenu ที่ชื่อ Data Analysis ที่มุมขวานบน (ก่อนหน้านี้ไม่มี)

Exercise 3 : การใช้ Excel ในการทดสอบ 2 samples t Test

1. ใส่ข้อมูล ตามตัวอย่าง โดยเป็นขนาดของก้อนมะเร็งของกลุ่มที่ Regular BSE และ Non Regular จาก Click ที่ menu Data เลือก submenu Data Analysis เลือกการทดสอบทางสถิติ t Test two samples equal variance

	A	B	C	D	E	F	G	H	I
1			Cancer size						
2	คนที่	regular BSE	Non Regular						
3	1	1.2	1.8						
4	2	1.3	1.9						
5	3	1.6	2.2						
6	4	1.8	2.4						
7	5	1.6	2.2						
8	6	1.8	2.4						
9	7	1.9	2.5						
10	8	2.1	2.7						
11	9	2.3	2.9						
12	10	2.5	3.1						
13	11	2.6	3.2						
14	12	2.7	3.3						
15	13	3.1	3.7						
16	14	3.5	4.1						
17	15	3.6	4.2						

not Assume

2. จากนั้นเลือกขอบเขตของข้อมูลให้ครอบคลุม cell ที่บันทึกข้อมูลทั้งกลุ่ม regular & Non Regular ตามตัวอย่าง จากนั้น click ok

	A	B	C	D	E	F	G	H	I	J
1			Cancer size							
2	คนที่	regular BSE	Non Regular							
3	1	1.2	1.8							
4	2	1.3	1.9							
5	3	1.6	2.2							
6	4	1.8	2.4							
7	5	1.6	2.2							
8	6	1.8	2.4							
9	7	1.9	2.5							
10	8	2.1	2.7							
11	9	2.3	2.9							
12	10	2.5	3.1							
13	11	2.6	3.2							
14	12	2.7	3.3							
15	13	3.1	3.7							
16	14	3.5	4.1							
17	15	3.6	4.2							

t-Test: Two-Sample Assuming Unequal Variances

Input

Variable 1 Range:

Variable 2 Range:

Hypothesized Mean Difference:

☒ Labels

Alpha:

Output options

☒ Output Range:

☐ New Worksheet Ply:

☐ New Workbook

3. การแปลผล output คือ กลุ่ม regular และ Non regular mean เท่ากับ 2.24 และ 2.84 ซม.ตามลำดับ p value (one tailed) = 0.019 ความแตกต่างอย่างมีนัยสำคัญทางสถิติ นั่นคือ กลุ่มที่ Regular BSE ขนาดก้อนเล็กกว่า กลุ่มที่ Nonregular BSE อย่างมีนัยสำคัญทางสถิติ

	A	B	C	D	E	F	G
1		Cancer size					
2	คนที่	regular BSE	Non Regular				
3	1	1.2	1.8	t-Test: Two-Sample Assuming Unequal Variances			
4	2	1.3	1.9				
5	3	1.6	2.2				
6	4	1.8	2.4				
7	5	1.6	2.2				
8	6	1.8	2.4				
9	7	1.9	2.5				
10	8	2.1	2.7				
11	9	2.3	2.9				
12	10	2.5	3.1				
13	11	2.6	3.2				
14	12	2.7	3.3				
15	13	3.1	3.7				
16	14	3.5	4.1				
17	15	3.6	4.2				

เฉลย Epidemiological Study 1.

Relative Risk ,Odds Ratio & Chi square in 2x2 Table									
Exposure		HRT							
Outcome		Breast Cancer							
ใส่ข้อมูลเฉพาะใน Cell ที่ระบุสีแดง					Statistic Analysis				
		Breast Cancer		Total			Lower	Upper	Sig
		D+	D-		RR(MH)	1.457	1.302	1.631	Sig
HRT	E+	300	400	700	OR (MH)	1.800	1.500	2.160	Sig
	E-	500	1,200	1,700			Chi Square	P-value	Sig
Total		800	1,600	2,400	Chi Square (Cochran)		40.34	0.001	Sig

1. การศึกษาเริ่มจาก Outcome หรือ disease (Breast Cancer) แล้วถามย้อนไปหาปัจจัยเสี่ยง คือ HRT เพราะฉะนั้นจึงมัน Case Control หรือ Retrospective study
2. ใช้ค่า OR
3. OR = 1.8 ค่า 95% CI = 1.50-2.16
4. แปลผลได้ว่า กลุ่มที่ได้รับ HRT เป็นมะเร็งเต้านม เป็น 1.8 เท่า เมื่อเทียบกับกลุ่มที่ไม่ได้รับ HRT (พยายามจะเลี้ยค่าว่า เสี่ยง เป็น 1.8 เท่า) และค่า 95% CI ไม่มีค่าใดผ่าน 1 แสดงว่ามีความแตกต่างอย่างมีนัยสำคัญทางสถิติ
5. chi square = 40.43 ค่า P value = .001 แสดงว่ามีความแตกต่างอย่างมีนัยสำคัญทางสถิติ
6. ค่า OR = 1.8 นั้นสามารถบอกทิศทางได้ว่ามากกว่า และสามารถอธิบายให้เข้าใจง่ายขึ้นว่า เป็นก็เท่าเมื่อเทียบกับกลุ่มที่ไม่ได้รับ HRT ซึ่ง chi square = 40.3 และ P = .001 ไม่สามารถสื่อความหมายให้กับประชาชนทั่วไปได้ดีเท่า OR

เฉลย Epidemiological Study 2.

Relative Risk ,Odds Ratio & Chi square in 2x2 Table									
Exposure		smoking							
Outcome		Lung Cancer							
ใส่ข้อมูลเฉพาะใน Cell ที่ระบายสีเหลือง					Statistic Analysis				
		Lung Cancer		Total			Lower	Upper	Sig
		D+	D-		RR(MH)	4.000	3.151	5.078	Sig
smoking	E+	200	8,800	9,000	OR (MH)	4.068	3.196	5.179	Sig
	E-	100	17,900	18,000			Chi Square	P-value	Sig
Total		300	26,700	27,000	Chi Square (Cochran)		151.69	0.001	Sig

1. การศึกษานี้เริ่มจากปัจจัยเสี่ยง(Smoking) แล้วติดตามไป 10 ปี ว่าเป็นมะเร็งปอดหรือไม่ (Outcome) จึงเป็น Cohort หรือ Prospective study
2. ใช้ค่า RR
3. OR = 4.0 ค่า 95% CI = 3.151-5.078
4. แปลผลได้ว่า กลุ่มที่สูบบุหรี่ เสี่ยงต่อการ เป็นมะเร็งปอด เป็น 4 เท่า เมื่อเทียบกับกลุ่มที่ไม่สูบบุหรี่ และค่า 95% CI ไม่มีค่าใดผ่าน 1 แสดงว่ามีความแตกต่างอย่างมีนัยสำคัญทางสถิติ
5. chi square = 151.69 ค่า P value =.001 แตกต่างอย่างมีนัยสำคัญทางสถิติ

ให้ [Download File Excel](#) เพื่อช่วยคำนวณ เพื่อตอบโจทย์ในข้อ 1-5

1. ติดตามหญิงที่มาฝากครรภ์ 44,978 ราย เป็นกลุ่ม Teenage 13,291 ราย และ Non Teenage 31,687 ราย จากนั้นติดตามจนกระทั่งคลอด พบว่า ในกลุ่ม Teenage เด็กมี LBW 7,868 ราย Normal Weight 5,423 ราย ส่วนแม่ Non Teenage เด็กมี LBW 19,200 Normal Weight 12,487 ราย ให้ตอบคำถามดังต่อไปนี้
 - 1.1. การศึกษานี้เป็นการศึกษาประเภทใด
 - 1.2. จะใช้ OR หรือ RR และค่า 95% CI เท่ากับเท่าไร มีนัยสำคัญทางสถิติหรือไม่ จะสรุปผลการศึกษาว่าอย่างไร
2. นำข้อมูลจาก Medical record มาทำการวิเคราะห์ พบว่า หญิงที่มาคลอดทั้งหมด 57,372 ราย พบ Low birth Weight (LBW) 5,200 ราย และ Normal Weight 52,172 ราย ในกลุ่ม LBW จำนวน 5,200 ราย พบ Asphyxia 382 ราย และ Non Asphyxia 4,818 ราย ในกลุ่ม Normal Weight 52,172 ราย พบ Asphyxia 893 ราย และ Non Asphyxia 51,279 ราย ให้ตอบคำถามดังต่อไปนี้
 - 2.1. การศึกษานี้เป็นการศึกษาประเภทใด
 - 2.2. จะใช้ OR หรือ RR และค่า 95% CI เท่ากับเท่าไร มีนัยสำคัญทางสถิติหรือไม่ จะสรุปผลการศึกษาว่าอย่างไร
3. ศึกษาการได้รับยาเม็ดเสริมธาตุเหล็กครบถ้วน กับภาวะโลหิตจาง โดยกลุ่มที่เป็นโลหิตจางจำนวน 16,210 ราย ส่วนกลุ่มควบคุมคือกลุ่มที่เป็นโลหิตจางจำนวน 62,888 คน โดยกลุ่มที่เป็นโลหิตจาง ได้รับยาเม็ดเสริมธาตุเหล็กครบถ้วน 14,031 ราย ไม่ครบถ้วน 2,179 ราย กลุ่มที่ไม่โลหิตจาง ได้รับ ได้รับยาเม็ดเสริมธาตุเหล็กครบถ้วน 55,104 ราย ไม่ครบถ้วน 7,784 ราย ให้ตอบคำถามดังต่อไปนี้
 - 3.1. การศึกษานี้เป็นการศึกษาประเภทใด
 - 3.2. จะใช้ OR หรือ RR และค่า 95% CI เท่ากับเท่าไร มีนัยสำคัญทางสถิติหรือไม่ จะสรุปผลการศึกษาว่าอย่างไร
4. นำข้อมูลจาก medical recprd ผู้ที่เป็นมะเร็งเต้านมจำนวน 1,758 ราย เป็นมะเร็งระยะแรก (Early Stage) จำนวน 1,758 ราย ระยะหลัง (Late Stage) จำนวน 799 ราย และดูจากบันทึกเกี่ยวกับประวัติการตรวจเต้านมสม่ำเสมอ ในช่วง 12 เดือน ก่อนการวินิจฉัยมะเร็งเต้านม พบว่า กลุ่ม early stage ตรวจเต้านมด้วยตนเองสม่ำเสมอ 1300 ราย ไม่สม่ำเสมอ 458 ราย กลุ่ม Late Stage ตรวจเต้านมด้วยตนเองสม่ำเสมอ 550 ราย ไม่สม่ำเสมอ 249 ราย ให้ตอบคำถามดังต่อไปนี้
 - 4.1. การศึกษานี้เป็นการศึกษาประเภทใด
 - 4.2. จะใช้ OR หรือ RR และค่า 95% CI เท่ากับเท่าไร มีนัยสำคัญทางสถิติหรือไม่ จะสรุปผลการศึกษาว่าอย่างไร
5. ทำการศึกษา โดยให้วิตามินเสริมธาตุเหล็กชนิดใหม่ กลุ่มแรกหญิงที่นัดมาฝากครรภ์ วันจันทร์ พุธ ศุกร์ กลุ่มที่ 2 สำหรับที่มารวัน อังคาร, พฤหัสบดี ติดตามทั้ง 2 กลุ่มจนถึงคลอด พบว่า กลุ่มที่ได้วิตามินเสริมธาตุเหล็กชนิดใหม่ จำนวน 80,289 ราย คลอดบุตร LBW 43,155 ราย Normal Weight 37,134 ราย สำหรับกลุ่มที่ได้ธาตุเหล็กธรรมดา 25,508 ราย คลอด บุตร LBW 9532 ราย Normal Weight 15976 ราย ตอบคำถามดังนี้
 - 5.1. การศึกษานี้เป็นการศึกษาประเภทใด
 - 5.2. จะใช้ OR หรือ RR และค่า 95% CI เท่ากับเท่าไร มีนัยสำคัญทางสถิติหรือไม่ จะสรุปผลการศึกษาว่าอย่างไร

Exercise 1

Relative Risk ,Odds Ratio & Chi square in 2x2 Table

Exposure	Teenage
Outcome	LBW

ใส่ข้อมูลเฉพาะใน Cell ที่ระบายสีเหลือง

		LBW		Total
		D+	D-	
Teenage	E+	7,868	5,423	13,291
	E-	19,200	12,487	31,687
Total		27,068	17,910	44,978

Statistic Analysis

		Lower	Upper	Sig
RR(MH)	0.977	0.961	0.993	Sig
OR (MH)	0.944	0.905	0.983	Sig
		Chi Square	P-value	Sig
Chi Square (Cochran)		7.60	0.02	Sig

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for teenage_1 (Teenage / Not Teenage)	.944	.905	.983
For cohort LBW = LBW	.977	.961	.993
For cohort LBW = Normal Birth weight	1.035	1.010	1.061
N of Valid Cases	44978		

Exercise 2

Relative Risk ,Odds Ratio & Chi square in 2x2 Table

Exposure	LBW
Outcome	Asphyxia

ใส่ข้อมูลเฉพาะใน Cell ที่ระบุสีเหลือง

		Asphyxia		Total
		D+	D-	
LBW	E+	382	4,818	5,200
	E-	893	51,279	52,172
Total		1,275	56,097	57,372

Statistic Analysis

		Lower	Upper	Sig
RR(MH)	4.292	3.820	4.822	Sig
OR (MH)	4.553	4.024	5.151	Sig
		Chi Square	P-value	Sig
Chi Square (Cochran)		690.88	0.001	Sig

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for LBW (LBW / Normal Birth weight)	4.553	4.024	5.151
For cohort asphyxia = Asphyxia	4.292	3.820	4.822
For cohort asphyxia = Not Asphyxia	.943	.935	.950
N of Valid Cases	57372		

Exercise 3

Relative Risk ,Odds Ratio & Chi square in 2x2 Table

Exposure	Fe
Outcome	Anemai

ใส่ข้อมูลเฉพาะใน Cell ที่ระบายสีเหลือง

		Anemai		Total
		D+	D-	
E	E+	14,031	55,104	69,135
	E-	2,179	7,784	9,963
Total		16,210	62,888	79,098

Statistic Analysis

		Lower	Upper	Sig
RR(MH)	0.928	0.892	0.966	Sig
OR (MH)	0.910	0.864	0.957	Sig
		Chi Square	P-value	Sig
Chi Square (Cochran)		13.27	0.001	Sig

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for "ได้รับวิตามินเสริม Fe (ได้ / ไม่ได้)"	.910	.864	.957
For cohort anemia_1 = Anemia	.928	.892	.966
For cohort anemia_1 = Normal	1.020	1.009	1.032
N of Valid Cases	79098		

Exercise 4

Relative Risk ,Odds Ratio & Chi square in 2x2 Table

Exposure	BSE
Outcome	Early Stage

ใส่ข้อมูลเฉพาะใน Cell ที่ระบายสีเหลือง

		Early Stage		Total
		D+	D-	
BSE	E+	1,300	550	1,850
	E-	458	249	707
Total		1,758	799	2,557

Statistic Analysis

		Lower	Upper	Sig
RR(MH)	1.085	1.020	1.154	Sig
OR (MH)	1.285	1.069	1.544	Sig
		Chi Square	P-value	Sig
Chi Square (Cochran)		7.18	0.02	Sig

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for regular (Regular / Non Regular)	1.285	1.069	1.544
For cohort earlyStage = Early stage	1.085	1.020	1.154
For cohort earlyStage = Late stage	.844	.747	.954
N of Valid Cases	2557		

Exercise 5

Relative Risk ,Odds Ratio & Chi square in 2x2 Table

Exposure	New drug
Outcome	LBW

ใส่ข้อมูลเฉพาะใน Cell ที่ระบายสีเหลือง

		LBW		Total
		D+	D-	
New drug	E+	43,155	37,134	80,289
	E-	9,532	15,976	25,508
Total		52,687	53,110	105,797

Statistic Analysis

		Lower	Upper	Sig
RR(MH)	1.438	1.414	1.463	Sig
OR (MH)	1.948	1.892	2.005	Sig
		Chi Square	P-value	Sig
Chi Square (Cochran)		2077.79	0.001	Sig

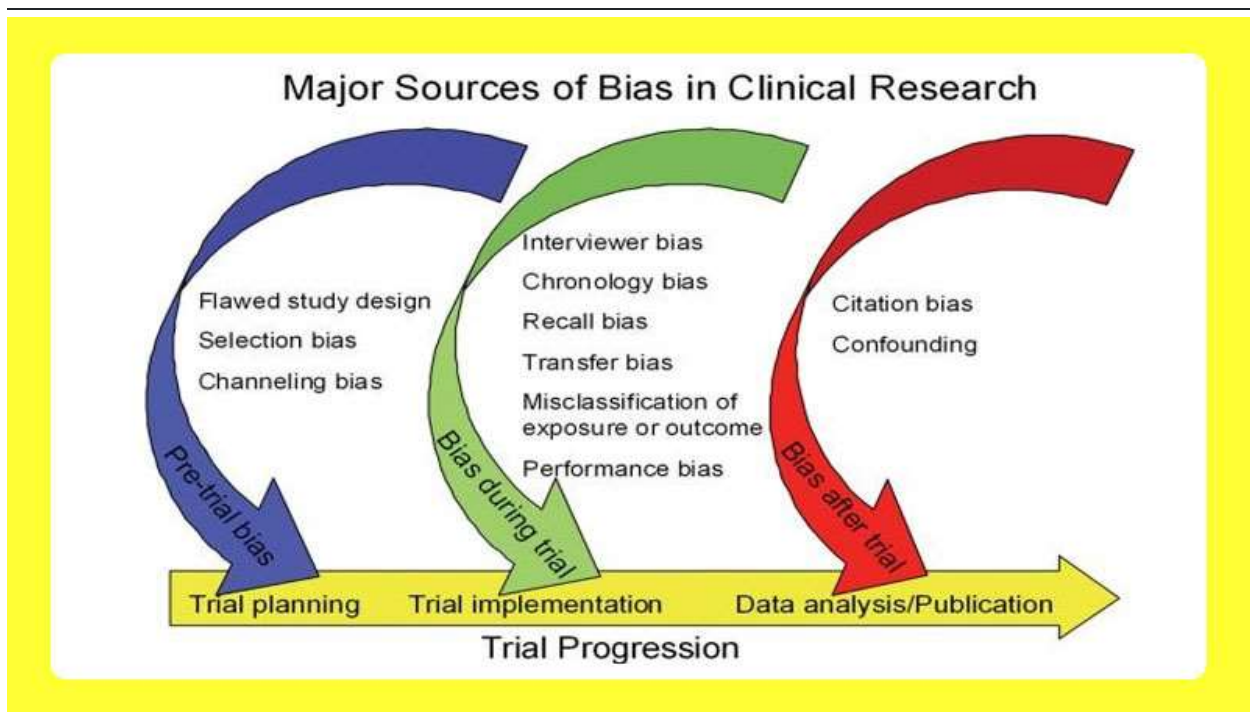
Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for ได้รับวิตามินเสริม Fe (ได้ / ไม่ได้)	1.948	1.892	2.005
For cohort LBW = LBW	1.438	1.414	1.463
For cohort LBW = Normal Birth weight	.738	.730	.747
N of Valid Cases	105797		

Bias in Clinical Research

แบ่งได้เป็น 3 ประเภท ได้แก่

1. [Bias during study design](#)
2. [Bias during clinical trail](#)
3. [Bias after a trial](#)
4. เอกสารอ่านประกอบ
 1. [Bias in clinical Research Bias in clinical Research](#)



Bias during study design

1. การออกแบบวิจัยที่ไม่ดี (Flawed study design) สาเหตุมักเกิดจาก
 1. นิยามของตัวแปรต่างๆ ไม่ชัดเจน โดยเฉพาะตัวแปรที่เป็น Exposure และ Outcome เช่น ไม่ชัดเจน เช่น เช่น ความสม่ำเสมอของการตรวจด้วยตนเอง ก้อนมะเร็งเต้านมขนาดเล็ก หรือ มะเร็งเต้านมระยะแรก เมื่อนิยามไม่ชัดเจนก็จะแปลงเป็นค่าตัวเลขได้ไม่ชัดเจนตามไปด้วย ทำการวิเคราะห์ทางสถิติมีปัญหาตามมา
 2. การเก็บข้อมูล ได้แก่ แบบสอบถาม การตรวจร่างกาย การทำ Lab หรือการตรวจทางรังสี ต่างๆ รวมถึงการฝึกอบรมผู้เก็บข้อมูล เพื่อชี้แจงให้เข้าใจนิยามต่างๆที่ใช้ในการวิจัยให้ตรงกัน รวมถึงการออกแบบว่ามี Blind หรือไม่ให้ผู้เก็บข้อมูล ผู้วิเคราะห์ข้อมูล ไม่ทราบว่ากลุ่มไหนเป็นกลุ่มศึกษา กลุ่มไหนเป็นกลุ่มควบคุม

2. Selection bias คือความลำเอียงในการเลือกกลุ่มตัวอย่างและกลุ่มควบคุม จึงต้องกำหนด inclusion & Exclusion Criteria ที่ชัดเจนเพื่อหลีกเลี่ยง selection bias
3. Channeling bias เกิดจากปัจจัยของการพยากรณ์โรคที่แตกต่างกันมีอิทธิพลต่อการตัดสินใจในการกำหนดกลุ่มศึกษาไปข้างหน้า ซึ่งจะเกิดบ่อยในกลุ่ม non randomized trial เช่น การผ่าตัดบริเวณมือนั้นในเด็ก ศัลยแพทย์มักจะตัดสินใจในการผ่าตัดมือในกลุ่มเด็กมากกว่ากลุ่มผู้สูงอายุ ทำให้เกิดความลำเอียงในการเลือกจัด channel ของกลุ่มศึกษาและกลุ่มควบคุม

Bias during clinical trial

Information bias เช่น ความผิดพลาดของข้อมูลในการวัดทั้ง exposure และ outcome

1. Interview bias โดยเฉพาะผู้สัมภาษณ์ที่รู้ว่าโรคนี้มีอะไรเป็นปัจจัยเสี่ยง ก็จะพยายามถามนำว่ามีปัจจัยเสี่ยงนั้นหรือไม่ ทำให้เกิดความลำเอียงขึ้น

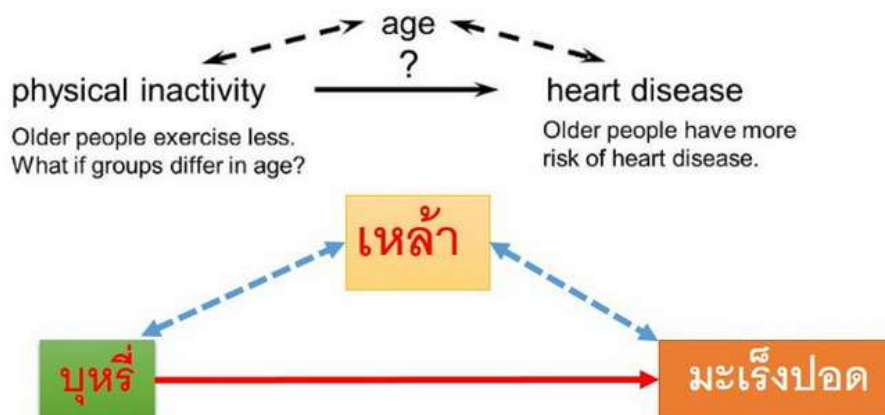
Chronology bias ส่วนใหญ่มักจะเกิดการศึกษาคือทำ intervention นี้เมื่อเทียบกับเมื่อก่อน โดยให้กลุ่มที่ทำ intervention เป็นกลุ่มศึกษา และกลุ่มที่ไม่ได้ทำ intervention ก่อนหน้านั้นเป็น Historic control ซึ่งบริบทของทั้ง 2 เวลามีความแตกต่างกัน ทำให้เกิด bias ขึ้นได้

2. Recall bias เกิดจากวิจัยแบบ retrospective เช่น จะหาความสัมพันธ์ระหว่าง ความพิการแต่กำเนิด กับการได้รับรังสีขณะตั้งครรภ์ ถ้าถามในกลุ่มแม่ที่คลอดลูกพิการ จะมีความลำเอียง ที่จะ recall ถึงการได้รับรังสี มากกว่า กลุ่มที่คลอดลูกไม่พิการ เช่นกลุ่มที่คลอดลูกพิการ เคยเดินผ่านห้องเอกซเรย์ก็จะ recall ถึงเหตุการณ์นั้น ทำให้ตอบว่าใช่ ซึ่งต่างกับกลุ่มที่คลอดลูกไม่พิการ มักจะไม่ recall ไปไกลขนาดนั้น
3. Transfer bias จะเกิดในการศึกษาแบบ cohort ที่มีอัตราการ loss Follow up ในกลุ่มศึกษาและกลุ่มควบคุมที่ไม่เท่ากัน ทำให้ไม่ทราบว่าการกลุ่มที่ตามไปได้นั้น ผลลัพธ์ที่แท้จริงเป็นอย่างไร
4. Bias from misclassification of exposure or outcome ในทางคลินิก บางครั้งไม่สามารถจัดว่า Exposure หรือ outcome นั้นควรจัดอยู่ที่กลุ่มไหน เนื่องจากเกณฑ์ในการจัดไม่ชัดเจน หรือเกณฑ์อาจจะชัดเจน แต่ประสบการณ์ของแพทย์ไม่เพียงพอที่จะจัดว่าอยู่ในกลุ่มใดได้
5. Performance bias เกิดจากคุณภาพของการทำ Intervention นั้นมีความแตกต่างระหว่างบุคคล เช่นผลการผ่าตัด เมื่อเปรียบเทียบกับ การไม่ผ่าตัด คุณภาพการทำผ่าตัดของศัลยแพทย์แต่ละคนไม่เท่ากัน ทำให้เกิด Performance bias ขึ้น ซึ่งแนวทางการแก้ไข อาจจะต้อง classified ศัลยแพทย์ เป็นกลุ่มๆ เช่น กลุ่มที่ผ่าตัดน้อยกว่า 100 ราย หรือมากกว่า 100 ราย เป็นต้น

Bias after a trial

1. Citation bias เกิดจากผู้วิจัยและผู้สนับสนุนการวิจัย มักไม่ยินดีที่จะเผยแพร่การวิจัยที่ผลลัพธ์ไม่เป็นไปตามความต้องการของผู้ทำวิจัย หรือผู้สนับสนุนการวิจัย เพราะฉะนั้น ผลวิจัยที่สอดคล้องกับความต้องการของผู้วิจัยหรือผู้สนับสนุนการวิจัยมักจะมีโอกาสที่จะตีพิมพ์มากกว่าผลการวิจัยที่ไม่สอดคล้อง
2. Confounding คือการที่ตัวแปรที่เป็น (Exposure) ที่เราทำการศึกษา (เช่น กินเหล้า) ไปสัมพันธ์กับตัวแปรอีกตัวที่อยู่นอกกรอบของการศึกษาและเป็นสาเหตุที่ทำให้เกิดผลลัพธ์ (Outcome) ที่ทำการศึกษา (เช่น มะเร็งปอด) ทำให้ตัวแปรที่เป็น Exposure นั้น (เหล้า) ไปมีความสัมพันธ์กับผลลัพธ์ที่ทำการศึกษา (มะเร็งปอด) แบบ Non Causal Relationship การป้องกันปัญหา Confounding นั้นโดยการตระหนักของ Confounding factors ดังกล่าว ในการวิเคราะห์ข้อมูลต้องกำหนดให้เป็น Covariate ในส่วนที่เป็น Cofounding Factors ที่ไม่ทราบมาก่อน การศึกษาแบบ randomization.

Confounding Variable/factor



Confounding Variable คือ ตัวแปรที่เราศึกษาไปสัมพันธ์กับ ตัวแปรซึ่งเป็นต้นเหตุให้เกิด outcome นั้น ทำให้ตัวแปรนั้นเกิดความสัมพันธ์เทียมกับ outcome

Confounding Variable คือ ตัวแปรที่เราศึกษา (ในที่นี้คือ Age) ไปสัมพันธ์กับ ตัวแปรซึ่งเป็นต้นเหตุให้เกิด outcome นั้น (ในที่นี้คือ Physical Inactivity) ทำให้ตัวแปรนั้นเกิดความสัมพันธ์เทียมกับ outcome (ในที่นี้คือ Heart Diseases) ขอยกตัวอย่างโดยดูรูปด้านล่างประกอบ

เพราะฉะนั้น เวลาไปทดสอบความสัมพันธ์แล้วพบว่า Exposure นั้นไม่น่าจะทำให้เกิด Outcome เช่น กินเหล้า กับมะเร็งปอด ไม่น่าจะเป็นสาเหตุของมะเร็งปอด ให้คิดถึงว่าน่าจะเกิดจากปัจจัยในเรื่อง Confounding Factor เนื่องจากกลุ่มที่กินเหล้ามักจะสูบบุหรี่ด้วย ซึ่งเป็นบุหรีเป็นสาเหตุโดยตรงที่ทำให้เกิดมะเร็งปอด เมื่อไปทดสอบความสัมพันธ์ระหว่างเหล้ากับมะเร็งปอด จึงเกิดความสัมพันธ์เทียม (Non Casual Relationship) ขึ้น ในทางสถิติ จึงมีการ fixed Covariate ขึ้น ในตัวอย่างนี้ ถ้าแยกกลุ่ม แล้วหาความสัมพันธ์ใหม่ เป็น 2 กลุ่มใหญ่ดังนี้

1. กลุ่มที่สูบบุหรี่ กินเหล้า ก็คน ไม่กินเหล้า ก็คน เป็นมะเร็งปอดก็คน ไม่เป็นมะเร็งปอดก็คน
2. กลุ่มที่ไม่สูบบุหรี่ กินเหล้า ก็คน ไม่กินเหล้า ก็คน เป็นมะเร็งปอดก็คน ไม่เป็นมะเร็งปอดก็คน

จากนั้นมาทดสอบความสัมพันธ์ ถ้าแยกกลุ่มเป็นสูบบุหรี่ กับไม่สูบบุหรี่ ทดสอบความสัมพันธ์ระหว่าง เหล้า กับมะเร็งปอด แล้ว พบว่าไม่มีความสัมพันธ์กันทั้ง 2 กลุ่ม แสดงว่า การมีความสัมพันธ์ในครั้งแรกที่ไม่มีการแยกกลุ่มเกิดจาก Confounding Factor

3. Interval Validity & External Validity

1. Internal Validity ผลการศึกษานำไปใช้ได้เฉพาะในเงื่อนไขอุดมคติและในกลุ่มประชากรที่เลือกสรรเฉพาะ ไม่สามารถนำไปใช้ในสถานการณ์ทั่วไปได้ หรือจะกล่าวได้ว่า ไม่สามารถนำไปขยายผลในพื้นที่อื่น และกลุ่มประชากรที่อื่นๆได้ (test efficacy of a specific intervention under idealized conditions in a highly selected population.)
2. External validity of research design ผลการศึกษาสามารถนำไปใช้ได้ในกลุ่มประชากรอื่นๆหรือพื้นที่อื่นๆ ได้ หรือสามารถนำไปขยายผลได้ (deals with the degree to which findings are able to be generalized to other groups or populations.)