Concepts of Association and Causation

ไพบลย์ โล่ห์สุนทร*

I. บทนำ

ในการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ นักระบาดวิทยามักจะพยายามวินิจฉัยแยกความสัมพันธ์ระหว่างตัวแปรว่าเป็นความสัมพันธ์ทางเหตุและผลต่อกันหรือไม่ ความสัมพันธ์อาจเป็นในรูปตัวแปรตัวหนึ่ง ทำให้เกิดผลอย่างหนึ่งหรือตัวแปรหลายตัว ทำให้เกิดผลอย่างเดียว ความสัมพันธ์ทางสถิติระหว่างตัวแปรต่าง ๆ เป็นจุดเริ่มต้นในการที่จะค้นหาสาเหตุของโรค การค้นพบความสัมพันธ์ทางสาเหตุนับเป็นสิ่งสำคัญในการป้องกันโรคอย่างมีประสิทธิภาพ ในการที่จะบอกว่าองค์ประกอบซึ่งสัมพันธ์กับโรคหนึ่งเป็นสาเหตุของโรคนั้น จำเป็นที่จะต้องมี **"หลักเกณฑ์"** ในการพิจารณา

การศึกษาเกี่ยวกับความสัมพันธ์ในทางระบาดวิทยามีจุดประสงค์ดังนี้คือ

1. ค้นหาองค์ประกอบต่าง ๆ ที่เลียงต่อการเกิดโรค
2. พิจารณาองค์ประกอบที่เป็นสาเหตุของโรค
3. วางมาตรการในการป้องกันและควบคุมโรค

บทความนี้จะอธิบายให้เห็นถึงความแตกต่างระหว่าง **"Association"** และ **"Causation"** ชนิดของความสัมพันธ์ การทดสอบความสัมพันธ์หลักเกณฑ์ในการพิจารณาความสัมพันธ์ทางสาเหตุและขั้นตอนในการค้นหาความสัมพันธ์ทางสาเหตุ

II. ความหมาย

Association หมายถึงความสัมพันธ์ระหว่างการเกิดของของสิ่งหนึ่งกับของอีกสิ่งหนึ่งหรือของหลายสิ่งกับของอีกสิ่งหนึ่ง ความสัมพันธ์อาจศึกษาได้จากการเปรียบเทียบบุคคลที่มีลักษณะอย่างหนึ่งหรือได้รับองค์ประกอบอย่างหนึ่งว่ามีอัตราการเกิดโรคสูงกว่าบุคคลที่ไม่มีลักษณะอย่างนั้นหรือไม่ได้รับองค์ประกอบอย่างนั้น ในทางตรงข้ามอาจเปรียบเทียบจำนวนองค์ประกอบที่ได้รับในกลุ่มบุคคลที่เป็นโรคและไม่เป็นโรค นอกจากนี้อาจดูได้จากระดับของความสัมพันธ์ระหว่างตัวแปรต่าง ๆ เช่น สหสัมพันธ์ (Correlation) เป็นต้น

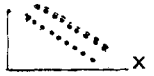
ความสัมพันธ์ระหว่างของสองสิ่งอาจเป็นในแบบ

ก. ความสัมพันธ์บวกคือ

ของสองสิ่งมักพบเกิดร่วมกัน

*แผนกวิชาเวชศาสตร์ป้องกันและสังคม คณะแพทยศาสตร์ จุฬาลงกรณมหาวิทยาลัย

ข. ความสัมพันธ์ลบคือ ของสองสิ่งมักพบเกิดในทางตรงข้ามกัน



ค. ไม่มีความสัมพันธ์คือ ของสองสิ่งต่างเกิดขึ้นเองเป็นอิสระต่อกัน

เมื่อพบมีความสัมพันธ์ระหว่างองค์ประกอบและโรคเกิดขึ้น ความสัมพันธ์นั้นอาจเป็นจริงหรือไม่จริงก็ได้

1. อาจไม่เป็นจริง (Artifactual) โดยเนื่องมาจาก

ก. เกิดขึ้นโดยบังเอิญ

ข. เกิดขึ้นจากอิทธิพลขององค์ประกอบอื่นซึ่งไม่ได้ควบคุม

ค. เกิดขึ้นจากความลำเอียงหรืออคติ

ง. เกิดขึ้นจากความผิดพลาดหรือข้อสมมุติที่ผิดพลาดในการศึกษาโดยตั้งใจหรือไม่ก็ตาม ทำให้สิ่งต่าง ๆ ที่ศึกษาได้มานั้นคลาดเคลื่อนไปจากความเป็นจริง

2. อาจเป็นจริง โดยเนื่องมาจาก

ก. องค์ประกอบนั้นเป็นสาเหตุทำให้เกิดโรค

ข. โรคเป็นสาเหตุทำให้เกิดองค์ประกอบนั้น

ค. องค์ประกอบที่สามทำให้เกิดเหตุการณ์ทั้งสองอย่างที่สัมพันธ์กัน

Causation หมายถึงความสัมพันธ์ระหว่างเหตุการณ์สองอย่างหรือหลายอย่างโดยเหตุการณ์แรกเป็นสาเหตุและเหตุการณ์หลังเป็นผล

การที่พบเหตุการณ์สองอย่างมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ (Statistical association) ไม่ได้หมายความว่าเหตุการณ์สองอย่างนั้นมีความสัมพันธ์กันในเชิงสาเหตุและผล (Causal association) ความแตกต่างกันจำนวนน้อยซึ่งไม่มีความสำคัญในทางชีวภาพหรือในทางปฏิบัติ อาจพบมีนัยสำคัญในทางสถิติได้ในเมื่อเพิ่มขนาดของจำนวนตัวอย่างมากขึ้นเช่นในการทดสอบความมีนัยสำคัญทางสถิติ โดยใช้ Chi square test(X^2) ถ้าเพิ่มขนาดของจำนวนตัวอย่างเป็น k เท่าของจำนวนตัวอย่างเดิม

ค่าของ Chi square ใหม่ $=$ k x (ค่าของ Chi square เก่า)[6]

ตัวอย่าง สมมุติในการทดสอบ Chi square ครั้งแรกใช้ขนาดตัวอย่าง 100 ได้ค่า $X^2 = 3.00$ ไม่มีนัยทางสถิติที่ $\alpha = 0.05$ เมื่อเพิ่มขนาดตัวอย่างเป็น 200 จะได้ค่าของ X^2 ใหม่ $= 6.00$ ซึ่งมีนัยสำคัญทางสถิติที่ α 0.05

ความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent variables) และตัวแปรพึ่งพิง (Dependent variables) ในการเกิดโรคอาจจำแนกออกเป็น 4 ชนิด[21] ตามลักษณะของตัวแปรอิสระ (X) ที่จำเป็นและพอเพียงในการเกิดโรค (Y)

แบบที่	"X" จำเป็นในการเกิดโรค	"X" เพียงพอที่ทำให้เกิดโรค (Y)
1	+	+
2	+	−
3	−	+
4	−	−

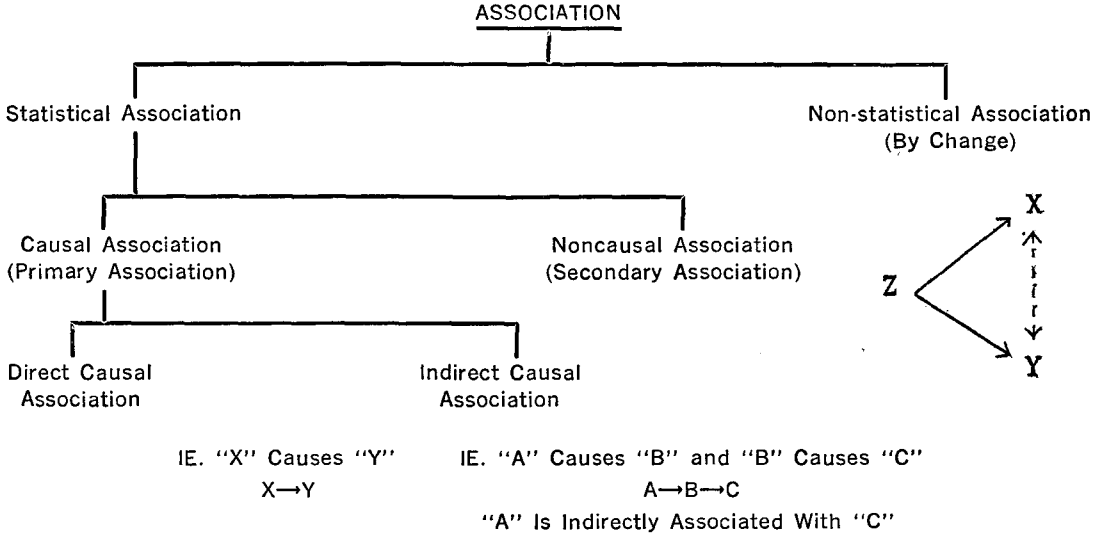
แบบที่ 1 "X" จำเป็นและเพียงพอที่ทำให้เกิด "Y" (X⟶Y) "X" และ "Y" มักพบเกิดร่วมกันเสมอ องค์ประกอบ "X" เท่านั้นที่จำเป็นที่ทำให้เกิดโรค "Y"

แบบที่ 2 "X" จำเป็นแต่ไม่เพียงพอที่ทำให้เกิดโรค "Y" (X+A⟶Y) "X" จะต้องพบเสมอเมื่อมีโรค "Y" เกิดขึ้น แต่เมื่อมีองค์ประกอบ "Y" อยู่จะไม่พบโรค "Y" เสมอไป

แบบที่ 3 "X" ไม่จำเป็นแต่เพียงพอที่จะทำให้เกิดโรค "Y" (X⟶Y, Z⟶Y) "X" อาจจะพบหรือไม่พบก็ได้เมื่อมีโรค "Y" เกิดขึ้น เพราะโรค "Y" อาจเกิดจากสาเหตุอื่น (Z) ได้

แบบที่ 4 "X" ไม่จำเป็นและไม่เพียงพอที่จะทำให้เกิดโรค "Y" (X+P⟶Y, Z+P⟶Y) "X" อาจจะพบหรือไม่พบก็ได้ เมื่อมีโรค "Y" เกิดขึ้น "X" เปรียบเสมือนองค์ประกอบที่ส่งเสริมต่อการเกิดโรค

รูปที่ 1 แสดงชนิดต่าง ๆ ของ Association



Criteria of Judgement

- Strength
- Consistency
- Specificity
- Time Relationship
- Coherence
- Biological Plausibility

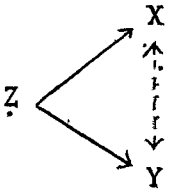
III. ชนิดของความสัมพันธ์

MacMahon[16] ได้แบ่งแยกชนิดของความสัมพันธ์ออกเป็นชนิดต่าง ๆ เพื่อสะดวกในการพิจารณาเกี่ยวกับความสัมพันธ์ในเชิงสถิติและความสัมพันธ์ในเชิงเหตุและผลต่อกัน การแบ่งแยกชนิดของความสัมพันธ์นี้นับเป็นประโยชน์ในการค้นหาสาเหตุของโรค โดยได้แบ่งชนิดของความสัมพันธ์ตามรูปที่ 1

1. ความสัมพันธ์ที่ปราศจากนัยสำคัญทางสถิติ เป็นความสัมพันธ์ที่เกิดขึ้นโดยบังเอิญ ตัวแปรเป็นอิสระต่อกัน

2. ความสัมพันธ์ที่มีนัยสำคัญทางสถิติความสัมพันธ์ทางสถิติเป็นความสัมพันธ์ที่เกิดขึ้นจริงโดยองค์ประกอบนั้นอาจเป็นสาเหตุหรือไม่ใช่สาเหตุของโรคก็ได้ สมมุติประชากรกลุ่มหนึ่งมีสัดส่วนขององค์ประกอบอันหนึ่ง (X) เป็น "a" และประชากรกลุ่มเดียวกันนี้มีสัดส่วนของคนเป็นโรค "Y" เท่ากับ "b" จะมีประชากรบางส่วนที่มีลักษณะทั้งสองอย่างดังกล่าว (XY) ในสัดส่วนเท่ากับ "ab" สัดส่วนนี้เมื่อนำไปเปรียบเทียบกับกลุ่มควบคุมซึ่งไม่มีองค์ประกอบ "X" แต่เป็นโรค "Y" ถ้าสูงหรือต่ำกว่าอย่างมีนัยสำคัญ จึงจะถือว่ามีความสัมพันธ์ทางสถิติ

2.1 ความสัมพันธ์ที่ไม่ใช่ทางสาเหตุ(Non-causal or secondary association) ความสัมพันธ์ทางสถิติที่เกิดจากองค์ประกอบที่สาม โดยองค์ประกอบที่สาม (Z) สัมพันธ์กับตัวแปรอิสระ (X) และตัวแปรพึ่งพิง (Y) ทำให้ "X" และ "Y" สัมพันธ์กันทางสถิติ



2.2 ความสัมพันธ์ทางสาเหตุ (Causal or primary association)

ความสัมพันธ์ทางสถิติระหว่างองค์ประกอบต่าง ๆ โดยองค์ประกอบอันหนึ่งหรือหลายอันเป็นสาเหตุของผลหรือของโรคหนึ่ง ทั้งนี้เป็นไปตามหลักเกณฑ์การพิจารณาองค์ประกอบที่เป็นสาเหตุของโรค เช่น มีขนาดกำลังความสัมพันธ์สูง ผลการศึกษาจากรายงานต่าง ๆ คล้ายคลึงกัน และความสัมพันธ์นั้นเป็นไปได้ในทางชีวภาพเป็นต้น

2.2.1 ความสัมพันธ์ทางตรงเกี่ยวกับสาเหตุ

องค์ประกอบที่ศึกษาเป็นสาเหตุที่ทำให้เกิดผลนั้นโดยตรง เช่น ยา Digitalis ทำให้การเต้นของหัวใจช้าลง

X ⟶ Y
Digitalis ชีพจรช้าลง

2.2.2 ความสัมพันธ์ทางอ้อมเกี่ยวกับสาเหตุ

องค์ประกอบที่ศึกษาเป็นสาเหตุที่ทำให้เกิดผลนั้นโดยทางอ้อม เช่น ยาปฏิชีวนะ ทำให้ไข้ในโรคติดเชื้อลดลง

X_1 ⟶ X_2 ⟶ X_3 ⟶ Y

X_1	X_2	X_3	Y
ยาปฏิชีวนะ	ทำลายเชื้อหรือหยุดการเจริญเติบโตของเชื้อ	ลดปฏิกริยาระหว่าง agent และ host	อุณหภูมิลดลง

IV. หลักเกณฑ์ที่ใช้พิจารณาถึงความสัมพันธ์ทางสาเหตุ*

หลักเกณฑ์ที่ใช้พิจารณาถึงความสัมพันธ์ทางสาเหตุ ประกอบด้วยหัวข้อสำคัญต่อไปนี้คือ

1. กำลังของความสัมพันธ์ (Strength of association)

2. ความเสมอต้นเสมอปลายของความสัมพันธ์ (Consistency of association)

3. ความจำเพาะของความสัมพันธ์ (Specificity of association)

4. การเรียงลำดับเวลาของความสัมพันธ์ (Time sequence of association)

5. ความสอดคล้องของความสัมพันธ์ (Coherence of association)

6. ความเป็นไปได้ทางชีวภาพ (Biological possibility)

1. กำลังความสัมพันธ์

กำลังความสัมพันธ์บอกถึงขนาดของความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรพึ่งพิงว่ามีมากน้อยแค่ไหน ถ้าขนาดกำลังความสัมพันธ์ยิ่งมาก โอกาสที่ความสัมพันธ์นั้นจะเป็นเหตุและผลต่อกันก็ยิ่งมีมากขึ้น แต่ถ้าความสัมพันธ์อยู่ในระดับต่ำ มักเป็นความสัมพันธ์ที่ไม่จริง โดยมากมักเนื่องมาจากองค์ประกอบที่สาม ขนาดกำลังของความสัมพันธ์นี้ไม่ได้หมายถึงความมีนัยสำคัญทางสถิติ ความสัมพันธ์อาจอยู่ในระดับต่ำ แต่มีนัยสำคัญทางสถิติได้ ทั้งนี้เนื่องจากขนาดของจำนวนตัวอย่างมากถ้าขนาดกำลังความสัมพันธ์ต่ำ โอกาสที่ความสัมพันธ์นั้นจะเป็นเหตุและผลต่อกันก็มีน้อย

ขนาดกำลังความสัมพันธ์วัดได้โดย

1.1 Relative risk (R.R.)

เป็นอัตราส่วนระหว่างอัตราการเกิดโรคในกลุ่มประชากรที่มีองค์ประกอบที่เสี่ยงต่อการเกิดโรคต่ออัตราการเกิดโรคในกลุ่มประชากรที่ไม่มีองค์ประกอบนั้น

$$\text{Relative risk} = \frac{Ie}{Io}$$

Ie = อัตราการเกิดโรคในกลุ่มประชากรที่มีองค์ประกอบ

Io = อัตราการเกิดโรคในกลุ่มประชากรที่ไม่มีองค์ประกอบนั้น

ถ้าขนาดของ relative risk สูง และมีนัยสำคัญทางสถิติ ความสัมพันธ์นั้นก็มีโอกาสที่จะเป็นความสัมพันธ์ทางเหตุและผลมาก ในการเปรียบเทียบกันควรจะได้ปรับความแตกต่างในเรื่องอายุและเพศ เป็นต้น

ตัวอย่าง Doll[7] พบว่าคนที่ทำงานเก็บกวาดเขม่าควันดำหรือน้ำมันแร่ในปล่องไฟ พบเป็นมะเร็งของลูกอัณฑะมากเป็น 200 เท่าของผู้ที่ไม่ได้รับเขม่าควันดำ จะเห็นได้ว่า relative risk เท่ากับ 200 นับว่าสูงมาก

สำหรับความสัมพันธ์ระหว่างการสูบบุหรี่และโรคมะเร็งในปอด ปรากฏว่าคนสูบบุหรี่มีโอกาสเป็นโรคมะเร็งของปอดมากเป็น 10 เท่าของคนที่

ไม่ได้สูบบุหรี่ relative risk เท่ากับ 10 นับว่าอยู่ในเกณฑ์สูง

1.2 อัตราส่วนตายมาตรฐาน (Standard mortality ratio "SMR")

อัตราส่วนตายมาตรฐาน หมายถึงอัตราส่วนของจำนวนคนตายจากสาเหตุใดสาเหตุหนึ่ง (หรือจากสาเหตุตายทั้งหมด) ต่อ จำนวนคนตายที่คาดว่าจะเกิดจากสาเหตุนั้น (หรือจากสาเหตุตายทั้งหมด) คิดเป็นอัตราร้อยละ

$$\text{Standard mortality ratio (SMR)} = \frac{\text{Observed deaths} \times 100}{\text{Expected deaths}}$$

การศึกษาเกี่ยวกับ SMR ในอาชีพเฉพาะบางอย่างกับการเกิดโรคหรือการตายจากโรคบางอย่างช่วยเป็นแนวทางในการศึกษาถึงความสัมพันธ์ระหว่างอันตรายจากอาชีพและโรคที่เกิดจากอาชีพนั้นได้ ถ้า SMR มีค่าสูงโอกาสของความสัมพันธ์ระหว่างองค์ประกอบและโรคก็มีมาก

1.3 สัมประสิทธิ์แห่งสหสัมพันธ์ (Correlation coefficient)[1,20]

แสดงถึงขนาดของความสัมพันธ์ระหว่างตัวแปร (Variables) สองตัวหรือหลายตัว โดยวัดระดับแห่งสหสัมพันธ์ระหว่างข้อมูล ใช้อักษรย่อว่า "r" มีค่าตั้งแต่ 0.0 ถึง $\pm$ 1.0 เมื่อสัมประสิทธิ์แห่งสหสัมพันธ์ระหว่างข้อมูลมีค่าเข้าใกล้ 1.0 หมายความว่าความสัมพันธ์มีลักษณะสมบูรณ์เป็นเส้นตรง ข้อมูลมีความสัมพันธ์ต่อกันในระดับสูง และเมื่อสัมประสิทธิ์แห่งสหสัมพันธ์ระหว่างข้อมูลเข้าใกล้ศูนย์หมายความว่าข้อมูลมีความสัมพันธ์ต่อกันในระดับต่ำ และเมื่อมีค่าเท่ากับศูนย์ แสดงว่าไม่มีความสัมพันธ์กันหรือเป็นอิสระต่อกัน

$$r = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{[n\Sigma X^2 - (\Sigma X)^2][n\Sigma Y^2 - (\Sigma Y)^2]}}$$

n = จำนวนตัวอย่างที่ทำการศึกษา

X = ตัวแปรที่หนึ่ง เช่น องค์ประกอบที่ทำให้เกิดโรค

Y = ตัวแปรที่สอง เช่น โรคที่กำลังศึกษา

1.4 สัมประสิทธิ์แห่ง regression (Regression coefficient)[1,20]

หมายถึงอัตราการเปลี่ยนแปลงเฉลี่ยที่คาดไว้ในตัวแปรพึ่งพิง (Y) ต่อหนึ่งหน่วย การเปลี่ยนแปลงของตัวแปรอิสระ (X)

เมื่อ Y = ตัวแปรพึ่งพิง X = ตัวแปรอิสระ

Regression coefficient

$$(b_{y.x}) = \frac{\text{Increment in Y}}{\text{Increment in X}}$$

$$b_{y.x} = \frac{\text{Covariance (X, Y)}}{\text{Variance X}} = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})/(n-1)}{(X-\bar{X})^2/(n-1)} = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{(X-\bar{X})^2}$$

$$b_{x.y} = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\Sigma(Y-\bar{Y})^2}$$

$$r^2 = (b_{y.x})(b_{x.y})$$

1.5 Gradient of risk

ถ้าอัตราการเกิดโรคเป็นสัดส่วนกับขนาดขององค์ประกอบที่ได้รับ ความสัมพันธ์นี้เรียกว่า "Dose-response relationship" (Gradient of risk) ถ้าขนาดขององค์ประกอบที่จะทำให้เกิดโรคมีน้อย โรคที่เกิดขึ้นก็มีจำนวนน้อย ถ้าองค์

ตารางที่ 1 อัตราส่วนตายมาตรฐานของโรคมะเร็งของปอดตามจำนวนการสูบบุหรี่ในปัจจุบัน จากรายงานการศึกษา 2 รายงาน

Doll และ Hill[7,8]			Hammond และ Horn[11,12]		
1. ตัวอย่างที่ศึกษา นายแพทย์ชาวอังกฤษ จำนวน 34,494 คน			ชาวอเมริกันชายจาก 9 รัฐ ในสหรัฐอเมริกา จำนวน 187,783 คน		
2. อายุ 35–75 ปี			50–69 ปี		
3. ระยะติดตามผล 120 เดือน			44 เดือน		
4. Gradient of risk					
จำนวนมวนที่สูบต่อวัน	Age-standardized death rate per 1,000	SMR	จำนวนมวนที่สูบต่อวัน	Age-standardized death rate per 1,000	SMR
ไม่เคยสูบ	0.07	1.0	ไม่เคยสูบ	0.12	1.0
			น้อยกว่า 10	0.95	7.9
1–14	0.47	6.7	10–19	1.08	9.0
15–24	0.86	12.3	20–39	2.29	19.1
≥25	1.66	23.7	≥40	2.64	22.0

ประกอบมีขนาดมาก จำนวนโรคที่เกิดขึ้นก็มีมาก ทั้งนี้ขนาดขององค์ประกอบจะต้องมากกว่า threshold level จึงจะเกิดสภาพการณ์ดังกล่าว เช่น ตัวอย่างในตารางที่ 1 จะเห็นว่าอัตราส่วนตายมาตรฐานเพิ่มขึ้นตามขนาดของการสูบบุหรี่ ทั้งนี้กำหนดให้อัตราส่วนตายของคนที่ไม่สูบบุหรี่มีค่าเท่ากับหนึ่ง

การที่มี Dose-response relationship ช่วยแสดงว่าความสัมพันธ์นั้นเป็นความสัมพันธ์ทางสาเหตุและผลต่อกันมากขึ้น

2. *ความเสมอต้นเสมอปลายของความสัมพันธ์*

การค้นพบความสัมพันธ์ระหว่างองค์ประกอบของโรคในลักษณะเดียวกันหลายครั้ง โดยผู้ทำการศึกษาหลายคนด้วยวิธีการศึกษาต่างๆกัน และทำการศึกษาในประชากรต่างท้องที่กัน นับเป็นเครื่องช่วยแสดงว่าความสัมพันธ์นั้นน่าจะเป็นสาเหตุและผลต่อกัน อย่างไรก็ตามควรต้องคำนึงไว้ด้วยเสมอว่าอคติแบบเดียวกันอาจจะเกิดขึ้นได้บ่อย ๆ ในการศึกษาต่างๆกัน

จากผลการทดลองเกี่ยวกับความสัมพันธ์

ระหว่างการสูบบุหรี่และโรคมะเร็งของปอด พบว่าในการศึกษาแบบ Case-control studies ประมาณ 29 รายงานและ Cohort studies ประมาณ 7 รายงาน ได้ข้อสรุปในลักษณะเดียวกันว่ามีความสัมพันธ์ระหว่างการสูบบุหรี่และโรคมะเร็งของปอด และมี relative risk ใกล้เคียงกัน[22] ผลการศึกษาต่างๆ ที่มีลักษณะคล้ายคลึงกันเช่นนี้ ทำให้ข้อสรุปเกี่ยวกับความสัมพันธ์ทางสาเหตุและผลมีน้ำหนักมากขึ้น

3. ความจำเพาะของความสัมพันธ์

ความจำเพาะของความสัมพันธ์ หมายถึงความแม่นยำในการพยากรณ์การเกิดขององค์ประกอบอันหนึ่งจากองค์ประกอบอีกอันหนึ่ง องค์ประกอบที่จะเป็นสาเหตุของโรคนั้นจะมีความสัมพันธ์จำเพาะกับโรคเพียงอย่างเดียวหรือน้อยอย่าง ถ้าองค์ประกอบอันหนึ่งจำเป็นและเพียงพอที่จะให้เกิดโรคอย่างหนึ่งความสัมพันธ์เป็นไปในรูปหนึ่งต่อหนึ่ง ทำให้เกิดโรคชนิดเดียว ความสัมพันธ์นี้ถือว่ามีความจำเพาะสมบูรณ์ (Complete specificity) คุณสมบัติข้อนี้สำคัญน้อยกว่าสองข้อแรก เนื่องจากว่าโรคอาจเกิดขึ้นจากสาเหตุเพียงอย่างเดียวหรือหลายอย่าง สาเหตุชนิดหนึ่งอาจทำให้เกิดโรคหลายชนิดก็ได้ ความจำเพาะของความสัมพันธ์ช่วยสนับสนุนความน่าจะเป็นไปได้ของความสัมพันธ์ในเชิงเหตุและผล แต่ความสัมพันธ์ทางสาเหตุอาจไม่มีคุณสมบัติในเรื่องความจำเพาะก็ได้

สำหรับการสูบบุหรี่ซึ่งถือเป็นสาเหตุของมะเร็งของโรคปอด มีข้อโต้แย้งเกี่ยวกับเรื่องความจำเพาะของความสัมพันธ์ เพราะการสูบบุหรี่มีความสัมพันธ์กับโรคอื่นๆ หลายโรค เช่น โรคหัวใจ โรคกระเพาะอาหารเป็นแผล โรคมะเร็งของกระเพาะปัสสาวะและโรคหลอดลมอักเสบเป็นต้น เหตุผลที่เป็นเช่นนี้ก็คือผู้สูบบุหรี่อาจมีลักษณะทางชีวภาพที่แตกต่างไปจากผู้ที่ไม่ได้สูบบุหรี่หลายอย่าง และควันบุหรี่ประกอบด้วยสารหลายอย่างที่เป็นอันตรายต่อสุขภาพของผู้สูบบุหรี่ เช่น benzpyrene สารที่ก่อให้เกิดมะเร็งอื่นๆ นิโคติน และคาร์บอนมอนนอกไซด์เป็นต้น

ถ้าพบมีความจำเพาะของความสัมพันธ์ในลักษณะหนึ่งต่อหนึ่ง ก็ช่วยสนับสนุนว่าความสัมพันธ์นั้นน่าจะเป็นความสัมพันธ์ทางสาเหตุ

4. การเรียงลำดับเวลาของความสัมพันธ์

ความสัมพันธ์ที่ถือเป็นเหตุและผลต่อกัน มักมีลักษณะการเรียงลำดับเวลาของเหตุการณ์ที่สัมพันธ์กัน เหตุการณ์ที่เป็นสาเหตุมักมาก่อนเหตุการณ์ที่เป็นผล เช่น องค์ประกอบ "X" ทำให้เกิดโรค "Y" "X" ควรจะพบก่อน "Y" การเรียงลำดับความสัมพันธ์ของตัวแปรนับว่ามีความสำคัญที่จะบอกว่าตัวแปรตัวหนึ่งเป็นสาเหตุของตัวแปรอีกตัวหนึ่งหรือไม่ ในโรคติดเชื้อและโรคที่มีระยะฟักตัวสั้นมักไม่ค่อยมีปัญหาเกี่ยวกับเรื่องการเรียงลำดับเวลาของความสัมพันธ์ ส่วนโรคปราศจากเชื้อหรือโรคที่มีระยะฟักตัวยาว บางครั้ง

อาจเกิดปัญหาได้ เพราะตัวแปรบางอย่างเปลี่ยนแปลงไปตามเวลา เช่น สภาวะ การสมรส อาชีพ เป็นต้น ทำให้เกิดปัญหาในการวินิจฉัยว่าอะไรมาก่อนหรือมาหลัง เช่น ความสัมพันธ์ระหว่างความผิดปกติทางจิตและการหย่า ความสัมพันธ์ระหว่างโรคหัวใจโคโรนารี่และบุคลิกภาพ เพราะโรคที่เกิดขึ้นอาจก่อให้เกิดการหย่าหรือเปลี่ยนแปลงในบุคลิกภาพได้

5. *ความสอดคล้องของความสัมพันธ์*

ความสัมพันธ์นั้นควรจะสอดคล้องกับความรู้ที่มีอยู่เกี่ยวกับสาเหตุของโรค เช่น คุณสมบัติทางเคมี พีสิคส์ และชีวภาพ และสอดคล้องกับความรู้ที่มีอยู่เกี่ยวกับโรค เช่น ทางคลีนิค พยาธิวิทยา วงจรธรรมชาติของโรค นอกจากนี้อาจทำการทดลองในสัตว์หรือในคนได้ผลสอดคล้องกัน

ตัวอย่าง ความสัมพันธ์ระหว่างการสูบบุหรี่และโรคมะเร็งของปอดสอดคล้องกับความรู้ทั่วไปที่มีอยู่ดังนี้ คือ

1. ควันบุหรี่มีสารที่ทำให้เกิดมะเร็ง ส่วนใหญ่เป็นสารประกอบไฮโดรคาร์บอน เช่น Benzpyrene สารเหล่านี้เกิดจากการเผาไหม้ที่อุณหภูมิสูงๆ ได้มีผู้ทดลองเอาขี้กล้อง (Tobacco tar) บ้ายที่หนังหนูตะเภาซ้ำที่เดียวอยู่เป็นเดือนๆ ปรากฏว่าบริเวณที่บ้ายเกิดเป็นมะเร็งเกิดขึ้น Wynder[24] ได้รวบรวมผลงานการวิจัยทางห้องปฏิบัติการและทางชีวภาพอื่นๆ ซึ่งสนับสนุนว่าควันบุหรี่ประกอบด้วยสารที่ทำให้เกิดมะเร็ง

2. การเปลี่ยนแปลงเยื่อบุหลอดลมในคนสูบบุหรี่ ตรวจพบโดย Auerbach และคณะ[2] ลักษณะการเปลี่ยนแปลงที่สำคัญคือ การเพิ่มจำนวนแถวของเซลล์ การทำลาย ciliated cells และการเพิ่มจำนวนเซลล์ที่มีลักษณะผิดปกติ พวกที่สูบบุหรี่น้อยกว่าวันละ 2 ซองจะพบเพียงร้อยละ 4.3 พวกที่สูบบุหรี่มากกว่าวันละ 2 ซองจะพบร้อยละ 11.4 และพวกที่เป็นมะเร็งของปอดจะพบร้อยละ 15.0 สำหรับพวกที่ไม่เคยสูบบุหรี่จะไม่พบการเปลี่ยนแปลงดังกล่าว

3. แนวโน้มของอัตราผู้ป่วยด้วยโรคมะเร็งของปอดมีลักษณะคล้ายคลึงกับแนวโน้มของอัตราการสูบบุหรี่ของประชากร โดยอัตราการเป็นมะเร็งของปอดจะเพิ่มขึ้นตามอัตราการสูบบุหรี่ของประชากรซึ่งปรากฏเพิ่มขึ้นก่อนหลายปี

ตารางที่ 2

รายการอาหาร	กลุ่มผู้รับประทานอาหารจำเพาะ			กลุ่มผู้ไม่ได้รับประทานอาหารจำเพาะ		
	จำนวนทั้งหมด	จำนวนผู้ป่วย	อัตราผู้ป่วยร้อยละ	จำนวนทั้งหมด	จำนวนผู้ป่วย	อัตราผู้ป่วยร้อยละ
1. แกงไก่	22	9	41	3	1	33
2. สลัดไก่	12	9	75	13	1	8
3. เนื้ออบ	15	6	40	10	4	40
4. กาแฟดำร้อน	10	9	90	15	1	7

6. *ความเป็นไปได้ทางชีวภาพ*

ความสัมพันธ์ที่ค้นพบจะต้องมีลักษณะที่เป็นไปได้ในทางชีวภาพและอธิบายกลไกต่าง ๆ ได้ เช่น การระบาดของโรคอาหารเป็นพิษจากเชื้อ samonella ในงานเลี้ยงแห่งหนึ่ง (ตารางที่ 2) พบว่ากาแฟดำร้อนไม่ใส่นมมีความแตกต่างของอัตราผู้ป่วยในกลุ่มคนที่รับประทานและไม่รับประทานอาหารดังกล่าว (ร้อยละ 83) สลัดไก่มีความแตกต่างรองลงมา (ร้อยละ 63) จากคุณสมบัติทางชีวภาพของเชื้อจะถูกทำลายที่อุณหภูมิ 60°ซ ภายในเวลา 15–20 นาที[23] จึงอาจสรุปได้ว่าสาเหตุการระบาดของโรคอาหารเป็นพิษดังกล่าวน่าจะเนื่องมาจากสลัดไก่มากกว่าอาหารชนิดอื่น

V. การทดสอบความสัมพันธ์[1,18,20]

ในการทดสอบความสัมพันธ์ระหว่างองค์ประกอบและโรค โดยมากมักทำการเปรียบเทียบอัตราการเกิดโรคในกลุ่มประชากรที่มีองค์ประกอบและไม่มีองค์ประกอบ หรือเปรียบเทียบอัตราการมีองค์ประกอบในกลุ่มบุคคลที่เป็นโรค และกลุ่มบุคคลที่ไม่เป็นโรค ทั้งนี้เพื่อเปรียบเทียบดูว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ ถ้าทดสอบพบมีความแตกต่างกันก็แสดงว่าโรคและองค์ประกอบนั้นมีความสัมพันธ์กัน

การทดสอบนัยสำคัญทางสถิติเกี่ยวกับความสัมพันธ์ที่ใช้กันบ่อยคือ

1. *Chi square test* สำหรับ *2×2 contingency tables*

		Disease: Positive	Disease: Negative	Total	Proportion
Facfor	Positive	A	B	$A + B = n_1$	$P_1 = A/n_1$
	Negative	C	D	$C + D = n_2$	$P_2 = C/n_2$
		$A + C$	$B + D$	$N = n_1 + n_2$	$P = (A + C)/N$

$$\text{Chi square } X^2 = \frac{\Sigma(O-E)^2}{E} \quad d,f = (r-1)(k-1)$$

d,f = Degree of freedom

r = Number of rows

k = Number of columns

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

การใช้ Chi square ในกรณีที่ d.f = 1 (Modified Siegel's[13]

1. เมื่อ N มีค่าระหว่าง 20 ถึง 40 และค่า Expected value เกิน 5 ทุกช่องใช้ x^2 corrected for continuity,

$$x^2 = \frac{\sum(|O-E| - 0.5)^2}{E}$$

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

2. เมื่อ N = 20 ถึง 40 ค่า Expected value ต่ำกว่า 5 ในบางช่องให้ใช้ Fisher test. หรือถ้า N น้อยกว่า 20 ใช้ Fisher test ทุกกรณี Fisher exact probability test (P)

$$p = \frac{(A+B)!\ (C+D)!\ (A+C)!\ (B+D)!}{N!A!B!C!D!}$$

P เป็นค่าของ probability

3. ในกรณีที่ N มากกว่า 40 จะใช้ Chi square with correction หรือไม่ก็ได้นอกจากค่า expected value ต่ำกว่า 5 ในบางช่อง จึงใช้ Chi square with correction

ใน Chi square test, ถ้าผลการทดสอบมีนัยสำคัญทางสถิติ แสดงว่ามีความสัมพันธ์ระหว่างองค์ประกอบและโรค แต่ไม่ได้บอกถึงระดับของความสัมพันธ์ว่ามีมากน้อยแค่ไหน

2. Test of two proportions by t-test

$$t = \frac{P_1 - P_2}{\sqrt{\frac{PQ}{n_1} + \frac{PQ}{n_2}}} \quad \text{with } n_1 + n_2 - 2 \text{ d, f}$$

ในการทดสอบนอกจากดูว่ามีนัยสำคัญทางสถิติแล้ว จำเป็นที่จะต้องพิจารณาดูว่า มีความสำคัญทางชีวภาพหรือไม่ โดยดูว่าความแตกต่างระหว่างสัดส่วนทั้งสองว่ามีความสำคัญในทางปฏิบัติหรือไม่ ทั้งนี้เนื่องจากถ้าจำนวนตัวอย่างมีมาก ถึงแม้ความแตกต่างของสัดส่วนมีน้อย การทดสอบนั้นก็อาจให้ผลแบบมีนัยสำคัญทางสถิติได้

3. Correlation and regression

ใน Correlation $t = \frac{r}{\sqrt{\frac{1-r_2}{n-2}}}$ with n−2 d,f,

r = Correlation coefficient

n = Sample size

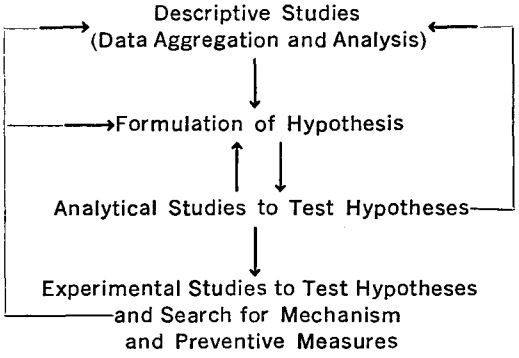
ใน Regression $t = \frac{b}{S_b}$ with n−2 d,f

b = Regression coefficient

S_b = Standard error of the estimate for the slope.

ถ้าผลทดสอบมีนัยสำคัญทางสถิติ แสดงว่าความเอียงของเส้นกราฟไม่เท่ากับศูนย์ตัวแปรที่กำลังศึกษามีความสัมพันธ์กัน

รูปที่ 2 แสดงถึงการหาความสัมพันธ์ทางสาเหตุ



VI. ขั้นตอนในการค้นหาความสัมพันธ์ทางสาเหตุ

ในการศึกษาทางระบาดวิทยาเพื่อค้นหาความสัมพันธ์ของเหตุการณ์หรือองค์ประกอบต่าง ๆ ที่

เป็นเหตุและผลต่อกัน ควรมีขั้นตอนต่าง ๆ ในการค้นหาความสัมพันธ์ การศึกษาทางระบาดวิทยาด้านพรรณนาเกี่ยวกับลักษณะทางบุคคล สถานที่และเวลา ช่วยในการชี้ให้เห็นถึงลักษณะของกลุ่มประชากรที่มีอัตราการเกิดโรคสูง อัตราการเกิดโรคที่แตกต่างกันในกลุ่มประชากรที่มีลักษณะต่าง ๆ กัน เป็นประโยชน์ในการสร้างสมมุติฐาน การยอมรับและไม่ยอมรับสมมุติฐานนั้นขึ้นอยู่กับผลการศึกษาทางระบาดวิทยาด้านวิเคราะห์และด้านการทดลอง ผลการศึกษาอาจช่วยก่อให้เกิดสมมุติฐานขึ้นใหม่ หรือช่วยดัดแปลงสมมุติฐานเดิมให้ถูกต้องและเหมาะสมมากขึ้น ขั้นตอนในการค้นหาความสัมพันธ์ทางสาเหตุประกอบด้วย

1. การรวบรวมข้อเท็จจริงและข้อมูลต่างๆ

ในการศึกษาทางระบาดวิทยาเกี่ยวกับโรคที่ยังไม่ทราบสาเหตุ หรือโรคที่พบใหม่ ข้อมูลขั้นพื้นฐานอาจได้มาโดย

1.1 การสังเกตเริ่มแรกจากทางคลีนิค พยาธิวิทยา และห้องปฏิบัติการ

ข้อมูลหรือลักษณะจำเพาะบางอย่างที่เป็นประโยชน์อาจเริ่มสังเกตได้จากผลตรวจทางห้องปฏิบัติการผลทางพยาธิวิทยา และทางคลินิค ตัวอย่างเช่น มีผู้ให้ข้อสังเกตว่าผู้ป่วยโรคจิตเป็นโรค Pellagra กันมาก แต่พยาบาลซึ่งสัมผัสกับผู้ป่วย Pellagra ในตึกกลับไม่เป็น จึงตั้งสมมุติฐานว่า สาเหตุของโรคน่าจะมาจากอาหารมากกว่าเชื้อโรคติดต่อ

1.2 การศึกษาระบาดวิทยาด้านพรรณนา

ศึกษาลักษณะการกระจายของโรคจากสถิติผู้ป่วย ประวัติผู้ป่วย หรือ จากการสำรวจโดยศึกษาเกี่ยวกับ

ก. ลักษณะการกระจายของโรคตามเวลา โดยดูว่ามีผู้ป่วยมากกว่าปกติในช่วงของเวลา วัน เดือน ปี หรือตามฤดูกาลหรือไม่การเปลี่ยนแปลงความถี่ของโรคโดยเฉพาะในระยะเวลาอันสั้น จะมีประโยชน์ในการช่วยตั้งสมมุติฐานเกี่ยวกับสาเหตุของโรคนั้น เช่น ในการระบาดของโรคติดเชื้อ ถ้าพบการกระจายของโรคเป็นแบบปกติ ผู้ป่วยทั้งหมดเกาะกลุ่มกันในช่วงเวลาของระยะฟักตัวของโรค การระบาดของโรคจะเป็นแบบ common source epidemic

ข. ลักษณะการกระจายของโรคตามสถานที่ โดยดูว่ามีความแตกต่างกันเกี่ยวกับการกระจายของโรคในท้องที่ต่าง ๆ ของประเทศหรือไม่ โดยเฉพาะความแตกต่างระหว่างในเมืองและในชนบท ท้องที่ใดที่มีอัตราป่วยของโรคสูง ท้องที่นั้นอาจมีองค์ประกอบหรือสาเหตุของโรคนั้นอยู่

ค. ลักษณะการกระจายของโรคตามบุคคล โดยดูการกระจายของโรคตามอายุ เพศ เชื้อชาติ อาชีพ ระดับการศึกษา เศรษฐานะ เป็นต้น เพื่อค้นหากลุ่มประชากรที่มีอัตราป่วยของโรคสูง และดูว่าองค์ประกอบใดที่เสี่ยงต่อการเกิดโรคมากขึ้น

2. การสร้างสมมุติฐาน

การสร้างสมมุติฐานทางระบาดวิทยา มีจุด

ประสงค์เพื่อชี้ถึงสาเหตุของโรค หรือชี้ถึงสาเหตุที่ทำให้ลักษณะการกระจายของโรคในชุมชนเป็นเช่นนั้น สมมุติฐานที่ดีควรประกอบด้วยสิ่งเหล่านี้

1. ประชากรที่อ้างถึงในสมมุติฐาน
2. องค์ประกอบที่ถูกพิจารณาว่าเป็นสาเหตุ
3. ผลหรือโรคที่คาดไว้
4. ความสัมพันธ์ระหว่างขนาดของสาเหตุและผล
5. ความสัมพันธ์ระหว่างเวลาและผล

สมมุติฐานเป็นเพียงข้อสรุปชั่วคราวที่ตั้งขึ้นจากข้อมูลที่มีอยู่ในขณะนั้น เพื่อเป็นแนวทางในการศึกษาและพิสูจน์ข้อสรุปต่อไป สมมุติฐานจะเปลี่ยนแปลงไปได้ตามผลของการพิสูจน์และตามข้อมูลที่ได้ศึกษาเพิ่มเติมมาใหม่ ในกรณีที่สมมุติฐานตั้งขึ้นได้รับการพิสูจน์แล้วว่าเป็นจริง สมมุติฐานนั้นก็จะเป็นประโยชน์ในการอธิบายปรากฏการณ์ต่าง ๆ ที่เกิดขึ้น ก่อให้เกิดความรู้ใหม่นำไปใช้ประโยชน์ในการศึกษาค้นคว้าต่อไป สมมุติฐานที่ระบุถึงองค์ประกอบที่เป็นสาเหตุและโรคที่เป็นผลตามมา เรียกว่า สมมุติฐานทางสาเหตุ (Causal hypothesis)

มีผู้แนะนำวิธีการต่าง ๆ ในการแสดงถึงความสัมพันธ์ที่เป็นสาเหตุและผลต่อกัน วิธีการเหล่านี้ได้นำมาดัดแปลงใช้ในการสร้างสมมุติฐานประกอบด้วยหลักต่าง ๆ ดังต่อไปนี้ คือ

1. วิธีของความแตกต่าง (Method of different)

ถ้าความถี่ของโรคแตกต่างกันมากในสองชุมชน หรือ สองภาพ องค์ประกอบบางอย่างสามารถพบได้ในสภาพหนึ่ง แต่ไม่พบในอีกสภาพหนึ่ง องค์ประกอบนั้นอาจเป็นสาเหตุของโรค ข้อยุ่งยากของวิธีนี้ก็คือโรคที่กำลังศึกษาอยู่อาจมีสาเหตุจากองค์ประกอบหลายอย่างร่วมกัน ทำให้เกิดปัญหาในการค้นหาสาเหตุของโรค

2. วิธีของความเหมือนกัน (Method of agreement)

ถ้ามีองค์ประกอบอันหนึ่งพบร่วมไปกับสภาพต่าง ๆ จำนวนมาก โดยที่สภาพเหล่านั้นมีโรคชนิดเดียวกันนั้นอยู่ด้วย องค์ประกอบอันนี้อาจจะเป็นสาเหตุของโรคได้

3. วิธีของการเปลี่ยนแปลงที่ไปด้วยกัน (Method of concomitant variation)

ในกรณีขนาดขององค์ประกอบอันหนึ่งเปลี่ยนแปลงไปตามจำนวนความถี่ของโรคที่เกิดขึ้นในชุมชนเมื่อขนาดขององค์ประกอบนั้นมีน้อย จำนวนความถี่ของโรคที่เกิดขึ้นก็พบน้อย เมื่อขนาดขององค์ประกอบนั้นมีมาก จำนวนความถี่ของโรคก็พบมาก ในลักษณะเช่นนี้องค์ประกอบนั้นก็อาจจะเป็นสาเหตุของโรค

ตัวอย่าง การหาความสัมพันธ์ระหว่างปริมาณของอาหารพวกไขมันกับอัตราการเกิดโรคของหลอดเลือดหัวใจ ในประเทศต่าง ๆ ประเทศที่มีปริมาณอาหารพวกไขมันสูง จะมีอัตราการเกิดโรคดังกล่าวสูง

4. วิธีของการอุปมาน (Method of analogy)

การกระจายของโรคชนิดหนึ่งที่ยังไม่ทราบสาเหตุ อาจมีลักษณะคล้ายคลึงกับการกระจายของ

โรคอีกชนิดหนึ่งซึ่งทราบสาเหตุแล้ว โรคทั้งสองอาจมีสาเหตุร่วมกัน

ตัวอย่าง การกระจายของโรคผู้ใหญ่ตามอายุและเพศในโรคมะเร็งของปอดและวัณโรคมีลักษณะคล้ายคลึงกัน มะเร็งของปอดมีความสัมพันธ์กับการสูบบุหรี่ ดังนั้นอาจสร้างสมมุติฐานได้ว่า progressive pulmonary tuberculosis มีส่วนสัมพันธ์กับการสูบบุหรี่

5. วิธีการของสิ่งที่เหลืออยู่ (Method of residues)

ในกรณีที่ได้กำจัดองค์ประกอบที่ทราบแน่ว่าเป็นสาเหตุของโรคแล้ว ปรากฏว่ายังมีโรคนั้นเหลืออยู่อีก แสดงว่าโรคนั้นอาจเกิดจากสาเหตุอย่างอื่นด้วย เช่น บุคคลที่ไม่เคยสูบบุหรี่เลยก็มีโอกาสเป็นโรคมะเร็งของปอดได้ โรคนี้อาจมีสาเหตุมาจากการสูบบุหรี่ประมาณ 90 เปอร์เซนต์ ที่เหลืออีก 10 เปอร์เซนต์ อาจมีสาเหตุมาจากอากาศเป็นพิษ เป็นต้น

3. *การทดสอบสมมุติฐานทางระบาดวิทยา*

สมมุติฐานในทางระบาดวิทยา โดยมากบ่งถึงความสัมพันธ์ระหว่างสิ่งสองสิ่ง คือ สาเหตุและผลในการที่พิจารณาว่าความสัมพันธ์นั้นเป็นสาเหตุและผลต่อกัน จะต้องทำการทดสอบดูว่า

1. มีความสัมพันธ์ทางสถิติระหว่างสิ่งสองสิ่งหรือไม่ โดยทดสอบความสัมพันธ์ทางสถิติ

2. เมื่อพบว่ามีความสัมพันธ์ทางสถิติแล้ว ต้องพิจารณาต่อไปว่าความสัมพันธ์เป็นสาเหตุและผลต่อกันหรือไม่

แบบการศึกษาที่ใช้ทดสอบสมมุติฐาน โดยมากใช้

1. *ระบาดวิทยาด้านวิเคราะห์* *Analytical epidemiology)*

ก. Case-control studies (Retrospective studies) การศึกษาไปข้างหลังเปรียบเทียบปริมาณขององค์ประกอบที่สงสัยเป็นสาเหตุของโรคในกลุ่มผู้ป่วยและกลุ่มควบคุม

ข. Cohort studies (Prospective studies) การศึกษาไปข้างหน้าเปรียบเทียบจำนวนโรคที่เกิดขึ้นในกลุ่มศึกษาซึ่งมีองค์ประกอบที่สงสัยเป็นสาเหตุของโรคและกลุ่มควบคุม ซึ่งไม่มีองค์ประกอบนั้นอยู่

2. *ระบาดวิทยาด้านทดลอง (Experimental epidemiology)*

การศึกษาระบาดวิทยาด้านการทดลองเพื่อทดสอบสมมุติฐาน ผู้ศึกษาจะเป็นผู้กำหนดว่ากลุ่มศึกษาและกลุ่มควบคุมจะได้รับองค์ประกอบที่สงสัยจะทำให้เกิดโรคในขนาดเท่าใดแล้วเปรียบเทียบผลในสองกลุ่ม การศึกษาแบบนี้อาจทำในคนหรือในสัตว์ก็ได้

สำหรับ Cohort และ Experimental studies นับว่าเป็นหัวใจสำคัญในการทดสอบความสัมพันธ์ทางสาเหตุ ส่วนการทดลองในสัตว์ช่วยอธิบายสิ่งที่ค้นพบจากการศึกษาในคน แต่ไม่สามารถพิสูจน์สมมุติฐานทางสาเหตุได้โดยตรง ทั้งนี้เนื่องจากโรคต่าง ๆ อาจมีความแตกต่างกันในชนิดของโฮสท์ (Species differences)

4. การค้นหากลไก (Searching for a mechanism)

การค้นหากลไกเพื่อช่วยอธิบายความสัมพันธ์ทางเหตุและผล อาจทำได้โดยการศึกษาจากความรู้ทางสรีรวิทยา ชีวเคมี จุลชีววิทยา พยาธิวิทยาและอื่น ๆ หรือทำการทดลองในสัตว์ทดลองเพื่อสนับสนุนให้เห็นความสัมพันธ์ทางเหตุและผลนั้นอีก การค้นพบกลไกช่วยอธิบายความสัมพันธ์ทางสาเหตุให้แจ่มชัดยิ่งขึ้น ทำให้รู้ถึงขั้นตอนต่าง ๆ จากเหตุไปถึงผล ทำให้สมมุติฐานนั้นเป็นที่ยอมรับมากขึ้น

สรุป

ในการศึกษาเกี่ยวกับความสัมพันธ์ขององค์ประกอบหรือเหตุการณ์ต่าง ๆ ในขั้นแรกจะต้องดูว่าองค์ประกอบเหล่านั้นมีความสัมพันธ์กันทางสถิติหรือไม่ การทดสอบความสัมพันธ์ในทางสถิติช่วยขจัดความสัมพันธ์ที่เกิดขึ้นโดยบังเอิญ ในขั้นต่อไปจะต้องดูว่าความสัมพันธ์ทางสถิตินั้นเป็นความสัมพันธ์ทางสาเหตุและผลต่อกันหรือไม่ โดยพิจารณาดูว่าองค์ประกอบนั้นเป็นสาเหตุที่ทำให้เกิดโรคหรือไม่ หลักเกณฑ์ที่ใช้ในการพิจารณาความสัมพันธ์ทางสาเหตุประกอบด้วย ขนาดของกำลัง ความสัมพันธ์สูง ผลการศึกษาจากรายงานต่าง ๆ คล้ายคลึงกัน เหตุการณ์ที่เป็นสาเหตุจะต้องพบก่อนเหตุการณ์ที่เป็นผล ความสัมพันธ์นั้นมีความจำเพาะและสอดคล้องกับความรู้เดิมที่มีอยู่ และประการสุดท้ายความสัมพันธ์นั้นจะต้องเป็นไปได้ในทางชีวภาพ

เอกสารอ้างอิง

1. Armitage P: Statistical methods in medical research. Oxford, Blackwell Scientific Publication, 1973 pp 99-146, 147-166
2. Auerbach O, Stout AP, Hammond EC, et al: Changes in bronchial epithelium in relation to cigarette smoking and in relation to lung cancer. N Engl J Med 265: 253-67, 61
3. Austin DF, Werner SB: Epidemiology for the health sciences. New York, Charles C. Thomas, 1974 pp 40-49
4. Benenson AS: Control of communicable diseases in man, New York, American Public Health Association, 1970 pp 213-216
5. Bross ID: Spurious effects from an extraneous variable. J Chronic Dis 19: 637-47, 66
6. Ching Chun Li: Human genetics principles and methods. New York, University of Pittsburgh, 1975 pp 79-101
7. Doll R, Hill AB: Mortality of doctors in relation to their smoking habits; preliminary report. Br Med J 1: 1451-5, 54
8. Doll R, Hill AB: Mortality in relation to smoking; ten years' observations of British doctors. Br Med J 1: 1399-410, 1460-7, 64
9. Fox JP, Hall CE, Elverback: Epidemiology; man and disease. New York, Macmillan Company, 1971 pp 306-314
10. Friedman GD: Primer of epidemiology. New York, McGraw-Hill Book Company, 1974 pp 150-168
11. Hammond EC, Horn D: Relationship between human smoking habits and death rates; follow-up study of 187, 766 men. JAMA 155: 1316-28, 54
12. Hammond EC, Horn D: Smoking and death rates-report on fourty-four months of follow-up of 187, 783 men. JAMA 166: 1159-72, 1294-08, 58
13. Hill AB: The environment and disease, association or causation? Proc Roy Soc Med 58: 295-300, 65
14. Lee DHK, Kotin P: Multiple factors in the causation of environmentally induced disease. New York, Academic Press, 1972 pp 2-12
15. Lilienfeld AM: Epidemiological methods in inferences in studies of noninfectious diseases. Public Health Rep 72: 51-60, 57
16. Macmahon B, Pugh TF: Epidemiology-principles and methods. Boston, Little, Brown and Company, 1970 pp 17-27
17. Mausner JS, Bahn AK: Epidemiology; an introductory text. Philadelphia, W.B. Saunders Company, 1974 pp 91-110

18. Siegel S: Nonparametric statistics for the behavioral sciences. The X^2 test for two independent samples. New York, McGraw-Hill Book Company, 1956 pp 104-110

19. Simon HA: Spurious correlation-A causal interpretation. J Am Statist Assoc 49: 467-79, 54

20. Sokal RR, Rohlf FJ: The principles and practice of statisties in biological research. New York, W.H, Freeman and Company, 1969 pp 404-490, 494-540

21. Susser M: Causal thinking in the health sciences-concepts and strategies of epidemiology. Oxford University Press, 1973 pp 41-47, 64-72, 140-162

22. United States Department of Health, Education and Welfare. Smoking and health-report of the advisory committee to the surgeon general. USPHS Pub no. 1103, 1964 p 83

23. Wilson GS, Miles AA: Topley and Wilson's principles of bacteriology and immumity Vol. I. New York, The Williams and Wilkins Company, 1964 p 870

24. Wynder EL: Laboratory contributions to tobacco-cancer problem. Br Med J 1: 317-22, 59