

สถิติที่ใช้ในการวิเคราะห์ปัจจัยการเกิดโรค

อรุณ จิรวัดเนกุล

ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

การศึกษาปัจจัยที่มีความสัมพันธ์กับการเกิดโรคเป็นขั้นตอนเบื้องต้นของการหาสาเหตุของการเกิดโรค การวิเคราะห์ความสัมพันธ์ของการเกิดโรคเป็นการดูว่าเมื่อข้อมูลของตัวแปรอิสระ (ปัจจัย) เปลี่ยนจะมีผลทำให้ข้อมูลของตัวแปรตาม (การเกิดโรค) เปลี่ยนแปลงหรือไม่

การเกิดโรคเส้นโลหิตในสมองแตกกับความดัน

	D	\bar{D}
ความดันโลหิตสูง	๖๐ (๕.๕%)	๕๕๐ (๕๔.๑%)
ความดันโลหิตปกติ	๖ (๐.๔%)	๑,๕๐๐ (๕๕.๖%)

จากข้อมูลแสดงการเกิดโรคเส้นโลหิตในสมองแตกกับระดับความดันโลหิต ถ้าความดันโลหิตสูงแล้วทำให้มีการเกิดโรคเส้นโลหิตในสมองแตกมากกว่าคนที่มีความดันโลหิตปกติ แสดงว่าปัจจัยความดันโลหิตสูงทำให้เกิดโรคเส้นโลหิตในสมองแตกเพิ่มขึ้นจะเรียกปัจจัยแบบนี้ว่าปัจจัยเสี่ยง (risk factor)

การเกิดโรคหลอดเลือดหัวใจกับการออกกำลังกาย

	D	\bar{D}
ออกกำลังกายสม่ำเสมอ	๗๘ (๑๓.๕%)	๕๐๐
ไม่ออกกำลังกายสม่ำเสมอ	๑๑๒ (๒๐.๗%)	๔๓๐

จากข้อมูลแสดงการเกิดโรคหลอดเลือดหัวใจพบว่าถ้าออกกำลังกายสม่ำเสมอทำให้เกิดโรคหลอดเลือดหัวใจน้อยกว่าคนที่ไม่ได้ออกกำลังกายสม่ำเสมอ แสดงว่าปัจจัยการออกกำลังกายทำให้เกิดโรคหลอดเลือดหัวใจลดลง จะเรียกปัจจัยแบบนี้ว่าปัจจัยป้องกัน (protective factor)

การขับซึรดผิวกุจรรจาจรจําแนกตามระดับการศึกษา

	D	\bar{D}
ปริญญาตรีและสูงกว่า	๗๘ (๑๕.๖%)	๓๒๐
ต่ำกว่าระดับปริญญาตรี	๑๐๘ (๒๐.๑%)	๔๓๐

จากข้อมูลแสดงการขับซึรดผิวกุจรรจาจรพบว่ามีอัตราการศึกษาในระดับใด มีอัตราการขับซึรดผิวกุจรรจาจรเท่า ๆ กัน แสดงว่าปัจจัยระดับการศึกษาไม่มีอิทธิพลหรือไม่มีความสัมพันธ์ต่อการขับซึรดผิวกุจรรจาจร ปัจจัยดังกล่าวจะไม่ใช่ปัจจัยเสี่ยง

สำหรับการศึกษาเพื่อหาปัจจัยเสี่ยงจะต้องออกแบบงานวิจัยเป็นแบบการศึกษาเชิงวิเคราะห์ (Analytical study) ซึ่งแบ่งย่อยได้ ๓ แบบคือ การศึกษาแบบ cross-sectional การศึกษาแบบ case-control และการศึกษาแบบ cohort แบบงานวิจัยทั้ง ๓ แบบใช้กันในสถานการณ์ที่แตกต่างกันตามสภาพปัญหาและข้อจำกัดของการศึกษารายละเอียดหาอ่านเพิ่มเติม

จากหนังสือวิธีวิทยาการวิจัยทางการแพทย์ และ สาธารณสุข หรือหลักการศึกษาด้านระบาดวิทยา

การคำนวณค่าสถิติเพื่อแสดงความสัมพันธ์ ระหว่างปัจจัยกับการเกิดโรค การเกิดโรคเป็นตัวแปรแบบ dichotomous (เป็นโรคหรือไม่เป็นโรค) ปัจจัยส่วนใหญ่จะเป็นตัวแปรแบบ dichotomous ด้วย เช่น สวมหมวกหรือไม่สวมหมวกนิรภัย สัมผัสหรือไม่สัมผัสอาหารติดเชื้อ หรือปริมาณโลหะหนักเกินหรือไม่เกิน ระดับมาตรฐาน

สถิติที่ใช้แสดงความสัมพันธ์ระหว่างการเกิดโรค กับปัจจัยที่เป็นตัวแปรแบบ dichotomous จะมีการคำนวณค่าสถิติได้ ๒ แบบ คือ

๑. ใช้สถิติ Chi-square แสดงว่าปัจจัยมีความสัมพันธ์กันหรือไม่ ค่า P-value ของค่าสถิติ Chi-square ที่คำนวณได้เมื่อเทียบกับค่าระดับนัยสำคัญที่กำหนดในการทดสอบสมมุติฐาน ใช้สรุปว่าปัจจัยที่มีความสัมพันธ์กับการเกิดโรคหรือไม่ ในกรณีพบว่ามีความสัมพันธ์กัน อย่างมีนัยสำคัญ จะสรุปได้แต่เพียงว่าปัจจัยนั้นมีความสัมพันธ์ โดยไม่สามารถระบุขนาด และทิศทางความสัมพันธ์ว่าเป็นปัจจัยเสี่ยงหรือปัจจัยป้องกัน

๒. ใช้สถิติ Relative Risk (RR) หรือ Odds Ratio (OR) แสดงความสัมพันธ์และขนาดความสัมพันธ์ จะใช้ RR ในกรณีการศึกษาเป็นแบบ cohort และใช้ OR กับแบบการศึกษาเป็นแบบ cross-sectional หรือ case-control

การพิจารณาสรุปปัจจัยเสี่ยงจากค่า RR หรือ OR พิจารณาจากช่วงเชื่อมั่นของ RR หรือ OR ถ้าช่วงเชื่อมั่นครอบคลุม ๑ (เช่น ๙๕% ช่วงเชื่อมั่นของ OR อยู่ระหว่าง ๐.๘๗ - ๑.๒๑) แสดงว่าปัจจัยไม่มีความสัมพันธ์กับการเกิดโรค ในกรณีช่วงเชื่อมั่นมีค่าน้อยกว่า ๑ แสดงว่าเป็นปัจจัยป้องกัน ถ้าช่วงเชื่อมั่นมีค่ามากกว่า ๑ แสดงว่าเป็นปัจจัยเสี่ยง เช่น ค่า ๙๕% RR อยู่ระหว่าง ๓.๗-๕.๘ จะสรุปว่าผู้ที่ได้รับปัจจัยดังกล่าว มีโอกาสเกิดโรคมมากกว่า ๓.๗ ถึง ๕.๘ เท่าของผู้ที่ไม่ได้รับปัจจัย

การคำนวณค่า OR นอกจากคำนวณจากสูตร ad/bc จากข้อมูลในตาราง ๒ x ๒ แล้วยังสามารถคำนวณค่า OR จากตัวแบบ logistic regression ค่า OR ที่ได้จากคำนวณทั้งสองวิธีมีค่าเท่ากัน

ในกรณีที่ตัวแปรปัจจัยเป็นตัวแปรแบบ polytomous ภายในตัวแปรมีมากกว่าสองกลุ่ม เช่น ความหนักเบาของโรคแบ่งเป็น น้อย ปานกลาง มาก หรือกรณีที่ตัวแปรต่อเนื่อง จะต้องคำนวณค่า OR จากตัวแบบ logistic regression ค่า OR ที่ได้จะเป็นความเสี่ยงที่เพิ่มขึ้นทุกหนึ่งหน่วยวัดของตัวแปรปัจจัย เช่น ความสัมพันธ์ของการเกิดโรคเส้นเลือดในสมองแตกกับระดับความดันโลหิต ค่า OR เท่ากับ ๑.๒ จะหมายความว่าทุก ๆ ๑ mmHg ที่ความดันสูงขึ้นไปจะมีโอกาสเสี่ยงเพิ่มขึ้น ๑.๒ เท่า ในทางปฏิบัติจะไม่นิยมนำตัวแปรปัจจัยที่เป็นตัวแปรต่อเนื่องมาคำนวณค่า OR โดยตรง เพราะความเสี่ยงไม่ได้เพิ่มขึ้นทุก ๆ หน่วยวัดที่เพิ่มขึ้น แต่ความเสี่ยงจะเกิดในบางช่วงของหน่วยวัด เช่น ความดันโลหิตซิสโตลิสที่เกิน ๑๔๐ mmHg ถือเป็นปัจจัยเสี่ยง ถ้าต่ำกว่าไม่เสี่ยง จึงนิยมแปลงข้อมูลให้เป็นกลุ่มก่อนจึงนำมาคำนวณค่า OR

ข้อควรระวังในการแปลความหมายของความสัมพันธ์ คือความบังเอิญที่ค่าสถิติที่คำนวณได้จากข้อมูลที่เก็บมา แสดงว่ามีความสัมพันธ์ แต่ไม่สามารถหาเหตุผลทางวิทยาศาสตร์มาอธิบายได้ว่ามีความสัมพันธ์กันอย่างไร เพื่อจะลดความผิดพลาดดังกล่าว ตัวแปรที่จะนำมาศึกษาปัจจัยเสี่ยงควรเป็นตัวแปรที่มีเหตุผลทางทฤษฎีสนับสนุนว่าน่าจะมีความสัมพันธ์ จึงนำมาศึกษา เพื่อเก็บข้อมูลมาคำนวณค่าสถิติเพื่อพิสูจน์ความสัมพันธ์ดังกล่าว

สรุป

๑. การใช้สถิติสำหรับการวิเคราะห์หาปัจจัยเสี่ยงของการเกิดโรคที่สามารถให้ข้อมูลครบถ้วนทั้งขนาด และทิศทางของความสัมพันธ์ คือ ค่า RR และ OR จึงนิยมใช้ค่าสถิตินี้ในการวิเคราะห์ข้อมูลการวิจัยเชิง-

วิเคราะห์

๒. การแปลผลปัจจัยเสี่ยงให้แปลผลจากค่า ๙๕% ช่วงเชื่อมั่นของ RR หรือ OR

๓. ตัวแปรในการศึกษาจะต้องมีเหตุผลทางวิทยาศาสตร์อธิบายได้ว่ามีความสัมพันธ์กันอย่างไร ก่อนที่จะเก็บข้อมูลมาวิเคราะห์สรุปผล

บรรณานุกรม

๑. อรุณ จีรวัดน์กุล. ชีวิตสถิติสำหรับงานวิจัยทางวิทยาศาสตร์ สุขภาพ. พิมพ์ครั้งที่ ๑. ขอนแก่น: คณะสาธารณสุขศาสตร์; ๒๕๔๗.
๒. Freidman LM, Purberg CD, De Mets DL. Fundamentals of clinical trials. Boston: John Wright; 1981.
๓. Schlesselman JJ. Case-control studies: design, conduct, analysis. New York: Oxford University Press; 1982.