

## Assignment #5

### Analysis in Case-Control Study

#### Reproductive and menstrual risk factors for pancreatic cancer: a population-based study of San Francisco Bay Area women.

Duell EJ, Holly EA. [Am J Epidemiol. 2005 Apr 15;161(8):741-7.]

1. จากบทความ จงอธิบายกลุ่มประชากรที่ทำการศึกษาของ Case-control Study นี้
  - ⇒ เป็น Population-based Case-control Study โดยกลุ่มประชากรที่ทำการศึกษาเป็นหญิงที่อาศัยอยู่ใน San Francisco Bay Area จำแนกเป็น Cases ที่ป่วยด้วย Pancreatic Adenocarcinoma จำนวน 241 ราย และ Controls ที่ไม่ป่วย จำนวน 818 ราย กลุ่มประชากรเหล่านี้มีอายุอยู่ในช่วง 21-85 ปี ส่วนใหญ่อยู่ใน 6 ย่านของของ San Francisco Bay Area (Alameda, Contra Costa, Marin, San Francisco, San Mateo, and Santa Clara) ระหว่างปี 1995 ถึง 1999
2. การศึกษานี้ มีการกำหนดกลุ่มที่เป็น Cases และกลุ่ม Controls อย่างไร ?
  - ⇒ Cases เป็นผู้ป่วยมะเร็งชนิดต่อมของตับอ่อน (Adenocarcinoma of Exocrine Pancreas) ที่ได้รับการวินิจฉัยระหว่างปี 1995 ถึง 1999 โดยเป็นผู้ที่อาศัยอยู่ใน 6 ย่านของ San Francisco Bay Area ซึ่งอยู่ภายใต้โครงการวิจัยด้านมะเร็งตับอ่อนที่กำลังดำเนินการอยู่ กลุ่ม Cases เหล่านี้มีอายุ 21-85 ปี ยังมีชีวิตอยู่ในช่วงที่ทำการศึกษาและสามารถให้สัมภาษณ์เป็นภาษาอังกฤษได้ ทั้งนี้ การวินิจฉัยโรคได้รับการยืนยันโดยแพทย์ประจำตัวของผู้เข้าร่วมโครงการ นอกจากนี้ ยังมี Cases ที่มีคุณสมบัติตามข้อกำหนด แต่อาศัยอยู่นอกพื้นที่ กรณีนี้จะอ้างอิงข้อมูลจากระเบียนประวัติผู้ป่วยของศูนย์การแพทย์ University of California, San Francisco
  - ⇒ Controls เป็นผู้ที่ไม่ป่วยเป็นมะเร็งตับอ่อน และมีคุณลักษณะเดียวกับ Cases คือ เป็นผู้ที่อาศัยอยู่ใน 6 ย่านของ San Francisco Bay Area มีอายุ 21-85 ปี ยังมีชีวิตอยู่ในช่วงที่ทำการศึกษา และสามารถให้สัมภาษณ์เป็นภาษาอังกฤษได้ ทั้งนี้ Controls ได้มาจากการสุ่มคัดเลือกทางโทรศัพท์ (Random Digit Dialing) โดยอ้างอิงจากรหัสพื้นที่โทรศัพท์ จากนั้นจับคู่กับ Cases ให้สอดคล้องกันตามเพศและช่วงอายุ
3. จงระบุปัจจัยด้านการรับสัมผัส (Exposures) และวิธีการประเมิน
  - ⇒ Exposures ที่สนใจศึกษาในงานวิจัยนี้คือ ปัจจัยด้านระบบสืบพันธุ์และการมีประจำเดือน ซึ่งเก็บข้อมูลจากแบบสอบถามที่ให้ตอบด้วยตนเอง (Self-report Questionnaire) และข้อมูลประวัติการเคยผ่าตัดมดลูก-รังไข่ ทั้งนี้ ตัวแปรที่ทำการศึกษาได้แก่
    - อายุที่เริ่มมีประจำเดือน
    - อายุที่หมดประจำเดือน
    - อายุเมื่อมีบุตรคนแรก
    - จำนวนการให้กำเนิดบุตร

- การกินยาเม็ดคุมกำเนิด
- ระยะเวลาการใช้ยาเม็ดคุมกำเนิด
- การใช้ยาทดแทนฮอร์โมนเอสโตรเจน (ERT)
- ระยะเวลาการใช้ ERT
- การผ่าตัดรังไข่ (Oophorectomy)
- อายุที่ผ่าตัดรังไข่
- การผ่าตัดมดลูก (Hysterectomy)
- อายุที่ผ่าตัดมดลูก

นอกจากนี้ ยังมีการพิจารณาตัวแปรร่วม หรือตัวแปรกวนที่อาจมีผลกระทบต่อ ได้แก่ อายุ เชื้อชาติ ระดับการศึกษา ดัชนีมวลกาย การสูบบุหรี่ การดื่มแอลกอฮอล์ โรคเบาหวานและโรคอื่นๆ

4. จงระบุประเด็นความสัมพันธ์ที่ทำการศึกษา และอธิบายวิธีการที่ใช้ในการประเมิน/วิเคราะห์

⇒ การศึกษานี้มุ่งประเด็นความสัมพันธ์ระหว่างการเกิดมะเร็งตับอ่อน กับปัจจัยเกี่ยวกับระบบสืบพันธุ์ (เช่น อายุที่เริ่มมีประจำเดือน อายุที่หมดประจำเดือน อายุเมื่อมีบุตรคนแรก การผ่าตัดรังไข่-มดลูก การใช้ยาคุมกำเนิด/ERT ฯลฯ) และปัจจัยอื่นๆ ที่อาจเป็นตัวแปรกวน (เช่น อายุ เชื้อชาติ การสูบบุหรี่ การดื่มแอลกอฮอล์ ค่าดัชนีมวลกาย โรคเบาหวาน ฯลฯ) โดยใช้ Odds ratios ในการวิเคราะห์ความสัมพันธ์

5. จากข้อมูลในตารางที่ 2 ของบทความ จงคำนวณ Odds Ratios ของกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ 13 และ  $\leq 12$  ปี โดยใช้กลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ  $\geq 14$  ปี เป็นกลุ่มอ้างอิง

⇒ กลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ 13 ปี:

$$\text{Odds Ratio} = \frac{61/211}{85/248} = 0.84$$

กลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ  $\leq 12$  ปี:

$$\text{Odds Ratio} = \frac{93/358}{85/248} = 0.76$$

⇒ ค่าโอกาสความเสี่ยงของการเกิดมะเร็งตับอ่อนในกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ 13 ปี และกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ  $\leq 12$  ปี คิดเป็น 0.84 เท่า และ 0.76 เท่า ตามลำดับ เมื่อเปรียบเทียบกับกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ  $\geq 14$  ปี

6. จากตารางที่ 2 จงแปลผล Odds Ratios สำหรับแต่ละกลุ่มซึ่งจำแนกตามอายุที่เริ่มมีประจำเดือน

⇒ โอกาสในการเป็นมะเร็งตับอ่อนในกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ 13 ปี อยู่ที่ 0.89 เท่าเมื่อเทียบกับกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ  $\geq 14$  ปี

⇒ โอกาสในการเป็นมะเร็งตับอ่อนในกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ  $\leq 12$  ปี อยู่ที่ 0.82 เท่าเมื่อเทียบกับกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ  $\geq 14$  ปี

7. จงเปรียบเทียบค่า Odds Ratios ที่คำนวณได้ในข้อ 5 กับค่าที่รายงานในตารางที่ 2 และให้เหตุผลว่า เพราะเหตุใดจึงมีความแตกต่าง ข้อมูลใดที่มีความถูกต้อง ความน่าเชื่อถือ หรือสะท้อนความสัมพันธ์ ระหว่างอายุที่เริ่มมีประจำเดือนกับความเสี่ยงในการเกิดมะเร็งตับอ่อนมากกว่า (อธิบายมาพอสังเขป)
  - ⇒ ค่าที่คำนวณได้ในข้อ 5 เป็น Crude Odd Ratios ที่ไม่ได้มีการพิจารณาตัวแปรร่วมอื่นๆ พบว่ามีค่า เท่ากับ 0.84 สำหรับกลุ่มที่เริ่มมีประจำเดือนตั้งแต่อายุ 13 ปี และ 0.76 สำหรับกลุ่มที่เริ่มมีประจำเดือน ตั้งแต่อายุ < 12 ปี ส่วนค่าที่รายงานในตารางที่ 2 เป็น Adjusted Odd Ratios ซึ่งคำนวณโดยมีการ พิจารณาปัจจัยด้านอายุ ระดับการศึกษา และการสูบบุหรี่ เท่ากับ 0.89 และ 0.82 ตามลำดับ โดย ค่าที่คำนวณได้ที่ต่ำกว่าเล็กน้อยนี้เป็นผลมาตัวแปรกวนที่ยังไม่ได้ถูกนำมาวิเคราะห์ ดังนั้น จึงมีความ ถูกต้อง ความน่าเชื่อถือที่น้อยกว่า ในขณะที่ค่าจากตาราง ซึ่งมีการวิเคราะห์ตัวแปรเหล่านี้ด้วย จะ สะท้อนความสัมพันธ์ระหว่างปัจจัยที่ศึกษาได้ใกล้เคียงความเป็นจริงมากกว่า
8. จากตารางที่ 3 ค่า Odds Ratio สำหรับกลุ่มที่มีการผ่าตัดรังไข่ในช่วงอายุ 44-51 ปี มีค่าเท่ากับ 1.1 จง แปลผลว่าหมายความว่าอย่างไร ?
  - ⇒ กลุ่มที่มีการผ่าตัดรังไข่ในช่วงอายุ 44-51 ปี มีโอกาสในการเป็นมะเร็งตับอ่อนเป็น 1.1 เท่า เมื่อเทียบ กับกลุ่มที่ไม่มีการผ่าตัดรังไข่ออก (กลุ่มอ้างอิง)
9. จากที่ผู้วิจัยระบุในบทความว่ามีการคัดเลือก Cases และ Controls ที่มีความสอดคล้องกันในด้านอายุและ เพศ โดยจำแนกเป็น 5 กลุ่ม นักศึกษาเข้าใจว่าหมายความว่าอย่างไร และมีวัตถุประสงค์เพื่ออะไร ?
  - ⇒ เนื่องจากผู้วิจัยตระหนักว่าอาจมีผลกระทบจากตัวแปรกวน (เช่น อายุ เชื้อชาติ ฯลฯ) จึงออกแบบ งานวิจัยให้มีการควบคุมตัวแปรกวนเหล่านี้ ซึ่งในกรณี Case-control Study แนวทางที่สำคัญคือ การคัดเลือก Controls ให้มีคุณลักษณะเฉพาะเหมือนกับ Cases และอยู่ภายใต้กรอบเวลาเดียวกัน เพื่อให้สามารถจับคู่เทียบเคียงกันได้ และช่วยปรับปรุงความแม่นยำ (ประสิทธิภาพ) ของการศึกษา
10. จงอธิบายว่าการหากกลุ่ม Controls โดยวิธี Random Digit Dialing ในการศึกษาวิจัยนี้ สามารถก่อให้เกิด ปัญหาหรือกระทบต่อผลการศึกษายังไงได้บ้าง
  - ⇒ อาจพิจารณาเป็นความลำเอียงในการคัดเลือก (Selection Bias) เนื่องจากไม่ใช่ทุกคนที่มีโทรศัพท์หรือมี หมายเลขที่อยู่ในสารบบของผู้วิจัย การใช้วิธีการนี้จึงจำกัดโอกาสของคนบางกลุ่ม (ผู้มีคุณสมบัติ เหมาะสม แต่ไม่มีหมายเลขโทรศัพท์ จะไม่ถูกคัดเลือกเข้ามา) นอกจากนี้ การใช้รหัสพื้นที่อาจไม่ น่าเชื่อถือ 100% เนื่องจากบางคนในพื้นที่เป้าหมายอาจใช้โทรศัพท์มือถือที่ไม่สามารถบ่งชี้รหัสพื้นที่ ดังนั้น การใช้หมายเลขโทรศัพท์แบบสุ่ม (Random Digit Dialing) อาจทำให้เกิดอคติในการเลือก Controls ที่มีสถานะทางเศรษฐกิจและสังคมที่ต่างกันอย่าง