

ตัวแปรผลลัพธ์แบบทวิภาค จำเป็นต้องใช้ Logistic Regression เสมอไป หรือไม่?

รองศาสตราจารย์ ดร. พงษ์เดช สารการ

สาขาวิชาวิทยาการระบาดและชีวสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

E-mail : spongde@kku.ac.th

Website : www.pongdechonline.com

1

สิ่งที่พบบ่อยในงานวิจัย



แนวทางปฏิบัติที่พบได้บ่อย



Outcome
มีโรค / ไม่มีโรค



เลือก
Logistic
regression
ทันที



รายงาน
OR หรือ aOR

2

คำถามคือ...



ค่าที่รายงานนั้น

ตอบคำถามวิจัยที่ต้องการจริงหรือไม่?



อยากรู้
ความชุก?



อยากรู้
ความเสี่ยง?



อยากรู้
ค่า odds?

3 ปัญหาที่ตามมา



?
OR = RR ?
OR = PR ?
ตีความ OR เป็น RR หรือ PR
โดยไม่ตั้งใจ



เลือกตัวชี้วัด
ไม่สอดคล้องกับ
Study design



สื่อสารผลลัพธ์
คลาดเคลื่อน
จากคำถามวิจัย



4 ประเด็นชวนแลกเปลี่ยนวันนี้



เริ่มจาก
คำถามวิจัย



เลือก
Study design



กำหนดตัวชี้วัด
ที่เหมาะสม



จึงค่อยเลือก
Statistical model



ดังนั้น Binary outcome จำเป็นต้องใช้ Logistic regression เสมอไปหรือไม่?

Binary outcome เหมือนกัน แต่ต้องการคำตอบเหมือนกันหรือไม่?

ตัวอย่างงานวิจัยสุขภาพ: ความอ้วนกับโรคความดันโลหิตสูง

Outcome: เป็นโรค = 1, ไม่เป็นโรค = 0



1 จุดตั้งต้นเหมือนกัน



กลุ่มอ้วน
(Exposed)



โรคความดันโลหิตสูง
(Disease)

Outcome เหมือนกัน:
มีโรค / ไม่มีโรค



เป็นโรค = 1



ไม่เป็นโรค = 0

แต่สิ่งที่อยากรู้
อาจไม่เหมือนกัน

2 คำถามวิจัยอาจต่างกัน

คำถามที่ 1: ความชุก (Prevalence)



ความชุกของโรคความดัน
โลหิตสูงในกลุ่มอ้วน
เป็นกี่เท่าของกลุ่มไม่อ้วน?



ตัวชี้วัดที่สนใจ: **PR**

คำถามที่ 2: ความเสี่ยง (Risk)



คนอ้วนมีความเสี่ยงเกิด
โรคความดันโลหิตสูงใน 5 ปี
เป็นกี่เท่าของกลุ่มไม่อ้วน?



ตัวชี้วัดที่สนใจ: **RR**

คำถามที่ 3: Odds



odds ของการมีโรคความดัน
โลหิตสูงในกลุ่มอ้วน
เป็นกี่เท่าของกลุ่มไม่อ้วน?



ตัวชี้วัดที่สนใจ: **POR**

3 สิ่งที่เปลี่ยนไปคืออะไร?

ความชุก
(Prevalence)



PR

ความเสี่ยงในอนาคต
(Risk)



RR

Prevalence odds
(Odds)



POR



Outcome เหมือนกัน

≠

คำถามวิจัยเหมือนกัน

ข้อมูลชุดเดียว คำนวณได้ต่างกัน

ตัวอย่างสมมติจากการศึกษาภาคตัดขวาง

กลุ่ม	เป็นโรค	ไม่เป็นโรค	รวม
กลุ่มอ้วน	40	60	100
กลุ่มไม่อ้วน	20	80	100



$$PR = \frac{(40/100)}{(20/100)} = 2.00$$



$$OR = \frac{(40/60)}{(20/80)} \approx 2.67$$



ผลต่างความชุก =
 $0.40 - 0.20 = 0.20$
เท่ากับ 20 จุดร้อยละ

ข้อมูลชุดเดียว คำนวณได้ต่างกัน

ตัวอย่างสมมติจากการศึกษาภาคตัดขวาง

PR



ความชุกเป็นกี่เท่า?

ตัวอย่าง:

กลุ่มอ้วนมีความชุกเป็น
2 เท่าของกลุ่มไม่อ้วน

RR



ความเสี่ยงเป็นกี่เท่า
ในช่วงเวลาที่กำหนด?

ตัวอย่างสมมติ cohort:

ความเสี่ยง 1 ปี 40%
เทียบกับ 20%
 $RR = 2.00$

ผลต่างความชุก / RD



ต่างกันกี่จุดร้อยละ?

ตัวอย่าง:

$40\% - 20\% =$
20 จุดร้อยละ

OR



odds เป็นกี่เท่า?

ตัวอย่างภาคตัดขวางเดิม:

odds ของการมีโรคเป็น
2.67 เท่า

คำถามวิจัยต่างกัน แล้ว Study design มีผลต่อการเลือกตัวชีวิตอย่างไร?

ตัวอย่างงานวิจัยสุขภาพ: ความอ้วนกับโรคความดันโลหิตสูง
โจทย์เดียวกัน แต่รูปแบบการศึกษาแตกต่างกัน อาจต้องใช้ตัวชีวิตต่างกัน



นอกจาก
คำถามวิจัยแล้ว
ต้องดู
Study design
ร่วมด้วยนะค่ะ

1

เริ่มจากการออกแบบการศึกษา



Study design มีผลต่อสิ่งที่คำนวณได้โดยตรง

Cross-sectional



- เก็บข้อมูล ณ เวลาใดเวลาหนึ่ง
- เหมาะกับคำถามเรื่องความชุก

Case-control



ผู้ป่วย
(Cases)

กลุ่มควบคุม
(Controls)

- เริ่มจากผลลัพธ์ แล้วจึงย้อนดูการสัมผัสปัจจัย
- เหมาะกับการเปรียบเทียบการสัมผัส

Cohort / RCT



กลุ่มสัมผัสปัจจัย
(Exposed)



ติดตาม
ไปในอนาคต



กลุ่มไม่สัมผัสปัจจัย
(Unexposed)

- เริ่มจากการสัมผัสปัจจัย แล้วติดตามผลในอนาคต
- เหมาะกับคำถามเรื่องความเสี่ยง

2 ตัวชี้วัดที่เหมาะสมในแต่ละแบบ

Cross-sectional



คำถามที่พบบ่อย:
ความชุกต่างกันเท่าไร?

PR

PD

POR

- PR = เปรียบเทียบความชุกระหว่างกลุ่ม
- PD = ความแตกต่างของความชุก
- POR = prevalence odds ratio

Case-control



คำถามที่พบบ่อย:
ผู้ป่วยมีโอกาสเคยสัมผัสปัจจัย
มากกว่ากลุ่มควบคุมหรือไม่?

OR

- OR = ค่าหลักของ case-control

⚠ โดยทั่วไปไม่สามารถคำนวณ
RR หรือ RD โดยตรง

Cohort / RCT



คำถามที่พบบ่อย:
ความเสี่ยงเกิดโรคในอนาคต
ต่างกันหรือไม่?

RR

RD

OR
(บางกรณี)

- RR = เปรียบเทียบความเสี่ยง
- RD = ความแตกต่างของความเสี่ยง
- OR อาจรายงานได้ในบางกรณี แต่ RR/RD มักสื่อความหมายตรงกว่า

3 ตัวอย่างคำถามเดียวกัน แต่ต่างกันตาม Study design

Cross-sectional



วันนี้ คนอ้วนมีความชุกของ
โรคความดันโลหิตสูงเป็นกี่เท่า
ของคนไม่อ้วน?

PR / PD / POR



Case-control



ในกลุ่มผู้ป่วยโรคความดัน
โลหิตสูง มีโอกาสเคยอ้วนมากกว่า
กลุ่มควบคุมหรือไม่?

OR



Cohort / RCT



คนอ้วนมีความเสี่ยงเกิดโรค
ความดันโลหิตสูงใน 5 ปี
สูงกว่าคนไม่อ้วนเท่าไร?

RR / RD





ตัวชี้วัด (Measure)	ใช้กับ Study design	คำถามวิจัยตัวอย่าง	Statistical model ที่พิจารณาได้ (สำหรับ binary outcome)
PR Prevalence Ratio อัตราส่วนความชุก	Cross-sectional (ภาคตัดขวาง)	ความชุกของโรคในกลุ่มอ้วน เป็นกี่เท่าของกลุ่มไม่อ้วน?	- Log-binomial regression - Modified Poisson regression (with robust SE) - Logistic regression (marginal standardization)
PD Prevalence Difference ผลต่างความชุก	Cross-sectional (ภาคตัดขวาง)	ความชุกของโรคในกลุ่มอ้วน สูงกว่ากลุ่มไม่อ้วน กี่ร้อยละ (กี่เปอร์เซ็นต์พอยต์)?	- Binomial regression (identity link) - Linear probability model (with robust SE) - Logistic regression (marginal standardization)
POR Prevalence Odds Ratio อัตราส่วน odds ของความชุก	Cross-sectional (ภาคตัดขวาง)	Odds ของการมีโรคในกลุ่มอ้วน เป็นกี่เท่าของกลุ่มไม่อ้วน?	Logistic regression
OR Odds Ratio อัตราส่วน odds	Case-control (ย้อนหลัง)	Odds ของการมีปัจจัยเสี่ยงในกลุ่มผู้ป่วยเป็นกี่เท่าของกลุ่มควบคุม?	Logistic regression (แบบมีหรือไม่มีการปรับตัวแปร)
RR Risk Ratio อัตราส่วนความเสี่ยง	Cohort / RCT (ไปข้างหน้า)	ความเสี่ยงเกิดโรคในกลุ่มอ้วน เป็นกี่เท่าของกลุ่มไม่อ้วน ในช่วงเวลา 5 ปี?	- Log-binomial regression - Modified Poisson regression (with robust SE) - Logistic regression (marginal standardization)
RD Risk Difference ผลต่างความเสี่ยง	Cohort / RCT (ไปข้างหน้า)	ความเสี่ยงเกิดโรคในกลุ่มอ้วน สูงกว่ากลุ่มไม่อ้วน กี่ร้อยละ (กี่เปอร์เซ็นต์พอยต์) ในช่วงเวลา 5 ปี?	- Binomial regression (identity link) - Linear probability model (with robust SE) - Logistic regression (marginal standardization)

บทสรุป

- **Logistic regression** ไม่ถือเป็นวิธีที่ผิดสำหรับตัวแปรผลลัพธ์ทวิภาค (Binary outcome)
- แต่การเลือกใช้เพียงเพราะ Outcome มีสองค่า **โดยไม่พิจารณาคำถามวิจัยและตัวชี้วัดที่ต้องการ** อาจนำไปสู่การรายงานและตีความผลที่ไม่ตรงกับวัตถุประสงค์ของงานวิจัยได้

Thank you

E-mail : spongd@kku.ac.th

<https://pongdechonline.com>

