

การประเมินและการจัดการตัวแปรกวน

สิริมา	มงคลลัมฤทธิ์	พย.บ, วท.ม ระบาดวิทยา*
ชยันตร์ธธ	ปทุมานนท์	พ.บ., วท.ม อายุรศาสตร์เขตธอนคลินิก, ปส.ด ระบาดวิทยาคลินิก**
เพชร	รอดอารีย์	พ.บ., ว.ว. อายุรศาสตร์, วท.ม ระบาดวิทยา***

บทคัดย่อ

เมื่อนักวิจัยพยายามอธิบายขนาดความสัมพันธ์ของการสัมผัสปัจจัยกับโอกาสเกิดเหตุการณ์ ซึ่งในความเป็นจริงจะมีตัวแปรตัวที่สามที่เรียกว่า “ตัวแปรกวน” เข้ามาทำให้ผลการศึกษาก่อเกิดความคลุมเครือหรือทำให้ผลการศึกษาคิดไปจากความเป็นจริง เนื่องจากวิธีการทางสถิติไม่สามารถช่วยในการควบคุมอิทธิพลของตัวแปรกวนที่ไม่ได้เก็บมาหรือตัวแปรที่ไม่ทราบว่าเป็นตัวแปรกวนได้ ดังนั้นจึงมีความจำเป็นที่นักวิจัยต้องค้นหาและควบคุมอิทธิพลของตัวแปรกวน สิ่งสำคัญอันดับแรกที่ต้องทำคือการทบทวนวรรณกรรมอย่างละเอียดก่อนเริ่มดำเนินการวิจัยเพื่อให้สามารถวางแผนในการเก็บข้อมูลได้ครบถ้วน ทั้งนี้รวมถึงการเก็บข้อมูลที่เป็นตัวแปรกวนด้วย เมื่อทราบว่าตัวแปรใดเป็นตัวแปรกวนแล้ว การจัดการสามารถทำได้หลายวิธีตั้งแต่การออกแบบการศึกษา เช่น randomization, restriction และ matching หรือจัดการในขั้นตอนวิเคราะห์ข้อมูลโดยวิธีการ stratification หรือทำ multivariable analysis

Abstract

How to Detect and Handle Confounding Factors

Sirima	Mongkolsomlit	BSc, MSc Epidemiology*
Jayanton	Patumanond	MD, DTM&H, MSc Clin Trop Med, DSc Clinical Epidemiology**
Petch	Rawdaree	MD, MSc Epidemiology, DLSHTM***

* Faculty of Public Health, Thammasart University.

** Faculty of Medicine, Chiangmai University.

*** Department of Medicine, BMA Medical College and Vajira Hospital.

When a researcher investigates the association between an exposure of factors and an occurrence of outcome, there are other influencing factors to the test result so called “confounding factors”. Confounding factors may cause indistinct or inaccurate results of the study. Since statistical analysis cannot clear out the effect of confounders which

* คณะสาธารณสุขศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต

** ภาควิชาเวชศาสตร์ชุมชน คณะแพทยศาสตร์ มหาวิทยาลัยเชียงใหม่

*** ภาควิชาอายุรศาสตร์ วิทยาลัยแพทยศาสตร์กรุงเทพมหานครและวชิรพยาบาล

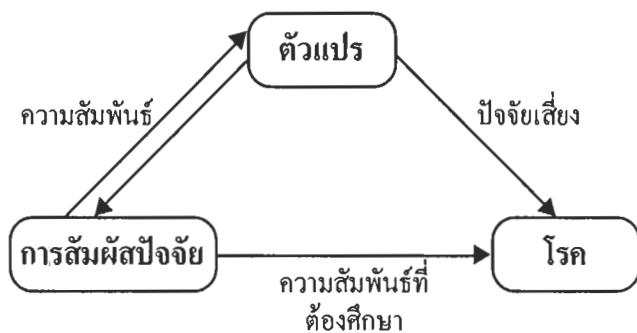
have not been collected or being unidentified, it is crucial for the researcher to identify and control the impact of these confounding factors. In order to collect all essential data including any potential confounding factors, a thorough literature review prior to conducting a research is the first and important step. If confounders are identified, there are many approaches to deal with confounders: by randomization, restriction, or matching in the research design or methodology process; or by stratification or multivariable analysis in the statistical analysis process.

Keywords: confounding factors, randomization, restriction, matching, stratification

บทนำ

การศึกษาทางระบาดวิทยาส่วนใหญ่เป็นการศึกษาความสัมพันธ์ของปัจจัยเสี่ยงกับการเกิดโรค สิ่งที่นักวิจัยต้องการคือผลที่แท้จริงเมื่อสัมผัสปัจจัยนั้น แต่สิ่งที่ไม่สามารถควบคุมได้ในการศึกษาทางระบาดวิทยาเชิงสังเกต คือ ตัวแปรกวน¹⁻³ และในความเป็นจริงไม่มีใครสามารถบอกได้ว่าตัวแปรใดควรถูกนำเข้ามาวิเคราะห์เพื่อควบคุมอิทธิพลไม่ให้กระทบต่อผลที่แท้จริงของปัจจัยที่สนใจกับการเกิดโรค นอกจากมีการศึกษาก่อนหน้านี้ที่ค้นพบว่าตัวแปรนั้นเป็นตัวแปรกวนของความสัมพันธ์ระหว่างปัจจัยกับเหตุการณ์ที่สนใจ จึงมีความจำเป็นที่นักวิจัยต้องทบทวนวรรณกรรมอย่างละเอียดก่อนเริ่มการวิจัย ในเอกสารฉบับนี้ได้อธิบายแนวคิดเกี่ยวกับตัวแปรกวน การพิจารณาตัวแปรกวน และวิธีการควบคุม ซึ่งจะกล่าวถึงต่อไป

ตัวแปรกวนที่มีอิทธิพลต่อผลการศึกษา (Potential confounder)
ปรากฏการณ์การเกิดขึ้นของตัวแปรกวน (รูปที่ 1) เป็นสิ่งที่นักวิจัยจำเป็นต้องพิจารณาเมื่อต้องการวิเคราะห์หาความสัมพันธ์ระหว่างการสัมผัสปัจจัยเสี่ยง (determinant) กับโอกาสการเกิดโรคหรือเหตุการณ์ที่สนใจ (outcome)

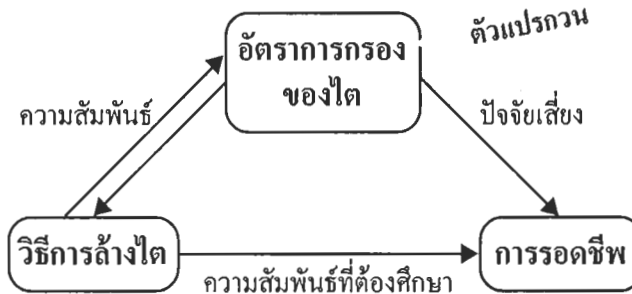


รูปที่ 1 ปรากฏการณ์การเกิดขึ้นของตัวแปรกวน

การประเมินว่าตัวแปรใดเป็นตัวแปรกวนสามารถพิจารณาจาก 3 ประเด็นต่อไปนี้ร่วมกัน^{4,5} คือ

1. ตัวแปรนั้นต้องมีความสัมพันธ์กับการเกิดโรค ซึ่งหมายความว่าตัวแปรนั้นควรจะเป็นปัจจัยเสี่ยงของโรคนั้น
2. ตัวแปรนั้นต้องมีความสัมพันธ์กับการสัมผัสปัจจัย (determinant, exposure) หมายความว่าจะต้องพบตัวแปรที่เป็นตัวแปรกวน กระจายไม่เท่ากันในกลุ่มที่สัมผัสปัจจัย (determinant, exposure) และ ไม่ได้สัมผัสปัจจัย (non determinant, non exposure)
3. ตัวแปรนั้นต้องไม่มีผลต่อการสัมผัสปัจจัย และไม่เป็นสาเหตุหนึ่งที่เป็นเส้นทาง (pathway) ของการเกิดโรคหรือเหตุการณ์ที่สนใจ

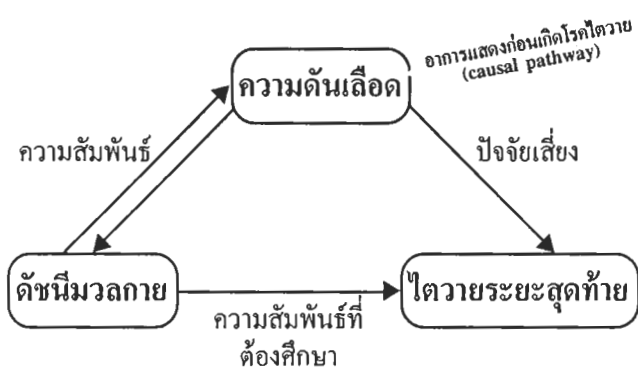
เพื่อให้มองเห็นภาพที่ชัดเจนขึ้น ขออธิบายโดยยกตัวอย่างการศึกษาความสัมพันธ์ระหว่างการเลือกวิธีการรักษาในการล้างไตกับการรอดชีพภายใน 2 ปีของผู้ป่วยสูงอายุที่เป็นโรคไตวาระยะสุดท้าย⁶ ซึ่งเป็นรูปแบบการศึกษาแบบติดตามไปข้างหน้า (cohort study) ในกลุ่มตัวอย่าง 3,512 ราย นักวิจัยได้ควบคุมอิทธิพลของตัวแปรอัตราการกรองของไต (glomerular filtration rate, GFR) และปัจจัยอื่น ๆ ดังรูปที่ 2



รูปที่ 2 ความสัมพันธ์ระหว่างวิธีการล้างไตกับการรอดชีพของผู้ป่วย

การศึกษานี้ได้รายงานผลหลังจากที่มีการควบคุมอิทธิพลของตัวแปรอัตราการกรองของไต (GFR) และปัจจัยอื่น ๆ แล้วพบว่า การล้างไตด้วยวิธีการฟอกเลือดโดยไม่ได้อาหารแผนล่วงหน้า (unplanned hemodialysis) มีความเสี่ยงต่อการเสียชีวิตเพิ่มขึ้นร้อยละ 50 เมื่อเทียบกับการฟอกเลือดโดยมีการวางแผนล่วงหน้า และการล้างไตทางหน้าท้อง (peritoneal dialysis) มีความเสี่ยงต่อการเสียชีวิตเพิ่มขึ้นร้อยละ 30 เมื่อเทียบกับการฟอกเลือดโดยมีการวางแผนล่วงหน้า สิ่งที่น่าสนใจคือนักวิจัยมีเหตุผลอะไรในการเลือกควบคุมอิทธิพลของตัวแปร GFR ในระหว่างการวิเคราะห์ข้อมูล เหตุผลข้อแรกคือ จากการทบทวนวรรณกรรมนักวิจัยทราบว่า GFR เป็นปัจจัยเสี่ยงของการเสียชีวิตในการล้างไต เหตุผลที่สองนักวิจัยทราบว่าผู้ป่วยที่เริ่มรักษาด้วยการล้างไตทางหน้าท้องมีแนวโน้มของ GFR สูงกว่าผู้ป่วยที่เริ่มรักษาด้วยวิธีการฟอกเลือด ค่า GFR ที่แตกต่างกันในเวลาเริ่มต้นของการรักษาอาจจะมีอิทธิพลต่อการเลือกวิธีการรักษาจากแพทย์ ดังนั้น GFR จึงถูกพิจารณาว่าเป็นตัวแปรกวนในการศึกษานี้

อีกหนึ่งตัวอย่างเรื่องการศึกษาความสัมพันธ์ระหว่างดัชนีมวลกาย (body mass index, BMI) กับการเกิดโรคไตวายเรื้อรังระยะสุดท้าย (end stage renal disease, ESRD)⁷ โดยใช้ขนาดตัวอย่างมากกว่า 320,000 ราย ผลการศึกษาหลังจากควบคุมตัวแปรกวนต่าง ๆ เช่น อายุ เพศ และเชื้อชาติ โดยที่ไม่ควบคุมตัวแปรความดันเลือด พบว่า BMI ที่เพิ่มขึ้นมีความสัมพันธ์กับการเพิ่มความเสี่ยงของการเกิดโรคไตวายเรื้อรังระยะสุดท้าย ทำให้ในการศึกษานักวิจัยจึงพิจารณาว่า “ความดันเลือด” ไม่ใช่ตัวแปรกวน เหตุผลคือความดันเลือดเป็นส่วนหนึ่งที่ปรากฏอยู่ในเส้นทางของการเกิดโรคนี้ (pathway) ซึ่งเป็นธรรมชาติของการเกิดโรคไตวายเรื้อรัง (natural of outcome) จึงไม่เข้าข่ายการเป็นตัวแปรกวน ดังนั้นจึงไม่มีความจำเป็นต้องควบคุมอิทธิพลของความดันเลือดในการศึกษานี้ (รูปที่ 3)



รูปที่ 3 ความสัมพันธ์ระหว่างดัชนีมวลกายกับการเกิดโรคไตวายเรื้อรังระยะสุดท้าย

การทบทวนวรรณกรรมโดยละเอียดเกี่ยวกับกลไกการเกิดโรค และพยาธิสภาพของโรคก่อนทำการวิจัยจะทำให้นักวิจัยทราบว่าตัวแปรตัวใดที่ควรพิจารณาเป็นตัวแปรกวน ซึ่งช่วยในการวางแผนเก็บข้อมูลตัวแปรกวนนั้นเข้ามาควบคุมอิทธิพลได้นั้นหมายความว่าเราสามารถควบคุมอิทธิพลของตัวแปรกวนที่เราทราบและเก็บเข้ามาในการศึกษาได้ แต่ไม่สามารถควบคุมอิทธิพลของตัวแปรกวนที่เราไม่ทราบรวมถึงตัวแปรกวนที่ไม่ได้เก็บมา

การควบคุมอิทธิพลของตัวแปรกวน

ด้วยเหตุที่ตัวแปรกวนทำให้ผลการวิจัยที่แท้จริงถูกบิดเบือนไป จึงจำเป็นต้องป้องกันหรือกำจัดถ้าเป็นไปได้^{4,8,9} นักวิจัยสามารถวางแผนควบคุมอิทธิพลของตัวแปรกวนได้ทั้งในช่วงก่อนและหลังเก็บข้อมูล วิธีการควบคุมอิทธิพลของตัวแปรกวนก่อนเก็บข้อมูลทำได้โดยเลือกใช้ระเบียบวิธีวิจัย (methodology) ที่เหมาะสม เช่น การทำ randomization, restriction และ matching สำหรับการควบคุมตัวแปรกวนหลังเก็บข้อมูลสามารถทำได้ในระหว่างการวิเคราะห์ข้อมูล (statistical analysis) เช่น การทำ stratification และ multivariable analysis

การสุ่ม (randomization)

การสุ่มสามารถช่วยลดอิทธิพลของตัวแปรกวนต่อผลของการรักษาหรือสิ่งทดลองได้ เนื่องจากกระบวนการสุ่มเป็นการเลือกกลุ่มตัวอย่างและกลุ่มทดลองโดยอาศัยความน่าจะเป็น (probability) ซึ่งช่วยป้องกันการเกิดอคติจากการเลือกกลุ่มศึกษาของนักวิจัย (selection bias) บางครั้งเรียกอคติชนิดนี้ว่า confounding by indication ทั้งนี้การสุ่มในผู้ป่วยกลุ่มใหญ่ ๆ ทำให้มีการกระจายของตัวแปรกวนทั้งที่ทราบและไม่ทราบ มีลักษณะคล้าย ๆ กันทั้งในกลุ่มควบคุมและกลุ่มทดลอง แต่ไม่ได้หมายความว่าลักษณะพื้นฐานของกลุ่มควบคุมและกลุ่มทดลองต้องเหมือนกันเสมอไป การสุ่มไม่ได้เป็นการทำให้เกิดความสมดุลในลักษณะที่สมบูรณ์แบบหรือกลุ่มควบคุมเหมือนกับกลุ่มทดลองทุกอย่างโดยสิ้นเชิง เพียงแต่เป็นกระบวนการที่รับรองว่าความแตกต่างที่เกิดขึ้นระหว่างทั้งสองกลุ่มเกิดขึ้นเนื่องจากความบังเอิญ (chance) ไม่ใช่เกิดขึ้นเนื่องจากการเลือกของนักวิจัย (choice) ดังนั้นความแตกต่างของตัวแปรกวนทั้งสองกลุ่มอาจจะยังคงอยู่หลังการสุ่มก็ได้ แต่หลงเหลือน้อยลงเท่าที่จะเป็นไปได้ แต่ถ้าเป็นการสุ่มในตัวอย่างกลุ่มเล็ก ๆ อาจจะไม่มีประโยชน์ในการช่วยควบคุมตัวแปรกวนเพราะไม่ช่วยให้การกระจายของตัวแปรในกลุ่มควบคุมและกลุ่มทดลองเท่ากันได้

ตัวอย่างการทดลองแบบ randomized control trial (RCT) ขนาดใหญ่ของโครงการ HEMO study group¹⁰ เพื่อ

ตารางที่ 1 ลักษณะข้อมูลพื้นฐานจากการสุ่มในกลุ่มศึกษาขนาดใหญ่ ในโครงการ HEMO study group¹⁰

ปัจจัย	จำนวนผู้ป่วย (n = 1846)	วิธีการล้างไต			
		Standard dose (n = 926)	High dose (n = 920)	Low flux (n = 925)	High flux (n = 921)
อายุ (ปี)*	57.6 ± 14.0	57.8 ± 13.9	57.4 ± 14.1	57.6 ± 14.2	57.7 ± 13.9
เพศหญิง**	1037 (56.2)	521 (56.3)	517 (56.2)	516 (55.8)	522 (56.7)
ประวัติโรคเบาหวาน**	823 (44.6)	414 (44.7)	409 (44.5)	411 (44.4)	412 (44.7)
คะแนน ICED*	2.0 ± 0.8	2.0 ± 0.8	2.0 ± 0.8	2.0 ± 0.8	2.0 ± 0.8
โรคหัวใจ**	1479 (80.1)	755 (81.5)	724 (78.7)	745 (80.5)	734 (79.7)
ระยะเวลาล้างไต (ปี)*	3.7 ± 4.4	3.9 ± 4.5	3.6 ± 4.3	3.7 ± 4.2	3.8 ± 4.5
ความดันเลือดตัวบน (มม.ปรอท)*	151.8 ± 22.1	151.7 ± 22.0	151.9 ± 22.3	151.7 ± 22.7	151.9 ± 21.6
ความดันเลือดตัวล่าง(มม.ปรอท)*	81.4 ± 13.0	81.4 ± 12.8	81.5 ± 13.1	81.5 ± 13.3	81.9 ± 13.0
ครีเอตินิน (มก./ดล.)*	10.3 ± 2.9	10.3 ± 2.8	10.3 ± 3.0	10.3 ± 2.9	10.2 ± 2.9

* รายงานเป็น ค่าเฉลี่ย ± ค่าเบี่ยงเบนมาตรฐาน

** รายงานเป็นจำนวน (ร้อยละ)

หมายเหตุ: จาก “Effect of dialysis dose and membrane flux in maintenance hemodialysis” Eknoyan G, et al. N Engl J Med 2002; 347: 2010–9. ลิขสิทธิ์ปี ค.ศ. 2010 โดย Massachusetts Medical Society คัดลอกและดัดแปลงเมื่อวันที่ 22 ม.ค. 2553.

ศึกษาการรอดชีวิตของผู้ป่วยโรคไตวายที่ได้รับการรักษาด้วยวิธีการล้างไตที่ต่างกัน ซึ่งมีขนาดของกลุ่มตัวอย่างที่มากพอจะช่วยให้วิธีการสุ่ม (randomization) ประสบความสำเร็จ คือ มีการกระจายของตัวแปรกวนที่ทราบเท่า ๆ กันทั้งในกลุ่มควบคุมและกลุ่มทดลอง เช่น อายุเฉลี่ยของการล้างไตทั้ง 4 วิธี (ตารางที่ 1)

แต่ในตัวอย่างโครงการ NECOSAD study group¹¹ ซึ่งเป็นการศึกษาแบบ randomized control trial (RCT) โดยเปรียบเทียบการพยากรณ์โรคในกลุ่มที่ได้รับการล้างไตโดยการฟอกเลือด (hemodialysis) กับกลุ่มที่ล้างไตทางหน้าท้อง (peritoneal dialysis) ด้วยขนาดตัวอย่างเพียง 38 ราย ทำให้การกระจายของตัวแปรกวนทั้งสองกลุ่มมีลักษณะไม่เหมือนกัน พบว่าผู้ป่วยที่ถูกสุ่มให้ได้รับการรักษาด้วยวิธีการล้างไตทางหน้าท้องมีอายุเฉลี่ยน้อยกว่ากลุ่มที่ได้รับการฟอกเลือดประมาณ 7 ปี และยังมีประวัติเบาหวานรวมถึงโรคอื่น ๆ น้อยกว่ากลุ่มที่ได้รับการฟอกเลือด (ตารางที่ 2) ด้วยปัจจัยต่าง ๆ เหล่านี้ทำให้ผู้ป่วยที่ล้างไตทางหน้าท้องมีการพยากรณ์โรคที่ดีกว่ากลุ่มที่ได้รับการฟอกเลือด (HR = 3.8; p-value = 0.03)

ถึงแม้การศึกษานี้จะใช้วิธีการสุ่มเพื่อพยายามควบคุมอิทธิพล

ของตัวแปรกวนแล้วก็ตาม แต่ก็ไม่สามารถประสบความสำเร็จในการควบคุมตัวแปรกวนเนื่องจากขนาดตัวอย่างที่น้อยเกินไป จึงจำเป็นต้องควบคุมตัวแปรกวนโดยใช้วิธีการทางสถิติช่วยคือ multivariable analysis ซึ่งผลการศึกษาเมื่อตัวแปรกวนต่าง ๆ เหล่านั้นถูกควบคุมอิทธิพลแล้ว พบว่ายังไม่มีหลักฐานเพียงพอที่จะสรุปได้ว่าผู้ป่วยที่ได้รับการล้างไตทางหน้าท้องมีการพยากรณ์โรคที่ดีกว่ากลุ่มที่ได้รับการฟอกเลือด (HR = 3.6; p-value = 0.09) ดังตารางที่ 3

การจำกัดกลุ่มศึกษา (Restriction)

เป็นการควบคุมตัวแปรกวนในขั้นตอนของการออกแบบการศึกษาอีกวิธีหนึ่งโดยการจำกัดตัวแปรที่ใช้ในการควบคุมจำนวนหนึ่งหรือมากกว่าหนึ่งตัวแปร เช่น เพศ อาจจำกัดการศึกษาเฉพาะเพศชายหรือเพศหญิง หรือ การจำกัดอายุที่ศึกษาเฉพาะช่วงอายุ 40–50 ปี ซึ่งข้อดีของการใช้วิธีนี้ คือ ไม่สามารถอ้างอิงผลการศึกษาไปสู่กลุ่มอื่นได้ และในกรณีที่มีตัวแปรกวนหลาย ๆ ตัวเราสามารถที่จะควบคุมตัวแปรได้ไม่เกิน 2 ตัวแปร แต่ข้อดีของวิธีนี้คือ ไม่ต้องการขนาดตัวอย่างที่ใหญ่ ดังตัวอย่างการศึกษากาเวซิมเศร้ากับการเกิดโรคหัวใจในหญิงที่เป็นเบาหวาน¹² ซึ่งจำกัดกลุ่มศึกษา

ตารางที่ 2 ลักษณะข้อมูลพื้นฐานจากการสุ่มในกลุ่มศึกษาขนาดเล็กในโครงการ NECOSAD study group¹¹

ปัจจัย	Hemodialysis (n=18)		Peritoneal dialysis (n=20)	
	จำนวน	ร้อยละ	จำนวน	ร้อยละ
อายุ (ปี)*	62 ± 11		55 ± 12	
เพศชาย	11	61	11	55
สาเหตุการเกิดโรคไต				
เบาหวาน	5	28	4	20
กรวยไตอักเสบ	0	0	2	15
โรคหลอดเลือดที่ไต	4	22	2	15
อื่น ๆ	9	50	10	50
คะแนน Khan's comorbidity				
ต่ำ	9	50	9	45
กลาง	4	22	7	35
สูง	5	28	4	20
ยูเรีย (มิลลิโมลต่อลิตร)*	32.8 ± 10.2		29.0 ± 9.5	
Creatinine clearance (มล./นาที/1.73 ม ²)*	11.0 ± 9.2		11.6 ± 11.2	

* รายงานเป็นค่าเฉลี่ย ± ค่าเบี่ยงเบนมาตรฐาน

หมายเหตุ: จาก “Effect of starting with hemodialysis compared with peritoneal dialysis in patients new on dialysis treatment: a randomized controlled trial” Korevaar JC, et al. Kidney Int 2003; 64: 2222–8. ลิขสิทธิ์ ปี ค.ศ. 2010 โดย Nature Publishing Group ถัดลอกและดัดแปลงเมื่อวันที่ 20 ม.ค. 2553.

ตารางที่ 3 การรอดชีพของผู้ป่วยโรคไตด้วยวิธีการรักษาแบบ hemodialysis เทียบกับ peritoneal dialysis จากกลุ่มกลุ่มศึกษาขนาดเล็กในโครงการ NECOSAD study group¹¹

วิธีการรักษาแบบ hemodialysis เทียบกับ peritoneal dialysis	Hazard ratio (95% CI)	p-value
โมเดล 0: เมื่อไม่มีการควบคุมอิทธิพลของตัวแปรอื่น ๆ	3.8 (1.1–12.6)	0.03
โมเดล 1: เมื่อควบคุมอิทธิพลของตัวแปรอายุ	2.9 (0.8–10.1)	0.10
โมเดล 2: เมื่อควบคุมอิทธิพลของตัวแปรอายุ โรคร่วม และโรคที่เป็นสาเหตุของโรคไต	3.6 (0.8–15.4)	0.09

หมายเหตุ: จากแหล่งเดียวกับตารางที่ 2

เฉพาะเพศหญิง โดยนักวิจัยให้เหตุผลว่ามีการศึกษาก่อนหน้านี้ รายงานอัตราการเกิดโรคหัวใจสูงในผู้ป่วยเบาหวานเพศหญิง และภาวะซึมเศร้าที่พบในเพศชายมีจำนวนน้อย รายละเอียดเหล่านี้ นักวิจัยจำเป็นต้องเขียนชี้แจงไว้ในระเบียบวิธีการวิจัยให้ชัดเจน

การจับคู่ (matching)
การจับคู่ตัวแปรกวน (confounder matching) สามารถทำได้ทั้งในรูปแบบการศึกษาแบบ cohort study และ case control study เพื่อทำให้อิทธิพลของตัวแปรกวนลดลง

ในกรณีที่เป็นการศึกษาแบบ cohort study เป็นการจับคู่ของตัวแปรความให้เท่ากันในกลุ่มสัมผัสปัจจัย (exposed) และกลุ่มไม่ได้สัมผัสปัจจัย (unexposed) เช่น การศึกษาอัตราการเสียชีวิตในผู้ป่วยสูงอายุที่เข้ารับการรักษานในหอผู้ป่วยหนัก (intensive care unit, ICU)¹³ เป็นการศึกษาจับคู่แบบติดตามไปข้างหน้า (a prospective matched cohort study) นักวิจัยคัดเลือกผู้ป่วยที่มีอายุมากกว่า 70 ปี จำนวน 100 คน และผู้ป่วยที่มีอายุน้อยกว่า 70 ปี อีกจำนวน 100 คน ที่เข้ารับการรักษานในเดือนเดียวกัน มีเกณฑ์ในการจับคู่ คือ ระดับความรุนแรงของความเจ็บป่วยเมื่อแรกรับเข้ารักษาที่หอผู้ป่วยหนักโดยใช้แบบประเมินระดับความรุนแรงของโรกระบบ Simplified Acute Physiology Score II (SAPS II) โดยไม่คำนวนอายุร่วมด้วยในการประเมิน ผลจากการจับคู่ด้วยเกณฑ์ดังกล่าวทำให้คะแนนเฉลี่ยระดับความรุนแรงของโรกระบบ SAPS II ในกลุ่มผู้ป่วยที่มีอายุมากกว่า 70 ปี และกลุ่มผู้ป่วยที่มีอายุน้อยกว่า 70 ปี มีค่าใกล้เคียงกันมากคือ 31 ± 14 และ 32 ± 15 คะแนนตามลำดับ

สำหรับการศึกษารูปแบบ case control study เป็นการจับคู่ของกลุ่มศึกษา (case) คือกลุ่มที่มี outcome ที่ศึกษา และกลุ่มเปรียบเทียบ (control) คือกลุ่มที่ไม่มี outcome ที่ศึกษา ให้มีลักษณะตัวแปรความที่เหมือนกัน ดังการศึกษาเรื่องแอนติบอดีที่เป็นปัจจัยที่มีความสัมพันธ์กับการแท้งในอายุครรภ์น้อย¹⁴ โดยกำหนดกลุ่มศึกษา คือ หญิงที่มีภาวะแท้งบุตรเองตามธรรมชาติ (spontaneous abortion) เมื่ออายุครรภ์ 8-9 สัปดาห์ จำนวน 743 ราย และ หญิงที่ทำแท้ง (provoked abortion) เมื่ออายุครรภ์ 8-9 สัปดาห์ จำนวน 743 ราย เป็นกลุ่มเปรียบเทียบ เนื่องจากมีหลักฐานบ่งบอกว่าระดับของแอนติบอดีชนิด antiphospholipid มีความสัมพันธ์กับการแท้งบุตรเองตามธรรมชาติ ซึ่งนักวิจัยวางแผนควบคุมอิทธิพลของตัวแปรความ 3 ตัว ได้แก่ อายุของหญิงตั้งครรภ์ ระยะเวลาที่ท้องว่างหลังจากการแท้ง และจำนวนการตั้งครรภ์ ซึ่งการควบคุมอิทธิพลของทั้งสามตัวแปรนี้พร้อมกันทำให้กลุ่มหญิงที่มีภาวะแท้งบุตรเองตามธรรมชาติและกลุ่มหญิงที่ทำแท้งมีลักษณะของตัวแปรความทั้งสามตัวเหมือนกัน คือ อายุของหญิงตั้งครรภ์ 24 (19-31) ปี ระยะเวลาที่ท้องว่างหลังจากการแท้ง 9 (6-12) เดือน จำนวนการตั้งครรภ์ 1, 2 และ 3 ครั้งเท่ากับ 212, 266, 265 คนตามลำดับ¹⁴

ส่วนใหญ่เราจะเห็นรูปแบบการศึกษาแบบ case control study โดยใช้วิธีการจับคู่อยู่เป็นประจำ แต่การเลือกตัวแปรมาจับคู่จำเป็นต้องอธิบายให้ได้ว่าเพราะเหตุใด เพราะการจับคู่ที่ไม่เหมาะสมก็เป็นสาเหตุของความผิดพลาดได้เช่นกัน ซึ่งจะกล่าวต่อไปในตอนท้าย

ที่กล่าวมาทั้ง 3 วิธีข้างต้น คือ randomization, restriction และ matching เป็นวิธีการควบคุมตัวแปรความในขั้นตอนออกแบบการศึกษาหรือก่อนการเก็บข้อมูล แต่การจัดการหรือควบคุมตัวแปรความยังสามารถทำได้ในขั้นตอนของการวิเคราะห์ข้อมูล คือ stratification analysis และ multivariable analysis

การวิเคราะห์จำแนกชั้นภูมิ (Stratification analysis)

เป็นกระบวนการในการแบ่งกลุ่มตัวอย่างออกเป็นระดับชั้นตามลักษณะพื้นฐานประชากรหรือตามลักษณะของตัวแปรที่นักวิจัยพิจารณาเป็นตัวแปรความ จุดประสงค์ของการจำแนกชั้นภูมิเพื่อให้กลุ่มตัวอย่างมีลักษณะของตัวแปรความที่คล้าย ๆ กัน เช่น การศึกษาความสัมพันธ์ระหว่างอายุที่เริ่มมีประจำเดือนครั้งแรกกับความเสี่ยงในการเป็นเบาหวานชนิดที่ 2 ในโครงการ Nurses' Health Study (NHS)¹⁵ ทำการศึกษาในผู้หญิงจำนวน 101,415 ราย โดยปัจจัยที่สนใจคืออายุที่เริ่มมีประจำเดือนครั้งแรก ≤ 11 , 12, 14 และ ≥ 15 ปี เทียบกับเมื่อเริ่มมีประจำเดือนครั้งแรกตอนอายุ 13 ปี ซึ่งเป็นกลุ่มอ้างอิง (reference group) โดยวิเคราะห์ค่าความเสี่ยงสัมพัทธ์ (relative risk, RR) จำแนกตามกลุ่มอายุที่เข้าร่วมในการศึกษาของอาสาสมัครคืออายุ < 45 และ ≥ 45 ปี (ตารางที่ 4)

จากตารางที่ 4 พบว่าในกลุ่มอายุ < 45 ปี ผู้ที่เริ่มมีประจำเดือนครั้งแรกเมื่ออายุ ≤ 11 , 12, 14 และ ≥ 15 ปี มีความเสี่ยงต่อการเกิดเบาหวานชนิดที่ 2 เป็น 1.31, 1.20, 1.05 และ 1.00 เท่าของผู้ที่เริ่มมีประจำเดือนครั้งแรกเมื่ออายุ 13 ปี ส่วนในกลุ่มอายุ ≥ 45 ปี ผู้ที่เริ่มมีประจำเดือนครั้งแรกเมื่ออายุ ≤ 11 , 12, 14 และ ≥ 15 ปี มีความเสี่ยงต่อการเกิดเบาหวานชนิดที่ 2 เป็น 1.11, 1.03, 0.85 และ 0.90 เท่าของผู้ที่เริ่มมีประจำเดือนครั้งแรกเมื่ออายุ 13 ปี เราเรียกค่าความเสี่ยงสัมพัทธ์ของการวิเคราะห์จำแนกชั้นภูมิว่า stratum-specific relative risks ในกรณีนี้คือจำแนกตามกลุ่มอายุ < 45 ปี และ ≥ 45 ปี ซึ่งมีค่า stratum-specific relative risks ไม่เท่ากัน จึงจำเป็นต้องนำเสนอผลการศึกษจำแนกตามกลุ่มอายุ เพื่อให้ทราบว่าความสัมพันธ์ของการมีประจำเดือนครั้งแรกกับการเกิดเบาหวานชนิดที่ 2 ขึ้นอยู่กับตัวแปรกลุ่มอายุอีกหนึ่งปัจจัย

บ่อยครั้งมีผู้ต้องการทราบว่าตัวแปรความนั้นเป็นตัวแปรความที่แท้จริงของความสัมพันธ์ที่สนใจจริงหรือไม่ และการควบคุมอิทธิพลของตัวแปรความมีความแตกต่างจากการไม่ควบคุมอย่างไร ขอยกตัวอย่างจากข้อมูลที่สมมติขึ้นของ Jager และคณะ¹⁶ ในการศึกษาความสัมพันธ์ของการเป็นโรคหัวใจขาดเลือดกับประวัติโรคเบาหวานในผู้ป่วยที่เปลี่ยนไต สมมติให้เป็นการศึกษารูปแบบ cohort study (ตารางที่ 5)

ตารางที่ 4 การวิเคราะห์จำแนกชั้นภูมิตามลักษณะของอายุที่เข้าในการศึกษา และประวัติเบาหวานในครอบครัวกับความเสี่ยงในการเป็นเบาหวานชนิดที่ 2¹⁶

		อายุเมื่อมีประจำเดือนครั้งแรก (ปี)								
		≤ 11		12		13	14		≥ 15	
		RR	95% CI	RR	95% CI	กลุ่มอ้างอิง	RR	95% CI	RR	95% CI
อายุที่เข้าในการศึกษา (ปี)										
< 45		1.31	1.16-1.47	1.20	1.07-1.34	1	1.05	0.89-1.24	1.00	0.80-1.24
≥ 45		1.11	1.01-1.22	1.03	0.94-1.13	1	0.85	0.75-0.96	0.90	0.79-1.03
p for interaction		0.008								
ประวัติเบาหวานในครอบครัว										
ไม่มี		1.25	1.13-1.38	1.13	1.02-1.24	1	0.92	0.80-1.04	1.01	0.89-1.17
มี		1.11	0.99-1.24	1.06	0.95-1.18	1	0.92	0.80-1.07	0.87	0.73-1.04
P for interaction		0.20								

RR = relative risk; 95% CI = 95% confidence interval
หมายเหตุ: จาก “Age at menarche and risk of type 2 diabetes: results from 2 large prospective cohort studies.”
Chunyah He, et al. Am J Epidemiol 2010; 171: 334-344. ลิขสิทธิ์ปี ค.ศ. 2010 โดย Oxford University Press.
คัดลอกและดัดแปลง เมื่อวันที่ 24 ม.ค. 2553.

ตารางที่ 5 การพิจารณาตัวแปรกวนในการศึกษาความสัมพันธ์ของการเป็นโรคหัวใจขาดเลือดกับโรคเบาหวานในผู้ป่วยที่เปลี่ยนไต⁴

		โรคหัวใจขาดเลือด			
		รวม	ป่วย	ไม่ป่วย	
ผู้ป่วยที่เปลี่ยนไตทุกราย					
เป็นเบาหวาน		560	184	376	} crude RR = 1.7 (1.45-1.99)
ไม่เป็นเบาหวาน (กลุ่มอ้างอิง)		1440	278	1162	
ผู้ป่วยเปลี่ยนไตที่อายุ < 65 ปี					
เป็นเบาหวาน		150	36	114	} stratum-specific RR < 65 ปี = 1.5 (1.09-2.07)
ไม่เป็นเบาหวาน (กลุ่มอ้างอิง)		850	136	714	
ผู้ป่วยเปลี่ยนไตที่อายุ ≥ 65 ปี					
เป็นเบาหวาน		410	148	262	} stratum-specific RR ≥ 65 ปี = 1.5 (1.24-1.82)
ไม่เป็นเบาหวาน (กลุ่มอ้างอิง)		590	142	448	

RR = relative risk
หมายเหตุ: จาก “Confounding: what it is and how to deal with it” Jager KJ, et al. Kidney Int 2008; 73: 256-60.
ลิขสิทธิ์ปี ค.ศ. 2010 โดย Nature Publishing Group. คัดลอกและดัดแปลง เมื่อวันที่ 21 ม.ค. 2553.

เมื่อวิเคราะห์ความสัมพันธ์ดังกล่าวโดยไม่ควบคุมอิทธิพลของอายุ เราเรียกค่าความเสี่ยงสัมพัทธ์นี้ว่า crude relative risk (crude RR) มีค่าเท่ากับ 1.7 หมายความว่าผู้ป่วยเปลี่ยนไตที่เป็นเบาหวานมีความเสี่ยงต่อการเกิดโรคหัวใจขาดเลือดเป็น 1.7 เท่าของผู้ป่วยเปลี่ยนไตที่ไม่ได้เป็นเบาหวาน แต่เมื่อวิเคราะห์จำแนกตามอายุที่ทำการเปลี่ยนไตคือ อายุ < 65 ปี และ ≥ 65 ปี ค่าความเสี่ยงสัมพัทธ์ที่ได้เรียกว่า stratum-specific RR ซึ่งเท่ากับ 1.5 ทั้งสองกลุ่ม หมายความว่าผู้ป่วยเปลี่ยนไตที่เป็นเบาหวานมีความเสี่ยงต่อการเกิดโรคหัวใจขาดเลือดเป็น 1.5 เท่าของผู้ป่วยเปลี่ยนไตที่ไม่ได้เป็นเบาหวาน ซึ่งความเสี่ยงที่เกิดขึ้นนี้เหมือนกันทั้งในกลุ่มที่เปลี่ยนไตเมื่ออายุ < 65 ปี และ ≥ 65 ปี แต่พบว่า crude RR มีค่าสูงกว่า stratum-specific RR นั้นหมายความว่าอายุมีอิทธิพลต่อความสัมพันธ์ของการเป็นโรคหัวใจขาดเลือดกับประวัติโรคเบาหวานสามารถสรุปได้ว่าอายุเป็นตัวแปรกวนในการศึกษานี้

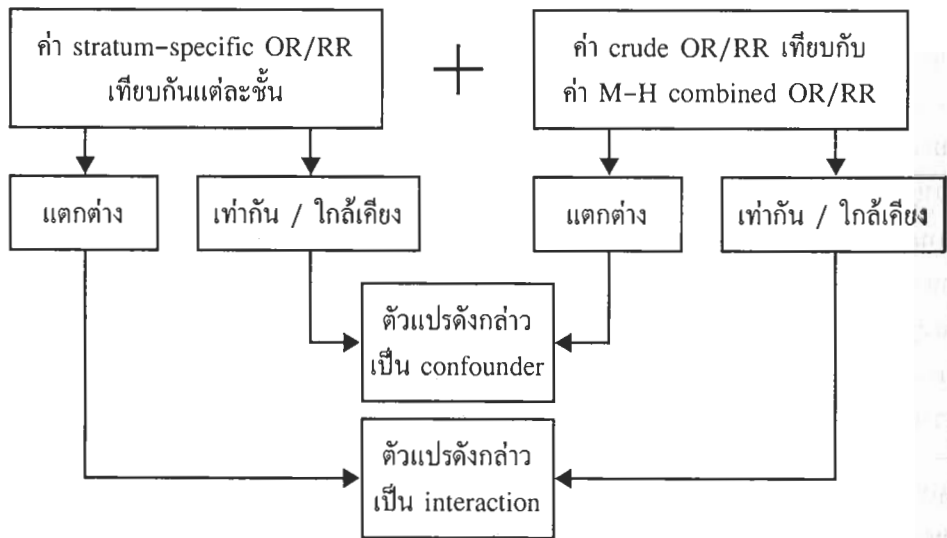
ข้อด้อยของการวิเคราะห์จำแนกชั้นภูมิ คือสามารถควบคุมอิทธิพลของตัวแปรกวนได้เพียงบางส่วน ไม่สามารถที่จะควบคุมอิทธิพลของตัวแปรกวนอื่น ๆ ได้ทั้งหมด นั่นคือยังคงมีตัวแปรกวนบางส่วนหลงเหลืออยู่ และเนื่องจากกลุ่มตัวอย่างถูกแบ่งออกเป็นกลุ่มย่อยทำให้ขนาดตัวอย่างในแต่ละชั้นภูมิลดลง ทำให้อำนาจในการวิเคราะห์ (power) ลดลง

การจำแนกชั้นภูมินอกจากสามารถควบคุมอิทธิพลของตัวแปรกวนได้เพียงบางส่วนแล้ว ยังสามารถช่วยในการประเมินอิทธิพลของตัวแปรตัวที่สามซึ่งมีอิทธิพลต่อปัจจัยที่นักวิจัยสนใจเช่นกันแต่ไม่ใช่ตัวแปรกวนเราเรียกว่า interaction หรือ effect modification ดังในตารางที่ 4 ที่แสดงค่า probability for

interaction (p for interaction) ซึ่งเป็นค่าที่บ่งบอกความน่าจะเป็นที่ผลของปัจจัยที่นักวิจัยสนใจไปขึ้นอยู่กับตัวแปรอีกตัวหนึ่งหรือตัวแปรทั้งสองตัวส่งผลต่อกันและกัน ดังนั้นเมื่อพิจารณาค่า RR ในตารางดังกล่าวจะเห็นว่ากลุ่มอายุเมื่อเข้าในการศึกษา ≥ 45 ปีขึ้นไปมีค่าความเสี่ยงต่อการเป็นเบาหวานชนิดที่ 2 แตกต่างจากกลุ่มอายุเมื่อเข้าในการศึกษา < 45 ปี และแสดงค่า p for interaction = 0.008 หมายความว่าความสัมพันธ์ของตัวแปรอายุเมื่อมีประจำเดือนครั้งแรกกับการเป็นเบาหวานชนิดที่ 2 ขึ้นอยู่กับอายุเมื่อเข้าในการศึกษา แต่ความสัมพันธ์ของอายุเมื่อมีประจำเดือนครั้งแรกกับการเป็นเบาหวานชนิดที่ 2 ไม่ขึ้นกับการมีประวัติเบาหวานในครอบครัว ดังจะเห็นจากค่า p for interaction = 0.20

ตัวแปรที่พิสูจน์แล้วว่าเป็น interaction มักเป็นตัวแปรที่เกิดอยู่ตามธรรมชาติไม่สามารถกำจัดออกไปได้ หากพบว่ามี interaction เกิดขึ้นจำเป็นต้องหาอธิบายแสดงไว้ในรายงานการศึกษา แต่ถ้าไม่สามารถอธิบายการเกิดขึ้นได้จะเป็นหนทางนำไปสู่การตั้งสมมติฐานการวิจัยต่อไป ซึ่งต่างจากตัวแปรกวนที่เป็นอคติ (bias) ที่เกิดขึ้นจำเป็นต้องทำการควบคุมอิทธิพลหรือกำจัด

ดังนั้นนักวิจัยจึงจำเป็นต้องทราบความแตกต่างระหว่างตัวแปรกวนกับ interaction ซึ่งพิจารณาอย่างง่ายจากค่า stratum-specific OR/RR ซึ่งเป็นค่าความเสี่ยงของปัจจัยที่สนใจกับโรคจำแนกตามปัจจัยอื่น สำหรับ crude OR/RR เป็นค่าความเสี่ยงของปัจจัยที่สนใจกับโรคโดยยังไม่ควบคุมอิทธิพลของตัวแปรอื่น และ M-H combined (Mantel-Haenszel estimate) ซึ่งเป็นค่าความเสี่ยงของปัจจัยที่สนใจกับโรคโดยได้ควบคุมอิทธิพลของตัวแปรอื่นแล้ว ดังรูปที่ 4



รูปที่ 4 การพิจารณาความแตกต่างระหว่าง confounder และ interaction อย่างง่าย

เพื่อให้มองเห็นภาพที่ชัดเจนขึ้นขอยกตัวอย่างข้อมูลที่สมมติขึ้น 2 เรื่องคือ การศึกษาความสัมพันธ์ระหว่างการสูบบุหรี่กับการเกิดโรคความดันเลือดสูง และการศึกษาความสัมพันธ์ระหว่างประวัติเบาหวานในครอบครัวกับการเกิดโรคจอประสาทตาเสื่อมซึ่งเป็นรูปแบบการศึกษาแบบภาคตัดขวาง (cross-sectional study) และวิเคราะห์ข้อมูลด้วยโปรแกรม STATA เพื่อทดสอบว่าเพศเป็นตัวแปรกวน หรือ interaction ในการศึกษาใด ซึ่งในการวิเคราะห์ความสัมพันธ์ระหว่างการสูบบุหรี่กับการเกิดโรคความดันเลือดสูงมีค่า stratum-specific OR เท่ากันทั้งเพศชายและเพศหญิง คือ 1.49 แต่พบว่า crude OR มีค่า 1.28 ซึ่งแตกต่างจาก M-H combined ซึ่งมีค่า 1.49 แสดงว่าเพศเป็นตัวแปรกวนในการศึกษานี้ (รูปที่ 5)

สำหรับการวิเคราะห์ความสัมพันธ์ระหว่างประวัติเบาหวานในครอบครัวกับการเกิดโรคจอประสาทตาเสื่อม stratum-specific OR ของเพศหญิงและเพศชายมีค่าแตกต่างกัน คือ 1.10 และ 0.87 ตามลำดับ แต่ค่า crude OR เท่ากับ M-H combined คือมีค่า 1.04 เท่ากัน แสดงว่าตัวแปรเพศเป็น interaction ของความสัมพันธ์ดังกล่าวในการศึกษานี้ (รูปที่ 6)

การวิเคราะห์ตัวแปรพหุ (multivariable analysis)

Multivariable analysis เป็นวิธีที่ใช้บ่อยมากในการลดบทบาทของตัวแปรกวนในระหว่างการศึกษาวิเคราะห์ข้อมูล การเลือกใช้วิธีการวิเคราะห์แบบ multivariable analysis¹⁶ ขึ้นอยู่กับว่าตัวแปรที่เป็นเหตุการณ์ที่สนใจ (outcome) เป็นข้อมูลชนิดใด ถ้าเป็นข้อมูล

SEX	OR	[95% Conf. Interval]		M-H Weight	Confounder
male	1.492062	1.278304	1.741671	138.892	(exact)
female	1.496444	1.065208	2.130953	30.58257	(exact)
Crude	1.284360	1.13201	1.457313		(exact)
M-H combined	1.492852	1.300392	1.713797		
Test of homogeneity (M-H)		Chi2 (1) =	0.00	Pr>chi2 =	0.9874
Test that combined OR = 1:					
		Mantel-Haenszel chi2 (1) =		32.57	
		Pr>chi2 =		0.0000	

รูปที่ 5 การพิจารณา confounder

SEX	OR	[95% Conf. Interval]		M-H Weight	Interaction
male	.8708134	.575385	1.317839	26.38889	(exact)
female	1.103714	.8691809	1.401558	67.89565	(exact)
Crude	1.038794	.846174	1.275273		(exact)
M-H combined	1.038529	.8503163	1.2684		
Test of homogeneity (M-H)		Chi2 (1) =	1.03	Pr>chi2 =	0.3106
Test that combined OR = 1:					
		Mantel-Haenszel chi2 (1) =		0.14	
		Pr>chi2 =		0.7110	

รูปที่ 6 การพิจารณา interaction

เชิงปริมาณ (numerical data) จำเป็นต้องใช้ linear regression ถ้าตัวแปรตามเป็นข้อมูลกลุ่ม 2 ระดับ (dichotomous data) เช่น เป็นโรค/ไม่เป็นโรค ต้องใช้วิธี dichotomous logistic regression หรือถ้าเป็นตัวแปรตามชนิดข้อมูลกลุ่ม 3 ระดับที่มีลักษณะเรียงลำดับ (ordinal data) เช่น ระดับความรุนแรงของโรคควรต้องวิเคราะห์ด้วย ordinal logistic regression แต่ถ้าตัวแปรตามเป็นการวัดการเกิดเหตุการณ์ตามเวลาที่เปลี่ยนแปลงไป เช่น ระยะเวลารอดชีพ ควรใช้ proportional hazards model ในการวิเคราะห์ข้อมูล แต่ทั้งนี้การวิเคราะห์ตัวแปรพหุแต่ละวิธีจำเป็นต้องศึกษารายละเอียดเพิ่มเติมต่อไปเนื่องจากมีข้อกำหนดบางประการที่ไม่ได้กล่าวถึงในบทความนี้

ตัวอย่างการใช้ multivariable analysis ในการควบคุมตัวแปรกวนดังตารางที่ 6 เป็นการศึกษาเรื่องการให้ยาทางจิตเวชกับความสัมพันธ์ของการเกิดกระดูกหัก¹⁷ นักวิจัยได้วิเคราะห์ผลการศึกษาโดยใช้ conditional logistic regression ซึ่งเป็นการวิเคราะห์ตัวแปรพหุสำหรับรูปแบบการศึกษาแบบ case control study ผลการศึกษาเมื่อไม่ได้ควบคุมอิทธิพลของตัวแปรหรือปัจจัยอื่น ๆ พบว่าการใช้ยาจิตเวชเสี่ยงต่อการเกิดกระดูกสะโพกหัก 1.8–2.3 เท่าของผู้ที่ไม่ได้ใช้ แต่เมื่อทำการควบคุมอิทธิพลของปัจจัยต่าง ๆ ได้แก่ โรคประจำตัว เช่น โรคหลอดเลือดสมอง ภาวะซึมเศร้า โรคหัวใจ ปวดหลัง เบาหวาน การใช้ยาอื่น ๆ เช่น ยาต้านภาวะซึมเศร้า วิตามินดี ยาขยายหลอดเลือด รวมถึงการสูบบุหรี่ และค่าดัชนีมวลกาย พบว่าการใช้ยาจิตเวชเป็นปัจจัยเสี่ยง

ของการเกิดกระดูกสะโพกหัก 1.2–1.3 เท่าของผู้ที่ไม่ใช้ ซึ่งระดับความเสี่ยงที่สูงเมื่อไม่ได้ทำการควบคุมอิทธิพลของตัวแปรหรือปัจจัยอื่น ๆ บ่งบอกว่าปัจจัยเหล่านั้นส่งผลร่วมกับการใช้ยาจิตเวชในการทำให้เกิดกระดูกสะโพกหัก

ความผิดพลาดที่เกิดขึ้นจากการพยายามควบคุมตัวแปรกวนเกินความจำเป็น

ปัจจุบันมีโปรแกรมทางสถิติช่วยในการวิเคราะห์ข้อมูลที่สามารถช่วยควบคุมอิทธิพลของตัวแปรกวนได้ง่ายขึ้น แต่บางครั้งก็มีแนวโน้มที่จะทำให้นักวิจัยได้ผลการศึกษาที่ผิดพลาดจากการพยายามควบคุมตัวแปรมากเกินไป (over adjustment)¹⁸ เช่น การควบคุมตัวแปรที่เป็นสาเหตุหนึ่งที่เป็นเส้นทางของการเกิดโรค (pathway) ระหว่างความสัมพันธ์ของการสัมผัสปัจจัย (determinant) กับการเกิดโรค (outcome) หรือการพยายามควบคุมอิทธิพลของตัวแปรร่วมอื่นที่ไม่ใช่ตัวแปรที่มีความเกี่ยวข้องกับการเกิดโรคซึ่งเป็นการควบคุมตัวแปรโดยไม่มีความจำเป็น (unnecessary adjustment)

ดังในตัวอย่างการศึกษาความสัมพันธ์ของดัชนีมวลกายกับการเกิดโรคไตวายเรื้อรังระยะสุดท้าย⁷ ที่กล่าวมาแล้วในตอนต้น ถ้านักวิจัยไม่ได้รับคำแนะนำตรวจสอบก่อนว่าความดันเลือดเป็นตัวแปรหนึ่งในเส้นทางของการเกิดไตวายเรื้อรังระยะสุดท้ายจากการทบทวนวรรณกรรม แล้วทำการควบคุมตัวแปรความดันเลือดด้วยวิธีการ multivariable analysis ก็จะทำให้บังผลที่แท้จริงของความ

ตารางที่ 6 ความสัมพันธ์ระหว่างการใช้ยาต้านทางจิตกับการเกิดกระดูกสะโพกหัก¹⁷

การใช้ยาจิตเวช	Case (n = 22,250)	Control (n = 22,250)	Crude odds ratio (95% CI)	Adjusted odds ratio* (95% CI)
ไม่ได้ใช้	19251 (86.5)	20702 (93.0)	1	1
ใช้	1495 (6.7)	751 (3.4)	2.2 (2.0–2.4)	1.3 (1.1–1.5)
เริ่มใช้	215 (1.0)	135 (0.6)	1.8 (1.4–2.2)	1.2 (0.9–1.6)
ใช้นาน ๆ ครั้ง	1280 (5.8)	616 (2.8)	2.3 (2.1–2.5)	1.3 (1.1–1.5)
ใช้ประจำ	1504 (6.8)	797 (3.6)	2.1 (1.9–2.3)	1.3 (1.2–1.5)

* ควบคุมอิทธิพลของโรคประจำตัว เช่น โรคหลอดเลือดสมอง ภาวะซึมเศร้า โรคหัวใจ ปวดหลัง เบาหวาน การใช้ยาอื่น ๆ เช่น ยาต้านภาวะซึมเศร้า วิตามินดี ยาขยายหลอดเลือด ประวัติการสูบบุหรี่ และดัชนีมวลกาย

หมายเหตุ: จาก “Risk of hip/femur fractures in patients using antipsychotics.” Hugenholtz GW, et al. Bone 2005; 37: 864–70. ลิขสิทธิ์ปี ค.ศ. 2010 โดย Elsevier ถัดลอกและดัดแปลง เมื่อวันที่ 21 ม.ค. 2553.

ตารางที่ 7 เปรียบเทียบขนาดของความสัมพันธ์เมื่อควบคุมตัวแปรที่เป็นเส้นทางของการเกิดโรค (pathway)^{*}

ดัชนีมวลกาย (กก./ม ²)	Adjusted RR (95% CI) (โมเดล 1 [*])	Adjusted RR (95% CI) (โมเดล 2 ^{**})
ปกติ (18.5-24.9)	1	1
น้ำหนักเกิน (25.0-29.9)	1.87 (1.64-2.14)	1.72 (1.50-1.96)
อ้วน ระดับ 1 (30.0-34.9)	3.57 (3.05-4.18)	2.98 (2.54-3.49)
อ้วน ระดับ 2 (35.0-39.9)	6.12 (4.97-7.54)	4.68 (3.79-5.79)
อ้วนมาก (≥ 40)	7.07 (5.37-9.37)	4.99 (3.77-6.60)

RR = Relative risk

* โมเดล 1; ไม่ได้ใส่ตัวแปรความดันเลือดเข้ามาควบคุมอิทธิพล, **โมเดล 2; ใส่แปรความดันเลือดเข้ามาควบคุมอิทธิพล

หมายเหตุ: จาก "Body mass index and risk for end-stage renal disease" Hsu Cy et al. Ann Intern Med 2006; 144: 21-8.

สัมพันธ์ระหว่างดัชนีมวลกายกับการเกิดโรคไตวายเรื้อรังระยะสุดท้าย
คือได้การประมาณค่าความเสี่ยงที่ต่ำกว่าความเป็นจริง (ตารางที่ 7)

ในบางครั้งเราอาจไม่ทราบตัวแปรใดเป็นตัวแปรกวน (confounder) หรือเป็นตัวแปรที่เป็นเส้นทางของการเกิดโรค (pathway) เนื่องจากความรู้ที่มีอยู่ในปัจจุบันยังไม่สามารถยืนยันได้ เราสามารถทำการค้นหาได้โดยทดลองใส่ตัวแปรที่เราตั้งข้อสงสัยเก่านั้น ในการวิเคราะห์และเปรียบเทียบผลระหว่างเมื่อใส่และไม่ใส่ตัวแปร นั้น โดยพิจารณาจากความแตกต่างระหว่างค่า crude OR (หรือ crude RR) และ adjusted OR (หรือ adjusted RR) ถ้าทั้งสองตัวมีค่าเท่ากันแสดงว่าตัวแปรนั้นไม่ใช่ตัวแปรกวนของความสัมพันธ์ นั้น แต่ถ้าพบมีความแตกต่างระหว่าง 2 ค่านี้แสดงว่าตัวแปรนั้น เป็นตัวแปรกวน ถ้าสามารถยืนยันได้ว่าตัวแปรนั้นเป็นตัวแปรกวน ก็เป็นความรู้ใหม่หรือเป็นข้อมูลที่เป็นประโยชน์ความรู้เดิม แต่ถ้าไม่สามารถยืนยันได้ว่าตัวแปรนั้นเป็นตัวแปรกวนก็อาจจะมียุทธศาสตร์ในการตรวจสอบว่าเป็นเส้นทางของการเกิดโรคนั้นหรือไม่ และนำไปสู่ การตั้งสมมติฐานการวิจัยต่อไป

ปัญหาในลักษณะของความพยายามที่จะควบคุมตัวแปร มากเกินไป (over adjustment) เช่นนี้ นอกจากพบในระหว่าง วิเคราะห์ข้อมูลแล้วอาจพบในระหว่างออกแบบการศึกษาได้ เช่น รูปแบบการศึกษาแบบ case control study เมื่อ case และ control ถูกจับคู่ด้วยตัวแปรที่เป็น pathway

สรุปได้ว่า การควบคุมตัวแปรกวนโดยไม่คำนึงถึงวิธีการใน การใช้ อาจจะทำให้เกิดอคติขึ้นแทนที่จะเป็นการป้องกันการเกิดอคติ ด้วยเหตุผลนั้นนักวิจัยจึงควรต้องระมัดระวังและตรวจสอบว่าตัวแปรนั้น เข้าเกณฑ์อย่างไรของการเป็นตัวแปรกวนทั้ง 3 เกณฑ์ตามที่กล่าว

ตั้งแต่แรก

การประเมินความเหมาะสมในการควบคุมอิทธิพลของตัวแปรกวน
สิ่งที่นักวิจัยควรตั้งข้อคำถามเพื่อประเมินผลการวิเคราะห์
ข้อมูลว่าเหมาะสมหรือไม่มีดังนี้^๑

1. การวิเคราะห์ทางสถิติสามารถอธิบายได้ชัดเจนหรือไม่

นักวิจัยควรอธิบายวิธีการทางสถิติอย่างชัดเจน เพื่อให้ผู้อ่าน
เข้าใจได้ว่าใช้วิธีการใดในการพยายามควบคุมอิทธิพลของตัวแปรกวน
เช่น ถ้าใช้วิธี multivariable analysis สิ่งสำคัญที่ต้องทราบคือ
ตัวแปรอะไรบ้างที่ใส่เข้าไปในระหว่างการวิเคราะห์ข้อมูล และ
ตัวแปรนั้นมีความสัมพันธ์กับเหตุการณ์ที่สนใจอย่างไร หรือถ้าใช้วิธี
การวิเคราะห์แบบแบ่งชั้นภูมิ (stratification analysis) นักวิจัยควร
ต้องเขียนให้ผู้อ่านทราบว่าใส่ตัวแปรใดในการแบ่งชั้นภูมิ และด้วย
เหตุผลใด

2. การวิเคราะห์ที่แตกต่างกันให้ผลการศึกษาที่เหมือนเดิมหรือไม่

การเปรียบเทียบผลการวิเคราะห์เมื่อยังไม่ควบคุมอิทธิพล
ของตัวแปรและเมื่อมีการควบคุมตัวแปร ถ้าผลการวิเคราะห์แตกต่าง
กันมาก ก็สามารถที่จะบอกได้ว่าความแตกต่างของผลการวิเคราะห์
นั้นเกิดขึ้นเนื่องจากตัวแปรบางตัวที่ส่งผลกระทบต่อขนาดของผล
การศึกษา ดังนั้นในรายงานผลการศึกษาคควรแสดงผลการเปรียบเทียบ
ไว้คู่กันเพื่อเป็นประโยชน์ในการประเมิน

3. ความเป็นไปได้ของผลการศึกษา

จากที่กล่าวในตอนต้นว่าเราไม่สามารถจัดการกับตัวแปรกวน
ที่ไม่ได้เก็บมาหรือตัวแปรกวนที่เราไม่ทราบได้ บางครั้งผลการศึกษา
อาจมีผลกระทบจากกรณีนี้ได้ซึ่งเป็นสิ่งที่นักวิจัยไม่ตั้งใจให้เกิดขึ้น

แต่เป็นไปได้ยากที่จะควบคุม ดังนั้นการพิจารณาถึงความเป็นไปได้ของผลการวิเคราะห์ก็เป็นสิ่งสำคัญที่ต้องพิจารณา เช่น การพิจารณาถึงความเป็นไปได้ทางด้านชีววิทยา ดังผลการศึกษาจากตัวอย่างการให้ยาทางจิตเวชกับความสัมพันธ์ของการเกิดกระดูกสะโพกหักที่แสดงว่าผู้รับประทานยาลดอาการทางจิตเวชมีความเสี่ยงต่อการเกิดกระดูกสะโพกหักเมื่อควบคุมอิทธิพลของปัจจัยอื่น ๆ แล้ว เมื่อพิจารณาถึงความเป็นไปได้แล้ว ขาดจิตเวชไม่สามารถทำให้กระดูกสะโพกหักได้โดยตรง แต่การที่กระดูกสะโพกหักนั้นขึ้นอยู่กับกลไกหลายอย่าง เช่น การได้รับยาป้องกันจิตประสาท อาจทำให้มีการเปลี่ยนแปลงความแข็งแรงของกล้ามเนื้อ หรือมีภาวะความดันเลือดต่ำ จึงทำให้ผู้ป่วยหกล้มและเกิดกระดูกสะโพกหักได้เป็นต้น

สรุป

ตัวแปรกลายเป็นสาเหตุทำให้ผลที่เกิดขึ้นระหว่างปัจจัยกับเหตุการณ์ที่สนใจไม่ใช่ผลที่แท้จริง การพิจารณาควบคุมอิทธิพลของตัวแปรใด ควรทบทวนวรรณกรรมเป็นอย่างดีเพื่อให้สามารถวางแผนครอบคลุมถึงตัวแปรกวนที่ทราบ และเข้าใจพยาธิสภาพของการดำเนินของโรคเพื่อให้ทราบว่าตัวแปรใดเป็น pathway ของการเกิดโรคนั้นและไม่เกิดปัญหา over adjustment จากการพยายามควบคุมอิทธิพลเนื่องจากเข้าใจว่าเป็นตัวแปรกวน ดังนั้นก่อนที่จะควบคุมตัวแปรกวนควรพิจารณาว่าเป็นตัวแปรกวนจริงหรือไม่ วิธีในการจัดการตัวแปรกวนสามารถจัดการได้ทั้งในขั้นตอนการออกแบบการศึกษา เช่น randomization, restriction และ matching และในขั้นตอนการวิเคราะห์ข้อมูล เช่น stratification หรือ multivariable analysis แต่ถึงแม้ว่าจะมีวิธีในการควบคุมตัวแปรทั้งก่อนและหลังการเก็บข้อมูล แต่ไม่มีวิธีใดที่สามารถกำจัดอคติที่เกิดขึ้นจากตัวแปรกวนที่ไม่ได้เก็บข้อมูลไว้ หรือตัวแปรกวนที่ไม่ทราบได้

เอกสารอ้างอิง

1. Clancy MJ. Overview of research designs. *Emerg Med J* 2002; 19: 546-9.
2. Gerstman BB. Epidemiology kept simple: an introduction to classic and modern epidemiology. New York: Wiley-Liss, Inc; 1998. p.183-7.
3. Rothman KJ, Greenland S, Lash TL. Modern epi-

demiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008. p.5-31.

4. Jager KJ, Zoccali C, Macleod A, Dekker FW. Confounding: what it is and how to deal with it. *Kidney Int* 2008; 73: 256-60.
5. Szklo M, Nieto J. Epidemiology: beyond the basics. 2nd ed. Sudbury: Jones and Barlett Publishers; 2007. p.151-82.
6. Couchoud C, Moranne O, Frimat L, Labeeuw M, Allot V, Stengel B. Associations between comorbidities, treatment choice and outcome in the elderly with end-stage renal disease. *Nephrol Dial Transplant* 2007; 22: 3246-54.
7. Hsu CY, McCulloch CE, Iribarren C, Darbinian J, Go AS. Body mass index and risk for end-stage renal disease. *Ann Intern Med* 2006; 144: 21-8.
8. Miettinen OS, Cook EF. Confounding: essence and detection. *Am J Epidemiol* 1981; 114: 593-603.
9. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002; 359(9302): 248-52.
10. Eknoyan G, Beck GJ, Cheung AK, Daugirdas JT, Greene T, Kusek JW, et al; Hemodialysis (HEMO) study Group. Effect of dialysis dose and membrane flux in maintenance hemodialysis. *N Engl J Med* 2002; 347: 2010-9.
11. Korevaar JC, Feith GW, Dekker FW, van Manen JG, Boeschoten EW, Bossuyt PM, et al; NECOSAD Study Group. Effect of starting with hemodialysis compared with peritoneal dialysis in patients new on dialysis treatment: a randomized controlled trial. *Kidney Int* 2003; 64: 2222-8.
12. Clouse RE, Lustman PJ, Freedland KE, Griffith LS, McGill JB, Carney RM. Depression and coronary heart disease in women with diabetes. *Psychosom Med* 2003; 65: 376-83.
13. Johnson A, Berg G, Fleegler E, Sauerbrun M. A matched-cohort study of selected clinical and utilization outcomes for an asthma care support program. *Dis Manag* 2005; 8: 144-54.

14. Gris JC, Perneger TV, Quere I, Mercier E, Fabbro-Peray P, Lavigne-Lissalde G, et al. Antiphospholipid/antiprotein antibodies, hemostasis-related autoantibodies, and plasma homocysteine as risk factors for a first early pregnancy loss: a matched case-control study. *Blood* 2003; 102: 3504-13.
15. Chunyan H, Cuilin Z, David JH, Susan EH, Germaine MBL, Mary LH, et al. Age at menarche and risk of type 2 diabetes: results from 2 large prospective cohort studies. *Am J Epidemiol* 2010; 171: 334-44.
16. Katz MH. Multivariable analysis: a practical guide for clinicians. 2nd ed. New York: Cambridge University Press; 2008. p.14-35.
17. Hugenholtz GW, Heerdink ER, van Staa TP, Nolen WA, Egberts AC. Risk of hip/femur fractures in patients using antipsychotics. *Bone* 2005; 37: 864-70.
18. VanderWeele TJ. On the relative nature of overadjustment and unnecessary adjustment. *Epidemiology* 2009; 20: 496-9.
19. Sharon GE. Readers' responses to "Doctors and patients working together". *MedGenMed* 2005; 7: 6.