

Odds Ratio ของปัจจัยที่เป็นตัวแปรต่อเนื่อง

อรุณ จิรวัดณ์กุล วท.บ. (อาชีวอนามัย), วท.ม. (ชีวสถิติ), M.Sc. (Clinical Epidemiology)

ค่า Odds Ratio หรือ OR เป็นสถิติที่ใช้แสดงความสัมพันธ์ระหว่างตัวแปรผลกับตัวแปรปัจจัย การคำนวณค่า OR ส่วนใหญ่จะคำนวณจากตารางการแจกแจง 2 ทางระหว่างกับตัวแปรปัจจัยกับตัวแปรผลที่เป็นตัวแปรกลุ่มแบบมีสองค่า ในกรณีที่ตัวแปรปัจจัยเป็นตัวแปรต่อเนื่อง จะคำนวณ OR ได้ 2 วิธีคือ วิธีแรกนำค่าตัวแปรต่อเนื่องไปคำนวณค่า OR โดยตรง วิธีที่ 2 แปลงค่าตัวแปรต่อเนื่องให้เป็นตัวแปรกลุ่มก่อนนำไปคำนวณค่า OR อาจแปลงเป็นตัวแปรกลุ่มแบบมีสอง 2 ค่า หรือมากกว่า 2 ค่า ค่า OR ที่คำนวณจากค่าของตัวแปรต่อเนื่องโดยตรง ค่าแปลงที่เป็นตัวแปรกลุ่มแบบมีสองค่า หรือค่าแปลงที่เป็นตัวแปรกลุ่มแบบมีมากกว่า 2 ค่า จะได้ค่า OR ที่ต่างกัน การเลือกว่าจะคำนวณค่า OR จากข้อมูลแบบใดขึ้นอยู่กับลักษณะความสัมพันธ์ระหว่างตัวแปรผลและตัวแปรปัจจัย ค่า OR ที่คำนวณด้วยวิธีที่สอดคล้องกับลักษณะความสัมพันธ์ จึงจะแสดงความสัมพันธ์ได้ถูกต้อง

ปัจจุบันโปรแกรมวิเคราะห์ทางสถิติและคอมพิวเตอร์ที่ใช้วิเคราะห์ผลมีความสามารถในการคำนวณสูง ทำให้การคำนวณค่า OR ทุกวิธีทำได้ง่าย ความไม่เหมาะสมของการคำนวณค่า OR ของตัวแปรต่อเนื่องที่พบบ่อยมีสาเหตุมาจากการเลือกวิธีคำนวณไม่สอดคล้องกับลักษณะความสัมพันธ์ และในกรณีคำนวณค่า OR ของตัวแปรปัจจัยที่แบ่งมากกว่า 2 กลุ่ม ถ้านักวิจัยไม่สั่งโปรแกรมว่าต้องการคำนวณค่า OR ของกลุ่มใดเปรียบเทียบกับกลุ่มใด โปรแกรมจะคำนวณตามค่าเบื้องต้นที่กำหนดไว้ (default) ค่า OR ที่ได้อาจไม่ใช่สิ่งที่นักวิจัยต้องการ

การคำนวณค่า OR ด้วยการจัดกลุ่มตัวแปรต่อเนื่องให้เป็นสองกลุ่ม มีสมมุติฐานว่าค่าตัวแปรต่อเนื่องแต่ละข้างของจุดตัดมีความสัมพันธ์เป็นแนวราบระดับเดียวกัน ค่าข้างหนึ่งของจุดตัดจะสัมพันธ์กับการเกิดผลลัพธ์ (outcome) ที่สนใจเช่น ป่วย หรือตาย ถ้าข้อมูลที่ศึกษามีความสัมพันธ์เช่นนี้ ค่า OR ที่คำนวณได้สามารถใช้ประเมินว่าถ้าตัวแปรปัจจัยสูง (หรือต่ำ) กว่าเกณฑ์ (จุดตัด) จะมีโอกาสเกิดผลลัพธ์เพิ่มขึ้นอีกเท่าเท่า เช่น ความสัมพันธ์การเกิด stroke กับระดับความดันโลหิต โดยกำหนดจุดตัดของความดันซิสโตลิก ไว้ที่ 140 mmHg (เป็นหรือไม่เป็นโรคความดันโลหิตสูง) พบว่ามีค่า OR = 2 แสดงว่าคนที่มีความดันซิสโตลิก >140 mmHg จะมีความเสี่ยงของการเกิด stroke 2 เท่าของผู้ที่มีความดันซิสโตลิกเท่ากับหรือน้อยกว่า 140 mmHg

ในการแปลผลควรทำความเข้าใจเหตุผลของการกำหนดจุดตัดด้วย เช่น ความดันซิสโตลิก >140 mmHg เป็นจุดที่ใช้แยกระหว่างการเป็นหรือไม่เป็นโรคความดันโลหิตสูง นำไปหาสัมพันธ์ของการเกิด stroke เพื่อจะดูว่าการเป็นโรคความดันโลหิตสูงเป็นปัจจัยเสี่ยงของการเกิด Stroke หรือไม่ ค่า OR ที่คำนวณได้จะเป็นค่าเฉลี่ยของทั้งกลุ่มที่มีความดันซิสโตลิก >140 mmHg ค่า OR ที่ได้ จะไม่สามารถนำไปสรุปว่า คนที่มีความดันซิสโตลิก 160 mmHg จะมีความเสี่ยงของการเกิด stroke เท่ากับคนที่มีความดันซิสโตลิก 142 mmHg เพราะลักษณะความสัมพันธ์ระหว่างความดันโลหิตกับการเกิด stroke ของผู้ที่เป็นโรคความดันโลหิตสูง (ความดันโลหิต >140 mmHg)

ความเสี่ยงจะเพิ่มขึ้นเป็นช่วง ๆ ตามระดับความดันโลหิตที่เพิ่มขึ้น

ถ้าต้องการดูความเสี่ยงตามระดับของปัจจัยต้องแบ่งตัวแปรต่อเนื่องออกช่วง ๆ (กลุ่ม) ตามระดับความเสี่ยง การคำนวณค่า OR ในกรณีนี้จะมีสมมติฐานว่า ค่าตัวแปรต่อเนื่องแต่ละช่วง มีความสัมพันธ์เป็นแนวราบระดับเดียวกัน เช่น ความดันซิสโตลิกแบ่งออกเป็น 4 กลุ่ม ปกติ (SBP < 140) stage 1 (SBP 141–159) stage 2 (SBP 160–179) stage 3 (SBP >179) ในการคำนวณค่า OR จะมีสมมติฐานว่าคนที่มีความดันซิสโตลิกต่างกัน แต่อยู่ในกลุ่มเดียวกันจะมีความเสี่ยงของการเกิด stroke เท่ากัน ในการคำนวณค่า OR นักวิจัยต้องรู้ว่าต้องการคำนวณค่า OR ระหว่างกลุ่มใดเทียบกับกลุ่มใด จึงจะสามารถส่งโปรแกรมวิเคราะห์ให้คำนวณค่า OR ที่ต้องการ

การแบ่งกลุ่มปัจจัยที่เป็นตัวแปรต่อเนื่อง ถ้ามีองค์ความรู้หรือมีการกำหนดจุดตัดทางวิชาการอยู่แล้ว ให้ใช้จุดตัดดังกล่าวกับงานวิจัยได้เลย ถ้าตัวแปรปัจจัยที่ศึกษา ยังไม่มีการกำหนดจุดตัดมาก่อน นักวิจัยต้องการกำหนดจุดตัดขึ้นมาเอง ควรศึกษาเพิ่มเติมเรื่องวิธีการหาจุดตัดที่เหมาะสม (optimum cutoff point) ที่ทำได้หลายวิธีเพื่อให้ได้วิธีการแบ่งกลุ่มที่สอดคล้องกับลักษณะความสัมพันธ์ระหว่างตัวแปรผลกับตัวแปรปัจจัย

กรณีที่คำนวณค่า OR จากค่าของตัวแปรต่อเนื่องโดยตรง จะต้องสมมติฐานว่าตัวแปรผลและตัวแปรปัจจัยมีความสัมพันธ์เชิงเส้นตรงทั้งช่วงข้อมูลตั้งแต่ค่าน้อยถึงค่ามาก วิธีการคำนวณจะคำนวณจากค่าสัมประสิทธิ์ของสมการ logistic regression ค่า OR ที่คำนวณได้จะแสดงความเสี่ยงที่เพิ่มขึ้นเมื่อค่าของปัจจัยเพิ่มขึ้น

หนึ่งหน่วย เช่น ค่า OR ของอายุ (ปี) กับการเป็นโรคเท่ากับ 1.2 หมายความว่า การเป็นโรคจะเพิ่ม 1.2 เท่าของอายุที่เพิ่มขึ้นในแต่ละปี

ค่า OR ตัวแปรต่อเนื่องที่คำนวณจากค่าข้อมูลโดยตรง ต้องระมัดระวังในการแปลผล เพราะลักษณะความสัมพันธ์ของการเกิดโรคกับตัวแปรปัจจัยที่เป็นตัวแปรต่อเนื่องส่วนใหญ่ ความเสี่ยงของการเกิดโรคจะเพิ่มขึ้นเป็นช่วง ๆ ของตัวแปรปัจจัยต่อเนื่อง เช่น อายุกับการเป็นโรคหัวใจ ความเสี่ยงจะเพิ่มขึ้นตามช่วงอายุดังนี้ <40, 40–59, 60–79 และ 80+ ทำให้ค่า OR ที่คำนวณโดยตรงจากค่าตัวแปรต่อเนื่อง แสดงความสัมพันธ์ไม่สอดคล้องกับความจริง

การแบ่งกลุ่มตัวแปรต่อเนื่องเพื่อหาค่า OR นอกจากจะเข้าใจลักษณะความสัมพันธ์แล้ว นักวิจัยควรเข้าใจถึงวัตถุประสงค์ของการนำผลไปใช้งานด้วย เช่น ในการป้องกันการเกิดโรค จะแบ่งเป็นสองกลุ่มคือ กลุ่มที่มีความเสี่ยงกับกลุ่มไม่มีความเสี่ยงโรคกับการเป็นโรค เพื่อคัดกรองผู้ที่มีความเสี่ยง กรณีต้องการทำนายผลลัพธ์ของการรักษา อาจแบ่งออกเป็นหลายกลุ่มตามระดับความเสี่ยง

สรุป

การวิเคราะห์ปัจจัยเสี่ยงที่เป็นตัวแปรต่อเนื่อง นักวิจัยต้องเข้าใจลักษณะความสัมพันธ์ของตัวแปรปัจจัยกับตัวแปรผล ว่ามีความสัมพันธ์ในลักษณะใด และต้องการค่า OR ไปใช้งานอย่างไร จึงจะสามารถเลือกวิธีการคำนวณค่า OR ได้สอดคล้องกับลักษณะความสัมพันธ์และการนำผลไปใช้งาน