

การวัดความสอดคล้องด้วยสถิติ Kappa

อรุณ จิรวัดนกุล

ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

ในการศึกษาเรื่องรูปแบบการค้นหาผู้ป่วยโรคเบาหวาน โดยใช้ระดับน้ำตาลในเลือด (FBS) เป็นตัวประเมินประสิทธิผลของรูปแบบการค้นหาผู้ป่วยใหม่โดยมีนักวิทยาศาสตร์การแพทย์ 2 คน เป็นผู้ตรวจวิเคราะห์ระดับน้ำตาลในเลือด

ประเด็นที่เกี่ยวข้องกับความถูกต้องของผลการตรวจเลือดเมื่อมีผู้ตรวจ 2 คน คือ ผู้ตรวจทั้งสองคนจะต้องสามารถตรวจวัดสิ่งเดียวกันให้ผลลัพธ์ที่เหมือนกัน

การประเมินว่าผู้วิเคราะห์ 2 คนวัดได้เหมือนกันหรือไม่ เรียกว่า ความสอดคล้อง (agreement) การประเมินความสอดคล้องของผู้วัดควรดำเนินการก่อนเริ่มเก็บข้อมูลวิจัย ตัวอย่างเช่น นำตัวอย่างเลือด 30 ตัวอย่าง โดยแต่ละตัวอย่างแบ่งให้ทั้ง 2 คนวิเคราะห์ ได้ผลการวิเคราะห์ดังตารางที่ 1

ในการประเมินความสอดคล้องระหว่างผู้ตรวจ 2 คน ไม่สามารถใช้ paired t-test (ทดสอบความแตกต่างของค่าเฉลี่ย) มาใช้ในการสรุปผลความสอดคล้องได้ ทั้งนี้เนื่องจาก

1. ผลการตรวจไม่เหมือนกัน แต่ให้ค่าเฉลี่ยที่เท่ากันได้ เช่น จากข้อมูลตัวอย่างระดับน้ำตาลในเลือดของคนที่ 4, 5 และ 6 ถ้าเปลี่ยนเป็น 84, 119 และ 96 ค่าเฉลี่ยที่ได้จะมีค่าเท่าเดิม (2 ตัวอย่างเพิ่มรวม 6 อีก 1 ตัวอย่างลดลง 6) แต่ถ้าพิจารณาค่าของข้อมูลแต่ละตัวอย่าง พบว่า ความเหมือนกันเปลี่ยนไปแล้ว

2. ในการตรวจวินิจฉัยโรคเบาหวาน ความถูกต้องของการตัดสินใจว่าเป็นโรค หรือไม่เป็นโรคอยู่ที่ผล

ตารางที่ 1

ตัวอย่าง	ผู้วิเคราะห์ 1	ผู้วิเคราะห์ 2
1	132	92
2	99	106
3	102	104
4	80	80
5	125	125
6	94	94
7	72	78
8	107	97
9	121	115
10	84	86
11	132	128
12	115	120
13	87	84
14	93	90
15	79	79
16	131	128
17	89	89
18	128	126
19	94	94
20	140	142
21	93	95
22	86	88
23	126	124
24	116	120
25	139	136
26	79	82
27	101	104
28	86	88
29	131	130
30	92	91

การตรวจช่วง 110 mg/dl ซึ่งเป็นเกณฑ์ที่ใช้ในการวินิจฉัยโรค ถึงแม้ผลการตรวจในช่วงอื่น ๆ แตกต่างกันบ้างแต่เมื่อนำมาสรุปผลการวินิจฉัยแล้ว ให้คำวินิจฉัยเหมือนกัน ก็ถือว่าการตรวจวิเคราะห์ของทั้ง 2 คนให้ผลสอดคล้องกัน

สถิติที่ใช้วัดความสอดคล้องในกรณีข้อมูลกลุ่ม คือสถิติ Kappa คำนวณจากค่าสัดส่วนความสอดคล้องที่พบกับสัดส่วนความสอดคล้องจากความบังเอิญตามสูตรข้างล่างนี้

$$K = \frac{(\text{สัดส่วนความสอดคล้องที่พบจริง} - \text{สัดส่วนความสอดคล้องจากความบังเอิญ})}{1 - \text{สัดส่วนความสอดคล้องจากความบังเอิญ}}$$

ค่าที่คำนวณได้จึงแสดงขนาดความสอดคล้องจริงที่หักโอกาสความสอดคล้องจากความบังเอิญออกแล้ว ค่า K ที่คำนวณได้จะมีค่าไม่เกิน 1

วิธีการวิเคราะห์ต้องนำข้อมูลที่นำมาเปลี่ยนเป็นผลการวินิจฉัยก่อนนำไปคำนวณค่า Kappa ได้ดังนี้

		ผู้วิเคราะห์ 2	
		ไม่เป็นเบาหวาน	เป็นเบาหวาน
ผู้วิเคราะห์ 1	ไม่เป็นเบาหวาน	19 a	2 b
	เป็นเบาหวาน	c 1	d 8

สูตรคำนวณค่าสถิติ Kappa กรณีตาราง 2 × 2

$$K = \frac{2(ad - bc)}{n^2 (p_1q_2 + p_2q_1)}$$

โดยที่ $p_1 = a + b/n$ และ $p_2 = (a + c)/n$

$$q_1 = 1 - p_1 \text{ และ } q_2 = 1 - p_2$$

ข้อมูลจากตารางคำนวณค่า K ได้เท่ากับ 0.77

Altman¹ เสนอแนวทางการแปลผลค่า Kappa (โดยนำแนวทางที่เสนอโดย Landis and Koch² มาปรับ

เล็กน้อย) ไว้ดังนี้

ค่า Kappa (K)	ระดับความสอดคล้อง
≤ 0.20	Poor
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81-1.00	Very Good

ค่า K ที่คำนวณจากตัวอย่างเท่ากับ 0.77 จะสรุปได้ว่าผลการวิเคราะห์ของผู้วิเคราะห์ทั้ง 2 คนมีความสอดคล้องดี

สถิติที่ใช้ในการวิเคราะห์ความสอดคล้องในกรณีมีผู้ประเมินมากกว่า 2 หรือมีผลการประเมินมากกว่า 2 กลุ่ม จะมีวิธีการคำนวณค่า Kappa ด้วยสูตรที่แตกต่างกัน ในกรณีที่ผลการวัดเป็นข้อมูลต่อเนื่องจะใช้สถิติ Intraclass correlation ในการประเมินความสอดคล้องในกรณีที่ต้องการประเมินความสอดคล้องระหว่างวิธีการตรวจ จะใช้วิธีการคำนวณค่า Kappa เหมือนกัน โดยเปลี่ยนจากผู้ประเมินเป็นวิธีการตรวจ

สรุป

การประเมินความสอดคล้องระหว่างผู้ตรวจในกรณีที่มีเกณฑ์การวินิจฉัยจะต้องนำผลที่วัดได้มาเทียบกับเกณฑ์แล้วเปลี่ยนเป็นผลการวินิจฉัย แล้วจึงนำผลการวินิจฉัยที่ได้ไปคำนวณค่าสถิติ Kappa เพื่อสรุปความสอดคล้อง

บรรณานุกรม

1. Altman DG. Practical Statistics for Medical Research. New York: Chapman & Hall/CRC; 1999.
2. Landis JR, Koch GG. The measurement of Observer agreement for categorical Data. Biometrics 1977; 33:159-74.
3. Looney SW. Biostatistical methods. New Jersey: Humana Press; 2002.