

การพิจารณางานวิจัยทางการแพทย์ ตอนที่ 3

กาญจนา จันทร์สูง

ภาควิชาอายุรศาสตร์ คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น ขอนแก่น 40002

Critical Appraisal of Scientific Paper (Part III)

Kanchana Chansung

Department of Medicine, Faculty of Medicine, Khon Kaen University.

ตอนที่ 3 หลักการพิจารณางานวิจัยเกี่ยวกับ การตรวจวินิจฉัย (Diagnostic test)

ในปัจจุบันแพทย์มักประสบปัญหาในการประเมินประสิทธิภาพของเครื่องมือที่ใช้ในการตรวจวินิจฉัยโรคชนิดใหม่ๆ ที่มีการคิดค้นพัฒนาอย่างมากมายเนื่องจากความก้าวหน้าอย่างรวดเร็วของพัฒนาการด้านวิทยาศาสตร์การแพทย์ ความรู้เกี่ยวกับหลักการประเมินงานวิจัยที่เกี่ยวข้องกับการตรวจวินิจฉัยจะช่วยให้แพทย์สามารถตัดสินใจได้ว่าควรเลือกใช้การตรวจใหม่ๆ ชนิดใดเพื่อช่วยในการวินิจฉัยโรคหรือไม่ และสามารถแปลผลข้อมูลที่ได้จากการตรวจวินิจฉัยได้อย่างถูกต้อง ครอบคลุมครบถ้วน เพื่อประโยชน์สูงสุดในการดูแลรักษาผู้ป่วย

ผู้เขียนใคร่ขอทำความเข้าใจถึงความหมายของคำว่า “เครื่องมือในการตรวจวินิจฉัย” เป็นเบื้องต้นก่อนว่าในที่นี้มีความหมายค่อนข้างกว้างขวาง โดยหมายถึงสิ่งซึ่งช่วยในการวินิจฉัยโรคหรือภาวะผิดปกติต่างๆ ของร่างกาย ซึ่งอาจเป็นการตรวจทางห้องปฏิบัติการ ข้อมูลทางคลินิกที่ได้จากการซักประวัติ ตรวจร่างกายผู้ป่วย หรือแม้แต่ภาพถ่ายทางรังสีต่างๆ ก็ได้ทั้งสิ้น

ในการพิจารณางานวิจัยเกี่ยวกับการตรวจวินิจฉัยก็เช่นเดียวกับการพิจารณางานวิจัยเกี่ยวกับประสิทธิภาพของการรักษาซึ่งจำเป็นต้องพิจารณาในสามหัวข้อหลัก คือ

1. ความถูกต้องเที่ยงตรงและน่าเชื่อถือของงานวิจัย
 2. ผลของงานวิจัย
 3. การนำไปใช้ประโยชน์ในการวินิจฉัยผู้ป่วย
- หลักในการพิจารณาทั้งสามหัวข้อโดยละเอียดมีดังนี้

การพิจารณาความถูกต้องเที่ยงตรงและน่า เชื่อถือของงานวิจัย

ข้อที่ควรคำนึงถึงในการพิจารณาความถูกต้องเที่ยงตรงของงานวิจัยเกี่ยวกับการตรวจวินิจฉัยมีดังต่อไปนี้

1. ได้มีการเปรียบเทียบผลของการตรวจวินิจฉัยโดยเครื่องมือที่ต้องการทดสอบกับผลของการตรวจวินิจฉัยด้วยเครื่องมือที่ถือเป็นมาตรฐานอ้างอิงโดยไม่มีอคติหรือไม่
2. กลุ่มตัวอย่างที่นำมาทดสอบนั้นเหมาะสมหรือไม่
3. ผลการตรวจโดยเครื่องมือที่ต้องการทดสอบมีอิทธิพลต่อการตัดสินใจเลือกส่งผู้ป่วยตรวจด้วยวิธีมาตรฐานที่ใช้อ้างอิงในงานวิจัยนั้นหรือไม่
4. มีการบันทึกรายละเอียดวิธีการตรวจวินิจฉัยโดยเครื่องมือที่ต้องการทดสอบไว้เพียงพอที่ผู้อ่านจะสามารถทำตามซ้ำได้หรือไม่

1. การเปรียบเทียบผลการตรวจวินิจฉัยโดยเครื่องมือที่ต้องการทดสอบกับผลของการตรวจวินิจฉัยด้วยเครื่องมือที่ถือเป็นมาตรฐานโดยไม่มีอคติ

มีสิ่งที่พึงพิจารณาอยู่สองข้อใหญ่ๆ ดังนี้

- การตรวจที่ถือเป็นมาตรฐานอ้างอิงที่เหมาะสมต้องสามารถให้การวินิจฉัยได้ถูกต้องตามความเป็นจริงมากที่สุด โดยไม่มีผลเสียหรืออันตรายต่อผู้ป่วย เรียกว่าเป็น “Gold Standard” ทั้งนี้ผู้อ่านพึงระลึกไว้เสมอว่าไม่มีสิ่งใดในโลกนี้ที่งามพร้อมสมบูรณ์ ในโลกแห่งความเป็นจริงไม่มีการตรวจวินิจฉัยใดที่สมบูรณ์แบบสามารถให้การวินิจฉัยผู้ป่วยได้ถูกต้องเต็มร้อยโดยไม่มีผลบวกหรือผลลบลงและไม่มีอันตรายต่อผู้ป่วย ดังนั้นจึงมักอนุโลมให้ใช้การตรวจที่ดีที่สุดเท่าที่

เป็นไปได้และได้รับการยอมรับจากกลุ่มผู้เชี่ยวชาญในสาขาวิชาที่เกี่ยวข้องเป็น Gold Standard แทน ตัวอย่างของการตรวจที่สามารถถือเป็น Gold Standard ได้เช่น การตรวจชิ้นเนื้อทางพยาธิวิทยา การตรวจศพ การวินิจฉัยจากการติดตามผู้ป่วยระยะยาว เป็นต้น

ในกรณีที่มีการตรวจที่ใช้เป็นเกณฑ์อ้างอิงในงานวิจัยนั้นไม่สามารถให้การวินิจฉัยได้แม่นยำเพียงพอที่ผู้อ่านจะยอมรับเป็นมาตรฐานอ้างอิงที่เหมาะสมได้ ผลของงานวิจัยนั้นย่อมไม่มีคุณค่าเพียงพอที่ผู้อ่านจะนำไปใช้ประโยชน์ต่อไปได้

● กลุ่มตัวอย่างที่นำมาศึกษาต้องได้รับการตรวจทั้งโดยวิธีการตรวจวินิจฉัยที่ต้องการทดสอบและวิธีการตรวจที่เป็นมาตรฐานอ้างอิงโดยไม่ขึ้นอยู่แก่กัน ผู้ทำการตรวจหรือแปลผลการตรวจของแต่ละวิธีต้องไม่ทราบผลการตรวจของอีกวิธีหนึ่งทั้งนี้เพื่อลดการเกิดอคติในการวัดเนื่องจากถ้าไม่มีการปกปิดการวินิจฉัยผู้อ่านผลย่อมมีแนวโน้มที่จะแปลผลการตรวจใกล้เคียงกับการตรวจด้วยวิธีอ้างอิง ซึ่งจะนำไปสู่การสรุปผลที่ไม่ถูกต้องในภายหลัง

2. ความเหมาะสมของกลุ่มตัวอย่างที่นำมาทดสอบ

การตรวจวินิจฉัยที่มีประโยชน์ต้องสามารถช่วยแยกแยะผู้ป่วยโรคนั้นๆ จากผู้ที่มีความผิดปกติอื่นที่มีอาการคล้ายคลึงกัน โดยทั่วไปแล้วการตรวจเกือบทุกชนิดมักสามารถแยกแยะคนปกติกับผู้ป่วยที่มีอาการรุนแรงออกจกันได้ แต่คุณสมบัตินี้มีประโยชน์ในทางคลินิกน้อยเนื่องจากการแยกแยะผู้ที่มีความแตกต่างกันสุดโต่งเช่นนี้มักทำได้โดยง่ายโดยไม่ต้องอาศัยการตรวจอื่นๆ เข้าช่วย คุณค่าที่แท้จริงของการตรวจวินิจฉัยใดๆ ก็ตามจะสามารถแสดงออกมาได้ชัดเจนในการศึกษาที่กลุ่มตัวอย่างที่นำมาทดสอบมีความคล้ายคลึงกับกลุ่มผู้ป่วยที่จะมีการใช้การตรวจนั้นเพื่อการวินิจฉัยในทางคลินิกจริงเท่านั้น

กลุ่มตัวอย่างที่เหมาะสมในการศึกษาประสิทธิภาพของการตรวจวินิจฉัยควรประกอบด้วยผู้ป่วยโรคนั้นๆ ที่มีระดับความรุนแรงต่างๆ กันตั้งแต่ที่น้อยที่สุดจนถึงรุนแรงที่สุดและผู้ป่วยโรคอื่นที่มีอาการคล้ายคลึงกันซึ่งมีความจำเป็นต้องการวินิจฉัยแยกโรคจากกันไปในทางคลินิก

ตัวอย่างอิทธิพลของการเลือกใช้กลุ่มตัวอย่างต่อผลการศึกษาประสิทธิภาพของการตรวจวินิจฉัยได้แก่การใช้ carcinoembryonic antigen (CEA) ในการวินิจฉัยผู้ป่วยมะเร็งลำไส้ใหญ่ (colorectal cancer) การศึกษาในผู้ป่วยมะเร็งลำไส้ใหญ่ระยะลุกลาม 36 รายพบว่า มีถึง 35 รายที่มีระดับของ CEA ในเลือดสูง ในขณะที่ในคนปกติและผู้ป่วยด้วยโรคระบบอื่นพบระดับของสารนี้ในเลือดต่ำกว่ามาก ผลของ

งานวิจัยนี้ทำให้มีความหวังว่าการตรวจหาระดับของ CEA น่าจะมีประโยชน์ในการวินิจฉัยโรคมะเร็งลำไส้ใหญ่หรือแม้แต่อาจใช้ในการตรวจคัดกรองผู้ป่วยได้ แต่เมื่อมีการศึกษาซ้ำโดยใช้กลุ่มตัวอย่างที่ใกล้เคียงกับผู้ป่วยที่จะนำการตรวจนี้ไปใช้จริงทางคลินิกคือผู้ป่วยมะเร็งลำไส้ใหญ่ระยะต่างๆ กันตั้งแต่ระยะต้นจนถึงระยะลุกลาม ผู้ป่วยโรคมะเร็งชนิดอื่น และผู้ป่วยโรคระบบทางเดินอาหารอื่นๆ ที่มีอาการคล้ายคลึงกับโรคมะเร็งลำไส้ใหญ่กลับพบว่าประสิทธิภาพของการตรวจนี้ในการวินิจฉัยโรคมะเร็งลำไส้ใหญ่ลดลงมาก จนต้องยกเลิกการใช้การตรวจนี้ในการวินิจฉัยและตรวจคัดกรองโรคมะเร็งลำไส้ใหญ่ ในปัจจุบันการตรวจนี้ใช้เป็นส่วนหนึ่งในการติดตามอาการของผู้ป่วยที่ได้รับการวินิจฉัยว่าเป็นมะเร็งลำไส้ใหญ่แน่นอนแล้วเท่านั้น

3. อิทธิพลของผลการตรวจวินิจฉัยที่ต้องการทดสอบต่อการเลือกส่งผู้ป่วยตรวจด้วยวิธีมาตรฐาน

ถ้าผลของการตรวจที่ต้องการทดสอบมีอิทธิพลต่อการเลือกส่งผู้ป่วยตรวจด้วยวิธีที่ใช้เป็นมาตรฐานอ้างอิงจะส่งผลให้ผลการทดสอบประสิทธิภาพของการตรวจวินิจฉัยนั้นบิดเบือนไปได้มาก สถานการณ์เช่นนี้เรียกว่า “verification bias” หรือ “work-up bias” ตัวอย่างเช่นผู้ป่วยที่สงสัยว่าจะเป็นโรคหลอดเลือดหัวใจที่การตรวจ exercise test ให้ผลบวกลบมีโอกาสถูกส่งตรวจ coronary angiography ซึ่งเป็นการตรวจมาตรฐานอ้างอิงมากกว่าผู้ป่วยที่มีอาการคล้ายคลึงกันแต่ผลการตรวจ exercise test ให้ผลเป็นลบ ทั้งนี้เนื่องจากแพทย์ผู้ดูแลรักษาผู้ป่วยย่อมรู้สึกกังวลที่จะส่งผู้ป่วยที่มีโอกาสเป็นโรคหลอดเลือดหัวใจต่ำไปทำการตรวจที่ยุ่งยากมีค่าใช้จ่ายสูง และเสี่ยงต่อการเกิดภาวะแทรกซ้อน

การศึกษาที่ดีต้องมีการวางแผนล่วงหน้าชัดเจนให้กลุ่มตัวอย่างที่นำมาทดสอบประสิทธิภาพของการตรวจมีโอกาสโดยเท่าเทียมกันที่จะได้รับการตรวจโดยวิธีที่เป็นมาตรฐานอ้างอิงจนถึงที่สุดให้ได้รับการวินิจฉัยขั้นสุดท้ายไม่ว่าผลการตรวจด้วยวิธีที่ต้องการทดสอบในเบื้องต้นให้ผลเป็นบวกหรือลบ ในกรณีที่ผู้ป่วยกลุ่มตัวอย่างไม่ได้รับการตรวจด้วยวิธีการตรวจที่ถือเป็นมาตรฐานอ้างอิงทุกรายผู้อ่านต้องพิจารณาว่าสัดส่วนของผู้ที่ไม่ได้รับการตรวจต่อมีมากจนยอมรับไม่ได้หรือไม่ และกลุ่มที่ได้รับการตรวจต่อมีความแตกต่างจากกลุ่มที่ไม่ได้รับการตรวจต่อหรือไม่ อย่างไร

4. การบันทึกรายละเอียดวิธีการตรวจวินิจฉัยโดยเครื่องมือที่ต้องการทดสอบ

หากผู้ทำการวิจัยสรุปผลการศึกษาว่าการตรวจวินิจฉัยนั้นมีประโยชน์และสมควรมีการนำไปใช้ในเวชปฏิบัติ ก็ย่อมสมควรให้รายละเอียดของการตรวจนั้นเพียงพอที่ผู้อ่าน

จะสามารถนำการตรวจนั้นไปใช้ได้จริง โดยควรมีรายละเอียดครอบคลุมสิ่งที่มีความสำคัญทั้งหมด ทั้งการเตรียมผู้ป่วยที่ถูกต้อง เทคนิควิธีการตรวจ ข้อควรระวัง ภาวะแทรกซ้อน การวิเคราะห์และแปลผลการตรวจ เป็นต้น

การพิจารณาผลของงานวิจัย และการนำไปใช้ประโยชน์

หลังจากที่ได้พิจารณาแล้วว่างานวิจัยมีความเที่ยงตรงน่าเชื่อถือผู้อ่านจึงค่อยทำการพิจารณาผลของงานวิจัย ในงานวิจัยเกี่ยวกับการตรวจวินิจฉัยสิ่งที่ควรพิจารณาคือผู้ทำการวิจัยได้ให้ข้อมูลที่เพียงพอจะนำมาคำนวณค่า Likelihood Ratios ของการตรวจวินิจฉัยไว้หรือไม่

ผลการศึกษาเกี่ยวกับประสิทธิภาพของการตรวจวินิจฉัยต้องสามารถจัดวางตัวเลขเข้าในตาราง 2 X 2 ดังภาพได้ทั้งนี้เพื่อใช้ในการคำนวณค่าความไว (Sensitivity) ความจำเพาะ (Specificity) ความถูกต้อง (accuracy) , predictive values และ Likelihood Ratios

		การวินิจฉัยจากการตรวจมาตรฐาน	
		เป็นโรค	ไม่เป็น
ผลของการตรวจ	ผลบวก	a	b
ที่ต้องการทดสอบ	ผลลบ	c	d

จากตาราง 2 X 2 จะสามารถคำนวณค่าต่างๆ ได้จากสูตรต่อไปนี้

- Sensitivity = $a / a + c$
- Specificity = $d / b + d$
- Prevalence = $[a + c] / [a + b + c + d]$
- Accuracy = $[a + d] / [a + b + c + d]$
- + Predictive value = $a / a + b$
- Predictive value = $d / c + d$
- Likelihood Ratios + = $[a / a + c] / [b / b + d]$
- Likelihood Ratios - = $[c / a + c] / [d / b + d]$

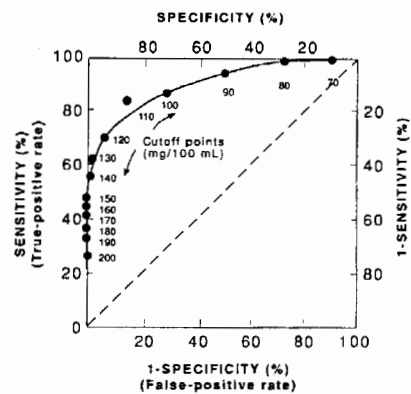
จากสูตรที่เห็นข้างต้นจะพบว่าค่า Sensitivity ของการตรวจคือสัดส่วนของผู้ป่วยที่ให้ผลการทดสอบเป็นบวกต่อผู้ป่วยทั้งหมด ในทางปฏิบัติควรเลือกใช้การตรวจที่มีค่า Sensitivity สูงๆ ในการตรวจกรองผู้ป่วยสำหรับโรคที่มีความรุนแรงมากแต่สามารถรักษาได้ ซึ่งหากผู้ป่วยไม่ได้รับการวินิจฉัยทั้งที่เป็นโรคจะทำให้ผู้ป่วยเสียประโยชน์มาก นอกจากนี้ยังเหมาะจะใช้เป็นการตรวจเบื้องต้นเพื่อลดจำนวนผู้ป่วยที่จะต้องทำการตรวจที่จำเพาะขึ้นเพื่อการ

วินิจฉัยต่อไป ผลการตรวจด้วยวิธีการที่มีค่า Sensitivity สูง จะมีความหมายมากในกรณีที่ผลการทดสอบออกมาเป็นลบ เนื่องจากหมายความว่าผู้ป่วยรายนั้นมีโอกาสป่วยด้วยโรคที่วินิจฉัยด้วยการตรวจนั้นน้อยมาก (มีค่า - predictive value สูง)

การตรวจที่มีค่า Specificity สูง หมายความว่าผู้ป่วยที่ให้ผลการตรวจเป็นบวกมีโอกาสที่จะเป็นโรคจริงสูง ดังนั้นจึงมีประโยชน์ในการยืนยันการวินิจฉัยในกรณีที่มิข้อมูลจากการตรวจอื่นชี้แนะมาบ้างแล้วว่าผู้ป่วยน่าจะป่วยด้วยโรคนั้น การตรวจที่มีคุณสมบัตินี้มีประโยชน์มากในกรณีที่ผลบวกลงจะก่อให้เกิดผลเสียแก่ผู้ป่วยอย่างมากทั้งด้านจิตใจ และการรักษาเสี่ยงต่ออันตรายเช่นกรณีของโรคมะเร็ง หรือทางด้านค่าใช้จ่ายในการรักษา เป็นต้น ดังนั้นการตรวจที่มี Specificity สูงจึงมีประโยชน์มากในกรณีที่ผลการทดสอบเป็นบวก

โดยทั่วไปเรามักจะคาดหวังให้วิธีการตรวจวินิจฉัยที่พัฒนาขึ้นมีทั้ง Sensitivity และ Specificity สูงสุดแต่มักเป็นไปได้ไม่ได้ เมื่อเพิ่ม Sensitivity ให้สูงขึ้น การตรวจนั้นมักมี Specificity ลดลง ในทางตรงกันข้ามการตรวจที่มี Specificity สูงขึ้นมักมี Sensitivity ต่ำ ตำแหน่งจุดตัด (cut - off point) ที่เหมาะสม ในการใช้แยกแยะระหว่างภาวะปกติกับความผิดปกติในกรณีที่การตรวจให้ผลเป็นข้อมูลต่อเนื่อง (continuous data) ขึ้นอยู่กับความเหมาะสมว่าต้องการให้การตรวจนั้นมีความไวสูงหรือมีความจำเพาะสูง

อีกวิธีการหนึ่งที่สามารถนำมาใช้เพื่อเลือกจุดตัดที่เหมาะสมได้ คือการสร้าง Receiver Operator Characteristic (ROC) curve คือสร้างกราฟความสัมพันธ์ระหว่าง true positive rate (Sensitivity) กับ false positive rate (1 - Specificity) โดยการแปรค่าจุดตัด (cut - off point) ที่ใช้ต่างๆ กัน ดังที่เห็นในภาพที่ 1



ภาพที่ 1 ตัวอย่าง ROC curve ของการใช้ 2-hr postprandial blood sugar เป็นการตรวจวินิจฉัยสำหรับโรคเบาหวาน

นอกจากนี้การสร้าง ROC curve ยังช่วยในการเปรียบเทียบประสิทธิภาพของการตรวจวินิจฉัยได้ด้วยโดยเปรียบเทียบพื้นที่ใต้เส้นโค้งของการตรวจแต่ละชนิด พื้นที่ใต้โค้งที่มากกว่าแสดงถึงประสิทธิภาพที่สูงกว่า

การแปลผลค่า predictive value

positive predictive value (PPV) มีความหมายถึงความน่าจะเป็นที่ผู้ป่วยจะเป็นโรคนั้นจริงเมื่อการตรวจให้ผลบวก การตรวจที่มีความจำเพาะสูงมักจะมีค่า PPV สูง

negative predictive value (NPV) หมายถึงความน่าจะเป็นที่ผู้ป่วยไม่ได้เป็นโรคนั้นเมื่อการตรวจให้ผลเป็นลบ การตรวจที่มีความไวสูงมักจะมีค่า NPV สูง

ในขณะที่ค่า Sensitivity และ Specificity ของการตรวจใดๆ มีค่าคงที่ ค่า predictive value กลับขึ้นอยู่กับความชุกของโรคนั้นในประชากรที่มีการนำการตรวจวินิจฉัยนั้นๆ ไปใช้ด้วย ดังนั้นแม้ว่าในงานวิจัยส่วนใหญ่มักคำนวณค่า predictive value ไว้ให้ด้วยผู้อ่านพึงระลึกไว้ว่าค่าเหล่านั้นมักนำไปใช้ในเวชปฏิบัติไม่ได้เนื่องจากความชุกของโรคนั้นในประชากรทั่วไปมักมีความแตกต่างจากกลุ่มประชากรที่นำมาใช้ในกลุ่มตัวอย่างในงานวิจัย และนอกจากนั้นความชุกของโรคยังอาจผันแปรตามกลุ่มอายุ เพศ ถิ่นที่อยู่ และปัจจัยอื่นๆ อีกมากอย่างไรก็ดีหากผู้อ่านทราบความชุกของโรคนั้นในประชากรและทราบค่า Sensitivity และ Specificity ของการตรวจก็ย่อมสามารถคำนวณหาค่า predictive value ได้ไม่ยากโดยอาศัยสมการ

$$PPV = (\text{sensitivity} \times \text{prevalence}) / (\text{sensitivity} \times \text{prevalence}) + (1 - \text{specificity}) \times (1 - \text{prevalence})$$

Likelihood Ratios : ความหมายและการนำไปใช้

Likelihood ratios เป็นวิธีการแสดงประสิทธิภาพของการตรวจวินิจฉัยอีกวิธีหนึ่งนอกเหนือจากค่าความไว ความจำเพาะ และค่า predictive value โดยมีความหมายถึงอัตราส่วนของความน่าจะเป็นของผลการตรวจในผู้ที่เป็นโรค เปรียบเทียบกับผู้ที่ไม่เป็นโรค

ในการใช้ค่า Likelihood Ratios จำเป็นต้องเข้าใจถึงหลักการของ odds เนื่องจากค่า sensitivity , specificity และ predictive values วางอยู่บนพื้นฐานของหลักการของความน่าจะเป็น (probability) แต่ odds เป็นการแสดงข้อมูลเดียวกันโดยวิธีการต่างกัน โดย odds เป็นอัตราส่วนของ probability สองค่า เราสามารถแปลงค่า ทั้งสองกลับไปกลับมาได้โดยอาศัยสูตรง่าย ๆ ดังนี้

$$\text{Odds} = \text{Probability of event} / [1 - \text{Probability of event}]$$

$$\text{Probability} = \text{Odds} / [1 + \text{Odds}]$$

ตัวอย่างเช่นเราอาจพูดว่าทีมฟุตบอลไทยมีโอกาสชนะการแข่งขันในคืนนี้ ร้อยละ 80 หรือ Odds ในการที่ทีมฟุตบอลไทยจะชนะในคืนนี้เท่ากับ 4 : 1 ก็ได้

ค่า Prevalence อาจเรียกอีกอย่างหนึ่งว่า Pretest probability ค่านี้สามารถนำมาแปลงเป็นค่า Pretest odds ได้ และจากนั้นก็สมารถนำไปหาค่า Posttest odds ได้โดยอาศัยสูตร

$$\text{Posttest odds} = \text{Pretest odds} \times \text{Likelihood ratio}$$

หลังจากนั้นก็สมารถนำเอาค่าที่ได้มาแปลงเป็น Posttest probability เพื่อบอกความน่าจะเป็นที่ผู้ป่วยรายนั้นจะเป็นโรคที่ทำการตรวจวินิจฉัยได้

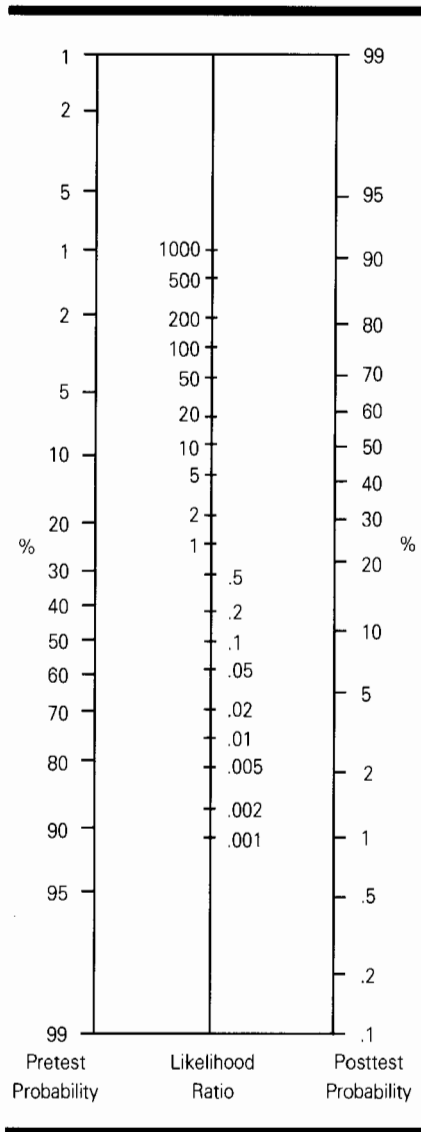
การคำนวณหาค่า posttest probability ตามวิธีที่แสดงไว้ข้างบนเป็นเรื่องค่อนข้างยุ่งยาก โชคดีที่ Fagan ได้ทำให้การแปลง pretest probability ไปสู่ posttest probability ทำได้ง่ายขึ้นมากโดยการจัดสร้าง nomogram ขึ้นดังภาพที่ 2 การใช้งาน nomogram นี้ทำได้โดยการใช้ไม้บรรทัดทาบลงไปที่ตัวเลขของ pretest probability ที่ต้องการโดยให้พาดผ่าน Likelihood Ratio ของการตรวจวินิจฉัยที่ใช้ด้วย ปลายอีกด้านหนึ่งของไม้บรรทัดจะพาดผ่านค่า posttest probability ให้อ่านค่าได้พอดี

จะเห็นว่าค่า Likelihood Ratios ของการตรวจวินิจฉัยสามารถให้ข้อมูลเกี่ยวกับการตรวจนั้นๆ ได้มากกว่า sensitivity, specificity และ predictive value เนื่องจากเราสามารถหาค่า Likelihood Ratios ได้ที่ cut-off point หลายๆ ระดับ ทำให้การแปลผลการตรวจวินิจฉัยนั้นทำได้ละเอียดขึ้น

Likelihood Ratios ที่มีค่ามากกว่า 10 หรือต่ำกว่า 0.1 มักทำให้เกิดการเปลี่ยนแปลงของค่า posttest probability ไปจากค่า pretest probability เป็นอย่างมาก ถ้าค่านี้อยู่ระหว่าง 5-10 หรือ 0.1-0.2 จะทำให้เกิดการเปลี่ยนแปลงปานกลาง ในขณะที่ถ้าอยู่ระหว่าง 2-5 หรือ 0.5-0.2 การเปลี่ยนแปลงจะเกิดขึ้นเพียงเล็กน้อย (แต่อาจมีความสำคัญได้ในบางภาวะ)

การตรวจที่ให้ค่า Likelihood Ratios ระหว่าง 1-2 หรือ 0.5-1 มักไม่ค่อยมีประโยชน์ในการวินิจฉัยทั้งนี้เนื่องจากทำให้ค่า Posttest probability แทบไม่แตกต่างจากค่า Pretest probability

โดยทั่วไปแพทย์มักต้องอาศัยประสบการณ์ที่มีในการคาดเดา Pretest probability ในผู้ป่วยแต่ละรายโดยตั้งอยู่บนข้อมูลเกี่ยวกับ เพศ อายุ อาการและการตรวจพบจากการตรวจร่างกาย และข้อมูลอื่นๆ ที่เป็นความรู้พื้นฐานเดิม ทั้งนี้เนื่องจากการเก็บข้อมูลจริงจากฐานประชากรขนาดใหญ่มี



Nonogram for interpreting diagnostic test results. Adapted from Fagan.⁹

ภาพที่ 2 nomogram ที่ใช้ในการหาค่า posttest Probability พัฒนาขึ้นโดย Fagan และคณะ

อยู่น้อย มีเพียงบางโรคเช่นโรคหลอดเลือดหัวใจที่มีการศึกษาจากฐานข้อมูลขนาดใหญ่ในระดับชาติของสหรัฐอเมริกาให้อ้างอิงได้ อย่างไรก็ตามการนำเอาข้อมูลนั้นมาใช้จะต้องคำนึงถึงความแตกต่างทางด้านเชื้อชาติและวิถีชีวิต จะมีผลทำให้ค่าเหล่านั้นมีความแตกต่างกันหรือไม่ อย่างไรด้วย

ช่วงเชื่อมั่น (Confidence interval)

การแสดงผลของการตรวจวินิจฉัยก็เป็นเช่นเดียวกับการแสดงผลในงานวิจัยชนิดอื่นๆ ค่าที่ได้คือค่าโดยการประมาณการจากการศึกษาในผู้ป่วยเพียงจำนวนหนึ่ง ดังนั้นย่อมมีความแปรปรวนของค่าต่างๆ เกิดขึ้นได้ ยิ่งขนาดของกลุ่มตัวอย่างที่ใช้มีขนาดเล็กเท่าใดอิทธิพลของความแปรปรวนนี้ต่อผลที่ได้ยิ่งมาก ดังนั้นการแสดงผลช่วงเชื่อมั่น 95% (95% confidence intervals) ไว้ในการแสดงผลจะช่วยให้อ่านทราบขอบเขตของผลการศึกษามากขึ้นว่ามีความแม่นยำมากน้อยเพียงใด ในกรณีที่ผู้วิจัยไม่ได้แสดงค่า 95% CI เอาไว้ในผลการศึกษาผู้อ่านสามารถประมาณค่าได้เองโดยอาศัย สูตร

$$95\% \text{ CI} = P \pm 2 \sqrt{\frac{P(1-P)}{N}}$$

P = observed proportion ในที่นี้คือค่าประมาณการที่ต้องการหาค่า 95% CI

N = number of people observed หรือ จำนวนของตัวอย่างที่ทำการศึกษา

การตัดสินใจนำการตรวจวินิจฉัยไปใช้ในผู้ป่วยของผู้อ่าน

ในการตัดสินใจว่าจะส่งผู้ป่วยไปทำการตรวจเพื่อการวินิจฉัยใดหรือไม่ปัจจัยหลายประการเข้ามาเกี่ยวข้องมากสิ่งที่แพทย์ควรคำนึงถึงเมื่อต้องตัดสินใจส่งผู้ป่วยตรวจ นอกเหนือจากประสิทธิภาพในการวินิจฉัยโรคของการตรวจนั้นยังมีดังต่อไปนี้

1. การตรวจนั้นสามารถทำได้ในสถานนั้นและการแปลผลการตรวจแม่นยำเป็นที่น่าพอใจหรือไม่

ในบางกรณีที่มีปัญหาทางด้านเทคนิคในการตรวจมากจนทำให้การทำการตรวจวินิจฉัยนั้นซ้ำให้เหมือนเดิมได้ยาก เช่น ความแปรปรวนของน้ำยาที่ใช้ในการตรวจหาระดับฮอโมนด้วยวิธี radioimmunoassay หรือการตรวจนั้นต้องมีการแปลผลโดยบุคลากรที่มีความเชี่ยวชาญมากแต่สถานของผู้อ่านไม่มีผู้ที่มีความชำนาญในเรื่องนั้น แพทย์ผู้ทำการรักษาผู้ป่วยย่อมน่าจะตัดสินใจเลือกการตรวจวินิจฉัยอื่นมาใช้แทน

2. ผลของการศึกษาเกี่ยวกับการตรวจวินิจฉัยนั้นสามารถใช้ได้กับผู้ป่วยของผู้่านหรือไม่

ข้อนี้สามารถพิจารณาได้โดยเปรียบเทียบลักษณะของผู้ป่วยของผู้่านว่าคล้ายคลึงกับกลุ่มตัวอย่างของการศึกษานั้นหรือไม่ และผู้ป่วยของผู้่านอยู่ในกลุ่มย่อยที่ค่า Likelihood Ratios ของการตรวจวินิจฉัยนั้นมีค่าระหว่าง 0.5-2 หรือไม่ ในกรณีที่ผู้ป่วยของผู้่านมีความแตกต่างจากกลุ่มตัวอย่างมากหรือค่า Likelihood Ratios ของการตรวจวินิจฉัยในกลุ่มตัวอย่างที่มีลักษณะคล้ายคลึงกับผู้ป่วยของผู้่านมีค่าอยู่ระหว่าง 0.5-2 การตรวจนั้นย่อมไม่มีประโยชน์สำหรับผู้ป่วยของผู้่าน

3. ผลของการตรวจที่ได้จะเปลี่ยนแปลงการรักษาของผู้่านหรือไม่

การตรวจวินิจฉัยใดๆ ก็ตามจะมีประโยชน์ในการวินิจฉัยผู้ป่วยก็ต่อเมื่อความน่าจะเป็นที่ผู้ป่วยจะเป็นโรคนั้นๆ อยู่ในระดับกลางๆ เท่านั้น ในกรณีที่ผู้รักษาค่อนข้างแน่ใจในการวินิจฉัยการส่งตรวจย่อมมีประโยชน์ต่อผู้ป่วยน้อย ดังนั้นหากผู้อ่านพบว่าผลของการตรวจนั้นจะไม่มีอิทธิพลต่อการตัดสินใจรักษาในผู้ป่วยรายนั้นก็ย่อมไม่มีความจำเป็นใดๆ ที่จะต้องส่งตรวจ

4. ผู้ป่วยจะได้รับประโยชน์จากผลการตรวจนั้นหรือไม่

จุดมุ่งหมายสูงสุดของการตรวจใดๆ ก็ตามย่อมอยู่ที่การได้ข้อมูลเพิ่มเติมซึ่งเป็นประโยชน์ต่อการดูแลรักษาผู้ป่วยและนำไปสู่การเปลี่ยนแปลงการรักษาแต่ทั้งนี้ต้องเอื้อประโยชน์ต่อผู้ป่วย การตรวจวินิจฉัยที่แม่นยำจะมีประโยชน์สูงในกรณีที่โรคหรือภาวะที่ต้องการวินิจฉัยจะก่อให้เกิด

อันตรายถ้าถูกละเลย การตรวจมีอันตรายน้อยยอมรับได้ และโรคหรือภาวะนั้นมีวิธีการรักษาที่มีประสิทธิภาพ ในบางภาวะที่ไม่มีการรักษาที่มีประสิทธิภาพในขณะนั้นการส่งตรวจเพียงเพื่อต้องการทราบผลเป็นสิ่งที่ไม่จำเป็นและไม่มีความประโยชน์ต่อผู้ป่วย

ด้วยหลักในการพิจารณาและนำผลการศึกษาที่เกี่ยวกับการตรวจวินิจฉัยไปใช้ที่ได้กล่าวมาแล้วผู้อ่านสามารถนำไปใช้ประโยชน์ทั้งในการประเมินคุณค่าของการตรวจวินิจฉัยและนำไปใช้ในเวชปฏิบัติได้ด้วย ในปัจจุบันค่า Likelihood Ratios ได้ทวีความสำคัญมากขึ้นเรื่อยๆ อันจะเห็นได้จากการที่มีรายงานวิจัยเกี่ยวกับการตรวจวินิจฉัยรายงานผลในรูปแบบขึ้นทุกที่ การเรียนรู้เกี่ยวกับเรื่องนี้จึงเป็นสิ่งจำเป็นสำหรับติดตามความก้าวหน้าทางด้าน การตรวจวินิจฉัยต่อไป

Recommended Readings

1. Jaeschke R, Guyatt G, Sackett D. Users' Guides to the Medical Literature III. How to Use an Article About a Diagnostic Test : A. Are the Results of the Study Valid? JAMA 1994 ; 271(5) : 389-91.
2. Jaeschke R, Guyatt G, Sackett D. Users' Guides to the Medical Literature III. How to Use an Article About a Diagnostic Test : B. What Are the Results and Will They Help Me in Caring for My Patients ? JAMA 1994; 271(9) : 703-7.
3. Fletcher R, Fletcher S, Wagner E. Clinical Epidemiology the Essentials. 3 rd eds. Baltimore: Williams & Wilkins, 1996.

