

การประเมินความแตกต่างของผลจากงานวิจัยในการวิเคราะห์เมตา

เชษฐา งามจรัส

ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

Assessing Heterogeneity Among studies' Results in Meta-analysis

Chetta Ngamjarus

Department of Biostatistics and Demography, Faculty of Public Health, Khon Kaen University, Thailand, 40002.

การพิจารณาความแตกต่างของผลจากงานวิจัยเป็นขั้นตอนหนึ่งที่สำคัญในการวิเคราะห์เมตาเพื่อการตัดสินใจเลือกตัวแบบ (model) ที่เหมาะสมในการวิเคราะห์เพื่อหาผลรวมของค่าสถิติที่ได้จากงานวิจัยที่นำเข้ามาศึกษา ซึ่งวิธีการทางสถิติที่นิยมใช้ ได้แก่ การสำรวจความแตกต่างของผลจากรายงานวิจัยด้วยสายตาโดยใช้ Forest plot Galbraith plot และ L'Abbé plot รวมทั้งการประเมินด้วยวิธีสถิติทดสอบ ได้แก่ สถิติทดสอบ Q และ I^2 อย่างไรก็ตามในทางปฏิบัตินิยมใช้ Forest plot สถิติทดสอบ Q และ I^2 ในการพิจารณาความแตกต่างของผลจากรายงานวิจัยร่วมกับความแตกต่างที่อาจเป็นไปได้ในเชิงเนื้อหา

คำสำคัญ: ความแตกต่างของผลจากงานวิจัย การวิเคราะห์เมตา Forest plot Galbraith plot L'Abbé plot สถิติทดสอบ Q I^2

Assessing heterogeneity among studies' results is the most important step in meta-analysis for the purpose of selecting an appropriate model for combining the results to produce an overall statistic from the included studies. The popular statistics methods are the visual investigation of the heterogeneity such as Forest plot, Galbraith plot and L'Abbé plot. Additionally, the Q statistics and I^2 can be used to assess the heterogeneity. In practice it is suggested that Forest plot, Q statistics and I^2 should be used together with the possibility of content heterogeneity in order to assess the heterogeneity between studies' results.

Keywords: heterogeneity, meta-analysis, Forest plot, Galbraith plot, L'Abbé plot, Q statistics, I^2

สรินกรินทร์เวชสาร 2552; 24(2): 142-7 • Srinagarind Med J 2009; 24(2): 142-7

บทนำ

ปัจจุบันการวิเคราะห์เมตา (meta-analysis) เป็นวิธีการทางสถิติที่มีบทบาทสำคัญในการทำวิจัยที่เป็นการทบทวนวรรณกรรมอย่างเป็นระบบ (systematic review) ซึ่งเป็นรูปแบบการวิจัยที่สามารถให้ข้อมูลเชิงประจักษ์ทางด้านวิทยาศาสตร์การแพทย์ที่มีความน่าเชื่อถือมากที่สุด เนื่องจากงานวิจัยดังกล่าวเป็นการหาข้อสรุปหรือการสังเคราะห์ข้อมูลจากงานวิจัยที่ทำการศึกษาในปัญหาวิจัยคล้ายๆ กัน เพื่อให้ได้ข้อสรุปที่เป็นหลักฐานทางวิชาการซึ่งนำไปสู่การกำหนดนโยบายและการตัดสินใจในการเลือกวิธีการที่เหมาะสมในการรักษาและดูแลผู้ป่วย

ในกระบวนการวิเคราะห์เมตานั้นนักวิจัยต้องมีการประเมินความแตกต่าง (heterogeneity) ของผลจากงานวิจัยที่นำเข้า

มาศึกษา ซึ่งอยู่ในรูปของค่าสถิติ ถ้าผลลัพธ์ (outcome) ของงานวิจัยเป็นข้อมูลแบบทวินาม (binary data) ค่าสถิติจะถูกรายงานออกมาเป็นค่า OR (Odds Ratio) RR (Risk Ratio หรือ Relative Risk) หรือ RD (Risk Difference) เป็นต้น แต่ถ้าผลลัพธ์ของงานวิจัยเป็นข้อมูลแบบต่อเนื่อง (continuous data) ค่าสถิติที่รายงานออกมาส่วนใหญ่เป็นค่าเฉลี่ย (mean) คู่กับส่วนเบี่ยงเบนมาตรฐาน (standard deviation) ค่าเฉลี่ยคู่กับช่วงความเชื่อมั่น (confidence interval) เป็นต้น ผลของการประเมินความแตกต่างของผลจากงานวิจัยที่นำเข้ามาศึกษาจะนำไปใช้เพื่อประกอบการตัดสินใจเลือกตัวแบบทางสถิติในการวิเคราะห์ข้อมูลเพื่อหาค่าผลรวมของค่าสถิติโดยตัวแบบในการวิเคราะห์ข้อมูลจะมีด้วยกัน 2 ตัวแบบ คือ ตัวแบบคงที่ (fixed-effect model) และตัวแบบสุ่ม (random-

effects model) ซึ่งทั้ง 2 ตัวแบบมีการคำนวณความแตกต่างของค่าสถิติจากงานวิจัยต่างกัน ตัวแบบคงที่จะคำนวณความแตกต่างดังกล่าวจากความแปรปรวนภายในงานวิจัย (within study variance) เท่านั้น² ซึ่งเกิดจากความคลาดเคลื่อนจากการสุ่ม (sampling error)³ ในขณะที่ตัวแบบสุ่มมีการคำนวณความแตกต่างดังกล่าวจากความแปรปรวน 2 แห่งคือ ความแปรปรวนภายในงานวิจัยและความแปรปรวนระหว่างงานวิจัย (between studies variance) จากเหตุผลดังกล่าวทำให้นักวิจัยที่ทำการทบทวนวรรณกรรมอย่างเป็นระบบจำเป็นต้องมีความเข้าใจในวิธีการทางสถิติเกี่ยวกับการประเมินความแตกต่างของผลจากงานวิจัย

วัตถุประสงค์ของบทความนี้เพื่อทบทวนวิธีการทางสถิติที่ใช้ในการประเมินความแตกต่างทางสถิติ (statistical heterogeneity) ของผลจากงานวิจัยซึ่งประกอบด้วยแผนภาพต่างๆ และการทดสอบด้วยสถิติทดสอบ

วิธีการสำหรับประเมินความแตกต่างของผลจากงานวิจัย

วิธีการทางสถิติที่ใช้ในการประเมินความแตกต่างของผลจากงานวิจัยประกอบด้วยการใช้แผนภาพและการทดสอบด้วยสถิติทดสอบ

1. การสำรวจด้วยแผนภาพ

การสำรวจความแตกต่างของผลจากงานวิจัยด้วยแผนภาพมีด้วยกันหลายแผนภาพประกอบด้วย Forest plot Galbraith plot และ L'Abbé plot⁴

1.1 Forest plot

โดยทั่วไป Forest plot จะใช้เพื่อนำเสนอผลลัพธ์จากงานวิจัยที่นำเข้ามาศึกษาและผลลัพธ์จากการทำการวิเคราะห์เมตา⁵ รวมทั้งสามารถตรวจสอบความแตกต่างของผลที่ได้จากงานวิจัยที่นำเข้ามาทำการทบทวนวรรณกรรมอย่างเป็นระบบได้ ซึ่ง Forest plot จะมีส่วนประกอบดังนี้ แกนนอนจะแสดงค่าสถิติ เส้นตรงในแกนตั้ง ณ ตำแหน่งที่แสดงว่าผลลัพธ์จากงานวิจัยไม่แตกต่างกันระหว่างกลุ่มทดลอง (treatment group) กับกลุ่มควบคุม (control group) คือ ณ ตำแหน่งที่ค่าสถิติเท่ากับ 1 เมื่อผลลัพธ์จากงานวิจัยเป็นข้อมูลแบบทวินามหรือแกนตั้งจะอยู่ ณ ตำแหน่งที่ค่าสถิติเท่ากับ 0 เมื่อผลลัพธ์จากงานวิจัยเป็นข้อมูลแบบต่อเนื่อง เส้นตรงแนวนอนที่อยู่ภายใน Forest plot แสดงช่วงความเชื่อมั่นของค่าพารามิเตอร์ของผลจากงานวิจัยต่างๆ ที่นำเข้ามาศึกษา กล่องสี่เหลี่ยมที่อยู่ตำแหน่งกึ่งกลางบนช่วงความเชื่อมั่นเป็นค่าสถิติจากงานวิจัย ขนาดของกล่องคือส่วนกลับของความแปรปรวนของค่าสถิติ⁶ ที่บอกถึงน้ำหนักของข้อมูลจากงานวิจัยนั้นๆ สี่เหลี่ยมขนมเปียกปูนคือ ผลรวมของค่าสถิติที่ได้จากการวิเคราะห์เมตา

การใช้ Forest plot ในการตรวจสอบความแตกต่างของผลที่ได้จากงานวิจัยจะสังเกตจากช่วงความเชื่อมั่นจากงานวิจัยทั้งหมด ถ้าช่วงความเชื่อมั่นไปในทิศทางเดียวกันและครอบคลุมเส้นตรงแนวตั้งที่ลากจากสี่เหลี่ยมขนมเปียกปูน อาจกล่าวได้ว่าไม่พบความแตกต่างระหว่างผลจากงานวิจัย แต่หากช่วงความเชื่อมั่นไม่ไปในทิศทางเดียวกันและไม่ครอบคลุมเส้นตรงแนวตั้งที่ลากจากสี่เหลี่ยมขนมเปียกปูน อาจกล่าวได้ว่ามีความแตกต่างระหว่างผลจากงานวิจัย จากข้อมูลในตารางที่ 1 ซึ่งเป็นข้อมูลจากการทบทวนวรรณกรรมอย่างเป็นระบบเกี่ยวกับการเปรียบเทียบประสิทธิผลของการใช้ Topical Fluoride กับการใช้ยาหลอก (placebo) หรือไม่ใช้ยา (no treatment) ผลลัพธ์คือการเกิดฟันผุใหม่ตั้งแต่ 1 ซี่ขึ้นไป⁷ สามารถนำมาสร้าง Forest plot (รูปที่ 1)

จากรูปที่ 1 พบว่าช่วงความเชื่อมั่นของ RR จาก 13 งานวิจัย ส่วนใหญ่ไปในทิศทางเดียวกันคือค่า RR น้อยกว่า 1 และสี่เหลี่ยมขนมเปียกปูนอยู่ ณ ตำแหน่งที่ RR น้อยกว่า 1 แสดงว่า topical fluoride สามารถช่วยป้องกันการเกิดฟันผุใหม่ตั้งแต่ 1 ซี่ขึ้นไปได้ แต่พบว่ามี 3 งานวิจัย คือ study ที่ 2 6 และ 12 ที่ให้ช่วงความเชื่อมั่นไม่ครอบคลุมค่าผลรวมของค่าสถิติที่ได้จากการวิเคราะห์เมตา จึงกล่าวได้ว่าอาจมีความแตกต่างของผลจากงานวิจัย

1.2 Galbraith plot

Galbraith plot สร้างจากการลงจุดระหว่างค่า z (แกน y) กับ 1/se (แกน x) โดย $z = \frac{T_i}{\sqrt{v_i}}$ เมื่อ T_i คือค่าสถิติจากงานวิจัยที่ i v_i คือความแปรปรวนจากงานวิจัยที่ i และ se_i คือความคลาดเคลื่อนมาตรฐาน (standard error) ที่ i ซึ่งเท่ากับ $\sqrt{v_i}$ เส้นตรงที่ลาดเอียงซึ่งอยู่กึ่งกลางจะเป็นเส้นค่าประมาณของผลรวมของผลลัพธ์จากการวิจัย⁶ (pooled treatment effect) ส่วนเส้นตรงลาดเอียงที่อยู่ด้านบนและล่างคือเส้นแสดงขอบเขตบนและขอบเขตล่างของค่า z ที่ความน่าจะเป็นประมาณ 0.95 ถ้ามีจุดกระจายอยู่นอกขอบเขตบนและขอบเขตล่างดังกล่าวแสดงว่ามีความแตกต่างของผลจากงานวิจัย⁶ เมื่อนำข้อมูลจากตารางที่ 1 มาสร้าง Galbraith plot จะได้ผล (รูปที่ 2) และพบว่าอาจมีความแตกต่างของผลจากงานวิจัยเกิดขึ้น เนื่องจากมี 3 งานวิจัย คือ study ที่ 2 6 และ 12 ที่อยู่นอกขอบเขตบนและขอบเขตล่างของค่า z ที่ความน่าจะเป็นประมาณ 0.95 ใน Galbraith plot

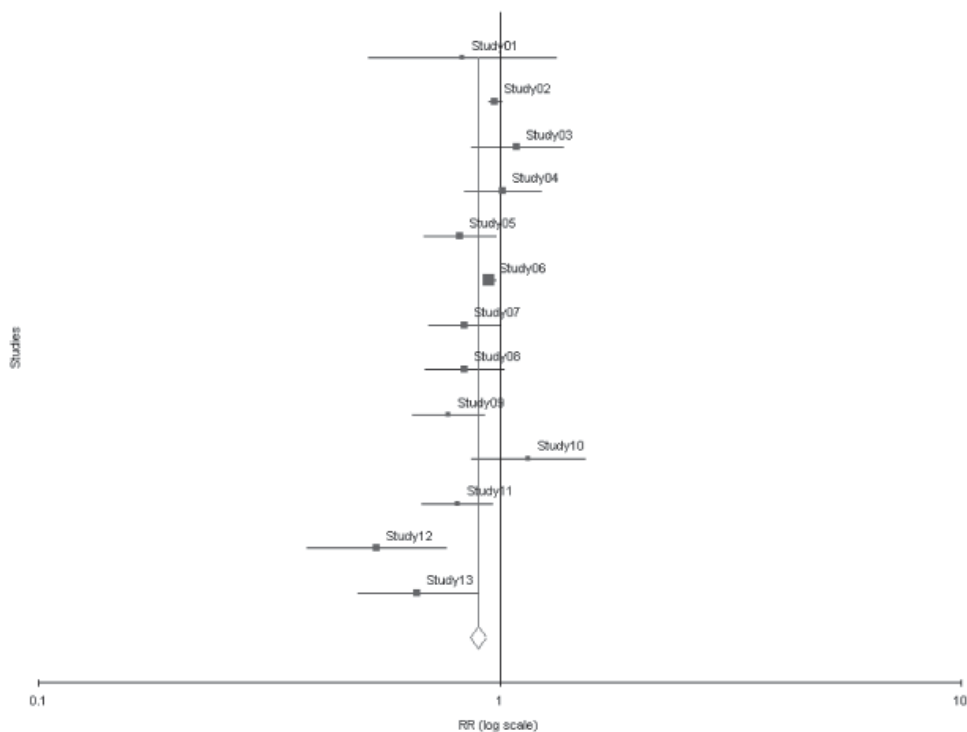
1.3 L'Abbé plot

L'Abbé plot จะใช้สำหรับสำรวจความแตกต่างของผลจากงานวิจัย⁵ และจะใช้กับผลลัพธ์จากงานวิจัยที่เป็นข้อมูลทวินามเท่านั้น โดยสร้างจากการลงจุดระหว่างสัดส่วนของคนที่เกิดเหตุการณ์ที่สนใจในกลุ่มทดลอง (แกน y) กับสัดส่วนของคนที่เกิดเหตุการณ์ที่สนใจในกลุ่มควบคุม

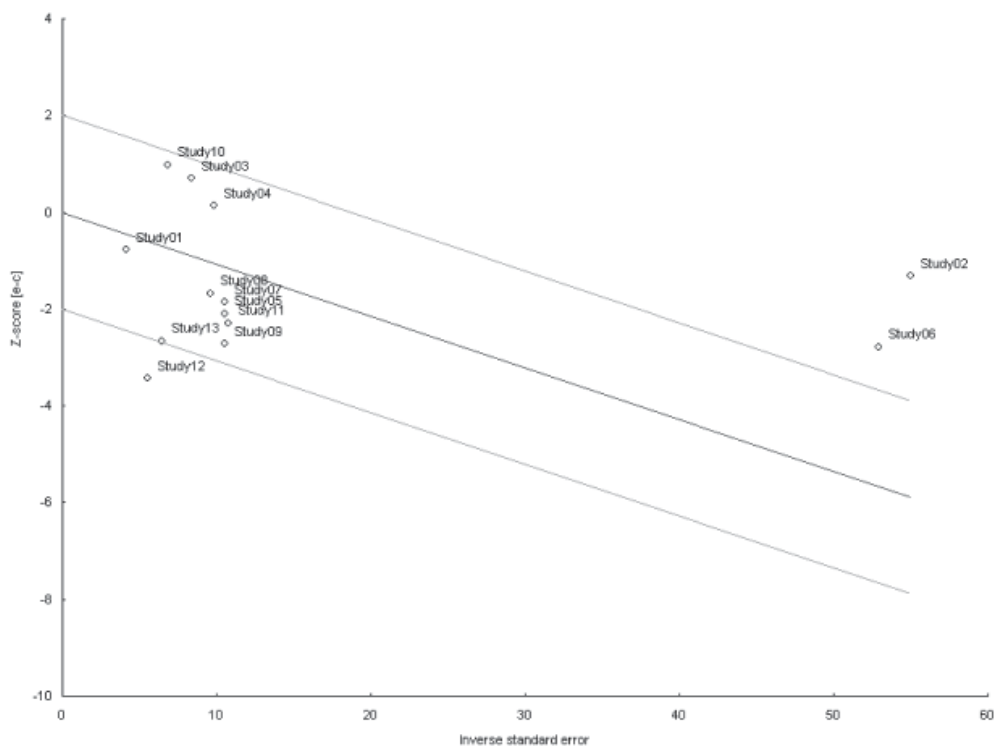
ตารางที่ 1 จำนวนการเกิดฟันผุใหม่ตั้งแต่ 1 ซี่ขึ้นไปในกลุ่มทดลองและกลุ่มควบคุม

งานวิจัย (study)	ชื่อนักวิจัย	ปีที่ตีพิมพ์	กลุ่มทดลอง (ใช้ Topical Fluoride)		กลุ่มควบคุม (ใช้ยาหลอกหรือไม่ใช้ยา)	
			จำนวนคนที่ เกิดฟันผุ	จำนวนตัวอย่าง ทั้งหมดในกลุ่ม ทดลอง	จำนวนคนที่ เกิดฟันผุ	จำนวนตัวอย่าง ทั้งหมดในกลุ่ม ควบคุม
1	Dolles	1980	13	24	15	23
2	Finn	1975	278	292	157	161
3	Forsman	1974	174	414	56	145
4	Forsman	1974a	139	262	69	132
5	Gisselsson	1999	102	182	67	98
6	Hanachowicz	1984	425	473	447	472
7	Heidman	1992	134	426	162	432
8	Holm	1979	64	112	77	113
9	Holm	1984	36	50	55	59
10	Kelber	1996	45	77	40	79
11	Marthaler	1974	37	50	54	59
12	Torell	1965	51	332	46	162
13	Torell	1965a	56	335	84	333

แหล่งที่มาของข้อมูล: Marinho Valeria CC, Higgins Julian PT, Logan S, Sheiham A. Topical fluoride (toothpastes, mouthrinses, gels or varnishes) for preventing dental caries in children and adolescents. Cochrane Database of Systematic Reviews. 2003(4).



รูปที่ 1 Forest plot



รูปที่ 2 Galbraith plot

(แกน x)⁶ จากจุดที่ถูกลงไว้ใน L'Abbé plot จะมีลักษณะเป็นวงกลม ขนาดของวงกลมจะขึ้นอยู่กับความเที่ยง (precision) ของผลลัพธ์ของแต่ละงานวิจัย (รูปที่ 3) เส้นทแยงมุมเส้นยาวเป็นเส้นที่แสดงความเสี่ยงของการเกิดผลลัพธ์ในกลุ่มทดลองเท่ากับความเสี่ยงของการเกิดผลลัพธ์ในกลุ่มควบคุม เส้นทแยงมุมเส้นสั้นในรูปเป็นเส้นของสมการถดถอยแบบไม่ถ่วงน้ำหนัก (unweighted regression) ที่สร้างจากความเสี่ยงดังกล่าวของงานวิจัยที่ศึกษา การพิจารณาความแตกต่างของผลจากงานวิจัยจะดูจากการกระจายตัวของวงกลม หากกระจายใกล้ๆ เส้นทแยงมุมเส้นสั้นแสดงว่าไม่พบความแตกต่างของผลจากงานวิจัย แต่ถ้าวงกลมกระจายไกลจากเส้นทแยงมุมเส้นสั้นก็แสดงว่ามีความแตกต่างของผลจากงานวิจัย จากรูปที่ 3 พบว่ามีวงกลมบางวงกลมที่อยู่ห่างจากเส้นทแยงมุมเส้นสั้นจึงกล่าวได้ว่าอาจมีความแตกต่างของผลจากงานวิจัย

จากการสำรวจความแตกต่างของผลจากงานวิจัยด้วย Forest plot Galbraith plot และ L'Abbé plot เป็นการสำรวจด้วยสายตาซึ่งอาจจะยากต่อการตัดสินใจของนักวิจัยทางเลือกหนึ่งที่ควรพิจารณาคือการใช้สถิติในการทดสอบความแตกต่างของผลจากงานวิจัย ซึ่งจะกล่าวถึงในหัวข้อต่อไป

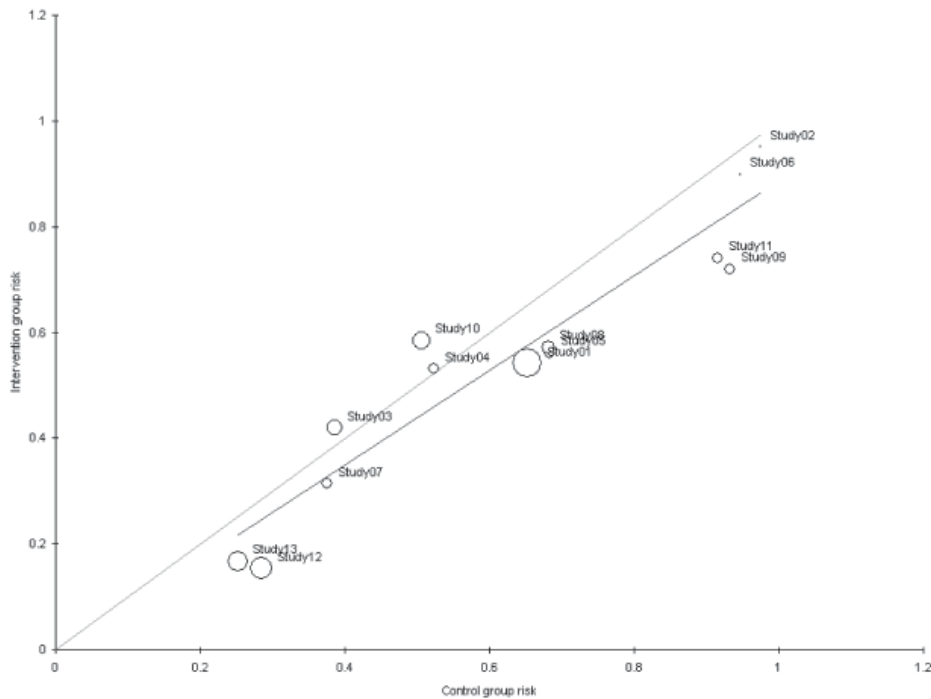
2. สถิติทดสอบสำหรับการประเมินความแตกต่างของผลจากงานวิจัย

สถิติที่ใช้ในการประเมินความแตกต่างของผลจากงานวิจัยมีด้วยกันหลายตัว แต่ที่นิยมใช้กันมากในการทบทวนวรรณกรรมอย่างเป็นระบบคือ สถิติทดสอบ Q และ I^2 ดังนั้นผู้เขียนจะขอกล่าวถึงเฉพาะสถิติสองตัวนี้เท่านั้น ซึ่งมีรายละเอียดดังนี้

2.1 สถิติทดสอบ Q

สถิติทดสอบ Q หรือ Cochran's X^2 ถูกพัฒนาโดย Cochran ในปี ค.ศ. 1954³ มีสูตรดังนี้ $Q = \sum_{i=1}^k w_i (T_i - \bar{T})^2$ เมื่อ i คืองานวิจัยที่ i โดย i จะมีค่าตั้งแต่ 1 ถึง k โดยที่ k คือจำนวนงานวิจัยทั้งหมดที่ทำการศึกษา w_i คือค่าถ่วงน้ำหนักสำหรับงานวิจัยที่ i คำนวณได้จาก $w_i = 1/v_i$ เมื่อ v_i คือความแปรปรวนจากงานวิจัยที่ i T_i คือค่าสถิติจากงานวิจัยที่ i และ \bar{T} คำนวณได้จาก $\bar{T} = \sum_{i=1}^k w_i T_i / \sum_{i=1}^k w_i$ สำหรับสมมติฐานของการทดสอบกำหนดได้ดังนี้ H_0 : ค่าพารามิเตอร์ของผลจากงานวิจัยต่างๆ ไม่มีความแตกต่างกัน H_A : มีค่าพารามิเตอร์ของผลจากงานวิจัยอย่างน้อยหนึ่งงานวิจัยที่แตกต่างจากค่าพารามิเตอร์ของผลจากงานวิจัยอื่น

สถิติทดสอบ Q จะมีการแจกแจงแบบกำลังสอง (Chi-square distribution) ด้วยองศาเสรี (degree of freedom:



รูปที่ 3 L'Abbé plot

df เท่ากับ $k-1$ ในการทดสอบสมมติฐานดังกล่าวนิยมกำหนดระดับนัยสำคัญ (level of significance: α) เท่ากับ 0.1 เนื่องจากสถิติทดสอบ Q มีอำนาจในการทดสอบความแตกต่างของผลจากงานวิจัยต่ำ โดยเฉพาะเมื่อมีจำนวนรายงานการวิจัยน้อย³ หากผลการทดสอบปฏิเสธสมมติฐาน H_0 แสดงว่าพบนัยสำคัญทางสถิติของความแตกต่างของผลจากงานวิจัย นักวิจัยควรพยายามค้นหาสาเหตุของความแตกต่างดังกล่าวในกรณีที่ไม่สามารถสรุปสาเหตุของความแตกต่างดังกล่าวได้ อาจใช้ตัวแบบสุ่มในการวิเคราะห์ข้อมูลเพื่อหาค่าผลรวมของค่าสถิติจากงานวิจัยที่นำเข้ามาศึกษา ในทางกลับกัน หากผลการทดสอบไม่ปฏิเสธสมมติฐาน H_0 แสดงว่าความแตกต่างของผลจากงานวิจัยไม่มีนัยสำคัญทางสถิติ นักวิจัยอาจใช้ตัวแบบคงที่ในการวิเคราะห์ข้อมูลเพื่อหาค่าผลรวมของค่าสถิติจากงานวิจัยที่นำเข้ามาศึกษา

เมื่อนำข้อมูลจากตารางที่ 1 มาวิเคราะห์ด้วยโปรแกรม Mix^{9, 10} จะพบว่าม้งานวิจัยที่นำเข้ามาทำการวิเคราะห์เมตาทั้งหมด 13 งานวิจัย มีจำนวนผู้เข้าร่วมในงานวิจัยดังกล่าวทั้งหมด 5,297 คน เมื่อทำการวิเคราะห์ด้วยตัวแบบคงที่ได้ค่าผลรวมของ RR เท่ากับ 0.898 ด้วยช่วงความเชื่อมั่นร้อยละ 95 อยู่ระหว่าง 0.865 ถึง 0.933 แต่เมื่อทำการทดสอบความแตกต่างของผลจากงานวิจัยด้วยสถิติทดสอบ Q พบว่ามีความแตกต่างของผลจากงานวิจัย อย่างมีนัยสำคัญทางสถิติ ($p\text{-value} < 0.0001$) ดังนั้นนักวิจัยควรพยายามค้นหาสาเหตุของ

ความแตกต่างดังกล่าว หากไม่สามารถหาข้อสรุปได้จึงเลือกใช้ตัวแบบสุ่มในการวิเคราะห์ข้อมูลเพื่อหาค่าผลรวมของค่าสถิติจากงานวิจัย

ตารางที่ 2 การแปลผลของ I^2

ค่า I^2	การแปลผล
0% ถึง 40%	ความแตกต่างของผลจากงานวิจัยที่พบอาจไม่มี ความสำคัญ
30% ถึง 60%	อาจมีความแตกต่างของผลจากงานวิจัยใน ระดับปานกลาง
50% ถึง 90%	อาจมีความแตกต่างของผลจากงานวิจัยใน ระดับมาก
75% ถึง 100%	พิจารณาได้ว่ามีความแตกต่างของผลจาก งานวิจัยอย่างชัดเจน

2.2 สถิติ I^2

สถิติ I^2 ถูกนำเสนอโดย Higgins และ Thompson ในปี ค.ศ. 2002¹¹ ซึ่งมีสูตรในการคำนวณดังนี้ $I^2 = 100\% \times (Q - df)/Q$ เมื่อ Q คือสถิติทดสอบ Q และ df คือองศาเสรี เท่ากับ $k-1$ เมื่อ k คือจำนวนงานวิจัย ค่า I^2 ใช้อธิบายร้อยละของความผันแปรของค่าสถิติที่ประมาณได้ซึ่งเกิดจากความแตกต่างของผลจากงานวิจัยมากกว่าที่จะเกิดขึ้นโดยบังเอิญ¹²

ถ้า I^2 ที่คำนวณได้มีค่าลบจะถูกกำหนดให้มีค่าเป็นศูนย์แทน ดังนั้น I^2 จะมีค่าตั้งแต่ 0 ถึง 100 โดย $I^2 = 0$ หมายถึงไม่มีความแตกต่างของผลจากงานวิจัย หาก I^2 มีค่ามากแสดงว่าความแตกต่างของผลจากงานวิจัยจะมากด้วย คู่มือการทำ การทบทวนวรรณกรรมอย่างเป็นระบบของงานวิจัยเชิงทดลอง เวอร์ชัน 5.0.1 และเป็นเวอร์ชันล่าสุดที่จัดทำโดย Cochrane Collaboration ซึ่งเป็นองค์กรที่ไม่หวังผลกำไรและให้การสนับสนุน ในการจัดทำและเผยแพร่การทบทวนวรรณกรรมอย่างเป็นระบบ รวมทั้งมี Cochrane library ซึ่งเป็นฐานข้อมูลที่ใหญ่ที่สุดในโลก ที่เก็บรวบรวมและเผยแพร่การทบทวนวรรณกรรมอย่างเป็นระบบ ได้แนะนำการแปลผลของ I^2 อย่างคร่าวๆ (ตารางที่ 2)¹²

เนื่องจากค่า I^2 จะขึ้นอยู่กับหลายปัจจัยได้แก่ ขนาด และทิศทางของค่าสถิติจากงานวิจัย รวมทั้งความชัดเจนของหลักฐานที่บ่งบอกถึงความแตกต่างของผลจากงานวิจัย เช่น p-value จากสถิติทดสอบ Q หรือช่วงความเชื่อมั่นของ I^2 ดังนั้นการแปลผลในตารางที่ 2 จึงเป็นเพียงคำแนะนำคร่าวๆ ซึ่งการแปลผลดังกล่าวอาจทำให้เข้าใจผิดได้¹² แนวทางปฏิบัติคือควรใช้ I^2 ควบคู่กับสถิติทดสอบ Q

เมื่อนำข้อมูลจากตารางที่ 1 มาวิเคราะห์ด้วยโปรแกรม Mix^{9,10} ได้ค่า $I^2 = 77.73\%$ กล่าวได้ว่าอาจมีความแตกต่างของผลจากงานวิจัยในระดับมากหรือมีความแตกต่างของผลจากงานวิจัยอย่างชัดเจน (เมื่อพิจารณาตามตารางที่ 2 พบว่าช่วงของค่า I^2 มีบางส่วนที่ซ้อนทับกัน) อย่างไรก็ตามเพื่อความชัดเจนของหลักฐานที่บ่งบอกถึงความแตกต่างของผลจากงานวิจัยควรพิจารณาควบคู่กับสถิติทดสอบ Q ซึ่งผลการทดสอบด้วยสถิติทดสอบ Q พบว่ามีความแตกต่างของผลจากงานวิจัยอย่างมีนัยสำคัญทางสถิติ (p-value < 0.0001) ดังนั้นนักวิจัยอาจเลือกใช้ตัวแบบสุ่มในการวิเคราะห์ ข้อมูลเพื่อหาค่าผลรวมของค่าสถิติจากงานวิจัย ถ้าไม่สามารถอธิบายสาเหตุของความแตกต่างดังกล่าวได้ เมื่อนำข้อมูลจากตารางที่ 1 มาวิเคราะห์เมตาด้วยตัวแบบสุ่มในโปรแกรม Mix^{9,10} ได้ค่าผลรวมของ RR เท่ากับ 0.891 ด้วยช่วงความเชื่อมั่นร้อยละ 95 อยู่ระหว่าง 0.834 ถึง 0.950

สรุป

การพิจารณาความแตกต่างของผลจากงานวิจัยเป็นขั้นตอนที่สำคัญซึ่งนักวิจัยที่ทำการวิจัยแบบการทบทวนวรรณกรรมอย่างเป็นระบบต้องดำเนินการเพื่อประกอบการตัดสินใจดำเนินการขั้นตอนต่อไปของการวิเคราะห์เมตาซึ่งอาจใช้ในการเลือกตัวแบบในการวิเคราะห์เพื่อหาผลรวมของค่าสถิติที่ได้จากงานวิจัยที่นำเข้ามาศึกษา แม้ว่าจะมีหลายวิธีทางสถิติในการพิจารณาความแตกต่างดังกล่าวแต่ยังไม่มีวิธีใดที่จะสามารถให้คำตอบได้ชัดเจน ดังนั้นนักวิจัยควรพิจารณาจาก Forest plot ควบคู่กับสถิติทดสอบ Q I^2 และ

ความแตกต่างที่อาจเป็นไปได้ในเชิงเนื้อหาประกอบการตัดสินใจ

กิตติกรรมประกาศ

ผู้เขียนขอขอบคุณ รศ.ดร.มาลินี เหล่าไพบุลย์ และ อาจารย์พอใจ พิทยานิถยธรรม ที่กรุณาอ่านและให้คำแนะนำที่เป็นประโยชน์ในการเขียนบทความนี้

เอกสารอ้างอิง

1. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ* 2001; 323:334-6.
2. มาลินี เหล่าไพบุลย์. การทบทวนวรรณกรรมอย่างเป็นระบบเชิงปริมาณ: การวิเคราะห์เมตา. *วารสารวิชาการสาธารณสุข* 2544; 10:151-71.
3. Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I^2 Index? *Psychol Methods*. 2006; 11:193-206.
4. Cleophas TJ, Zwinderman AH. *Meta-Analysis*. *Circulation* 2007; 115:2870-5.
5. Xu H, Platt RW, Luo ZC, Wei S, Fraser WD. Exploring heterogeneity in meta-analyses: Needs, resources and challenges. *Paediatr Perinat Epidemiol* 2008; 22(Suppl.1): 18-28.
6. Sutton AJ. *Methods for meta-analysis in medical research*. Chichester: Wiley; 2000.
7. Marinho VC, Higgins JP, Logan S, Sheiham A. Topical fluoride (toothpastes, mouthrinses, gels or varnishes) for preventing dental caries in children and adolescents. *Cochrane Database of Systematic Reviews*. 2003(4).
8. Song F. Exploring heterogeneity in meta-analysis: Is the L'Abbe plot useful? *J Clin Epidemiol* 1999; 52:725-30.
9. Bax L, Yu LM, Ikeda N, Tsuruta H, Moons KG. Development and validation of MIX: comprehensive free software for meta-analysis of causal research data. *BMC Med Res Methodol*. 2006; 6:50.
10. Bax L, Yu LM, Ikeda N, Tsuruta H, Moons KGM. MIX: comprehensive free software for meta-analysis of causal research data. Version 1.7. Available from: <http://mix-for-meta-analysis.info> [cited; February 16, 2009].
11. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-58.
12. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated September 2008]. The Cochrane Collaboration. 2008.

