

State of the Art

บทความทบทวนพิเศษ

การทบทวนวรรณกรรม ระบบเชิงปริมาณ: การวิเคราะห์ เมตา

มาลินี เหล่าไพบูลย์

ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์, มหาวิทยาลัยขอนแก่น, จังหวัดขอนแก่น ๔๐๐๐๒

บทคัดย่อ ปัจจุบันการทบทวนวรรณกรรมอย่างเป็นระบบเป็นวิธีการที่ใช้กันอย่างแพร่หลายในงานวิจัยทางด้านบริการ การดูแลสุขภาพ รายงานวิจัยที่ใช้วิธีการของการทบทวนวรรณกรรมอย่างมีระบบได้ปรากฏเป็นจำนวนมากในวารสารการวิจัยทางด้านบริการการดูแลสุขภาพ บทความนี้ต้องการนำเสนอวิธีการดังกล่าวเชิง ปริมาณที่นิยมเรียกกันว่า “การวิเคราะห์เมตา” วัตถุประสงค์หลักของการนำเสนอครั้งนี้เพื่อเสนอแนวความคิดโดยทั่วไปของการวิเคราะห์เมตา

บทนำ

ปัจจุบันวารสารวิชาการที่เผยแพร่ผลงานวิจัยด้าน วิทยาศาสตร์การแพทย์มีอยู่ทั่วโลกมากกว่า ๕๐,๐๐๐ วารสาร. แต่ละปีวารสารเหล่านี้จะตีพิมพ์ผลงานวิจัย เผยแพร่ออกมาเป็นจำนวนมากมากเกินกว่าล้านเรื่อง

และก่อเหตุที่สำคัญประการหนึ่ง คือมีการศึกษาซ้ำใน เนื้อหาเรื่องเดียวกันบ่อยๆอาทิมีรายงานผลการวิจัยเชิง ทดลองที่ศึกษาประสิทธิภาพของโปรแกรมการป้องกัน การสูบบุหรี่ในหญิงมีครรภ์ที่ลงตีพิมพ์ในวารสารการ แพทย์และสาธารณสุขมากกว่า ๑๐๐ เรื่อง ในช่วงปี ๒๕๓๔-๒๕๔๑ (ข้อมูลดังกล่าวสืบค้นมาจากฐานข้อมูล

Medline). จำนวนผลงานวิจัยที่ปรากฏมากมายนั้นเกินกว่าที่จะควบคุมคุณภาพได้อย่างทั่วถึง, จึงมีรายงานวิจัยเป็นจำนวนมากที่ทำให้คนอ่านเกิดความสับสนในผลการวิจัย ดังเช่นพบว่ามึหลายรายงานวิจัยที่ทำการศึกษาในเรื่องเดียวกัน แต่เสนอผลการวิจัยตรงข้ามกันหรือไม่สอดคล้องกัน. ยิ่งกว่านั้นรายงานวิจัยบางเรื่องเสนอผลการวิจัยจากจำนวนขนาดตัวอย่างที่น้อยมากจนผู้อ่านไม่มั่นใจว่าผลการวิจัยที่น่าเสนอจะเชื่อถือได้.

การทบทวนอย่างมีระบบ (systematic reviews) เป็นการวิจัยซึ่งเป็นทางเลือกใหม่และเป็นที่ยอมรับว่าจะสามารถให้ข้อสรุปที่น่าเชื่อถือและนำไปใช้เป็นหลักฐานประกอบการตัดสินใจเพื่อนำไปสู่การวางแผนรักษาโรค ป้องกันโรค และบริการสาธารณสุขต่างๆ ให้แก่ชุมชนได้ถูกต้องมากกว่าข้อสรุปที่ได้จากแต่ละงานวิจัย.

การทบทวนอย่างมีระบบเป็นวิธีการทบทวนเอกสารผลการวิจัยที่อาศัยหลักการทางวิทยาศาสตร์ มีการวางแผนล่วงหน้าและกำหนดเป้าหมายสำคัญไว้ที่การสังเคราะห์ผลลัพธ์ของหลายๆ รายงานวิจัยโดยใช้กลวิธีต่างๆ เพื่อจำกัดอคติ (bias) และข้อผิดพลาดเชิงสุ่ม (random error). ผลการวิจัยหลายๆ เรื่องที่นำมาทบทวนอย่างเป็นระบบที่ให้ผลสรุปโดยใช้วิธีการทางสถิติในการรวมผลลัพธ์อาจจะเรียกว่า **“การทบทวนอย่างมีระบบเชิงคุณภาพ” (qualitative systematic reviews)**. ส่วนการทบทวนที่ใช้วิธีการทางสถิติในการรวมผลลัพธ์อาจจะเรียกว่า **“การทบทวนอย่างมีระบบเชิงปริมาณ” (quantitative systematic reviews) หรือการวิเคราะห์เมตา (meta-analysis)** ซึ่งเป็นประเด็นเนื้อหาที่จะนำเสนอในหัวข้อต่อไป

ในบทความนี้ จะนำเสนอหลักการเบื้องต้นของการวิเคราะห์เมตา, วิธีการทางสถิติเบื้องต้นของการรวมผลลัพธ์ของการศึกษาเชิงควบคุมแบบสุ่ม (randomized controlled trials), และเสนอข้อสรุปและประเด็นปัญหาของการวิเคราะห์เมตาที่ควรระมัดระวัง, ตามด้วยบทสรุป.

หลักการวิเคราะห์เมตา

(principle and procedure of meta-analysis)

การวิเคราะห์เมตามีลักษณะคล้ายคลึงกับการทำวิจัยทั่วไป นั่นคือ ต้องกำหนดแผนล่วงหน้าไว้อย่างชัดเจน, ซึ่งประกอบด้วยข้อคำถามที่ต้องการหาคำตอบ โดยการวิเคราะห์, เจาะไขในการพิจารณาเลือกรายงานวิจัยที่จะนำมาวิเคราะห์, แหล่งและกลวิธีที่จะใช้ค้นหารายงานวิจัย, วิธีการประเมินและวิเคราะห์รายงานวิจัยที่เลือกเข้ามาศึกษา, และวิธีการทางสถิติที่ใช้ในการรวมผลลัพธ์ของรายงานวิจัย.

๑. การกำหนดเงื่อนไขในการพิจารณาเลือกรายงานวิจัย

โดยปรกติการทำวิจัยไม่ว่าจะมีรูปแบบของการวิจัยเป็นแบบใด จำเป็นจะต้องกำหนดเงื่อนไขหรือคำจำกัดความของหน่วยที่สนใจศึกษาหรือประชากรไว้อย่างชัดเจนเพื่อนักวิจัยจะได้ทำการศึกษาในกลุ่มประชากรที่เหมาะสม และสามารถให้คำตอบของการวิจัยได้อย่างชัดเจน การวิเคราะห์เมตาก็เช่นเดียวกันจะต้องกำหนดเงื่อนไขเอาไว้ล่วงหน้าโดยมีวัตถุประสงค์เพื่อนำไปใช้พิจารณาว่าจะเลือกรายงานวิจัยเรื่องใดมาทบทวนในการวิเคราะห์ ซึ่งจะเห็นว่าหน่วยที่ศึกษาในการวิเคราะห์เมตาคือรายงานวิจัยแต่ละรายงาน. ฉะนั้นเงื่อนไขจะต้องระบุให้ชัดเจนครอบคลุมประเด็นต่างๆ ซึ่งมักเกี่ยวกับลักษณะและรูปแบบของวิธีการวิจัย (characteristic and methodology of research) ของรายงานวิจัยที่สนใจจะทบทวน อันประกอบด้วยรูปแบบการวิจัย (research design), ลักษณะของประชากรภายในรายงานวิจัยที่จะเลือก, ชนิดของสิ่งทดลองที่ศึกษา (type of interventions) และการจัดสิ่งทดลองให้กับหน่วยศึกษา (treatment allocation to the subjects) ของการวิจัยเชิงทดลอง หรือปัจจัยเสี่ยงที่สนใจของการวิจัยเชิงวิเคราะห์, ผลลัพธ์ และการวัดค่าผลลัพธ์ (outcome measurement) ของประสิทธิภาพของสิ่งทดลอง.

ตัวอย่างการวิเคราะห์เมตาของการทดลองทางเวชกรรมเพื่อศึกษาประสิทธิภาพของยาต้านจุลชีพป้องกันการติดเชื้อโรคที่อาจได้รับการผ่าตัด^(๑). ในการศึกษาเรื่องนี้ผู้วิจัยได้กำหนดเงื่อนไขของรายงานวิจัยที่ต้องการทบทวนไว้ดังนี้ ๑) รายงานวิจัยต้องเป็นการทดลองในผู้ป่วยที่ได้รับการผ่าตัด, ๒) การทดลองต้องมีการสุ่มผู้ป่วยที่รับยาต้านจุลชีพ หรือกลุ่มควบคุม, และ ๓) มีการติดเชื้อที่ไขสันหลังเป็นตัววัดประสิทธิภาพของการให้ยาต้านจุลชีพก่อนการผ่าตัด.

๒. กลวิธีในการค้นหางานวิจัย (search strategy for identification of studies)

การค้นหารายงานวิจัยที่ดีจะต้องพยายามสืบค้นจากหลายๆ แหล่ง เพื่อให้ได้รายงานวิจัยครอบคลุม **ทั้งที่ได้รับการตีพิมพ์ และที่ไม่ได้รับการตีพิมพ์**. แหล่งค้นหาของเอกสารที่ได้รับการตีพิมพ์ซึ่งควรพิจารณาประกอบด้วย

- ก) ฐานข้อมูลอิเล็กทรอนิกส์ (electronic database)
- ข) เอกสารอ้างอิง และบรรณานุกรมของรายงานวิจัยที่ค้นได้
- ค) ผู้เชี่ยวชาญและผู้สนใจในเนื้อหาที่ต้องการทบทวน
- ง) การค้นหาด้วยมือ
- จ) บทความจากการประชุมสัมมนา
- ฉ) รายงานของหน่วยงานต่างๆ

สำหรับฐานข้อมูลอิเล็กทรอนิกส์ถือว่าเป็นแหล่งรวบรวมงานวิจัยที่ได้รับการตีพิมพ์ใหญ่ที่สุดแห่งหนึ่ง และฐานข้อมูลที่สำคัญของงานวิจัยทางการแพทย์ในปัจจุบันได้แก่ Medline, Embase, Sciearch และ Cochrane controlled trials register. การจะค้นหาจากฐานข้อมูลดังกล่าวจำเป็นต้องเรียนรู้ถึงการกำหนดดัชนีค้นหาหรือคำสำคัญ (search index or key words) ของแต่ละฐานให้เข้าใจก่อน มิฉะนั้นอาจทำให้ค้นไม่พบซึ่งอาจเป็นผลเนื่องจากการกำหนดดัชนีที่ไม่เหมาะสม. บรรณารักษ์ห้องสมุดเป็นบุคคลที่ควรขอคำแนะนำในการกำหนดดัชนีค้นหาของฐานข้อมูลต่างๆ. อย่างไรก็ตามฐานข้อมูลเหล่านี้มีรายงานวิจัยที่ซ้ำซ้อนกันอยู่

เป็นจำนวนมาก ดังนั้นจะต้องตรวจสอบให้ดี มิฉะนั้นอาจทำให้ได้รายงานวิจัยซ้ำซ้อนกันเป็นผลให้การวิเคราะห์เมตาไม่ถูกต้อง.

แม้ว่าฐานข้อมูลอิเล็กทรอนิกส์จะเป็นแหล่งรวบรวมงานวิจัยที่ได้รับการตีพิมพ์ใหญ่ที่สุดก็ตาม แต่พบว่าครอบคลุมงานวิจัยที่ได้รับการตีพิมพ์เพียงประมาณร้อยละ ๔๐ เท่านั้น. ดังนั้นจึงต้องค้นหาจากแหล่งอื่นๆ อาทิ บทความจากการประชุมสัมมนาต่างๆ การค้นหาด้วยมือ และรายชื่อเอกสารอ้างอิงของรายงานวิจัยที่ค้นได้ เป็นต้น. สำหรับการค้นต่อจากรายชื่อเอกสารอ้างอิงของรายงานวิจัยที่ค้นได้นี้แม้ว่าอาจจะพบรายงานวิจัยเพิ่มเติมอีกจำนวนหนึ่ง แต่ข้อระมัดระวังของการค้นหาจากแหล่งนี้คือ อคติของผลการวิจัยที่มักจะไปในทิศทางเดียวกันซึ่งอาจส่งผลให้เกิดข้อสรุปที่เกินความเป็นจริง.

การสอบถามจากผู้เชี่ยวชาญนับว่าเป็นแหล่งค้นหาสำคัญที่อาจทำให้ได้รายงานวิจัยซึ่งไม่ได้รับการตีพิมพ์. ฉะนั้นจึงควรพยายามสอบถามจากผู้รู้เพื่อให้ได้รายงานวิจัยครอบคลุมมากที่สุด.

การตัดทิ้งรายงานวิจัยในขณะค้นหาควรพิจารณา รายงานวิจัยให้รอบคอบก่อนเพราะโอกาสที่จะกลับไปพิจารณารายงานวิจัยที่ได้ตัดทิ้งไปแล้วมีน้อยมาก. ฉะนั้นเมื่อพบว่ารายงานวิจัยที่กำลังพิจารณานำเข้ามาทบทวนมีข้อมูลไม่เพียงพอต่อการตัดสินใจควรพยายามหาข้อมูลเพิ่มเติมเช่นสอบถามจากผู้รู้หรือติดต่อไปยังผู้เขียนรายงานวิจัยเรื่องนั้นโดยตรง.

๓. การประเมินและรวบรวมรายงานวิจัย

ในขั้นตอนนี้ควรมีนักวิจัยในกลุ่มที่ทำกรวิเคราะห์เมตาอย่างน้อย ๒ ท่าน พิจารณาเลือกรายงานวิจัยที่ค้นได้ตามเงื่อนไขที่กำหนดไว้อย่างอิสระต่อกันจากแหล่งสืบค้นต่างๆ ที่ได้กำหนดไว้ว่าต้องการจะค้นหา รายงานวิจัย. หลังจากที่เลือกรายงานวิจัยเข้ามาได้ นักวิจัยแต่ละท่านจะประเมินรายงานวิจัยเหล่านั้นว่ามีคุณภาพที่ดีและน่าเชื่อถือเพียงพอที่จะนำมาศึกษาในการวิเคราะห์เมตาหรือไม่. ดังนั้นจึงควรพิจารณาความ

สอดคล้องของการประเมินรายงานวิจัยของบุคคลดังกล่าวก่อนรวบรวมข้อมูลจริง.

การประเมินรายงานวิจัยมีวัตถุประสงค์เพื่อพิจารณาความถูกต้องของผลการวิจัยและค้นหาสาเหตุของความแตกต่างของผลการวิจัยที่ปรากฏ. ในความเป็นจริงจะไม่สามารถทราบว่าการวิจัยนั้นๆ ถูกต้องหรือไม่ เนื่องจากงานวิจัยส่วนใหญ่เป็นการศึกษาจากกลุ่มตัวอย่าง แต่คาดหวังว่ารายงานวิจัยที่มีคุณภาพของระเบียบวิธีวิจัยดีจะสะท้อนถึงความถูกต้องของผลการวิจัยได้. ดังนั้นจึงประเมินผลการวิจัยโดยคำนึงถึงความชัดเจนของแต่ละส่วนของระเบียบวิธีวิจัย ตัวอย่างเช่น ในการวิเคราะห์เมตารายงานวิจัยเชิงทดลองจะประเมินคุณภาพของระเบียบวิธีวิจัยจากส่วนประกอบที่เกี่ยวข้องกับการดำเนินการเพื่อศึกษาประสิทธิภาพของสิ่งทดลองซึ่งได้แก่

ก. การสุ่มสิ่งทดลองให้กับหน่วยศึกษา (randomized allocation of treatment to subjects)

ข. การปกปิดผู้ให้สิ่งทดลอง (blinding treatment givers)

ค. การปกปิดผู้ประเมินผลสิ่งทดลอง (blinding outcome evaluators)

ง. การปกปิดผู้รับสิ่งทดลอง (blinding subjects)

จ. ผลข้างเคียงของสิ่งทดลอง (adverse effects of treatment)

ฉ. ข้อมูลเกี่ยวกับระยะติดตามผลของสิ่งทดลองและการขาดหายของหน่วยศึกษา

ช. วิธีการวิเคราะห์ทางสถิติ

ประเด็นต่างๆ เหล่านี้จะต้องพิจารณาว่ามีกล่าวไว้ชัดเจนเพียงใด เพื่อประเมินว่ารายงานวิจัยแต่ละรายงานมีอคติในระเบียบวิธีหรือไม่เพียงใด. หลังจากการประเมินดังกล่าวอาจจะพบว่าบางรายงานที่มีข้อมูลของระเบียบวิธีการวิจัยบ่งชี้ถึงอคติต่างๆเกิดขึ้นซึ่งอาจนำไปสู่การตัดบางรายงานวิจัยออกจากการวิเคราะห์เมตาได้.

เมื่อประเมินและรวบรวมรายงานวิจัยได้จำนวน

หนึ่งที่เป็นไปตามต้องการเรียบร้อยแล้ว นักวิจัยจะบันทึกข้อมูลต่างๆ ที่ต้องการ ซึ่งนำมาจากแต่ละรายงานวิจัยดังกล่าวลงในแบบวิเคราะห์มาตรฐานที่จัดทำขึ้นโดยทีมผู้วิจัยและจะต้องมีการทำความเข้าใจในการบันทึกข้อมูลลงในแบบวิเคราะห์ดังกล่าวก่อนการปฏิบัติจริงเสมอ. ข้อมูลเหล่านี้จะนำมาหาหรือกันเพื่อพิจารณาตรวจสอบและหาข้อสรุปในกรณีที่ผลประเมินไม่สอดคล้องกัน. กรณีที่บางรายงานแสดงวิธีวิจัยและผลการวิจัยไม่ชัดเจนจะต้องติดต่อขอความร่วมมือจากเจ้าของงานวิจัยนั้นๆ เพื่อทำความเข้าใจ และถ้ามีรายงานวิจัยใดจำเป็นต้องตัดออกจะต้องให้เหตุผลไว้อย่างชัดเจน และนำเสนอในรายงานที่จะเขียนในขั้นตอนสุดท้ายด้วย.

๔. สรุปผลของการวิเคราะห์เมตา

หลังจากที่ได้รวบรวมผลการวิจัยของรายงานวิจัยทั้งหมดที่ผ่านการประเมินเรียบร้อยแล้ว ขั้นตอนต่อไปเป็นการใช้วิธีการทางสถิติมาวิเคราะห์เพื่อหาข้อสรุปเชิงปริมาณที่น่าเชื่อถือ ซึ่งขั้นตอนนี้จะกล่าวรายละเอียดทั้งหมดในตอนต่อไป.

สำหรับข้อมูลต่างๆ จากแต่ละรายงานวิจัยที่จะนำไปใช้ในขั้นตอนการวิเคราะห์ทางสถิติจะได้แก่ ค่าผลสรุปสถิติซึ่งอธิบายประสิทธิภาพของสิ่งทดลอง เช่น ความแตกต่างของค่าเฉลี่ย หรืออัตราส่วนของ odds (odds ratio), และความแปรปรวนของค่าสรุปสถิติดังกล่าว เป็นต้น.

สถิติเบื้องต้นในการวิเคราะห์เมตา

ขั้นตอนต่างๆ ต่อไปนี้ จะเป็นเรื่องของการใช้วิธีการทางสถิติ โดยจะเรียงลำดับจากจุดเริ่มต้นตั้งแต่การพิจารณาตัววัด จนกระทั่งการนำเสนอผล.

๑. ตัววัด

การวิเคราะห์เมตาของรายงานวิจัยเชิงทดลองมีวัตถุประสงค์ที่สำคัญเพื่อหาข้อสรุปเชิงปริมาณที่น่าเชื่อถือของประสิทธิภาพของสิ่งทดลอง (treatment). ตัววัดประสิทธิภาพของสิ่งทดลองที่นิยมใช้อาจจะอยู่ใน

ลักษณะของตัวแปรแบบต่อเนื่อง (continuous outcome) และตัวแปรผลลัพธ์แบบสองกลุ่ม (binary outcome).

ตัวแปรแบบต่อเนื่อง เป็นค่าของผลลัพธ์ที่ถูกวัดออกมาในรูปของตัวเลขซึ่งมีความเป็นไปได้ในลักษณะต่อเนื่องกัน เช่น การศึกษาประสิทธิภาพของวิธีการเรียนรู้ด้วยคอมพิวเตอร์ ซึ่งใช้ค่าคะแนนเป็นตัววัด. การสรุปค่าสถิติของตัววัดประเภทนี้นิยมใช้ "ความแตกต่างของค่าเฉลี่ย ($X_T - X_C$)" หรือความแตกต่างของค่าเฉลี่ยมาตรฐาน ($\frac{X_T - X_C}{S}$ เมื่อ S คือความผิดพลาดมาตรฐานของ $X_T - X_C$) อธิบายความแตกต่างของประสิทธิภาพของสิ่งทดลอง ๒ กลุ่ม. ค่าความแตกต่างของค่าเฉลี่ยมาตรฐานเป็นค่าที่ไม่มีหน่วย จึงนิยมนำมาใช้กรณีที่แต่ละรายงานวิจัยมีหน่วยนับของตัววัดที่แตกต่างกัน เช่น หน่วยของน้ำหนักในแต่ละรายงานวิจัย อาจจะเป็นกิโลกรัม หรือปอนด์ แตกต่างกันได้.

ตัวแปรผลลัพธ์แบบสองกลุ่ม เป็นตัววัดอีกลักษณะหนึ่งที่นิยมมากในการทดลองทางเวชกรรมซึ่งให้ค่าออกมาอย่างใดอย่างหนึ่งใน ๒ กลุ่ม อาทิ หาย/ไม่หาย, รอด/ไม่รอด เป็นต้น. การสรุปค่าสถิติของตัววัดประเภทนี้อาจจะอยู่ในรูป "ความแตกต่างของความเสี่ย

(risk difference), "ความเสี่ยงสัมพัทธ์หรืออัตราส่วนความเสี่ยง (relative risk or risk ratio)" หรือ "อัตราส่วนของ odds (odds ratio)". ค่าสถิติเหล่านี้แสดงในรูปสัญลักษณ์ได้ดังต่อไปนี้

ตาราง 2x2 แสดงจำนวนของหน่วยศึกษาจำแนกตามกลุ่มสิ่งของสิ่งทดลองและผลลัพธ์

	ผลลัพธ์		ผลรวม
	+	-	
กลุ่มทดลอง (T)	a	b	a+b
กลุ่มควบคุม (C)	c	d	c+d

จากข้อมูลในตาราง 2x2 ดังกล่าวจะได้ว่า สัดส่วนของหน่วยศึกษาในกลุ่มทดลองที่ให้ผลลัพธ์ทางบวก (P_T)

$$= a/(a+b)$$

สัดส่วนของหน่วยศึกษาในกลุ่มทดลองที่ให้ผลลัพธ์ทางลบ

$$= b/(a+b) = 1 - P_T$$

ในการทำงานเดียวกันจะหาค่า P_C และ $1 - P_C$ ของกลุ่มควบคุมได้ และจากข้อมูลเหล่านี้สามารถหาค่าสถิติทั้ง ๓ ตัวของค่าผลลัพธ์แบบ ๒ กลุ่ม ได้ดังนี้

	Risk difference	Relative risk	Odds ratio
Parameter	$D = P_T - P_C$	$RR = P_T / P_C$	$OR = \frac{P_T / (1 - P_T)}{P_C / (1 - P_C)}$
Estimator	$d_i = \hat{p}_{ii} - \hat{p}_{ci}$	$rr_i = \hat{p}_{ii} / \hat{p}_{ci}$	$or_i = \frac{\hat{p}_{ii} (1 - \hat{p}_{ci})}{\hat{p}_{ci} (1 - \hat{p}_{ii})}$
Standard error	$S_{d_i} = \sqrt{\frac{p_{ii}(1-p_{ii})}{n_{ii}} + \frac{p_{ci}(1-p_{ci})}{n_{ci}}}$	$S_{rr_i} = \sqrt{\frac{(1-p_{ii})}{n_{ii} p_{ii}} + \frac{(1-p_{ci})}{n_{ci} p_{ci}}}$	$S_{or_i} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$

การตัดสินใจจะเลือกใช้ค่าสถิติใดในการวิเคราะห์ เมตาจะต้องพิจารณาจากวัตถุประสงค์ ซึ่งโดยทั่วไปมักจะเป็นไปตามค่าสถิติของตัววัดจากรายงานวิจัยที่รวบรวม, จากค่าสถิติทั้ง ๓ ของตัววัดแบบ ๒ กลุ่ม, ค่าความแตกต่างของความเสี่ย ($P_T - P_C$) เป็นค่าสถิติที่เข้าใจง่ายที่สุดเนื่องจากค่านี้เกิดจากความแตกต่าง

ของอัตราเสี่ยงของข้อมูล ๒ กลุ่ม ดังเช่นการศึกษาผลของยาแอสไพรินเปรียบเทียบกับยาหลอก ในการป้องกันการตายด้วยโรคหัวใจ พบว่าอัตราตายโรคหัวใจในกลุ่มคนที่ได้รับยาแอสไพรินเท่ากับร้อยละ ๘ เปรียบเทียบกับร้อยละ ๑๑ ของกลุ่มที่ได้รับยาหลอก. จากอัตราตายดังกล่าวมีความแตกต่างเท่ากับร้อยละ ๓

(๑๑%-๘%) อธิบายได้ว่าความเสี่ยงของการตายด้วยโรคหัวใจในคนที่ได้ยาแอสไพรินน้อยกว่าของคนที่ได้ยาหลอกอยู่ร้อยละ ๓ แม้ว่าความแตกต่างของความเสี่ยงดังกล่าวเข้าใจง่าย แต่ค่านี้ไม่ค่อยนิยมใช้เป็นค่าสถิติวัดผลของประสิทธิภาพระหว่างสิ่งทดลอง เนื่องจากความแปรปรวนของค่าความแตกต่างของความเสี่ยงถูกกำหนดด้วยอัตราความเสี่ยงของสิ่งทดลองแต่ละกลุ่ม (P_T, P_C). ดังนั้นจึงอาจพบว่าความแตกต่างของการเสี่ยงระหว่างกลุ่มของสิ่งทดลอง ($P_T - P_C$) ในรายงานวิจัยต่างๆ ที่ศึกษามีความแตกต่างกัน (heterogeneity) ซึ่งอาจจะเป็นผลที่เกิดจากข้อจำกัดทางคณิตศาสตร์ของค่าสถิติตัวนี้มากกว่าเป็นผลเนื่องจากเนื้อหาของสิ่งทดลองนั้นๆ ได้.

สำหรับความเสี่ยงสัมพัทธ์ ซึ่งเป็นการนำอัตราความเสี่ยงของ ๒ กลุ่มสิ่งทดลองมาเปรียบเทียบอัตราส่วนกัน (P_T/P_C) แม้ว่าจะเข้าใจได้ง่ายกว่าอัตราส่วนของ odds แต่กลับพบว่า odds ratio เป็นค่าสถิติที่ได้รับความนิยมสูงกว่า เนื่องจาก odds ratio สามารถคำนวณได้จากการศึกษาทั้งที่เป็นแบบกลุ่มผู้ป่วยควบคุม, cross-sectional, cohort study และการทดลอง, ในขณะที่ความเสี่ยงสัมพัทธ์ไม่สามารถคำนวณได้จากการศึกษาแบบ case-control และ cross-sectional. นอกจากนี้การใช้ค่าความเสี่ยงสัมพัทธ์เป็นตัววัดอาจทำให้พบความแตกต่าง (heterogeneity) ระหว่างรายงานวิจัยที่ไม่ถูกต้อง เนื่องจากเหตุผลทางคณิตศาสตร์ของค่าความเสี่ยงสัมพัทธ์เอง ในขณะที่เหตุการณ์เช่นนี้จะไม่ปรากฏในการใช้ค่า odds ratio^(๒)

๒ การประเมินความคล้ายคลึงระหว่างผลของรายงานวิจัย (assessment of homogeneity among results)

ในการวิเคราะห์เมตาเป้าหมายสำคัญอย่างหนึ่งคือ การประมาณค่าผลรวมที่อธิบายประสิทธิภาพของสิ่งทดลอง (overall estimate of treatments effects) แต่ก่อนจะนำผลการศึกษามาของแต่ละรายงานวิจัย ซึ่งโดยทั่วไปจะเป็นค่าสถิติ เช่น odds ratios มารวมกันเพื่อประมาณค่าผลรวมดังกล่าว จะต้องตรวจสอบก่อน

ว่าค่าสถิติของรายงานวิจัยที่จะนำมารวมกันเหล่านั้นมีความคล้ายคลึงกัน (homogeneous) หรือแตกต่างกัน (heterogeneous). ถ้าค่าสถิติระหว่างรายงานวิจัยมีความแตกต่างกันมาก การประมาณค่าผลรวมสถิติจากผลการศึกษาดังกล่าวอาจจะไม่เหมาะสม. ในกรณีเช่นนี้อาจจำเป็นต้องสรุปแยกตามกลุ่มย่อยของปัจจัยบางประการที่พบว่าเป็นสาเหตุทำให้เกิดความแตกต่างของผลการศึกษาดังกล่าว.

อย่างไรก็ตามประเด็นของความคล้ายคลึงกันหรือแตกต่างกันของผลการศึกษาระหว่างรายงานวิจัยนี้ยังไม่มีเครื่องมือใดที่จะสามารถบอกได้อย่างชัดเจนว่าต่างกันด้วยขนาดเท่าใดจึงจะถือว่าแตกต่างกัน. โดยทั่วไปจะใช้วิธีการทดสอบความคล้ายคลึงทางสถิติร่วมกับการพิจารณาเชิงอัตนัย (subjective) เข้ามาช่วยตัดสินใจ แม้ว่าวิธีดังกล่าวจะไม่ใช่วิธีที่ดีที่สุดก็ตาม แต่ก็ยังเป็นวิธีที่ปฏิบัติกันอยู่ในปัจจุบัน.

วิธีการทดสอบความคล้ายคลึงทางสถิติ (test of homogeneity) เป็นเครื่องมือชนิดหนึ่งที่น่าสนใจ ประเมินความคล้ายคลึงระหว่างผลของรายงานวิจัย. การทดสอบดังกล่าวกระทำภายใต้สมมติฐานของการทดสอบ (Null hypothesis) ที่กล่าวว่าผลการศึกษารายงานวิจัยมีค่าเท่ากัน นั่นคือ

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_k = \theta$$

เมื่อ θ_i ($i = 1, 2, \dots, k$) คือผลของประสิทธิภาพของสิ่งทดลองที่วัดออกมาอยู่ในรูปค่าสถิติซึ่งอาจจะมีความแตกต่างค่าเฉลี่ยของตัวแปรแบบต่อเนื่องหรือค่า odds ratio ของตัวแปรแบบ ๒ กลุ่ม ของรายงานวิจัย i โดยมีสมมติฐานทางเลือกอย่างน้อยที่สุดจะมีหนึ่งรายงานวิจัยที่ให้ผลของสิ่งทดลองแตกต่างไปจากรายงานวิจัยอื่นๆ ใน k รายงานวิจัย และสูตรที่ใช้คำนวณค่าสถิติเพื่อทำการทดสอบจะคิดดังนี้^(๓)

$$Q = \sum_{i=1}^k w_i (\hat{\theta}_i - \hat{\theta}_r)^2 \quad (1)$$

$$\text{หรือ} \quad Q = \sum_{i=1}^k w_i \hat{\theta}_i^2 - \frac{\left(\sum_{i=1}^k w_i \hat{\theta}_i\right)^2}{\sum w_i^2} \quad (2)$$

เมื่อ θ_i เป็นค่าประมาณสถิติของ θ_i ที่สังเกตได้จาก รายงานวิจัยที่ i ($i = 1, 2, \dots, k$)

θ_f เป็นค่าประมาณผลรวมที่แสดงถึงประสิทธิภาพ ของสิ่งทดลองที่สนใจ ซึ่งประมาณได้จาก วิเคราะห์โมเดลแบบคงที่ (fixed effect model) (รายละเอียดของการคำนวณค่า นี้ จะได้กล่าวในหัวข้อต่อไป)

และ w_i เป็นค่าถ่วงน้ำหนักของแต่ละรายงานวิจัยซึ่ง นิยมคิดจากค่าส่วนกลับของความแปรปรวน (inverse variance) ของค่า θ_i , v_i

นั่นคือ $w_i = 1/v_i$

ค่า Q จะมีการแจกแจงเข้าสู่ Chi-square ดังนั้น จึงตัดสินใจเพื่อสรุปความคล้ายคลึงข้อมูลภายใต้การ แจกแจง Chi-square ที่ระดับความเป็นอิสระ (df) เท่ากับ $k-1$ ณ ระดับนัยสำคัญที่กำหนด α เมื่อ k คือจำนวน รายงานวิจัย

ถ้าผลการทดสอบให้ ค่า $P > \alpha$ นั่นคือไม่ปฏิเสธ สมมติฐานของความคล้ายคลึงระหว่างผลของรายงาน วิจัย จะสรุปได้ว่าความแตกต่างระหว่างผลของรายงาน วิจัยที่ปรากฏน่าจะมีสาเหตุเนื่องจากความผิดพลาด แบบสุ่ม. ดังนั้นการวิเคราะห์โมเดลแบบคงที่ น่าจะเป็น วิธีที่เหมาะสมในการคำนวณหาค่าประมาณผลรวม ประสิทธิภาพของสิ่งทดลอง (สำหรับรายละเอียดของ การวิเคราะห์โมเดลแบบคงที่ จะมีการกล่าวในหัวข้อต่อไป)

ในทางตรงข้ามถ้าผลการทดสอบให้ ค่า $P \leq \alpha$ นั่น คือปฏิเสธสมมติฐานดังกล่าว จะสรุปได้ว่าพบความ แตกต่างระหว่างผลของรายงานวิจัย และวิธีการที่อาจ จะเหมาะสมในการหาค่าประมาณผลรวมประสิทธิภาพ ของสิ่งทดลองจะใช้วิธีการวิเคราะห์โมเดลแบบสุ่ม (random effect model) เมื่อไม่สามารถค้นพบสาเหตุที่อาจจะ ทำให้เกิดความแตกต่างดังกล่าวได้ (ซึ่งรายละเอียดของ โมเดลแบบสุ่มจะกล่าวถึงในหัวข้อต่อไป)

ข้อจำกัดของวิธีการทดสอบความคล้ายคลึงทางสถิติ เกิดจาก ความสามารถต่ำในการตรวจพบความแตกต่าง. ดังนั้นจึงมีผู้เสนอแนะให้เลือกใช้ α ที่ระดับ $0.05^{(๓)}$

“การทดสอบความคล้ายคลึง” และ “การทดสอบ ความแตกต่าง” ข้อความทั้งสองนี้มีการกล่าวถึงใน เอกสารบทความเกี่ยวกับวิเคราะห์เมตาอย่างมากมาย อาจจะทำให้เกิดความสับสน จึงขอทำความเข้าใจ ณ ที่ นี้ว่าโดยความเป็นจริงการทดสอบที่อ้างถึงโดยสอง ข้อความดังกล่าวมีหลักการเดียวกัน คือทดสอบ null hypothesis ของความคล้ายคลึง ($H_0 : \theta_i = \theta$) ดังที่กล่าว รายละเอียดไว้แล้ว และเนื่องจากยังไม่สามารถหาหลักฐานยืนยันได้ว่าข้อความใดถูกต้อง จึงขอใช้ “การ ทดสอบความคล้ายคลึง”

๓. โมเดลแบบคงที่และแบบสุ่ม

โมเดลในการวิเคราะห์เพื่อประมาณค่าผลรวมที่ อธิบายประสิทธิภาพของสิ่งทดลองในการวิเคราะห์ เมตามีอยู่ ๒ โมเดล ได้แก่

- โมเดลแบบคงที่ (fixed effect model)
- โมเดลแบบสุ่ม (random effect model)

๓.๑ โมเดลแบบคงที่

การวิเคราะห์ค่าผลรวมสถิติใดๆ ภายใต้โมเดล แบบคงที่จะยึดตามข้อสมมติของโมเดลที่ว่า “รายงาน วิจัยต่างๆ ที่ทำการศึกษาล้างทดลองเหมือนๆ กันจะให้ ผลลัพธ์เหมือนกัน และความแตกต่างที่ปรากฏระหว่าง รายงานวิจัย เป็นผลเนื่องจากข้อผิดพลาดเชิงสุ่ม เท่านั้น หรืออาจกล่าวได้ว่ารายงานวิจัยแต่ละรายงานที่ รวบรวมถือว่าเป็นแต่ละตัวอย่างที่สุ่มจากประชากรที่มี ค่าพารามิเตอร์เดียวกัน” ซึ่งความหมายนี้อาจเรียกว่าเป็น “ข้อสมมติของความคล้ายคลึงของผลลัพธ์ระหว่าง รายงานวิจัย”. ดังนั้นในการคำนวณค่าประมาณผล รวมและช่วงเชื่อมั่นของผลลัพธ์ของสิ่งทดลอง จึงคิด ว่าความแปรปรวนของค่าประมาณดังกล่าวเป็นผลมา จากความแตกต่างที่เกิดขึ้นเฉพาะภายในแต่ละรายงาน วิจัย (within-study variance) เท่านั้น

การคำนวณผลรวมของสิ่งทดลองแบบคงที่ (fixed effect methods)

หลักการเบื้องต้นของการประมาณค่าผลรวมของ สิ่งทดลอง (θ) คือ การประมาณค่าเฉลี่ย (θ_f) จากข้อมูล

ของ θ_i ที่มาจากรายงานวิจัย $i(i=1,2,\dots,k)$

ค่า θ_i อาจจะเป็นความแตกต่างค่าเฉลี่ยระหว่าง ๒ กลุ่มของตัววัดแบบต่อเนื่อง หรือค่า odds ratio ของตัววัดแบบ ๒ กลุ่มของรายงานวิจัย i

ภายใต้ข้อสมมุติของความคล้ายคลึงของผลระหว่างงานวิจัยซึ่งกล่าวว่า

$$\theta_i = \theta$$

เมื่อ θ_i คือ ค่าจริงของผลของสิ่งทดลองในรายงานวิจัย i (true treatment effect in each trial i)

θ คือ ค่าจริงของผลของสิ่งทดลองที่จะประมาณค่า (true overall treatment effect)

ในการประมาณ θ ค่าเฉลี่ยถ่วงน้ำหนัก (weighted average) θ_f ซึ่งเป็นค่าโดยประมาณที่ถือว่าไม่มีอคติ ภายใต้โมเดลแบบคงที่คำนวณจากสูตรดังนี้^(๓)

$$\hat{\theta}_f = \frac{\sum_{i=1}^k w_i \hat{\theta}_i}{\sum_{i=1}^k w_i} \quad (3)$$

เมื่อ w_i เป็นค่าถ่วงน้ำหนักดังที่ได้กล่าวไว้แล้วในการคำนวณค่า Q ดังสูตร (1) การที่นิยมให้ $w_i = 1/v_i$ เนื่องจากจะทำให้ได้ค่าประมาณ θ_f มีความถูกต้องมากที่สุด. ด้วยค่า w_i ดังกล่าวจะพบว่ารายงานวิจัยที่มีขนาดตัวอย่างโตๆ จะมีค่าถ่วงน้ำหนักมาก และในทางตรงกันข้ามจะพบค่าถ่วงน้ำหนักน้อย. ดังนั้นอาจกล่าวได้ว่าการถ่วงน้ำหนักด้วยวิธีนี้จะแปรตามขนาดตัวอย่าง

โดยความเป็นจริง w_i สามารถคำนวณได้จากข้อมูลอื่นๆ นอกเหนือจาก v_i เช่นจากคะแนนคุณภาพ ของระเบียบวิธีวิจัย แต่ไม่ค่อยนิยมใช้ เพราะอาจส่งผลให้ θ_f ไม่แม่นยำเนื่องจากค่าคะแนนคุณภาพดังกล่าวเป็นค่าที่กำหนดขึ้นมาในเชิงอัตนัย โดยทีมนักวิจัย

ภายใต้ข้อสมมุติเพิ่มเติมที่กำหนดว่า θ_f มีการแจกแจงแบบปกติ $N(\theta, v_i)$ เมื่อทราบค่า $w_i = 1/v_i$ และ

ด้วยการพิสูจน์ทางทฤษฎีสถิติจะได้ว่าความแปรปรวนของค่าประมาณผลรวมของสิ่งทดลองจะมีค่าเท่ากับ $1/\sum_{i=1}^k w_i$ หรือ

$$\text{Var}(\hat{\theta}_f) = \frac{1}{\sum_{i=1}^k w_i} \quad (4)$$

ดังนั้นจึงสามารถคำนวณช่วงเชื่อมั่นของค่าประมาณ θ ที่ระดับ $1-\alpha$ ได้เท่ากับ

$$\hat{\theta}_f \pm Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\theta}_f)} \quad (5)$$

ก) การคำนวณหาค่า θ_f สำหรับตัววัดแบบ ๒ กลุ่ม การคำนวณมีหลายวิธีแต่ละวิธีจะอาศัยหลักการดังที่แสดงไว้ในสมการที่ (3) - (5) แต่สิ่งที่แตกต่างกันคือ ที่มาของค่า θ_i และ V_i เท่านั้น, ในที่นี้ขอเสนอแนะวิธีพื้นฐานที่นิยมใช้ ๓ วิธีโดยอาศัยสัญลักษณ์ของตารางต่อไปนี้ ซึ่งสมมุติให้เป็นข้อมูลของรายงานวิจัย i

	ผลลัพธ์		ผลรวม
	+	-	
กลุ่มทดลอง	a_i	b_i	$a_i + b_i = n_{i1}$
กลุ่มควบคุม	c_i	d_i	$c_i + d_i = n_{i2}$
ผลรวม	$a_i + c_i = m_{i1}$	$b_i + d_i = m_{i2}$	N_i

จากข้อมูลในตาราง 2x2 ดังกล่าวจะหาค่า θ_i และ v_i ได้ดังนี้

วิธี	θ_i	V_i
Woolf's method (Inverse-variance method)	$\theta_i = \ln(\text{odds ratio})$	$v_i = \frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}$ $= \ln(a_i d_i / b_i c_i)$

วิธีนี้ถือว่าเป็นวิธีมาตรฐานที่ใช้กันทั่วไปในโมเดลแบบคงที่และแบบสุ่ม. ภายใต้วิธีนี้ค่าผลรวมที่ได้จากสูตร (๓) จะอยู่ในรูปค่าเฉลี่ยลอการิทึมของ odds ratios. ดังนั้นถ้าต้องการจะหา odds ratio ก็ต้องเอาค่าดังกล่าวมาใส่ exponential ซึ่งจะได้แสดงรายละเอียดในตัวอย่าง

Mantel - Haenszel Method	$\hat{\theta}_i = (\text{Odds ratio}) = \frac{a_i d_i}{b_i c_i}$ $\hat{\theta}_r = \frac{\left(\sum_{i=1}^k a_i d_i / N_i \right)}{\left(\sum_{i=1}^k b_i c_i / N_i \right)}$	<p>เนื่องจากในวิธีนี้</p> $W_i = b_i c_i / N_i$ $\text{Var}(\hat{\theta}_r) = \frac{\sum_{i=1}^k P_i R_i}{2 \left(\sum_{i=1}^k R_i \right)^2} + \frac{\sum_{i=1}^k (P_i S_i + Q_i R_i)}{2 \sum_{i=1}^k R_i \sum_{i=1}^k S_i} + \frac{\sum_{i=1}^k Q_i S_i}{2 \left(\sum_{i=1}^k S_i \right)^2}$ <p>เมื่อ</p> $P_i = (a_i + d_i) / N_i$ $Q_i = (b_i + c_i) / N_i$ $R_i = (a_i d_i) / N_i$ $S_i = (b_i c_i) / N_i$
--------------------------	---	--

วิธีการของ Haenszel นี้ค่า θ_f ที่ได้ตั้งสูตรที่ปรากฏจะให้ค่าเท่ากับกับสูตรที่ปรากฏใน (๓) และ $\text{Var}(\theta_f)$ ที่ได้จะใกล้เคียงกับที่ได้จากสูตร (๔)

Peto method เป็นวิธีที่เหมาะสมกับรายงานวิจัยที่มีขนาดตัวอย่างใหญ่เช่นกัน และขนาดตัวอย่างของแต่ละกลุ่มของแต่ละรายงานวิจัยจะต้องมีจำนวนใกล้เคียงกัน นอกจากนี้ ค่า odds ratio ควรมีค่าเข้าใกล้ ๑

วิธี	θ_i	V_i
Peto method	กำหนดให้ $O_i = a_i$ $E_i = n_{11} m_{11} / N_i$ และ $V_i = \frac{n_{11} n_{12} m_{11} m_{12}}{N_i^2 (N_i - 1)}$ $= w_i$ $\hat{\theta}_i = \ln (\text{Odds ratio})$ $= \frac{O_i - E_i}{V_i}$	เนื่องจาก $w_i = V_i$ ดังนั้นจะได้ว่า $\text{Var}(\hat{\theta}_r) = \frac{1}{\sum_{i=1}^k w_i}$

Mantel - Haenszel method เป็นวิธีที่สามารถใช้ได้กับข้อมูลจำนวนน้อย แม้แต่กรณีที่มีบางเซลล์เป็นศูนย์ก็ยังใช้ได้ แต่จะไม่เหมาะสมเช่นเดียวกับวิธี Peto ถ้าค่า odds ratio แตกต่างจาก ๑ มากๆ

ตัวอย่าง

จากการวิเคราะห์เมตาของ ๙ รายงานทดลองทางเวชกรรมที่ตีพิมพ์ในวารสาร Br Med J โดย Collins, Yusuf และ Peto^(๓) วัตถุประสงค์ข้อหนึ่งของการวิเคราะห์เมตานี้เพื่อศึกษาประสิทธิภาพของการใช้ยาขับปัสสาวะในหญิงมีครรภ์ที่มีต่ออุบัติการณ์ของโรคพิษแห่งครรภ์ระยะก่อนซั๊ก. จำนวนขนาดตัวอย่างทั้งหมด คือหญิงมีครรภ์ ๗,๐๐๐ คน และมีจำนวนผู้มีอาการโรคพิษแห่งครรภ์ระยะก่อนซั๊กทั้งหมด ๖๐๐ ราย ดังรายละเอียดในตารางที่ ๑ แสดงผลของแต่ละการศึกษาดังนั้น

Woolf's method เป็นวิธีที่เข้าใจง่ายและจะให้ผลน่าเชื่อถือเมื่อขนาดตัวอย่างของแต่ละรายงานวิจัยใหญ่ แต่จะนำมาใช้ไม่ได้กับข้อมูลที่มีค่าความถี่ในตาราง 2x2 ดังกล่าวในเซลล์ใดเซลล์หนึ่งเป็นศูนย์

ตารางที่ ๑ แสดงจำนวนโรคพิษแห่งครรภ์ระยะก่อนซั๊ก ขนาดตัวอย่างหญิงมีครรภ์ของกลุ่มศึกษาและกลุ่มควบคุม และค่า Odds ratio ของ ๕ รายงานวิจัย

รายงานวิจัย	จำนวน โรคพิษแห่งครรภ์ระยะก่อนซั๊ก จำนวนตัวอย่างหญิงมีครรภ์กลุ่ม		Odds ratio
	กลุ่มศึกษา	กลุ่มควบคุม	
๑	๑๔/๑๓๑ (๑๐.๗%)	๑๔/๑๓๖ (๑๐.๓%)	๑.๐๔
๒	๒๑/๓๘๕ (๕.๕%)	๑๗/๑๓๔ (๑๒.๗%)	๐.๔๐
๓	๑๔/๕๗ (๒๔.๖%)	๒๔/๔๘ (๕๐.๐%)	๐.๓๓
๔	๖/๓๕ (๑๕.๘%)	๑๘/๔๐ (๔๕.๐%)	๐.๒๓
๕	๑๒/๑,๐๑๑ (๑.๒%)	๓๕/๗๖๐ (๔.๖%)	๐.๒๕
๖	๑๓๘/๑,๓๗๐ (๑๐.๑%)	๑๗๕/๑,๓๓๖ (๑๓.๑%)	๐.๗๔
๗	๑๕/๕๐๖ (๓.๐%)	๒๐/๕๒๔ (๓.๘%)	๐.๗๗
๘	๖/๑๐๘ (๕.๖%)	๒/๑๐๓ (๑.๙%)	๒.๙๗
๙	๖๕/๑๕๓ (๔๒.๕%)	๔๐/๑๐๒ (๓๙.๒%)	๑.๑๔

จากข้อมูลในตารางดังกล่าว ถ้าต้องการหาค่าผลรวม odds ratio ของทั้ง ๙ การทดลองภายใต้ข้อสมมุติของความคล้ายคลึงกันของค่า odds ratios จะคำนวณหาค่าผลรวม $\hat{\theta}_f$ ภายใต้โมเดลแบบคงที่ โดยวิธี woolf's

method ซึ่งจะต้องเปลี่ยนค่า odds ratio ให้อยู่ในรูป $1n$ (odds ratio) ก่อน และ ค่า $\hat{\theta}_f$ ในสูตรที่ (๓) จะเท่ากับ $1n$ (ori) ดังแสดงรายละเอียดของแต่ละการทดลอง ดังนี้

รายงานวิจัย	Ori	$1n$ (Ori) ($\hat{\theta}_i$)	V_i	W_i	$wi\theta_i$
๑.	๑.๐๔	๐.๐๓๕๒๒๑	๐.๑๕๕๖๐๑	๖.๒๖๕๖๓	๐.๒๔๕๗๔๒
๒.	๐.๔	-๐.๕๑๖๒๕๑	๐.๑๑๗๗๓๗	๘.๔๕๓๕๑๘	-๗.๗๘๒๕๓๒
๓.	๐.๓๓	-๑.๑๐๘๖๖๓	๐.๑๗๘๐๑๘	๕.๖๑๗๔๑๘	-๖.๒๒๗๘๒๒
๔.	๐.๒๓	-๑.๔๖๕๖๗๖	๐.๒๕๘๕๒๗	๓.๓๔๕๓๐๑	-๔.๕๑๖๕๐๘
๕.	๐.๒๕	-๑.๓๘๖๒๕๔	๐.๑๑๔๒๘๕	๘.๗๕๐๐๔๕	-๑๒.๑๓๐๑๔
๖.	๐.๗๔	-๐.๓๐๑๑๐๕	๐.๐๑๔๖๓๔	๖๘.๓๓๕๕๒	-๒๐.๕๗๖๑๗
๗.	๐.๗๗	-๐.๒๖๑๓๖๕	๐.๑๒๐๖๘๗	๘.๒๘๕๘๖๕	-๒.๑๖๕๖๓๓
๘.	๒.๙๗	๑.๐๘๘๕๖๒	๐.๖๘๖๗๗๒	๑.๔๕๖๕๓๗	๑.๕๘๕๕๖๖
๙.	๑.๑๔	๐.๑๓๑๐๒๘	๐.๐๖๗๘๗๗	๑๔.๗๓๒๔๗	๑.๕๓๐๓๗
รวม				๑๒๕.๒๘๒๗	-๕๐.๐๓๖๗๖

จากข้อมูลดังกล่าวจะได้ว่า ค่าประมาณผลรวมของ Ln (OR)

$$= \frac{-50.03674}{125.2827}$$

$$= -0.39939$$

และความแปรปรวนของ Ln (OR)

$$= \frac{1}{125.2827}$$

$$= 0.00798$$

ดังนั้น ค่าประมาณผลรวม odds ratio

$$= \exp(-0.3999)$$

$$= 0.67$$

และ 95%CI ของ θ เท่ากับ

$$= \exp[-0.39939 + 1.96 (0.00798)^{1/2}]$$

$$= 0.56, 0.80$$

และเมื่อเปรียบเทียบกับวิธีอื่นๆ จะให้ผลดังนี้

วิธี	ค่าประมาณผลรวม odds ratio	๙๕% CI ของ odds ratio
Woolf's method	๐.๖๗	(๐.๕๖, ๐.๘๐)
Mantel-Haenszel	๐.๖๗	(๐.๕๖, ๐.๘๐)
Peto method	๐.๖๖	(๐.๕๖, ๐.๗๕)

ค่าประมาณผลรวม odds ratio และ ๙๕% ของ odds ratio ที่ได้จากทั้ง ๓ วิธีมีค่าใกล้เคียงกัน. จากข้อมูลดังกล่าวจะสรุปได้ว่ายาขับปัสสาวะมีผลในการป้องกันการเกิดโรคพิษแห่งครรภ์ระยะก่อนชักในหญิงมีครรภ์ได้อย่างมีนัยสำคัญทางสถิติ โดยลดอัตราเสี่ยงลงประมาณร้อยละ ๓๓

อย่างไรก็ตามสำหรับข้อมูลชุดนี้จากการตรวจสอบความแตกต่างของผลระหว่างรายงานวิจัยด้วยค่าสถิติ Q ดังที่ปรากฏในสูตร (๑) พบว่า

$$Q = \sum w_i [\ln(\hat{\theta}_i) - \ln(\hat{\theta}_R)]^2 = 27.27$$

$Q = 27.27$; ค่า $P < 0.0001$ ที่ระดับความอิสระเท่ากับ ๔ (๙-๑) จึงสรุปได้ว่าค่าลอการิทึมของ odds ratio ของรายงานวิจัยมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ. ฉะนั้นในการประมาณค่าผลรวม odds ratio ข้อมูลตัวอย่างดังกล่าวควรเลือกใช้โมเดลแบบสุ่มซึ่งจะได้กล่าวในหัวข้อ ๓.๒ ต่อไป

ข. การคำนวณ θ_i ของตัววัดแบบต่อเนื่อง

สำหรับในบทความนี้จะขอเสนอวิธีการประมาณค่าผลรวมเฉลี่ยที่คำนวณจากค่าความแตกต่างของค่าเฉลี่ยมาตรฐาน ๒ กลุ่ม จาก k รายงานวิจัยที่ศึกษาเปรียบเทียบสิ่งทดลอง ๒ กลุ่ม (T, C) โดยวัดผลออกมาในรูปตัวแปรแบบต่อเนื่อง เช่น ค่าความดันเลือด

น้ำหนัก ส่วนสูง ฯลฯ สมมติให้การกระจายของตัวแปรดังกล่าวมีความแปรปรวนไม่แตกต่างกัน ฉะนั้นในแต่ละงานวิจัยที่ i ค่าประมาณความแตกต่างของค่าเฉลี่ยมาตรฐาน ๒ กลุ่มที่อธิบายผลลัพธ์อันเนื่องมาจากสิ่งทดลองที่ศึกษา ($\hat{\theta}_i$) จะคำนวณได้จาก^(๓)

$$\hat{\theta}_i = \frac{\bar{X}_{iT} - \bar{X}_{iC}}{S_i} \quad (5)$$

เมื่อ \bar{X}_{iT} , \bar{X}_{iC} คือค่าเฉลี่ยของตัววัดแบบต่อเนื่องของสิ่งทดลองในกลุ่ม T และ C ของงานวิจัยที่ i ตามลำดับ

S_i คือค่าเบี่ยงเบนมาตรฐานรวม (pooled SD) ที่คำนวณได้จาก

$$\left[\frac{(n_{iT} - 1)S_{iT}^2 + (n_{iC} - 1)S_{iC}^2}{n_{iT} + n_{iC} - 2} \right]^{1/2}$$

ซึ่งสูตรนี้เหมือนกับในการคำนวณค่าสถิติทดสอบแบบ t

ค่า n_{iT} , n_{iC} และ S_{iT}^2 , S_{iC}^2 เป็นขนาดตัวอย่างและความแปรปรวนของกลุ่ม T และ C ตามลำดับ

ถ้า n_{iT} , n_{iC} มีขนาดใหญ่ และความแปรปรวนของตัววัดสิ่งทดลองทั้ง ๒ กลุ่มในประชากรเท่ากันแล้วจะได้ค่าความแปรปรวนของ $\hat{\theta}_i$ จะมีค่าเท่ากับ $(n_{iT} + n_{iC}) / (n_{iT}n_{iC})$

นั่นคือ

$$\text{Var}(\theta_i) = v_i = \frac{n_{iT} + n_{iC}}{n_{iT} n_{iC}}$$

และเป็นผลให้ค่าถ่วงน้ำหนัก

$$W_i = 1/v_i = \frac{n_{iT} n_{iC}}{n_{iT} + n_{iC}}$$

ตัวอย่าง จากตัวอย่างการศึกษาเมตาของ ๕ รายงานวิจัยที่คัดลอกมาจากตารางที่ ๑ ของ Fleiss J^(๓) เพื่อต้องการทราบประสิทธิภาพของวิธีการ รักษาทางจิตเวช เมื่อเปรียบเทียบกับกลุ่มควบคุมโดยบางรายงานวิจัยใช้จำนวนครั้งของการเข้ารับการรักษาคือเป็นตัววัดและบางรายงานวิจัยใช้จำนวนวันที่รับการรักษาอยู่ในโรงพยาบาลเป็นตัววัดได้ผลดังตารางที่ ๒

ตารางที่ ๒ ค่าสถิติที่ได้จากรายงานวิจัยทั้ง ๕

รายงานวิจัย	Psychotherapy			Control			
	n _{i1}	X ₁	S _{i1}	n _{i2}	X ₂	S _{i2}	s _i
๑ Davis	๑๓	๕.๐	๔.๗	๑๓	๖.๕	๓.๗	๔.๒๗
๒ Florell	๓๐	๔.๕๐	๑.๗๑	๕๐	๖.๑๐	๒.๓	๒.๑๐
๓ Gruen	๓๕	๒๒.๕	๓.๔๔	๓๕	๒๔.๕	๑๐.๖๕	๗.๕๑
๔ Hart	๒๐	๑๒.๕	๑.๔๗	๒๐	๑๒.๓	๑.๖๖	๑.๕๗
๕ Wilson	๗	๖.๕๐	๐.๗๖	๗	๗.๓๗	๑.๔๑	๑.๑๓

จากข้อมูลดังกล่าวจะได้ว่า

รายงานวิจัย	o _i	w _i
๑	๐.๓๕๑	๖.๕๐
๒	๐.๐๕๗	๑๗.๗๕
๓	๐.๓๐๓	๑๗.๕๐
๔	-๐.๑๒๗	๑๐.๐๐
๕	๐.๗๗๕	๔.๐๐
ผลรวม		๕๖.๗๕

ซึ่งจะได้ $\sum w_i \hat{\theta}_i = 20.137$ และ $\sum w_i \hat{\theta}_i^2 = 11.109$

จากข้อมูลชุดนี้จะทดสอบความคล้ายคลึงของความแตกต่างของค่าเฉลี่ยมาตรฐาน ๒ กลุ่ม จาก ๕ รายงานวิจัยได้ โดยคำนวณค่าสถิติ Q

$$Q = \sum w_i \hat{\theta}_i^2 - \frac{(\sum w_i \hat{\theta}_i)^2}{\sum w_i}$$

$$= 11.109 - \frac{(20.137)^2}{56.75}$$

$$= 3.96$$

ที่ df = 4 (k-1) จะได้ ค่าที่ > ๐.๑๐ จึงสรุปได้ว่าไม่มีข้อมูลสนับสนุนความแตกต่าง (heterogeneity) ของค่าความแตกต่างมาตรฐานดังกล่าวของรายงานวิจัยทั้ง ๕ รายงาน ดังนั้นจึงสามารถหาค่าผลรวมเฉลี่ยได้จากโมเดลแบบคงที่จากสูตรที่ (๓) ซึ่งจะได้ว่า

$$\hat{\theta}_r = \frac{\sum w_i \hat{\theta}_i}{\sum w_i}$$

$$= \frac{20.137}{56.75}$$

$$= 0.355$$

และมีค่า $SE(\hat{\theta}_r) = [1/(56.75)]^{1/2}$
= 0.133

เพราะฉะนั้น ๙๕% ช่วงเชื่อมั่นของ มีค่าอยู่ระหว่าง

$$0.355 \pm 0.09 (0.133)$$

$$0.09, 0.62$$

๓.๒ โมเดลแบบสุ่ม

การวิเคราะห์ภายใต้โมเดลแบบสุ่มจะยึดตามข้อสมมุติฐานของโมเดลที่ว่ารายงานวิจัยแต่ละรายงานจะให้ผลลัพธ์แตกต่างกัน หรืออาจกล่าวได้ว่าแต่ละรายงานวิจัย คือแต่ละตัวอย่างที่สุ่มมาจากประชากรที่มีผลลัพธ์ของสิ่งทดลองที่แตกต่างกัน ซึ่งเป็นสมมุติของความแตกต่างของผลระหว่างรายงานวิจัย. ดังนั้น ในการคำนวณค่าประมาณผลรวมและช่วงเชื่อมั่นของผลลัพธ์ของโมเดลนี้ จะมีความแปรปรวนเข้ามาเกี่ยวข้องกับ ๒ แหล่ง คือค่าแปรปรวนจากภายในแต่ละรายงานวิจัย (V_i) และค่าแปรปรวนระหว่างรายงานวิจัย (σ²)

การคำนวณผลรวมของสิ่งทดลองแบบสุ่ม

การวิเคราะห์ภายใต้โมเดลแบบสุ่มจะมีลักษณะการคำนวณ เช่นเดียวกับการวิเคราะห์ภายใต้โมเดลแบบคงที่ ยกเว้นในเรื่องของค่าความแปรปรวนของค่าประมาณ

ผลลัพธ์ของสิ่งทดลองที่ได้จากแต่ละรายงานวิจัย i ($\text{Var}[\theta_i]$) ซึ่งจะต้องเพิ่มความแปรปรวนระหว่างรายงานวิจัย (σ^2) เข้ามาด้วย ดังนั้นจึงต้องทราบถึงการประมาณค่า σ^2 สำหรับในบทความนี้จะเสนอวิธีการประมาณค่าผลรวม odds ratio ของ Dersimonian and Laird^(๕) ซึ่งเป็นวิธีการที่ได้รับความนิยมมาก เนื่องจากการคำนวณไม่ซับซ้อนโดยจะใช้การประมาณค่า σ^2 ด้วยวิธีของ moment estimate ซึ่งวิธีนี้จะต้องอาศัยค่า w_i ที่คำนวณได้จากโมเดลแบบคงที่ ซึ่งจะหาค่า σ^2 ดังนี้

$$\hat{\sigma}^2 = \frac{Q - (k-1)}{\sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i}} \quad (6)$$

เมื่อ Q คือค่าสถิติทดสอบความคล้ำยคลึงของผลลัพธ์ระหว่างรายงานวิจัย

k คือจำนวนรายงานวิจัย

และ w_i คือค่าถ่วงน้ำหนักตั้งที่กล่าวไว้ใน การคำนวณ $\hat{\theta}_i$ เนื่องจากค่า σ^2 ที่ประมาณได้ดังสมการที่ (6) มีโอกาสเป็นลบได้ในทางปฏิบัติจึงกำหนดว่าค่าประมาณ σ^2 ที่ต้องการเท่ากับ $\max [0, \hat{\sigma}^2]$

จากข้อมูลดังกล่าวภายใต้โมเดลแบบสุ่มจะได้ค่าประมาณ $\text{Var}[\theta_i]$ ซึ่งขอเขียนใหม่

$$\text{เป็น } \text{Var}^*[\hat{\theta}_i] = v_i + \hat{\sigma}^2$$

ดังนั้นจะได้ว่าค่าถ่วงน้ำหนัก $w_i^* = 1/\text{Var}^*[\theta_i]$ หรือ

$$w_i^* = 1/(v_i + \hat{\sigma}^2)$$

เพราะฉะนั้นค่าประมาณผลรวมสิ่งทดลอง

$$\hat{\theta}_r = \frac{\sum_{i=1}^k w_i^* \hat{\theta}_i}{\sum_{i=1}^k w_i^*} \quad (7)$$

และ

$$\text{Var}(\hat{\theta}_r) = 1/\sum_{i=1}^k w_i^* \quad (8)$$

จึงสามารถคำนวณช่วงเชื่อมั่นของค่าประมาณ θ ที่ระดับ $1-\alpha$ ได้เท่ากับ

$$\hat{\theta} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\theta}_r)} \quad (9)$$

จากข้อมูลตัวอย่างของตัวแปรแบบ ๒ กลุ่ม เรื่องเกี่ยวกับผลของยาขับปัสสาวะในหญิงมีครรภ์ในหัวข้อ ๓.๑ ซึ่งผลจากการทดสอบด้วยค่าสถิติ Q พบว่า รายงานวิจัยมีผลลัพธ์แตกต่างกัน ดังนั้นจึงประมาณค่าผลรวม odds ratio อาศัยโมเดลแบบสุ่มได้ ดังนี้

ในการประมาณค่าผลรวม odds ratio ภายใต้วิธีของ Dersimonian & Laird จะคำนวณมาจากค่าเฉลี่ยของ $\ln[\text{odds ratio}]$ จาก k รายงานวิจัย ดังนั้น ค่า θ_i ในสมการ (๗) เป็น $\ln(\text{odds ratio})$ ของรายงานวิจัยที่ i และ θ_r ในสมการ (๗) เป็นค่าเฉลี่ย $\ln(\text{odds ratio})$ ที่คำนวณจาก k รายงานวิจัย ค่าเฉลี่ย Odds ratio ที่ต้องการจะเท่ากับ $\exp(\theta_r)$ ซึ่งแสดงได้ด้วยตัวเลขจากตัวอย่างของการศึกษาประสิทธิภาพยาขับปัสสาวะในหญิงมีครรภ์ได้ดังนี้

จากค่า Q ที่คำนวณได้เท่ากับ ๒๗.๒๗ และค่า w_i ที่คำนวณจากโมเดลแบบคงที่ในวิธีการของ Woolf ซึ่งแสดงในตารางข้างล่างนี้ ทำให้สามารถประมาณค่า σ^2 ได้จากสูตร (๖) ดังนี้

$$\begin{aligned} \sigma^2 &= \frac{27.27 - (9-1)}{125.28 - \frac{5188.27}{125.28}} \\ &= 0.23 \end{aligned}$$

ดังนั้น $w_i = 1/(v_i + 0.23)$

เพราะฉะนั้นจะคำนวณค่าเฉลี่ย $\ln[\text{odds ratio}]$ ได้จาก

$$\begin{aligned} \theta_r &= (-12.405)/(24.08) \\ &= -0.5152 \end{aligned}$$

ดังนั้นค่าเฉลี่ย Odds ratio ที่ต้องการ

$$\begin{aligned} &= \exp(-0.5152) \\ &= 0.60 \end{aligned}$$

และ ๙๕% CI ของ odds ratio

และข้อมูลที่ใช้ในการประมาณค่า or จะแสดงได้ดังนี้

รายงานวิจัย	In (odds ratio) (θ_i)	w_i	$(w_i)^2$	w_i^* ($v_i + \sigma^2$)	$w_i^* \theta_i$
๑	๐.๐๓๕๒๒๑	๖.๒๖๕๖๓	๓๙.๒๕๙๑๒	๒.๕๖๖๗๒๙	๐.๑๐๐๖๖๙
๒	-๐.๙๑๖๒๙๑	๘.๕๙๓๕๑๘	๗๒.๑๓๙๘๕	๒.๘๗๕๗๓๘	-๒.๖๓๕๐๑๒
๓	-๑.๑๐๘๖๖๓	๕.๖๑๗๔๑๘	๓๑.๕๕๕๓๙	๒.๔๕๐๘๗๔	-๒.๗๑๗๑๙๒
๔	-๑.๔๖๙๖๗๖	๓.๓๕๕๓๐๑	๑๑.๑๙๑๐๔	๑.๘๙๐๖๒๑	-๒.๗๗๗๖๖
๕	-๑.๓๘๖๒๙๔	๘.๗๕๐๐๔๙	๗๖.๕๖๓๓๖	๒.๙๐๔๕๗	-๔.๐๒๖๕๘๙
๖	-๐.๓๐๑๑๐๕	๖๘.๓๓๕๕๒	๔๖๖๙.๗๔๔	๔.๐๘๗๗๔๕	-๑.๒๓๐๘๔๑
๗	-๐.๒๖๑๓๖๕	๘.๒๘๕๘๖๕	๖๘.๖๕๕๕๗	๒.๘๕๑๕๔๒	-๐.๗๕๕๒๙๓
๘	๑.๐๘๘๕๖๒	๑.๔๕๖๙๓๗	๒.๑๒๖๖๕	๑.๐๙๑๒๖	๑.๑๘๗๙๐๕
๙	๐.๑๓๑๐๒๘	๑๔.๗๓๒๔๗	๒๑๗.๐๔๕๗	๓.๓๕๗๐๘๗	๐.๔๓๙๘๗๓
ผลรวม		๑๒๕.๒๗	๕๑๘๘.๒๗	๒๔.๐๘	-๑๒.๔๐๕

และเมื่อนำค่า odds ratio จากทั้ง ๒ โมเดลเปรียบเทียบกัน ให้ผลดังนี้

โมเดล	θ^2	ค่า σ^2 ของ θ	Odds ratio	๙๕% CI ของ odds ratio
คงที่	-๐.๓๙๙๓	๐	๐.๖๗	(๐.๕๖, ๐.๘๐)
สุ่ม	-๐.๕๑๕๒	๐.๒๓๐	๐.๖๐	(๐.๔๐, ๐.๘๙)

$$= \exp [-0.5152 + 1.96(1/24.08)^{1/2}]$$

$$= 0.40, 0.89$$

จากข้อมูลตัวอย่างนี้จะพบว่าโมเดลสุ่มจะให้ช่วงความเชื่อมั่นกว้างกว่า ซึ่งเป็นผลเนื่องจากการเพิ่มความแปรปรวนระหว่างรายงานวิจัยเข้าป็นตัวเอง และปรากฏการณ์ดังกล่าวมีความหมายว่าโมเดลแบบสุ่มอาจให้ผลสรุปในเชิงอนุรักษ์มากกว่า

ปัจจุบันยังไม่สามารถบอกได้ว่าการเลือกใช้โมเดลใดถูกต้อง และประเด็นนี้ยังเป็นข้อถกเถียงกันอยู่ในการของผู้ศึกษาวิธีการวิเคราะห์เมตา. อย่างไรก็ตาม ช่วงเชื่อมั่นของผลรวมของสิ่งทดลองที่คำนวณภายใต้ข้อมูลสมมติทั้งสองจะมีค่าใกล้เคียงกัน ถ้าผลของสิ่งทดลองระหว่างรายงานวิจัยแตกต่างกันมาก. เหตุที่เป็นเช่นนี้เนื่องจากภายใต้ข้อสมมติแบบคงที่ค่า $\sigma^2=0$ ความแปรปรวนที่คิดมีเพียงเฉพาะค่า v_i แต่ภายใต้ข้อสมมติ

แบบสุ่มความแปรปรวนที่คิดมีค่าเท่ากับ $v_i + \sigma^2$ ฉะนั้น ถ้าผลของสิ่งทดลองระหว่างรายงานวิจัยแตกต่างกันมาก ค่านี้จะโตและส่งผลให้ความแปรปรวนมาก ช่วงเชื่อมั่นจึงต้องกว้างขึ้น

สำหรับการคำนวณค่าผลรวมของตัวแปรต่อเนื่องด้วยโมเดลแบบสุ่มจะมีลักษณะเช่นเดียวกับการคำนวณของตัวแปรแบบกลุ่ม

๓.๔ การนำเสนอผล

การนำเสนอรายงานการวิเคราะห์เมตาโดยพื้นฐานจะประกอบด้วย

ก. ตารางแสดงข้อมูลของลักษณะต่างๆ ทางระเบียบวิธีวิจัยของแต่ละรายงานวิจัยที่พบ

การนำเสนอตารางนี้มีวัตถุประสงค์เพื่อการอธิบายรายละเอียดของข้อมูลระเบียบวิธีวิจัยของแต่ละรายงานวิจัยให้ผู้อ่านทราบและผู้อ่านจะได้ใช้พิจารณาถึงความแตกต่างและคุณภาพของวิธีการที่ใช้

ตารางที่ ๓ Characteristics of Randomized Controlled Trials of Diuretic in Pregnancy

Study	Design	Criteria for entry	Treatment regimen	No followed up			No. withdrawn	Primary End points	Definition Of pre-Eclampsia
				Total	Diuretic Treated patients	Control patients			
Flowers et al	Double blind	< 30 th week, Mean = 19 th week	Chlorothiazide 250 mg/day, 500 mg/day, or 750 mg/day until Delivery	519 (445):	385 (335)+	134 (110)+	No details on Perinatal deaths For 50 treated Patients and 24 controls	pre-eclamsia Stillbirth Neonatal death	Systolic blood Pressure ≥ 140 mm Hg or diastolic blood pressure ≥ 90 mm Hg in previously normotensive patient, or appreciable change in hypertensive
Menzies	Open control: phenobar bitone	≥ 24 th week with systolic blood pressure ≥ 140 mm hg, diastolic Blood pressure ≥ 85 mm Gh, or ankle oedema, or weight gain ≥ i.8 kg in any tow weeks affer 24 th week	Chlorothiazide 100 mg/day plus potassium chloride 2 g/day for a week and continued if indications persist or return	105	57	48	Nil	Pre-eclampsia requiring admission Proteinuric Pre-eclampsia Stillbirth Neonatal death	Systolic blood Pressure >145 Mm Hg or Diastolic blood Pressure > 85 Mm Hg or weight Gain > 0.9 kg. In Week of Treatment; or Non-infective albuminura; or substantial increase in oedema
Cuadors And Tatum	Double Blind: "rotational" allocation	≥ 30 weeks	Bendoflumethiazide 5 mg/day until delivery	1771	1011	760	Nil	Pre-echampsia Proteinuria Eclampsia Stillbirt Neonatal death	Not available
Finnerty And Bepko	Open "altermale" allocation	<17 years. No history of renal discase or findings of oedema, increased blood pressure, or alburminuria	Thiazide diuretics Until delivery	3083	1340	1743	201 treated patients transferred to control group for "non-complicance"	Pre-Eclampsia Stillbirth Neonatal death	Oedema of Periorbital area And hands: > 10% rise in mean arterial pressure or non-infective aluminuria

ตารางที่ ๔ Characteristics of trial of Hemocult Screening for Colorectal Cancer

	Mandel et al (1993)	Hardcastle et al (1996)	Kronborg et al (1996)	Kewenter et al (1994)
Study population	Minnesota, USA, Volunteers aged 50-80	Nottingham, UK, Subjects aged 45-74	Funen, Denmark Inhabitants aged 45-75	Gothenburg, Sweden, Inhabitants aged 60-64
Study groups (No of Subjects allocated)	Annual screen (15 570) Biennial screen (15 587) Control (15 394)	Biennial screen (76 466) Control (76 384)	Biennial screen (30 967) Control (30 966)	2 screens, 16-22 months apart, (34 144) Control (34 164)
Hemocult screening :				
Type of tests	82.5% rehydrated	Unhydrated	Unhydrated	Most rehydrated (all of 2nd screen)
Screening periods	1975-82 rehydrated	Unhydrated	Unhydrated	Most rehydrated (all of 2 nd screen)
Screening periods	1975-82 and 1986-92	Recruitment Feb 1981 To Jan 1991, screening Ceased Feb 1995	Aug 1985 to Aug 1995, 5 screening rounds	Started 1982, rescreening offered 16-22 months later
Follow up	Continuing at last Report in 1993	To Jun 1995 (possibly Continuing)	Continues	Continues
Completed screening	Annual screen : 90% ≥ Iscreen, 46% all screens Biennial screen : 90% ≥ Iscreen, 60% all screens 2.4% unhydrated.	60% ≥ Iscreen, 38% all screens (3-6) 2.1% 1 st screen, 1.2%	67% 1 st screen, 46% all 5 screens (rescreening offered only to prior attenders) 1.0% 1 st screen, 0.8%	63% 1 st screen, 60% 2 nd screen, 68% ≥ 1 screen 1.4% regular
Positivity rate	9.8% rehydrated	2 nd screen	2 nd screen, 0.9% 3 rd screen, 1.3% 4 th screen, 1.8% 5 th screen	attenders 2.6% first

ซึ่งมักจะครอบคลุมประเด็นเกี่ยวกับลักษณะของประชากรศึกษา, ชนิดของสิ่งทดลอง, การเลือกชนิดของสิ่งทดลองให้แก่หน่วยศึกษา (treatment allocation to subjects), วิธีการให้สิ่งทดลองแก่หน่วยศึกษา, การประเมินผลสิ่งทดลอง ดังแสดงในตัวอย่างตารางต่อไปนี้

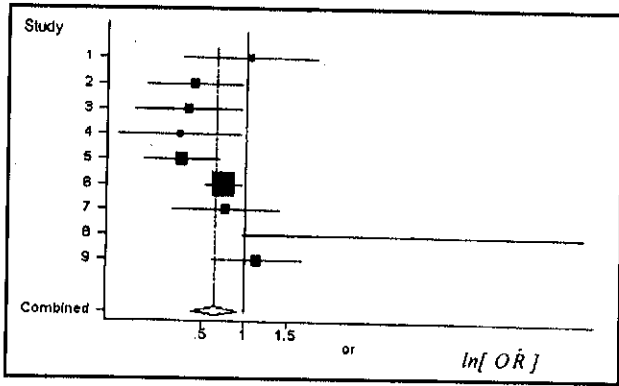
ตารางที่ ๓ เป็นตัวอย่างตารางของการนำเสนอการวิเคราะห์เมตารายงานวิจัยเชิงทดลอง ซึ่งคัดลอกบางส่วนมาจากตัวอย่างการศึกษาของ Collin, Yusuf และ Peto^(๔) เช่นเดียวกับตัวอย่างข้อมูลในหัวข้อ ๓.๑

ตารางที่ ๔ เป็นตัวอย่างตารางของการนำเสนอการวิเคราะห์เมตารายงานวิจัย การทดสอบคัดกรอง

ซึ่งคัดลอกบางส่วนมาจากการศึกษา "A systematic review of the effect of screening for colorectal cancer using the faecal occult blood test, Hemocult" ของ Towler, Irwing, et al^(๖)

ข. ตารางและกราฟแสดงผลลัพธ์ของตัววัดที่ศึกษา การนำเสนอขั้นตอนนี้ต้องการแสดงผลลัพธ์ของแต่ละรายงานซึ่งนิยมเสนอจำนวนขนาดตัวอย่างของแต่ละกลุ่ม และค่าสถิติของตัววัดพร้อมช่วงเชื่อมั่น ดังตัวอย่างตารางที่ ๑ ในหัวข้อ ๓.๑

นอกจากนี้อาจจะใช้นำเสนอด้วยกราฟแสดงค่าสถิติ เช่น odds ratio หรือค่าลอการิทึมของ odds



รูปที่ ๑ แสดงค่า $\ln[OR]$ และช่วงเชื่อมั่นของ $\ln[OR]$ ของแต่ละรายงานวิจัยของผลรวมจากการวิเคราะห์เมตาโมเดลแบบคงที่

ratio ($\ln[OR]$) และช่วงเชื่อมั่นของค่าสถิติของแต่ละรายงานเพื่อให้เห็นชัดเจนขึ้นพร้อมด้วยค่าผลรวมสถิติและช่วงเชื่อมั่นของค่าผลรวมที่วิเคราะห์ได้ จากการรวมแบบคงที่หรือแบบสุ่มกราฟเช่นนี้เรียกว่า “forest plots” ซึ่งแสดงไว้ในรูปที่ ๑ เส้นแต่ละเส้นที่ปรากฏภายในกราฟจะแสดงถึงช่วงเชื่อมั่นของค่าสถิติของผลลัพธ์ของสิ่งทดลอง เส้นที่สั้นแสดงให้ทราบว่ารายงานวิจัยมีค่าประมาณที่ถูกต้องมากกว่ารายงานวิจัยที่แสดงเส้นที่ยาว

บางครั้งโปรแกรมสถิติจะสามารถแสดงพื้นที่รูปสี่เหลี่ยมผืนผ้า ณ ตำแหน่งของค่าสถิติบนเส้นของช่วงเชื่อมั่นของแต่ละรายงานวิจัย ซึ่งสี่เหลี่ยมดังกล่าวจะบอกถึงความแม่นยำของรายงานวิจัยได้ด้วยขนาด นั่นคือสี่เหลี่ยมใหญ่จะบอกให้ทราบว่ารายงานวิจัยนั้นมีค่าประมาณที่แม่นยำมากกว่ารายงานวิจัยที่มีสี่เหลี่ยมเล็ก ดังตัวอย่างในรูปที่ ๑ งานวิจัยหมายเลข ๖ มีความถูกต้องมากที่สุดและหมายเลข ๘ มีความถูกต้องต่ำที่สุดและที่ล่างสุดของภาพมักจะแสดงข้อมูลของค่าผลรวมซึ่งอยู่ในรูปสี่เหลี่ยมขนมเปียกปูน

ข้อควรระวังในการวิเคราะห์เมตา

เนื่องจากจุดมุ่งหมายที่สำคัญของการวิเคราะห์เมตาเพื่อต้องการให้ได้ข้อสรุปเชิงปริมาณเกี่ยวกับผลของ

ประสิทธิภาพของสิ่งทดลองที่น่าเชื่อถือ ซึ่งจะนำไปใช้เป็นข้อมูลประกอบการตัดสินใจในงานหลายๆ ด้านทางการแพทย์และสาธารณสุข ฉะนั้นจึงจำเป็นต้องทราบถึงปัญหาบางประเด็นที่อาจเกิดขึ้นได้ในการวิเคราะห์เมตาเพื่อหาทางป้องกันให้เกิดขึ้นน้อยที่สุด

๑. การพิจารณาแหล่งของความแตกต่างระหว่างรายงานวิจัย

ในความเป็นจริงผลลัพธ์ของแต่ละรายงานวิจัยมีโอกาสจะแตกต่างกันเสมอด้วยหลายสาเหตุเช่น ในการวิเคราะห์เมตาของรายงานวิจัยเชิงทดลองอาจมีความแตกต่างกันในบางลักษณะที่เกี่ยวกับสิ่งทดลอง ได้แก่ ขนาดของสิ่งทดลองที่ให้, ระยะเวลาที่ให้สิ่งทดลองแก่หน่วยศึกษา (period of treatment), ระยะเวลาติดตาม เป็นต้น. นอกจากนี้อาจเป็นผลเนื่องจากลักษณะของประชากรที่ศึกษาหรือ การใช้ตัววัดที่มีสเกลแตกต่างกัน

ด้วยเหตุผลเหล่านี้ การแก้ปัญหาโดยอาศัยสมการทางสถิติที่เพิ่มค่าความแตกต่างระหว่างรายงานวิจัยเพียงค่าเดียว (σ^2) ลงในสมการเพียงวิธีเดียวอาจจะไม่ถูกต้อง วิธีที่ดีควรตรวจสอบว่าความแตกต่างที่ปรากฏระหว่างรายงานวิจัยเป็นผลเนื่องจากอะไร และพยายามหาเหตุผลอธิบายให้ได้มากที่สุด ซึ่งถ้าจำนวนรายงานวิจัยที่ทบทวนมีมากเพียงพออาจนำไปสู่การวิเคราะห์กลุ่มย่อย (subgroup analysis) ตามลักษณะของปัจจัยที่อาจเป็นสาเหตุ. อย่างไรก็ตามผลที่ได้จากการวิเคราะห์เช่นนี้อาจจะไม่ใช้ข้อสรุปที่น่าเชื่อถือเนื่องจากมิได้วางแผนที่จะวิเคราะห์ไว้ล่วงหน้า ฉะนั้นผลลัพธ์ที่ได้อาจนำมาตั้งเป็นสมมุติฐานเพื่อเสนอแนะให้มีการศึกษาถึงสาเหตุต่างๆ ที่พบในอนาคตต่อไป

ในการค้นหาสาเหตุของความแตกต่างดังกล่าวอาจใช้วิธีการทางสถิติที่เรียกว่า meta-regression ซึ่งมีหลักการวิเคราะห์เช่นเดียวกับการวิเคราะห์ความถดถอยทั่วไป แต่ในกรณีของการวิเคราะห์เมตานั้นแต่ละหน่วยที่ศึกษาคือรายงานวิจัยแต่ละรายงาน. ดังนั้นค่าสถิติของตัววัดเช่นค่าเฉลี่ยของความแตกต่างใน

ตัวแปรแบบต่อเนื่องหรือค่า In[RO] ของตัวแปรแบบ ๒ กลุ่ม จะแสดงถึงค่าตัวแปรตาม (dependent variable) ของแต่ละรายงานวิจัย และในทำนองเดียวกันค่าของตัวแปรอิสระ (independent variable) ก็จะต้องเป็นค่าสถิติของตัวแปรนั้นๆ เช่น ค่าเฉลี่ยน้ำหนักของผู้ที่ศึกษาในรายงานวิจัย เป็นต้น. ด้วยเหตุผลดังกล่าวจึงอาจพบว่าความสัมพันธ์ที่ปรากฏจากการวิเคราะห์ด้วย meta-regression ไม่สามารถใช้อธิบายความสัมพันธ์ของแต่ละหน่วยที่ศึกษาได้ ตัวอย่างเช่นอาจพบว่ายา ก. ให้ผลในการรักษาดีในกลุ่มผู้ป่วยที่มีอายุเฉลี่ย ๔๐ ปี แต่ข้อมูลนี้อาจไม่สามารถอธิบายได้ว่าถ้าผู้ป่วยคนหนึ่งที่มีอายุ ๔๐ ปี และใช้ยาดังกล่าวแล้วจะให้ผลการรักษาให้ดีเช่นกัน นอกจากนี้ในบางครั้งอาจพบว่าปัจจัยหรือสาเหตุที่สนใจไม่สามารถหาข้อมูลได้ครบถ้วนจากทุกๆ รายงานวิจัยที่ศึกษา ดังนั้นจึงควรระมัดระวังในการพิจารณาใช้ meta-regression เพื่อหาสาเหตุของความแตกต่างที่ปรากฏ

นอกจากนี้อาจจะใช้วิธีของการวิเคราะห์ความไว (sensitivity analysis) เพื่อตรวจสอบความคงที่ของข้อสรุป (robustness of the results) เช่นอาจจะลองวิเคราะห์ดูผลลัพธ์ที่ได้จากกลุ่มรายงานวิจัยที่ได้รับการตีพิมพ์เปรียบเทียบกับกลุ่มที่ไม่ได้รับการตีพิมพ์ หรือเปรียบเทียบผลลัพธ์เมื่อมีการตัดบางรายงานวิจัยที่ระเบียบวิธีวิจัยมีคุณภาพต่ำออกไป เป็นต้น. ถ้าผลการตรวจสอบความไวไม่ว่าจะคำนึงถึงปัจจัยใดๆ ที่สนใจแล้วให้ค่าผลลัพธ์ออกมาใกล้เคียงกัน ก็อาจอธิบายได้ว่าข้อสรุปที่ได้มีความน่าเชื่อถือ แต่ถ้าผลลัพธ์ที่ได้จากการวิเคราะห์ไวตามเงื่อนไขต่างๆ ที่กำหนดนั้นให้ผลสรุปแตกต่างกันจะต้องค้นหาสาเหตุให้พบ

๒. อคติ

ในการวิเคราะห์เมตาเหมือนกับการวิจัยโดยทั่วไป นั่นคือมีโอกาสเกิดอคติขึ้นได้เสมอ อคติซึ่งควรระมัดระวังที่สำคัญมีอยู่ ๒ แห่ง คือ อคติจากแหล่งตีพิมพ์รายงานวิจัย (publication bias) และอคติที่เกิดจากความไม่น่าเชื่อถือของรายงานวิจัยที่มีคุณภาพต่ำ.

ก. อคติที่เกิดจากแหล่งตีพิมพ์รายงานวิจัย

การเลือกรายงานวิจัยเข้ามาศึกษาในการวิเคราะห์เมตา มีลักษณะเช่นเดียวกับการพิจารณาเลือกผู้ป่วยเพื่อนำเข้าสู่การทดลองทางเวชกรรม ถ้ามีอคติเลือกได้ไม่เหมาะสมหรือไม่เป็นตัวแทนของประชากรก็อาจมีผลต่อค่าประมาณผลของสิ่งทดลองได้ การวิเคราะห์เมตาที่ทำการรวบรวมรายงานวิจัยที่ได้รับการตีพิมพ์เพียงอย่างเดียวอาจเป็นผลให้ค่าประมาณผลรวมของสิ่งทดลองสูงเกินกว่าที่เป็นจริงได้ สาเหตุเนื่องจากรายงานวิจัยที่ให้ผลลัพธ์เชิงบวก (หมายถึงรายงานวิจัยที่รายงานประสิทธิภาพของกลุ่มทดลองให้ผลดีกว่าของกลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติ) มักจะได้รับการพิจารณาตีพิมพ์มากกว่ารายงานวิจัยที่ให้ผลเชิงลบ และยิ่งกว่านั้นถ้ารายงานวิจัยที่ให้ผลเชิงลบศึกษาจากกลุ่มตัวอย่างขนาดเล็ก โอกาสของการได้รับการตีพิมพ์จะยิ่งน้อยลงไปอีก. อคติเช่นนี้บางครั้งไม่อาจจะหลีกเลี่ยงได้เนื่องจากการค้นหาเอกสารที่ไม่ได้รับการตีพิมพ์ทำได้ลำบาก หรืออาจทำไม่ได้เลยในบางกรณี. ในเมื่ออคติดังกล่าวเป็นสิ่งที่ยากจะกำจัดทิ้งได้ จึงควรพิจารณาตรวจสอบเสมอ

การเสนอกกราฟจุดรูปกรวย "funnel plot" เป็นวิธีการทางสถิติอย่างหนึ่งที่ช่วยให้ทราบว่าการวิเคราะห์เมตาที่ศึกษาอยู่มีอคติดังกล่าวมากน้อยเพียงใด ในการสร้างกราฟจุดรูปกรวย แต่ละจุดในกราฟคือแต่ละรายงานวิจัยที่แสดงความสัมพันธ์ระหว่างขนาดของผลลัพธ์ของสิ่งทดลอง เช่นค่า odds ratio เปรียบเทียบกับขนาดตัวอย่าง ซึ่งใช้เป็นค่าน้ำหนักของรายงานวิจัยนั้นๆ ถ้ารูปกรวยมีลักษณะสมมาตร นั่นคือจำนวนจุดที่เกิดในแต่ละด้านของค่าผลรวมมีความใกล้เคียงกันในแต่ละขนาดตัวอย่าง จะอธิบายได้ว่าการวิเคราะห์เมตาที่ทำอยู่รวบรวมรายงานวิจัยได้เหมาะสมในทางตรงกันข้าม ถ้ากราฟดังกล่าวมีลักษณะไม่สมมาตร นั่นคือ อาจพบจำนวนจุดของรายงานวิจัยที่ให้ผลเชิงลบมีค่อนข้างน้อยในกลุ่มรายงานวิจัยที่มีขนาดตัวอย่างต่ำ เมื่อเปรียบเทียบกับของรายงานวิจัยที่ให้ผลลัพธ์เชิงบวก

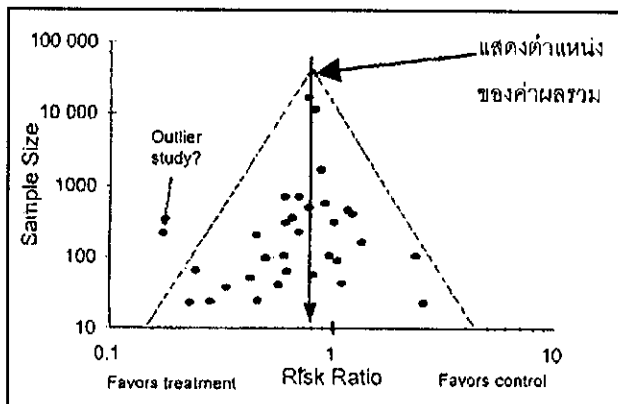
ซึ่งอธิบายได้ว่าการวิเคราะห์เมตาที่ทำอยู่มีอคติแหล่งพิมพ์เกิดขึ้น ตัวอย่างของกราฟจุดรูปกรวยคว่ำแสดงไว้ในรูปที่ ๒ ซึ่งเป็นภาพที่ให้ความสมมาตรในกลุ่มที่มีขนาดตัวอย่างที่มากกว่า ๑๐๐ ขึ้นไป

นอกจากนี้การเลือกรายงานวิจัยเรื่องเดียวกันซ้ำๆ ถือเป็นอคติแหล่งพิมพ์อย่างหนึ่งที่ควรระมัดระวัง บางครั้งอาจพบว่ารายงานวิจัยเรื่องเดียวกันมีการตีพิมพ์มากกว่าหนึ่งวารสาร ซึ่งแนวโน้มมักจะเป็นรายงานวิจัยที่ให้ผลลัพธ์เชิงบวก ฉะนั้น ถ้ารวบรวมรายงานวิจัยเหล่านี้ทั้งหมดเข้ามาศึกษาในการวิเคราะห์เมตาจะทำให้ค่าประมาณผลรวมของสิ่งทดลองไม่ถูกต้อง

จากที่กล่าวมาจึงควรพิจารณาตรวจสอบทุกครั้งเมื่อทำการวิเคราะห์เมตาและถ้ามีอคติแหล่งพิมพ์เกิดขึ้นจะต้องระมัดระวังการแปลผลที่วิเคราะห์ได้ มิฉะนั้นอาจจะทำให้ข้อสรุปที่ได้ขาดความน่าเชื่อถือ

ข. อคติที่เกิดจากคุณภาพของรายงานวิจัย

คุณภาพของรายงานวิจัยเป็นส่วนสำคัญอย่างหนึ่งที่สะท้อนถึงความแม่นยำของค่าประมาณผลรวมของประสิทธิผลสิ่งทดลอง รายงานวิจัยที่มีคุณภาพดีน่าจะเป็นเหตุทำให้เกิดผลการศึกษาที่น่าเชื่อถือ การวัดคุณภาพของรายงานวิจัยจะประเมินจากระเบียบวิธีวิจัยดังได้กล่าวไว้ในหัวข้อ ๒.๓ มีการศึกษาซึ่งอ้างถึงในบทความของ Lau, J, et al^(๑) พบว่า ในบางรายงานวิจัย



รูปที่ ๒ แสดงกราฟจุดรูปกรวยคว่ำของค่า risk ratio ที่สัมพันธ์กับขนาดตัวอย่างของแต่ละรายงานวิจัย

ของการวิเคราะห์เมตาของรายงานวิจัยเชิงทดลองทางเวชกรรม ซึ่งให้ค่าผลรวมของสิ่งทดลองเกินความเป็นจริง มีสาเหตุมาจากการรวบรวมรายงานการทดลองทางเวชกรรมไม่ใช่วิธีการเลือกสิ่งทดลองให้กับผู้ป่วยแบบสุ่มและไม่มีการปกปิดคนวัดผลเกี่ยวกับชนิดของสิ่งทดลองที่ผู้ป่วยได้รับ ฉะนั้นจึงต้องตรวจสอบคุณภาพของรายงานวิจัยที่นำมาวิเคราะห์เมตาด้วยความระมัดระวังก่อนตัดสินใจนำเข้ามาทบทวนในการวิเคราะห์เมตาเสมอ

สรุป

การวิเคราะห์เมตาปัจจุบันเป็นที่ยอมรับกันโดยทั่วไปว่าเป็นเครื่องมือที่มีประโยชน์ ซึ่งจะให้ข้อมูลที่น่าเชื่อถือเพื่อใช้พิจารณาประกอบการตัดสินใจในงานด้านการวางแผนรักษา ป้องกันและบริการสาธารณสุขต่างๆ ได้อย่างมีประสิทธิภาพ. สิ่งต่างๆ ดังกล่าวจะเกิดขึ้นได้ถ้าผู้วิเคราะห์เมตามีความรู้เกี่ยวกับวิธีการได้อย่างถูกต้อง และเข้าใจถึงข้อจำกัดต่างๆ ของการวิเคราะห์เมตาได้เป็นอย่างดี

บทความชิ้นนี้เป็นเพียงเนื้อหาเบื้องต้นของการวิเคราะห์เมตาซึ่งครอบคลุมวิธีการรวมข้อมูลทางสถิติระดับพื้นฐานเท่านั้นและเนื่องจากการวิเคราะห์เมตายังถือว่าเป็นวิธีการทางสถิติใหม่จึงอาจพบว่ามีหลายประเด็นของวิธีการที่ยังมีข้อโต้แย้งอยู่และปัจจุบันมีผู้ศึกษาและพัฒนาวิธีการเป็นจำนวนมาก ดังนั้นหากผู้อ่านท่านใดมีความประสงค์จะทราบรายละเอียดในระดับสูงของวิธีการวิเคราะห์เมตาควรติดตามจากบทความเชิงวิธีการของการวิเคราะห์เมตา ซึ่งมีปรากฏอยู่ในวารสารต่างๆ มากมาย. เอกสารที่ดีชิ้นหนึ่งคือบทความเรื่อง "Systematic reviews of trials and studies" ซึ่งตีพิมพ์ในวารสาร Health Technology Assessment 1998; Vol 2: No 19 ผู้สนใจสามารถเปิดดูได้จาก <http://www.hta.nhsweb.nhs.uk>

กิตติกรรมประกาศ

ผู้เขียนขอขอบคุณรองศาสตราจารย์เบญญา

มุกตพันธ์ และผู้ช่วยศาสตราจารย์นิคม หนองเมือง ที่ได้กรุณาอ่านต้นฉบับและให้ข้อเสนอแนะที่เป็นประโยชน์ และต้องขอขอบคุณ คุณรัตดา ปิติ และคุณนิสฎาพร ศรีแก้ว ที่กรุณาช่วยพิมพ์ต้นฉบับให้

เอกสารอ้างอิง

๑. Langley J, LeBlanc J, Drake J, Milner R. Efficacy of antimicrobial prophylaxis in placement of cerebrospinal fluid shunts: meta-analysis. *Clin Infect Dis* 1993; 17:98-103.
๒. Fleiss J. The statistical basis of meta-analysis. *Stat Methods Med Res* 1993; 2:121-45.
๓. Fleiss J. Statistical methods for rates and proportions. New York: John Wiley; 1981.
๔. Collins R, Yusuf S, Peto R. Overview of randomised trial of diuretics in pregnancy. *BMJ* 1985; 290:17-23.
๕. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7:177-88.
๖. Towler B, Irwig L, Glasziou P, Kewenter J, Weller D, Silagy C. A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, Hemoccult. *BMJ* 1998; 317:559-66.
๗. Lau J, Ioannidis J, Schmid C. Quantitative synthesis in systematic reviews. *Ann Int Med* 1997; 127:820-6.
๘. Abrams K, Jones D. Meta-analysis and synthesis of evidence. *J Math App Med Bio* 1995; 12:297-313.
๙. Berlin J, Laird N, Sacks H, Chalmers T. A comparison of statistical methods for combining event rates from clinical trials. *Stat Med* 1989; 8:141-51.
๑๐. Brand R, Kragt H. Importance of trends in the interpretation of an overall odds ratio in the meta-analysis of clinical trials. *Stat Med* 1992; 11:2077-82.
๑๑. Chene G, Thompson S. Methods for summarizing the risk associations of quantitative variables in epidemiologic studies in a consistent form. *Am J Epidemiol* 1996; 144:610-21.
๑๒. Collaboration TC. Topic 8: analysing and presenting results. *Cochrane Reviewers' Handbook* 4.0. 1999. p. 52-71.
๑๓. Egger M, Smith G, Phillips A. Meta-analysis: principle and procedure. *BMJ* 1997; 315:1533-7.
๑๔. Egger M, Smith G. Misleading meta-analysis. *BMJ* 1999; 310:752-4.
๑๕. Einarson T, Leeder J, Koren G. A method for meta-analysis of epidemiological studies. *Drug Intellig Clin Pharm* 1988; 22:813-24.
๑๖. Emerson JD. Combining estimates of the odds ratio: the state of the art. *Stat Method Med Res* 1994; 3:157-78.
๑๗. Hardy R, Thompson S. Detection and describing heterogeneity in meta-analysis. *Stat Med* 1998; 17:841-56.
๑๘. Houwelingen H, Zwinderman K. A bivariate approach to meta-analysis. *Stat Med* 1993; 12:2273-84.
๑๙. Jacobson R. Promises and pitfalls of meta-analysis in vaccine research. *Vaccine* 1999; 17:1628-1634.
๒๐. L'abbe' K, Detsky A, O'rourke K. Meta-analysis in clinical research. *Ann Int Med* 1987; 107:224-33.
๒๑. Laird N, Mostell F. Some statistical methods for combining experimental results. *Int J Tech Assess Health Care* 1990; 5:5-30.
๒๒. Lau J, Ioannidis J, Schmid C. Summing up evidence: one answer is not always enough. *Lancet* 1998; 351:123-7.
๒๓. Olkin I. Statistical and theoretical considerations in meta-analysis. *J Clin Epidemiol* 1995; 48:133-46.
๒๔. Olkin I. Meta-analysis: reconciling the results of independent studies. *Stat Med* 1995; 14:457-72.
๒๕. Pocock S. Editorials. *Stat Method Med Res* 1993; 2:117-9.
๒๖. Sacks H, Reitman D, Pagano D, Kupelnick B. Meta-analysis: an update. *Mt Sinai J Med* 1996; 63:216-24.
๒๗. Sutton A, Abrams K, Jones D, Shildon T, Song F. Systematic reviews of trials and other studies. *Health Tech Assess* 1998; 2:1-275.
๒๘. Thompson S. Meta-analysis of clinical trials. *Dictionary in Biostatistics* 1998; 2570-79.
๒๙. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardis* 1985; 335-71.

Abstract A Quantitative Systematic Review : Meta-analysis.

Malinee Laopaiboon

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

Journal of Health Sciencer 2000; 9:411-431.

Systematic review is a common approach in health care research. There are increasing number of systematic review papers published in health care service journals. This article presents the methodology of systematic review in quantitative approach, which has been often called "meta-analysis". The main purpose was to provide readers with general considerations in conducting a meta-analysis.