

ความเหมาะสมของการใช้การสุ่มตัวอย่างแบบ กลุ่มในการศึกษาปัจจัยเสี่ยง

อรุณ จิรวัดน์กุล

มหาวิทยาลัยขอนแก่น

มีคำถามเรื่องการใช้การสุ่มตัวอย่างแบบกลุ่มในการศึกษาปัจจัยเสี่ยงได้หรือไม่? ในการทำความเข้าใจคำตอบนักวิจัยต้องเข้าใจหลักที่ใช้ในการศึกษาปัจจัยเสี่ยงและการสุ่มตัวอย่าง⁽¹⁾

การสุ่มตัวอย่างแบบกลุ่มใช้กันมากในการศึกษาขนาดปัญหาสาธารณสุขในพื้นที่ เช่น การศึกษาอัตราป่วยด้วยโรคความดันโลหิตสูง หรืออัตราภาวะทุพโภชนาการในเด็กก่อนวัยเรียน เป็นต้น เหตุผลที่ต้องมีการศึกษาเพราะไม่สามารถทราบจำนวนผู้ป่วยทั้งหมดจากระบบรายงาน หรือมีผู้ป่วยบางส่วนมีอาการป่วยแต่ไม่ไปรับการรักษาในสถานบริการ ดังนั้นการศึกษาส่วนใหญ่เป็นการหาจำนวนผู้ป่วยในพื้นที่เพื่อนำมาคำนวณอัตราการป่วย

ในการศึกษาจะมีการกำหนดขอบเขตพื้นที่ทางภูมิศาสตร์ตามขอบเขตของระบบบริการที่ต้องการ ข้อมูลอัตราป่วยไปใช้ในการวางแผน หรือประเมินผลงาน ประชากรเป้าหมายคือประชาชนทุกคนที่พักอาศัยกระจายอยู่ตามหมู่บ้าน หรือชุมชนในพื้นที่ วิธีการสุ่มที่เหมาะสมคือการสุ่มแบบกลุ่ม จะใช้แบบหลายขั้นตอน (Multi-stage sampling) หรือขั้นตอนเดียว (Single stage sampling) ขึ้นอยู่กับขนาดของพื้นที่

วัตถุประสงค์หลักของการศึกษาเพื่อหาค่าสัดส่วนของการป่วย หรือความชุก ดังนั้นการคำนวณขนาดตัวอย่าง และการวิเคราะห์จึงใช้วิธีการประมาณค่าสัดส่วนที่ใช้วิธีสุ่มตัวอย่างแบบกลุ่ม

ในบางการศึกษาอาจมีวัตถุประสงค์รองเพื่อหาปัจจัยเสี่ยงของโรคดังกล่าวด้วยก็สามารถทำได้ เพราะผลการศึกษาคความชุกสามารถแยกตัวอย่างที่ศึกษาออกเป็นสองกลุ่มคือกลุ่มเป็นโรค และกลุ่มไม่เป็นโรค นักวิจัยสามารถนำตัวแปรปัจจัยเสี่ยงของทั้งสองกลุ่มมาคำนวณหาปัจจัยเสี่ยง ค่าความเสี่ยงที่ได้จากข้อมูลการศึกษาลักษณะนี้ควรแปรผลด้วยความระมัดระวัง เพราะการศึกษาประเภทนี้ส่วนใหญ่จะมีจำนวนผู้ป่วยไม่เพียงพอจะตอบคำถามเรื่องปัจจัยเสี่ยง ตัวอย่างเช่นในการศึกษาความชุกของผู้ป่วยเบาหวานขนาดตัวอย่างที่ใช้ศึกษา 1000 คนพบว่าอัตราการความชุกร้อยละ 1.6 จำนวนผู้ป่วยเบาหวานที่พบเพียง 16 คน อาจไม่เพียงพอสำหรับการตอบคำถามเรื่องปัจจัยเสี่ยง

ในการศึกษาเรื่องปัจจัยเสี่ยงจะมีประชากรสองกลุ่มคือกลุ่มเป็นโรค และกลุ่มไม่เป็นโรค ปัจจุบันอัตราการเป็นโรคต่าง ๆ ค่อนข้างต่ำ วิธีการศึกษาหาปัจจัยเสี่ยงจะใช้ case-control ซึ่งเริ่มจากคนเป็นโรคได้จากผู้ป่วยที่มารับบริการ ส่วนคนไม่เป็นโรค อาจจะเป็นผู้ป่วยโรคอื่นที่มารับบริการที่โรงพยาบาล (hospital control) หรืออาจไปสุ่มเลือกกลุ่มคนไม่เป็นโรคในชุมชน (community control) จะเห็นได้ว่า คุณสมบัติของประชากรที่ศึกษาจะเป็นเรื่องของการเป็นโรค ไม่ใช่ตัวแทนของพื้นที่ วิธีการศึกษาแบบ case-control จึงไม่ต้องใช้วิธีการสุ่มผู้ป่วยแบบกลุ่ม

ในกรณีที่ศึกษาปัจจัยเสี่ยงโดยการสุ่มตรวจคนใน

พื้นที่เพื่อหาคนเป็น และคนไม่เป็นโรคเช่นเดียวกับการศึกษาเรื่องความชุกก็ทำได้แต่การคำนวณขนาดตัวอย่างจะต้องใช้วิธีการคำนวณเพื่อหาปัจจัยเสี่ยง โดยขนาดตัวอย่างจะเป็นจำนวนคนเป็นโรค และจำนวนคนไม่เป็นโรค

ในการตอบคำถามหลักเรื่องปัจจัยเสี่ยงคนเป็นโรคไม่จำเป็นต้องเป็นตัวแทนของคนเป็นโรคทั้งหมดในพื้นที่ ทั้งนี้เพราะปัจจัยเสี่ยงเป็นปัจจัยที่เกี่ยวข้องกับตัวโรค ไม่ใช่ตัวแทนพื้นที่

การจะต้องตรวจคนกี่คนเพื่อให้ได้คนเป็นโรคจำนวนที่ต้องการ นักวิจัยต้องนำค่าความชุกมาคำนวณ เช่น ต้องการคนเป็นโรค 150 คน ถ้าอัตราความชุกร้อยละ 1.6 จะต้องตรวจคนทั้งหมดอย่างน้อย 9,375 คน ส่วนกลุ่มควบคุมก็สุ่มเอามาจากผู้ตรวจแล้วไม่เป็นโรค จะเห็นได้ว่าประชากรที่ต้องการคือคนเป็นและไม่เป็นโรค ดังนั้นการที่จะใช้การสุ่มแบบกลุ่มเพื่อให้ได้จำนวนผู้เป็นโรคตามจำนวนที่คำนวณได้ หรือจะทำการตรวจที่ละหมู่บ้านเรียงต่อกันไปเพื่อให้ได้คนเป็นโรค 150 คนก็ทำได้ โดยไม่มีอคติในการเลือกตัวอย่าง

ในการศึกษาปัจจัยเสี่ยงโดยตรวจประชาชนในพื้นที่เพื่อหาคนเป็นโรค การให้การวินิจฉัยว่าเป็นโรคจะต้องมีเกณฑ์มาตรฐานในการวินิจฉัย ตัวอย่างเกณฑ์

การวินิจฉัยโรคเบาหวานคือ พบระดับน้ำตาล FBS มากกว่าหรือเท่ากับ 126 มิลลิกรัมต่อเดซิลิตร อย่างน้อย 2 ครั้ง โดยตรวจซ้ำในวันถัดไป หรือ พบระดับน้ำตาลมากกว่าหรือเท่ากับ 126 มิลลิกรัมต่อเดซิลิตร และมีอาการที่เข้าได้กับโรคเบาหวาน ได้แก่ ปัสสาวะมาก ดื่มน้ำมาก น้ำหนักตัวลดลงโดยไม่ทราบสาเหตุ ดังนั้นการตรวจระดับน้ำตาล เพียงครั้งเดียว และไม่ได้รับการวินิจฉัยจากแพทย์อาจไม่ใช่ผู้ป่วยเบาหวาน การหาปัจจัยเสี่ยงที่ศึกษาจากคนกลุ่มนี้จะไม่ถูกต้อง

สรุป

การศึกษาปัจจัยเสี่ยงนักวิจัยต้องเข้าใจเรื่องโรคที่ศึกษา และแบบงานวิจัยที่ใช้หาปัจจัยเสี่ยง วิธีการสุ่มแบบกลุ่มเหมาะสำหรับการหาความชุก ในการตอบคำถามหลักเรื่องปัจจัยเสี่ยงไม่จำเป็นต้องสุ่มคนเป็นโรคให้เป็นตัวแทนของคนเป็นโรคในพื้นที่ศึกษา การใช้วิธีสุ่มแบบกลุ่มจึงไม่มีผลต่อคุณภาพของการการศึกษาถ้าใช้อาจทำให้การเก็บข้อมูลมีค่าใช้จ่ายมากขึ้น

เอกสารอ้างอิง

1. อรุณ จิรวัดน์กุล. สถิติทางวิทยาศาสตร์สุขภาพเพื่อการวิจัยที่ใช้ในงานวิจัยทางวิทยาศาสตร์สุขภาพ. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร: วิทย์พัฒน; 2552.