

## การเลือกใช้สถิติวิเคราะห์และปัญหาที่พบบ่อย

นพวรรณ เปี้ยชื่อ<sup>1</sup> PhD, RN, อพย. (การพยาบาลเวชปฏิบัติชุมชน)

**บทคัดย่อ:** การวิเคราะห์ข้อมูลมีผลต่อความตรงในการสรุปทางสถิติ การเลือกชนิดของสถิติที่ใช้ในการวิเคราะห์ข้อมูลจึงควรพิจารณาจากคำถามหรือวัตถุประสงค์การวิจัย แบบวิจัย และระดับการวัดของตัวแปร สิ่งที่คุกคามต่อความตรงในการสรุปทางสถิติ ได้แก่ 1) พลังในการทดสอบต่ำ เนื่องจากขนาดตัวอย่างไม่เพียงพอ ความไม่เหมาะสมในการกำหนดระดับนัยสำคัญ การกำหนดสมมติฐานหรือการเลือกชนิดของสถิติ 2) การละเมิดข้อตกลงเบื้องต้น และ 3) ความคลาดเคลื่อนที่เกิดจากการวิเคราะห์ข้อมูลโดยเลือกชนิดของสถิติไม่เหมาะสม ปัญหาที่พบบ่อยในการวิเคราะห์ข้อมูลคือ ปัญหาจากขั้นตอนของการเก็บรวบรวมข้อมูล การจัดการข้อมูล วิธีการวิเคราะห์ข้อมูล การแปลผล การนำเสนอหรือการเผยแพร่ผลการวิจัย

**คำสำคัญ:** ปัญหาที่พบบ่อย สถิติวิเคราะห์ การวิจัย

---

<sup>1</sup>รองศาสตราจารย์ โรงเรียนพยาบาลรามาธิบดี คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล

## บทนำ

การวิเคราะห์ข้อมูล (data analysis) เป็นส่วนหนึ่งของการออกแบบการวิเคราะห์ (analysis design) ซึ่งเป็นองค์ประกอบของการออกแบบวิจัย (research design) และเป็นขั้นตอนสำคัญของกระบวนการวิจัย (research process) ที่ผู้วิจัยกำหนดไว้ล่วงหน้าในโครงร่างการวิจัย โดยคำนึงถึงคำถามวิจัย แบบวิจัย ระดับการวัดของตัวแปรและข้อตกลงเบื้องต้นแต่อาจมีการปรับเปลี่ยนได้ หากไม่เป็นไปตามแผน เช่น มีการละเมิดข้อตกลงเบื้องต้น เป็นต้น การเลือกใช้สถิติในการวิเคราะห์ข้อมูลที่ไม่เหมาะสม มีผลให้เกิดความคลาดเคลื่อนของความตรงในการวิจัย (research validity) โดยเฉพาะอย่างยิ่ง ในส่วนของความตรงในการสรุปทางสถิติ (statistical conclusion validity)<sup>1</sup>

การวิเคราะห์ข้อมูลเป็นการดำเนินการลดข้อมูลอย่างเป็นระบบ ในรูปแบบของการจัดกลุ่ม จัดลำดับ จัดกระทำ แปลผลและสรุปผลเพื่อให้ได้คำตอบของคำถามวิจัยวัตถุประสงค์หรือสมมติฐานการวิจัยโดยลักษณะของคำถามวิจัย อาจเป็นการศึกษาสถานการณ์ ชนิด จำนวนลักษณะ การเปรียบเทียบความแตกต่าง การศึกษาความสัมพันธ์ การศึกษาอิทธิพล การพยากรณ์หรือการทำนายโดยข้อมูล que ทำการวิเคราะห์ได้จากการวัดตัวแปรระดับการวัดของตัวแปรแบ่งเป็น 4 ระดับคือ 1) นามสเกล (nominal scale) 2) อันดับสเกล (ordinal scale) 3) อินตรภาคสเกล (interval scale) และ 4) เรโซสเกล

(ratio scale) ในการวิเคราะห์ข้อมูลที่ศึกษาจากประชากร ค่าที่วิเคราะห์ได้เรียกว่าค่าพารามิเตอร์ สำหรับข้อมูลที่ศึกษาจากตัวอย่าง ค่าที่วิเคราะห์ได้เรียกว่าค่าสถิติ

การวิเคราะห์ข้อมูลเชิงปริมาณ ประกอบด้วยสถิติบรรยาย (descriptive statistics) และสถิติอ้างอิง (inferential statistics) สถิติบรรยายเป็นสถิติที่ใช้ในการสรุปข้อมูลและบรรยายลักษณะที่ศึกษาได้แก่การนับการวัด (การแจกแจงความถี่ ร้อยละ สัดส่วน) การวัดแนวโน้มเข้าสู่ส่วนกลาง (ค่าเฉลี่ย มัธยฐานฐานนิยม) และการวัดการกระจาย (พิสัย ค่าเบี่ยงเบนมาตรฐาน ความแปรปรวน สัมประสิทธิ์ความผันแปร) โดยสามารถศึกษาข้อมูลจากประชากรหรือข้อมูลจากตัวอย่าง การวิเคราะห์ข้อมูลไม่ต้องทดสอบสมมติฐานทางสถิติ เมื่อศึกษาข้อมูลจากตัวอย่างจึงอ้างอิงไปยังประชากรไม่ได้ สำหรับการวิเคราะห์ข้อมูลด้วยสถิติอ้างอิง ต้องมีการทดสอบสมมติฐานโดยแบ่งเป็นการทดสอบชนิดพารามेटริก (parametric test) ซึ่งคำนึงถึงลักษณะการแจกแจงของประชากรที่ตัวอย่างได้รับการสุ่มมาต้องมีการแจกแจงแบบปกติ (normal distribution) และการทดสอบชนิดนัยพารามेटริก (non parametric test) ซึ่งไม่คำนึงถึงลักษณะการแจกแจงของประชากร (distribution free)

การเลือกชนิดของสถิติที่ใช้ในการวิเคราะห์ข้อมูลพิจารณาจากคำถามหรือวัตถุประสงค์การวิจัยแบบวิจัย และระดับการวัดของตัวแปร ดังตารางที่ 1

ตารางที่ 1 แนวทางการเลือกชนิดของการวิเคราะห์ข้อมูล

วัตถุประสงค์	จำนวนกลุ่ม/ตัวแปร		ระดับการวัด	สถิติที่ใช้
ศึกษาลักษณะกลุ่ม	1 ตัวแปร	-	Nominal	Frequency (n), percentage, mode
			Ordinal	Median, percentage, quartile
			Interval/ratio	Mean
ทดสอบการกระจาย	1 ตัวแปร	-	Nominal	Range
			Ordinal	Range
			Interval/ratio	Range, SD, S <sup>2</sup> , Coefficient of variation (CV)
วิเคราะห์	2 ตัวแปร	-	Nominal	X <sup>2</sup>
ความสัมพันธ์			Ordinal	Spearman’s rank correlation
			Interval/ratio	Pearson’s Product moment correlation
เปรียบเทียบ	2 กลุ่ม	อิสระ	Nominal	X <sup>2</sup>
		ไม่อิสระ		
		อิสระ	Ordinal	Rank sum tes
		ไม่อิสระ		Signed rank test
		อิสระ	Interval/ratio	Independent t–test
		ไม่อิสระ		Paired t–test
	≥ 3 กลุ่ม	อิสระ	Interval/ratio	ANOVA
		ไม่อิสระ		ANOVA

### สิ่งคุกคามต่อความตรงในการสรุปทางสถิติ

สิ่งที่คุกคามหรือรบกวนความตรงในการสรุปทางสถิติมีดังนี้

1. พลังในการทดสอบต่ำ (low statistical power)<sup>1</sup> นั่นคือผลการวิเคราะห์ข้อมูล ไม่พบความแตกต่างหรือความสัมพันธ์ทั้งที่มีความแตกต่างกันหรือความสัมพันธ์อยู่จริง ทำให้ผลการวิเคราะห์ไม่ตรงกับความจริง (low precision) อาจเนื่องมาจากสาเหตุปัจจัยดังนี้

1) ขนาดตัวอย่างไม่เพียงพอ ทำให้เกิดความคลาดเคลื่อนจากการสุ่ม (random error) และความคลาดเคลื่อนมาตรฐานของความแปรปรวน (standard error of variance) เพิ่มขึ้น โอกาสปฏิเสธ

สมมติฐานนัลจึงยากขึ้นแนวทางการแก้ปัญหาคือ กำหนดขนาดตัวอย่างที่เพียงพอโดยอ้างอิงจากขนาดอิทธิพล (effect size) ของการศึกษาที่ผ่านมาตามหลักการวิเคราะห์อำนาจการทดสอบ (power analysis)<sup>2</sup> หลักการกำหนดตัวอย่างเมื่อทราบประชากร<sup>3</sup> หรือหลักการทางระบาดวิทยา<sup>4</sup>

2) การกำหนดระดับนัยสำคัญ (type I error:  $\alpha$ ) ที่น้อยมาก ถึงแม้จะทำให้การวิเคราะห์ข้อมูลมีความเชื่อมั่นสูง ( $1 - \alpha$ ) แต่มีผลให้เกิดความคลาดเคลื่อนชนิดที่ 2 (type II error:  $\beta$ ) สูง และ พลัง (power:  $1 - \beta$ ) ในการทดสอบต่ำ การวิจัยทางการแพทย์โดยทั่วไปจึงกำหนดระดับนัยสำคัญที่ .05

3) การกำหนดสมมติฐานแบบไม่มีทิศทาง (non directional hypothesis) มีโอกาสปฏิเสธ

สมมติฐานนัยได้ยากกว่าสมมติฐานแบบมีทิศทาง (directional hypothesis) อย่างไรก็ตาม การพิจารณา กำหนดสมมติฐานแบบมีทิศทางหรือไม่มีทิศทางขึ้นอยู่กับเหตุผลเชิงทฤษฎีหรือหลักฐานเชิงประจักษ์ เช่น การศึกษาผลของการนัดต่อความวิตกกังวลของผู้ป่วย ควรตั้งสมมติฐานแบบมีทิศทาง เนื่องจากสามารถอธิบายได้ตามหลักการเชิงทฤษฎีว่า การนัดมีกลไกที่ส่งผลให้ความวิตกกังวลของผู้ป่วยลดลง

4) การเลือกชนิดของสถิติที่ไม่เหมาะสมในการวิเคราะห์ข้อมูลทำให้เกิดความคลาดเคลื่อนในการสรุปผลซึ่งจะส่งผลกระทบต่อความถูกต้องในการตอบคำถามการวิจัย ผู้วิจัยจึงต้องศึกษาและเข้าใจอย่างถ่องแท้เกี่ยวกับหลักการและแนวทางในการวิเคราะห์ข้อมูลรวมทั้งพิจารณาความสอดคล้องกับคำถามวิจัยแบบวิจัยและระดับการวัดของตัวแปร

2. การละเมิดข้อตกลงเบื้องต้น (violation of assumption)<sup>1,5</sup> สถิติวิเคราะห์แต่ละชนิดมีข้อตกลงเบื้องต้นที่แตกต่างกัน ปัญหาที่พบได้บ่อยคือ การละเมิดข้อตกลงเบื้องต้นเรื่องการแจกแจงของข้อมูลเป็นแบบปกติ ซึ่งส่วนใหญ่ผู้วิจัยไม่ได้รายงานผลการทดสอบข้อตกลงเบื้องต้นไว้ในรายงานวิจัย อย่างไรก็ตามสามารถสังเกตได้จากการนำเสนอข้อมูลค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน ซึ่งหากพบว่า ค่าเบี่ยงเบนมาตรฐานใกล้เคียงกับครึ่งหนึ่งหรือมากกว่าครึ่งหนึ่งของค่าเฉลี่ย บ่งชี้ว่าสัมประสิทธิ์ความผันแปร (coefficient of variation: CV) มีค่ามาก แสดงว่าข้อมูลน่าจะมีความเบ้ (skewness) หรือความโด่ง (kurtosis) ไม่เหมาะสมกับการใช้สถิติชนิดพาราเมตริก หากผู้วิจัยยังคงวิเคราะห์ด้วยสถิติชนิดพาราเมตริก จะทำให้เกิดความคลาดเคลื่อนในการสรุป ดังนั้นผู้วิจัยควรนำเสนอผลการทดสอบข้อตกลงเบื้องต้นโดยสรุป ก่อนนำเสนอผลการวิเคราะห์เพื่อตอบวัตถุประสงค์หรือทดสอบสมมติฐาน

3. Fishing the error rate problem ทำให้เกิดความคลาดเคลื่อนชนิดที่ 1 (type I error)<sup>1</sup> เมื่อผู้วิจัยทำการเปรียบเทียบค่าเฉลี่ยหลาย ๆ ครั้ง หรือวิเคราะห์ความสัมพันธ์จากข้อมูลที่มีตัวแปรจำนวนมาก ตัวอย่างเช่น การวิจัยมีวัตถุประสงค์เพื่อเปรียบเทียบค่าเฉลี่ยของตัวแปรในตัวอย่าง 3 กลุ่ม ซึ่งควรวิเคราะห์ด้วยสถิติ ANOVA แต่หากผู้วิจัยทำการวิเคราะห์ข้อมูลด้วยสถิติ Independent t-test สามครั้ง เพื่อเปรียบเทียบค่าเฉลี่ยของตัวแปรระหว่างตัวอย่างกลุ่มที่ 1 และ 2, กลุ่มที่ 1 และ 3, กลุ่มที่ 2 และ 3 จะทำให้เกิดความคลาดเคลื่อนมากขึ้น จึงเรียกว่า Fishing the error rate

### ปัญหาที่พบบ่อยในการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลเป็นการจัดการข้อมูลอย่างเป็นระบบภายหลังเสร็จสิ้นการเก็บรวบรวมข้อมูล เพื่อจัดเตรียมให้พร้อมเข้าสู่การวิเคราะห์ การแปลผล และการนำเสนอหรือเผยแพร่ผลงานวิจัย ปัญหาที่พบบ่อยในการวิเคราะห์ข้อมูลมีดังนี้

1) การเก็บรวบรวมข้อมูล (data collection) ปัญหาที่พบบ่อยในขั้นตอนการเก็บรวบรวมข้อมูลคือ อัตราการไม่ตอบกลับ (non response rate) สูง ในกรณีที่เป็นการศึกษาแบบตัดขวาง (cross-sectional study) หรือมีการสูญหาย (drop out) ของตัวอย่างในกรณีที่มีการติดตามเก็บรวบรวมข้อมูลมากกว่าหนึ่งครั้งในการศึกษาระยะยาว (longitudinal study) หรือการศึกษาเชิงทดลอง นอกจากนี้ยังพบว่า การเก็บรวบรวมข้อมูลไม่สอดคล้องกับประชากรที่ศึกษาหรือไม่เป็นไปตามเกณฑ์คัดเข้า ตัวอย่างเช่น ประชากรที่ศึกษาเป็นผู้สูงอายุ แต่ผู้วิจัยเก็บรวบรวมข้อมูลในผู้ใหญ่ด้วย เป็นต้น

2) การจัดการข้อมูล (data management) ภายหลังที่การเก็บรวบรวมข้อมูลสิ้นสุดลง โดยปกติผู้วิจัยจะเริ่มจัดการข้อมูลดิบที่เก็บรวบรวมมาทำการประมวลข้อมูล (data processing) ขั้นตอนที่สำคัญในการประมวลข้อมูลคือ บรรณาธิกร (editing) เป็นการตรวจสอบความครบถ้วนสมบูรณ์ ความถูกต้อง และเป็นไปตามข้อตกลง รวมทั้งการกำหนดรหัสและการลงรหัส โดยมีประเด็นปัญหาที่พบบ่อยคือ ความสับสนในการลงรหัสและการกลับคะแนน ในกรณีที่เป็นการถามเชิงลบ ที่ผ่านมามีผู้วิจัยที่ทำการกลับคะแนนลงในข้อมูลดิบจากแบบสอบถามที่ได้ทำการเก็บรวบรวมข้อมูล แล้วจึงลงข้อมูลในฐานข้อมูลในคอมพิวเตอร์ ซึ่งบ่อยครั้งที่พบความผิดพลาดจากการกลับคะแนนตามวิธีการดังกล่าว แนวทางในการแก้ปัญหาคือ ลงข้อมูลในฐานข้อมูลตามรหัสที่ผู้วิจัยกำหนดแล้วจึงใช้คำสั่ง recode ในโปรแกรมวิเคราะห์ข้อมูลเพื่อทำการกลับคะแนน

นอกจากนี้แนวทางหนึ่งในการตรวจสอบความถูกต้องของการลงข้อมูลก่อนทำการวิเคราะห์สามารถทำได้โดยการใช้โปรแกรมวิเคราะห์ข้อมูลสำเร็จรูปวิเคราะห์ค่าความถี่ ค่าต่ำสุดและค่าสูงสุดจะช่วยให้ผู้วิจัยทราบว่าการลงข้อมูลดังกล่าวถูกต้องตามข้อมูลจริงหรือเป็นไปตามรหัสที่ผู้วิจัยกำหนดหรือไม่

3) การวิเคราะห์ข้อมูล (data analysis) การวิเคราะห์ข้อมูลต้องให้ครบตามวัตถุประสงค์หรือสมมติฐานที่กำหนดไว้ และในการเลือกใช้สถิติวิเคราะห์ต้องคำนึงถึง คำถามวิจัย ระดับการวัดของตัวแปร และแบบวิจัยดังตัวอย่างต่อไปนี้

การวิเคราะห์ข้อมูลโดยใช้สถิติบรรยายในเรื่องการนับการวัด เช่น การจัดกลุ่มข้อมูลรายได้ของตัวอย่าง หากผู้วิจัยแบ่งรายได้เป็น 2 กลุ่มคือ มากกว่า 2 หมื่นบาทต่อเดือน และ น้อยกว่า 2 หมื่นบาทต่อเดือน ผู้ที่มีรายได้เท่ากับ 2 หมื่นบาทจะไม่อยู่ในกลุ่มใด

ปัญหาที่พบบ่อยอีกประเด็นหนึ่งคือ ค่าร้อยละ 1) หากจำนวนตัวอย่างน้อยมาก อาจไม่จำเป็นต้องวิเคราะห์ค่าร้อยละ 2) การนำเสนอค่าร้อยละ ผลรวมต้องมีค่าไม่เกิน 100.0 และไม่น้อยกว่า 100.0 3) สำหรับข้อคำถามที่เปิดโอกาสให้ตอบได้มากกว่าหนึ่งคำตอบ ควรใช้จำนวนผู้ตอบทั้งหมดเป็นค่าเทียบ 100 4) ในการเปรียบเทียบค่าร้อยละของข้อมูลที่มีการจำแนกสองทาง จำนวนรวมที่ใช้เทียบ 100 ควรเป็นจำนวนรวมในแนวตัวแปรอิสระ

การวิเคราะห์แนวโน้มเข้าสู่ส่วนกลางหรือค่าตัวกลางเป็นการบอกลักษณะเบื้องต้นของข้อมูล การคำนวณค่าเฉลี่ยใช้ค่าสังเกตทุกค่า จึงเป็นตัวกลางที่มีความเสถียรมากที่สุด อย่างไรก็ตามในการนำเสนอผลในเชิงเปรียบเทียบค่าเฉลี่ยต้องอยู่บนฐานคะแนนเต็มเดียวกัน อีกประการหนึ่งคือ ลักษณะธรรมชาติของข้อมูลบางชนิด ไม่มีการแจกแจงเป็นแบบปกติ เช่น รายได้ และหากการทดสอบยืนยันว่าข้อมูลไม่มีการแจกแจงแบบปกติ การวิเคราะห์ค่าตัวกลางที่เหมาะสมคือ มัธยฐาน

การวิเคราะห์ข้อมูลโดยใช้สถิติอ้างอิงที่เป็นการทดสอบแบบพาราเมตริก ต้องมีการตรวจสอบข้อตกลงเบื้องต้น ในกรณีการทดสอบการแจกแจงของข้อมูลว่าเป็นแบบปกติหรือไม่ หากจำนวนตัวอย่าง 50 ขึ้นไป สามารถใช้วิธี Kolmogorov Smirnov test แต่หากจำนวนตัวอย่างน้อยกว่า 50 ควรใช้วิธี Shapiro Wilk test ในกรณีที่พบว่าข้อมูลไม่มีการแจกแจงเป็นแบบปกติ ในทางปฏิบัติหากพบว่ามีค่า outlier ที่สามารถพิจารณาตัดออก แล้วทำการทดสอบใหม่พบว่าข้อมูลมีการแจกแจงเป็นแบบปกติ และอาจพิจารณาร่วมกับ histogram จะเป็นข้อมูลที่ช่วยให้ผู้วิจัยสามารถตัดสินใจเลือกใช้สถิติแบบพาราเมตริกหรือนั้นพาราเมตริกได้อย่างเหมาะสม

ในบางครั้งผู้วิจัยต้องปรับเปลี่ยนแผนการวิเคราะห์ข้อมูล ซึ่งส่วนใหญ่เนื่องมาจากการละเมิดข้อตกลงเบื้องต้นของการวิเคราะห์สถิตินั้นๆ หรือพบว่าข้อมูลในระยะก่อนทดลองของกลุ่มควบคุมและกลุ่มทดลองแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ผู้วิจัยต้องใช้การวิเคราะห์ความแปรปรวนร่วม (ANCOVA) แทนการทดสอบที (t-test) อย่างไรก็ตาม ผู้วิจัยต้องทบทวนและปรับวัตถุประสงค์ของการศึกษาให้สอดคล้องกับการวิเคราะห์ข้อมูล

การวิเคราะห์เชิงสาเหตุ (causal analysis) เป็นอีกข้อจำกัดที่พบบ่อยในการศึกษาเชิงสำรวจ หรือ การศึกษาเชิงระบาดวิทยาแบบ Retrospective Case Control ซึ่งผู้วิจัยควรระบุในรายงานวิจัยว่า เป็นข้อจำกัดของแบบวิจัยที่หากพิจารณาตามหลักการแล้ว จะไม่สามารถสรุปความสัมพันธ์เชิงเหตุผลได้ เนื่องจากการเก็บรวบรวมข้อมูลเพียงครั้งเดียว หรือเป็นการศึกษาจากข้อมูลย้อนหลัง

การทดสอบแบบนันทราเมตริกที่ใช้บ่อยในการวิจัยทางการแพทย์คือการทดสอบไคสแควร์ (Chi-square test) หากมีการละเมิดข้อตกลงเบื้องต้น ผู้วิจัยต้องทำการยุบรวมกลุ่มย่อยที่มีความหมายไปในทางเดียวกันเข้าด้วยกันหรืออ้างอิงตามทฤษฎีไม่ใช่ ยุบรวมกลุ่มย่อยที่มีค่าความถี่คาดหวัง (expected frequency) น้อยเข้าด้วยกัน สำหรับการวิเคราะห์ Fisher's exact probability test เป็นการทดสอบโดยการเปรียบเทียบค่าความน่าจะเป็น ไม่มีการคำนวณค่าไคสแควร์ การนำเสนอผลการวิเคราะห์จึงไม่มีค่าสถิติไคสแควร์ปรากฏอยู่

4) การแปลผลการวิเคราะห์ข้อมูล (interpretation of data) ในการแปลผลการวิเคราะห์ข้อมูล ผู้วิจัยควรพิจารณาทั้งความมีนัยสำคัญทางสถิติ (statistical significance) และความมีนัยสำคัญทาง

คลินิก (clinical significance) กล่าวคือ ความมีนัยสำคัญทางสถิติที่พบจากผลการวิเคราะห์ข้อมูล อาจไม่ได้สะท้อนถึงความมีนัยสำคัญทางคลินิกเช่น ผลการวิเคราะห์ข้อมูลพบว่า กลุ่มทดลองมีความดันโลหิตภายหลังได้รับโปรแกรมน้อยกว่าก่อนได้รับโปรแกรมอย่างมีนัยสำคัญทางสถิติ แต่หากการเปลี่ยนแปลงระดับความดันโลหิตที่ลดลงอย่างมีนัยสำคัญทางสถิตินั้นยังอยู่ในเกณฑ์สูงอาจไม่แสดงถึงความมีนัยสำคัญทางคลินิก นอกจากนี้ผู้วิจัยต้องแปลผลการวิเคราะห์ข้อมูลให้เหมาะสมตามหลักการเชิงสถิติของสถิติแต่ละชนิด และสอดคล้องกับคำถามวิจัย วัตถุประสงค์หรือสมมติฐานของการวิจัย

5) การนำเสนอหรือการเผยแพร่ผลการวิจัย (presentation and dissemination of results)

การนำเสนอค่าของโอกาสในการตัดสินใจสรุปผลอ้างอิง (probability value: p-value) ในปัจจุบันนิยมรายงานผล p-value จากค่าจริง (มักใช้ทศนิยม 3 ตำแหน่ง) ที่เป็นผลจากการวิเคราะห์ข้อมูลเช่น  $p = .034$  มากกว่าการรายงานค่าที่แสดงระดับคือ  $p < .05$  สำหรับค่า  $p$  ที่เท่ากับ  $.000$  ให้รายงานผลเป็น  $p < .001$  เนื่องจากพื้นที่ในการปฏิเสธสมมติฐานนัยน้อยมาก

ความคลาดเคลื่อนที่พบบ่อยในการนำเสนอผลการวิเคราะห์ข้อมูล มีทั้งความคลาดเคลื่อนการสะกดคำ การใช้สัญลักษณ์ทางสถิติที่ไม่สอดคล้องกัน การนำเสนอข้อมูลตัวเลขและค่าสถิติที่ไม่สอดคล้องกัน ขาดข้อมูลสถิติที่จำเป็นในตารางนำเสนอผลการวิจัย หากมีการใช้อักษรย่อหรือรหัส ต้องระบุความหมายหรือคำเต็มท้ายตาราง หากเหตุท้ายตารางโดยทั่วไปเป็นการเสนอความหมาย การกำหนดรหัสตัวแปรหุ่น หรือระดับนัยสำคัญ ตัวอย่างความคลาดเคลื่อนในการเผยแพร่ผลการวิจัย เช่น ผู้วิจัยต้องการ



นำเสนอ “ค่าสถิติ t” แต่ระบุว่า “t-test” ซึ่งมีความหมายแตกต่างกัน ดังนั้นจึงจำเป็นต้องมีความละเอียดรอบคอบในทุกขั้นตอนของการวิเคราะห์ข้อมูล

โดยสรุป ผู้วิจัยควรเลือกชนิดของสถิติที่เหมาะสมในการวิเคราะห์ข้อมูล โดยพิจารณาจากคำถามหรือวัตถุประสงค์การวิจัย แบบวิจัย และระดับการวัดของตัวแปร และคำนึงถึงสิ่งคุกคามต่อความตรงในการสรุปทางสถิติ ที่สำคัญคือ พลังในการทดสอบต่ำ และการละเมิดข้อตกลงเบื้องต้น รวมทั้งจัดการปัญหาที่พบบ่อยในขั้นตอนของการเก็บรวบรวมข้อมูล การจัดการข้อมูล วิธีการวิเคราะห์ข้อมูล การแปลผล การนำเสนอหรือการเผยแพร่ผลการวิจัย

### เอกสารอ้างอิง

1. Burns N, Grove SK. **The Practice of nursing research: Appraisal, synthesis, and generation of evidence.** 7<sup>th</sup> ed. St. Louis: Saunders; 2013.
2. Cohen J. **Statistical power analysis for the behavioral sciences.** 2<sup>nd</sup> ed. New Jersey: Lawrence Erlbaum Associates, Publishers; 1988.
3. Cochran WG. **Sampling Techniques.** 3<sup>rd</sup> ed. New York: John Wiley & Sons; 1977.
4. Kelsey IL. **Methods in Observational Epidemiology and Biostatistics.** New York: Oxford University Press; 1986.
5. Plichta SB, Kelvin EA, Munro BH. **Statistical methods for health care research.** 6<sup>th</sup> ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013.

## Statistical Analysis and Common Problems

*Noppawan Piaseu<sup>1</sup> PhD, RN, Dip. ACNP*

**Abstract:** Data analysis results in statistical conclusion validity. Types of statistical analysis should be selected as appropriate for research questions or research objectives, research design, and levels of measurement. Threats to statistical conclusion validity include: 1) low statistical power due to inadequate sample size, inappropriate level of significance, hypothesis or type of statistical analysis, 2) violation of assumption, and 3) fishing the error rate problem. Common problems in data analysis are in the process of data collection, data management, data analysis method, interpretation, presentation or dissemination of research.

**Keywords:** common problems, statistical analysis, research

---

<sup>1</sup>Associate Professor, Ramathibodi School of Nursing, Faculty of Medicine, Ramathibodi Hospital, Mahidol University