

บทความวิชาการ

การเลือกใช้สถิติในการวิเคราะห์ข้อมูลสำหรับการวิจัยทางวิทยาศาสตร์สุขภาพ Choosing Statistics in Data Analysis for Health Science Research

บุญญพัฒน์ ไชยเมธ^{1*}
Bhunyabhadh Chaimay^{1*}

บทคัดย่อ

การเลือกใช้สถิติที่ถูกต้องและเหมาะสมสามารถลดความผิดพลาดที่อาจเกิดขึ้นจากกระบวนการวิเคราะห์ข้อมูล บทความนี้มุ่งเน้นการเลือกใช้สถิติในการวิเคราะห์ข้อมูลสำหรับการวิจัยทางวิทยาศาสตร์สุขภาพโดยเฉพาะบุคลากรทางการแพทย์และสาธารณสุข สำหรับการเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัยทางวิทยาศาสตร์สุขภาพมีองค์ประกอบในการพิจารณาที่สำคัญ คือ 1) ข้อมูล ตัวแปรและระดับการวัดของข้อมูล 2) การทดสอบทางสถิติ 3) คำถามวิจัย และ 4) ประเภทของสถิติ ซึ่งแต่ละองค์ประกอบมีความสำคัญและเชื่อมโยงสู่การเลือกใช้สถิติในการวิเคราะห์ข้อมูลทางวิทยาศาสตร์สุขภาพอย่างถูกต้องและเหมาะสม นอกจากนี้ สถิติในการวิเคราะห์ข้อมูลจำแนกออกเป็น สถิติพารามетริก และสถิตินอนพารามетริก ซึ่งสถิติทั้งสองประเภทมีข้อตกลงเบื้องต้นในการวิเคราะห์ที่แตกต่างกัน ดังนั้น การวางแผนที่ดีสำหรับการเลือกใช้สถิติที่เหมาะสมในการวิเคราะห์ข้อมูลส่งผลต่อการนำเสนอข้อมูลที่ต้องและเหมาะสม ต่อไป

คำสำคัญ : สถิติ การวิเคราะห์ข้อมูล การวิจัยทางวิทยาศาสตร์สุขภาพ

Abstract

Choosing an accurate and appropriate statistics in data analysis are to decrease bias during the process of analysis. The purpose of this article is to emphasize on choosing statistics in data analysis for health science research. In particular, it is benefit for medical and health personnel. The component of choosing an appropriate statistics in health science research comprises of 1) data, variables and scale of measurement 2) statistics testing 3) research question and 4) type of statistics. Those components are linked to choosing an appropriate and accrual statistics in health science. In addition, statistical testing is divided into two major type such as; parametric and non-parametric statistics which have different assumptions. Hence, well-planned and organized into using an appropriate statistics is resulting to appropriate data presentation.

Keywords: Statistics, Data Analysis, Health Science Research

¹ ผศ.ดร., สาขาวิชาสาธารณสุขศาสตร์ คณะวิทยาการสุขภาพและการกีฬา มหาวิทยาลัยทักษิณ พัทลุง 93110

*Corresponding: โทรศัพท์/โทรสาร 074-693-997 อีเมล bchaimay@gmail.com

บทนำ

ปัจจุบันการวิจัยทางวิทยาศาสตร์สุขภาพมีความก้าวหน้าและพัฒนาไปอย่างมาก โดยเฉพาะการนำองค์ความรู้จากการวิจัยมาประยุกต์ใช้ในการดูแลสุขภาพและพัฒนาคุณภาพชีวิต นอกจากนี้ องค์กรทางสุขภาพต่างมีนโยบายส่งเสริมให้มีการใช้ความรู้จากงานวิจัยในการพัฒนาระบบงานในทุกระดับ ดังนั้น การทำวิจัยให้สำเร็จผู้วิจัยต้องดำเนินการวิจัยตามกระบวนการอย่างถูกต้อง [1] อย่างไรก็ตาม การเลือกใช้สถิติที่ถูกต้องและเหมาะสมสามารถลดความผิดพลาดที่อาจเกิดขึ้นจากการวิเคราะห์ข้อมูล และส่งผลให้งานวิจัยมีความน่าเชื่อถือมากยิ่งขึ้น

“สถิติ” เป็นศาสตร์ที่ใช้ในการค้นหาข้อเท็จจริง [2] และมีความหมายหลายลักษณะ [3] ได้แก่ 1) ตัวเลขหรือข้อมูลสถิติ 2) ศาสตร์ที่เกี่ยวกับวิธีการศึกษาข้อมูลหรือระเบียบวิธีทางสถิติ [1] และ 3) วิชาสถิติ นอกจากนี้ สถิติเป็นเครื่องมือการวิจัยทางวิทยาศาสตร์ [4] มีขั้นตอนที่สำคัญในการวิจัย การทดสอบทางสถิติเป็นการแปลงข้อมูลดิบสู่ผลการศึกษาที่มีความหมายและสามารถประยุกต์ใช้ต่อไป อย่างไรก็ตาม การเลือกใช้สถิติได้อย่างถูกต้องและเหมาะสมส่งผลต่อการวิเคราะห์ข้อมูล ความแม่นยำและความน่าเชื่อถือจากการทบทวนวรรณกรรมงานวิจัยทางวิทยาศาสตร์สุขภาพพบว่า ประมาณ 2 ใน 3 มีข้อบกพร่องในการเลือกใช้สถิติทั้งสถิติเชิงพรรณนาและสถิติเชิงอนุมาน [5] และจากงานวิจัยที่เผยแพร่ตีพิมพ์ในวารสารทางวิทยาศาสตร์สุขภาพพบว่า มีการนำเสนอผลการวิเคราะห์ข้อมูลด้วยสถิติพรรณนาเพียงอย่างเดียว (47%) และสถิติเชิงอนุมาน (31%) และสถิติพหุตัวแปร (22%) และสถิติขั้นสูง (19%) [6]

อย่างไรก็ตาม การวิเคราะห์ด้วยสถิติที่ไม่เหมาะสม [4] ส่งผลต่อการวิเคราะห์ การแปลผลที่ผิดพลาดและเป็นการลดคุณค่างานวิจัย นอกจากนี้การเลือกใช้สถิติยังเป็นสิ่งสำคัญในการพิจารณาสนับสนุนทฤษฎี [7] การตัดสินใจเชิงนโยบายในกรณีที่ต้องใช้ข้อมูลหรือผลการศึกษาจากงานวิจัย (Evidence-based) ดังนั้น การเลือกใช้สถิติที่เหมาะสมผู้วิจัยควรเริ่มจากวางแผนตั้งแต่โครงร่างวิจัย โดยเฉพาะอย่างยิ่งขั้นตอนการวิเคราะห์ข้อมูล สำหรับบทความนี้ มุ่งเน้นการเลือกใช้สถิติในการวิเคราะห์ข้อมูลทางวิทยาศาสตร์สุขภาพโดยเฉพาะบุคลากรทางการแพทย์และสาธารณสุขเพื่อใช้ประโยชน์ในการเลือกใช้สถิติในการวิเคราะห์ข้อมูลสำหรับวิทยาศาสตร์สุขภาพต่อไป

ข้อควรพิจารณาการเลือกใช้สถิติ

สำหรับการพิจารณาการเลือกใช้สถิติสำหรับการวิเคราะห์ข้อมูลทางวิทยาศาสตร์สุขภาพ [4, 8-11] ควรคำนึงถึงประเด็นสำคัญ ดังนี้ 1) ตัวแปร ข้อมูล ระดับการวัดของข้อมูล 2) การทดสอบทางสถิติ 3) คำถามวิจัยหรือวัตถุประสงค์ และ 4) ประเภทของสถิติโดยมีรายละเอียด ดังนี้

ข้อมูลและตัวแปร และระดับของตัวข้อมูล

ข้อมูล คือ ข้อเท็จจริงที่ได้จากการวัด หรือสังเกต อาจอยู่ในรูปของตัวเลขหรือข้อความ [3, 12] ข้อมูลการวิจัยจำแนกเป็นข้อมูลเชิงคุณภาพ (Qualitative data) และข้อมูลเชิงปริมาณ (Quantitative data) ซึ่งข้อมูลดังกล่าวมีความสัมพันธ์กับระดับการวัดของข้อมูล [13] สำหรับตัวแปร (Variables) มีคุณลักษณะหรือคุณสมบัติที่สามารถเปลี่ยนแปลงค่าได้ [3, 15] ทัวไปแบ่งออกเป็น 1) ตัวแปรต้นหรือตัวแปรอิสระ (Independent variables) คือ ตัวแปรเหตุหรือก่อให้เกิดการแปรผันของปรากฏการณ์ และ 2) ตัวแปรตาม (Dependent variables) คือ ตัวแปรที่เป็นผลมาจากการเปลี่ยนแปลงค่าของตัวแปรต้นหรือตัวแปรอิสระ และเป็นตัวแปรที่ผู้วิจัยมุ่งวัดหรือผลลัพธ์ (Outcomes)

ในส่วนของระดับการวัดของข้อมูล (Scale of measurement) [3, 15-18] จำแนกเป็นระดับนามบัญญัติ (Nominal scale) ระดับเรียงลำดับ (Ordinal scale) ระดับอันตรภาค (Interval scales) และระดับสัดส่วน (Ratio scale) หรือจำแนกออกเป็น 2 ประเภท [1, 14] คือ ข้อมูลแจกแจง (Category) และข้อมูลต่อเนื่อง (Continuous) สำหรับข้อมูลข้อมูลแจกแจง [4] คือ ข้อมูลในระดับนามบัญญัติและระดับเรียงลำดับ ส่วนข้อมูลต่อเนื่อง คือ ข้อมูลในระดับอันตรภาคและระดับสัดส่วน

การทดสอบทางสถิติ

การทดสอบทางสถิติเป็นการสรุปหาคำตอบจากกลุ่มตัวอย่างสู่กลุ่มประชากร [19] มีขั้นตอนที่สำคัญ [10, 14] ดังนี้ 1) การกำหนดคำถามวิจัยหรือวัตถุประสงค์ 2) การกำหนดสมมติฐาน 3) การพิจารณาเลือกสถิติที่เหมาะสม 4) การกำหนดระดับนัยสำคัญทางสถิติ 5) การทดสอบทางสถิติและการพิจารณาผลการวิเคราะห์และ 6) การแปลผล ในส่วนการทดสอบทางสถิติจำแนกเป็นการประมาณค่า และการทดสอบสมมติฐาน

การประมาณค่า (Estimation) เป็นวิธีการทดสอบทางสถิติเพื่ออธิบายลักษณะของประชากร [4, 14] หรืออ้างอิงข้อมูลจากกลุ่มตัวอย่างสู่กลุ่มประชากร [12, 13, 20] ค่าที่ได้จากการประมาณค่าในกลุ่มประชากร คือ ค่าพารามิเตอร์ (Parameters: μ , σ , P) [21, 22] และค่าที่ได้จากการประมาณค่าในกลุ่มตัวอย่าง คือ ค่าสถิติ (Statistics: $\bar{X}$, SD, p) [1, 13, 21, 22] การประมาณค่าจำแนกออกเป็น 2 ประเภท คือ การประมาณค่าแบบจุด (Point estimation) และการประมาณค่าแบบช่วง (Interval estimation) สำหรับการประมาณค่าแบบจุดเป็นการประมาณค่าเพียงค่าเดียว ซึ่งผลจากการประมาณค่ามีโอกาสเท่ากับค่าจริงได้น้อย และเกิดโอกาสความคลาดเคลื่อนได้มาก สำหรับการประมาณค่าแบบช่วงเป็นการกำหนดช่วงความกว้างของค่าพารามิเตอร์ที่แสดงถึงค่าพารามิเตอร์ที่มีอยู่จริงในกลุ่มประชากร [14] นอกจากนี้การประมาณค่าแบบช่วงยังเป็นตัวบ่งชี้ความแม่นยำของการศึกษา สำหรับการคำนวณช่วงเชื่อมั่น (Confident interval) [13] มักกำหนดช่วงเชื่อมั่นที่ร้อยละ 95 (95 percent confident interval: 95%CI) [14] โดยที่ร้อยละ 95 ครั้ง ของการศึกษาจะปรากฏค่าเฉลี่ยของประชากรในช่วงดังกล่าว

การวิจัยเป็นการค้นหาคำตอบหรือความเป็นระบบ [12] เพื่ออ้างอิงผลการศึกษากลุ่มประชากร [14, 23] วิธีการดังกล่าวเรียกว่า “การทดสอบสมมติฐาน (Hypothesis testing)” [3] หรือ “การทดสอบนัยสำคัญ (Test of significance)” [13] โดยการตั้งคำถามเพื่อคาดคะเนคำตอบ [22] ในลักษณะ ข้อความที่คาดคะเนความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตาม ตามคำถามหรือวัตถุประสงค์การวิจัย [13] การสมมติฐานจำแนกออกเป็น สมมติฐานการวิจัย และสมมติฐานทางสถิติ

สำหรับสมมติฐานการวิจัย (Research hypothesis) เป็นข้อความที่เขียนขึ้นเพื่อเป็นการพยากรณ์คำตอบจากการวิจัย โดยนำเหตุผล ประสบการณ์ องค์ความรู้ หรือผลจากการวิจัยที่ผ่านมาประกอบการพิจารณาการ และการเขียนสมมติฐานการวิจัยเป็นการแสดงลักษณะของความสัมพันธ์ของตัวแปรตั้งแต่สองตัวขึ้นไป ในขณะที่สมมติฐานทางสถิติ (Statistical hypothesis) อยู่ในรูปของโครงสร้างทางคณิตศาสตร์และใช้สัญลักษณ์แทนข้อความหรือตัวแปรเพื่อใช้ในการทดสอบทางสถิติจำแนกเป็นสมมติฐานว่าง (Null hypothesis: H_0) และสมมติฐานทางเลือก (Alternative hypothesis: H_1) สมมติฐานว่างเป็นสมมติฐานมีลักษณะเป็นกลาง และใช้สัญลักษณ์ค่าพารามิเตอร์ของประชากรแทน เช่น ค่าเฉลี่ยของประชากร (μ) และความแปรปรวนของประชากร (σ^2) และในลักษณะ “ไม่มีความสัมพันธ์” หรือ “ไม่มีความแตกต่าง” [4] ในขณะที่สมมติฐานทางเลือกเป็นสมมติฐานตรงข้ามกับสมมติฐานว่าง และเขียนในลักษณะ “ความแตกต่าง” หรือ “ความสัมพันธ์”

การสรุปผลจากสมมติฐานโดยเปรียบเทียบกับระดับนัยสำคัญทางสถิติ (Alpha: α) ที่ 0.01 หรือ 0.05 อย่างไรก็ดีตาม การทดสอบสมมติฐานอาจมีความคลาดเคลื่อนเกิดขึ้นได้ เช่น ความคลาดเคลื่อนชนิดที่ 1 (Type 1 error) ความคลาดเคลื่อนที่เกิดขึ้นเมื่อปฏิเสธสมมติฐานว่างในขณะที่สมมติฐานว่างถูกต้อง และความคลาดเคลื่อนชนิดที่ 2 (Type 2 error) ความคลาดเคลื่อนที่เกิดขึ้นเมื่อไม่ปฏิเสธสมมติฐานว่าง ในขณะที่สมมติฐานว่างกล่าวผิด

ความสัมพันธ์ระหว่างการวิจัยและการเลือกใช้สถิติในการวิเคราะห์ข้อมูล

คำถามหรือวัตถุประสงค์การวิจัยเป็นสิ่งสำคัญในการดำเนินการวิจัย [8] และเป็นสิ่งกำหนดทิศทางในการทำงานเพื่อเชื่อมโยงสู่การพิจารณาการเลือกใช้สถิติที่เหมาะสมในการวิเคราะห์ข้อมูล [15, 16] คำถามหรือวัตถุประสงค์การวิจัย

จำแนกเป็น 1) วัตถุประสงค์เพื่อการพรรณนา เป็นวัตถุประสงค์ที่ต้องการพรรณนาลักษณะ ธรรมชาติของการเกิดโรค เช่น ความชุก อุบัติการณ์การเกิดโรค เป็นต้น 2) วัตถุประสงค์เพื่อการเปรียบเทียบ ศึกษาความสัมพันธ์ หรือเพื่อการทำนาย และ 3) วัตถุประสงค์เพื่อการศึกษาความสัมพันธ์เชิงเหตุและผล ซึ่งเป็นการศึกษาความสัมพันธ์เชิงเหตุและผลระหว่างตัวแปรต้นและตัวแปรตาม จากความสัมพันธ์ในวัตถุประสงค์ข้อที่ 1 ควรเลือกใช้สถิติเชิงพรรณนา (Descriptive statistics) และวัตถุประสงค์ข้อที่ 2 และ 3 ควรเลือกใช้สถิติเชิงอนุมาน (Inferential statistics) ในการวิเคราะห์ข้อมูล เป็นต้น

ประเภทของสถิติสำหรับการวิเคราะห์ข้อมูล

สถิติสำหรับการวิเคราะห์ข้อมูลจำแนกเป็น [1] 1) สถิติเชิงพรรณนา และ 2) สถิติเชิงอนุมาน สถิติเชิงพรรณนา [15] เป็นสถิติที่ใช้สำหรับการอธิบายคุณลักษณะของข้อมูล [24] เป็นสถิติที่ง่ายที่สุดในการวิเคราะห์และนำเสนอข้อมูล การแปลผลไม่ซับซ้อน สถิติเชิงพรรณนาจำแนกออกเป็น 2 ประเภท คือ การวัดแนวโน้มเข้าสู่ส่วนกลาง (Measure of central tendency) และการวัดการกระจาย (Measure of dispersion หรือ variability)

การวัดแนวโน้มเข้าสู่ส่วนกลางเป็นการวัดทางสถิติเบื้องต้น [21] และควรคำนึงถึง [17] 1) การเป็นตัวแทนของข้อมูล และ 2) การกระจายของข้อมูล ค่าสถิติที่ใช้ในการวัดแนวโน้มเข้าสู่ส่วนกลาง คือ ค่าเฉลี่ย (Mean) มัธยฐาน (Mean) และฐานนิยม (Mode) โดยที่ค่าเฉลี่ย คือ ค่ากลาง [14] หรือค่าที่ได้จากการนำข้อมูลทั้งหมดมารวมกันแล้วหารด้วยจำนวนข้อมูลทั้งหมด สำหรับค่ามัธยฐาน คือ ค่ากลางของชุดข้อมูลจากการเรียงลำดับ [14] ในขณะที่ฐานนิยม คือ ค่าความถี่ของข้อมูลที่ซ้ำกันมากที่สุด

การวัดแนวโน้มเข้าสู่ส่วนกลางมีความสัมพันธ์กับการกระจายของข้อมูล [21] ซึ่งข้อมูลที่มีค่าสูงหรือต่ำมากเกินไปกว่าปกติ (Outliers) มีโอกาสทำให้การวัดมีคลาดเคลื่อน [14]โดยที่ลักษณะการแจกแจงที่ดีต้องมี 1) การกระจายแบบโค้งปกติ (Normal distribution) ซึ่งค่าที่ได้จากการวัด [1, 4] ระหว่างค่าเฉลี่ย มัธยฐาน และฐานนิยมจะมีค่าเท่ากัน หรือลักษณะของการกระจายแบบที่สมมาตร (Normal distribution, Gaussian distribution หรือ bell-shaped) 2) ในกรณีที่ข้อมูลมีการกระจายแบบเบ้ [1, 25] จำแนกออกเป็น 2 ชนิด คือ เบ้ขวา (Positively skewed distribution หรือ right-skewed) และเบ้ซ้าย (Negatively skewed distribution หรือ left-skewed) ค่าที่ได้จากการวัดแนวโน้มเข้าสู่ส่วนกลางจะมีความแตกต่างกันในกรณีการวิเคราะห์ข้อมูลสำหรับข้อมูลระดับนามบัญญัติ และเรียงอันดับการนำเสนอด้วยร้อยละ เหมาะสมที่สุด เนื่องจากข้อมูลประเภทดังกล่าวไม่มีคุณสมบัติในเชิงคำนวณ สำหรับข้อมูลระดับอันตรภาคและระดับสัดส่วนควรนำเสนอด้วยค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน [15, 18]

สำหรับการวัดการกระจายเป็นการวัดการกระจายของข้อมูลดิบจากค่าเฉลี่ย ในกรณีที่ข้อมูลมีความแปรปรวนมาก หมายถึงข้อมูลมีการกระจายมาก [1, 17] การวัดการกระจายพิจารณาได้จากค่าสถิติ ดังนี้ 1) ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation: SD) คือ การวัดการกระจายหรือความแปรปรวนของชุดข้อมูลในกลุ่มตัวอย่าง (Data variability) หรือความแปรปรวนยกกำลังสอง 2) สัมประสิทธิ์ความแปรปรวน (Coefficient of variation: CV) คือ การวัดการกระจายของแต่ละตัวแปรเทียบกับค่าเฉลี่ย โดยที่ค่าสัมประสิทธิ์ความแปรปรวนมีค่าใกล้ “0” หมายถึง มีการกระจายน้อย ซึ่งส่วนใหญ่ใช้ในการเปรียบเทียบความแตกต่างระหว่างชุดข้อมูล และ 3) ค่า Interquartile range (IQR): คือ ค่าความแตกต่างของข้อมูลด้วยระยะห่างในลำดับที่ 3 และลำดับที่ 1 [17] ซึ่งแสดงเป็นลักษณะของพิสัยที่ร้อยละ 50 การนำเสนอด้วยค่า Interquartile range เหมาะสำหรับข้อมูลที่มีการกระจายแบบอิสระหรือมีชุดข้อมูลที่มีค่าสูงหรือต่ำมากเกินไปกว่าปกติ [25]

สถิติเชิงอนุมานเป็นสถิติที่ใช้สำหรับการประมาณค่าและทดสอบสมมติฐาน [21] เพื่อเป็นการยืนยันองค์ความรู้หรือค้นพบโดยการจัดข้อโต้แย้ง [15]] และนำข้อค้นพบไปอธิบายหรือสรุปผลสู่กลุ่มประชากร ที่ทำการศึกษา

ทำการศึกษา (Inference หรือ Generalization) [1, 26] โดยอาศัยหลักของความน่าจะเป็นในการทดสอบ [3, 14] สถิติเชิงอนุมานจำแนกเป็นสถิติพารามेटริก (Parametric statistics) และสถิตินอนพารามेटริก (Non-parametric statistics) โดยที่สถิติพารามेटริกมีอำนาจการทดสอบ (Power of test) สูงกว่าสถิตินอนพารามेटริก [14] ซึ่งสถิติทั้ง 2 ประเภทมีข้อตกลงเบื้องต้นในการนำไปใช้ที่แตกต่างกัน [26] กล่าวคือ การทดสอบด้วยสถิติพารามेटริกข้อมูลต้องมีการวัดของข้อมูลในระดับอัตราส่วนหรืออัตราส่วน และมีการกระจายแบบโค้งปกติ (Normal distribution) เท่านั้น [4, 14] สำหรับสถิตินอนพารามेटริกนั้น ข้อมูลอยู่ในระดับอัตราส่วนหรืออัตราส่วนซึ่งมีรูปแบบการกระจายของข้อมูลไม่จำกัดรูปแบบ หรือมีระดับการวัดของข้อมูลในระดับนามบัญญัติหรือเรียงอันดับเท่านั้น

นอกจากนี้ ข้อตกลงเบื้องต้นในการเลือกใช้สถิตินอนพารามेटริกในการวิเคราะห์ข้อมูลมีดังนี้ 1) กลุ่มตัวอย่างมีจำนวนน้อย 2) ข้อมูลมีระดับการวัดแบบเชิงอันดับ 3) ไม่ทราบรูปแบบการกระจายของข้อมูล 4) ชุดข้อมูลมีค่าที่สูงหรือต่ำผิดปกติ (Outliers) และ 5) ความแปรปรวนระหว่างกลุ่มของกลุ่มตัวอย่างไม่เท่ากัน

การเลือกสถิติสำหรับการวิเคราะห์ข้อมูลทางวิทยาศาสตร์สุขภาพ

การทดสอบทางสถิติจำแนกตามจำนวนชุดของข้อมูล คือ ข้อมูล 1 กลุ่ม และข้อมูล ≥ 2 กลุ่ม นอกจากนี้ การทดสอบยังสามารถจำแนกออกเป็น [26] การทดสอบความแตกต่าง (Test of difference) การทดสอบความสัมพันธ์ (Test of relationship) และสถิติขั้นสูง (Advanced statistics) สำหรับบทความนี้ มุ่งเน้น การเลือกสถิติโดยการจำแนกจำนวนชุดของข้อมูลตัวแปรตาม ดังแสดงในภาพที่ 1

การทดสอบทางสถิติสำหรับข้อมูล 1 กลุ่ม แบ่งตามลักษณะของข้อมูลออกเป็น ตัวแปรที่มีระดับการวัดแบบเชิงอันดับทดสอบด้วยสถิติ Pearson Chi-square test และ Fisher exact test ในกรณีที่ตัวแปรที่มีระดับการวัดแบบต่อเนื่อง ซึ่งก่อนทำวิเคราะห์ข้อมูลหรือทดสอบทางสถิติควรมีการทดสอบการกระจายของข้อมูลทุกครั้งว่ามีการกระจายแบบโค้งปกติหรือไม่ด้วยสถิติ Kolmogorov-Smirnov goodness-of-fit-test [4] เพื่อใช้ในการตัดสินใจ สำหรับชุดข้อมูลที่มีการกระจายแบบโค้งปกติทำการทดสอบด้วยสถิติ One sample t-test และสถิติ Wilcoxon signed-rank test สำหรับข้อมูลที่ไม่มีการกระจายแบบโค้งปกติ

สำหรับการวิเคราะห์ข้อมูล 2 กลุ่มการเลือกสถิติในการวิเคราะห์ข้อมูลพิจารณาจากระดับการวัดของตัวแปรตาม และความเป็นอิสระของข้อมูล ในกรณีที่ตัวแปรตามมีระดับการวัดแบบเชิงอันดับและมีความเป็นอิสระต่อกันสถิติที่ใช้ Pearson chi-square test และ Fisher exact test นอกจากนี้ ในกรณีที่ข้อมูล 2 ชุด ไม่มีอิสระต่อกันสถิติที่ควรใช้ คือ McNemar test

อย่างไรก็ตาม สำหรับการวิเคราะห์ข้อมูลด้วยสถิติ Pearson chi-square test และ Fisher's exact test มีข้อตกลงเบื้องต้นในการทดสอบที่แตกต่างกัน [27] กล่าวคือ การเลือกใช้สถิติ Pearson chi-square test ควรพิจารณาของขนาดกลุ่มตัวอย่าง (Observed value) ในแต่ละ Cell สำหรับตาราง 2×2 หรือ $M \times N$ ไม่น้อยกว่า 5 และควรมีค่าความคาดหวัง (Expected value) ไม่น้อยกว่าร้อยละ 20 ในกรณีที่ไม่เป็นไปตามข้อตกลงดังกล่าวข้างต้นควรยุบตารางให้เหลือตาราง 2×2 และสถิติที่เหมาะสมในการวิเคราะห์ คือ สถิติ Fisher exact test

สำหรับตัวแปรตามที่มีระดับการวัดต่อเนื่อง (Continuous data) เบื้องต้นควรทดสอบการกระจายของข้อมูลทุกครั้งว่ามีการกระจายแบบโค้งปกติหรือไม่ [14] ในกรณีที่มีการกระจายแบบโค้งปกติสามารถทดสอบด้วยสถิติพารามेटริก คือ Independent t-test สำหรับข้อมูลที่มีความอิสระต่อกัน และสถิติ Paired t-test สำหรับข้อมูลไม่อิสระต่อกัน นอกจากนี้ ข้อมูลที่ไม่มีการกระจายแบบโค้งปกติสถิตินอนพารามेटริกที่ใช้ในการวิเคราะห์ คือ สถิติ Mann-Whitney U test สำหรับข้อมูลที่มีอิสระต่อกัน และสถิติ Wilcoxon signed-rank test สำหรับข้อมูลที่ไม่อิสระต่อกัน

ในกรณีข้อมูลจำนวน ≥ 2 กลุ่ม การพิจารณาเลือกสถิติสำหรับตัวแปรที่มีระดับการวัดแรงแงนับมีลักษณะเช่นเดียวกันกับที่กล่าวในข้างต้น โดยพิจารณาจากระดับการวัดของข้อมูล และความเป็นอิสระต่อกัน [14] สำหรับข้อมูลมีความอิสระต่อกันสถิติที่ใช้ คือ Pearson chi-square test และข้อมูลที่ไม่อิสระต่อกัน คือ Cochran Q test ในกรณีประเภทข้อมูลต่อเนื่องและข้อมูลมีการกระจายแบบโค้งปกติและไม่อิสระต่อกันสถิติที่ใช้ในการวิเคราะห์ คือ สถิติ Repeated ANOVA ส่วนสถิติ One-way ANOVA ใช้ในกรณีกลุ่มตัวอย่างมีความอิสระต่อกันนอกจากนี้ ในกรณีที่ข้อมูลไม่มีการกระจายแบบโค้งปกติและข้อมูลมีความอิสระต่อกันควรใช้สถิติ Friedman test และข้อมูลที่ไม่อิสระต่อกันใช้สถิติ Kruskal-Wallis test ในการทดสอบ

นอกจากนี้ การเลือกใช้สถิติในการทดสอบความสัมพันธ์และการทำนาย ซึ่งสถิติดังกล่าวมีข้อควรพิจารณาในการเลือกใช้สถิติเช่นเดียวกับการเลือกใช้สถิติในกรณีข้อมูล 1 กลุ่ม และ ≥ 2 กลุ่ม [14] โดยพิจารณาจากประเภทระดับการวัดของตัวแปรตาม และตัวแปรต้น ในกรณีการเลือกใช้สถิติในการทดสอบความสัมพันธ์และการทำนายสำหรับตัวแปรตามที่มีระดับการวัดประเภทแรงแงนับ [14] ส่วนตัวแปรต้นสามารถเป็นได้ทั้งตัวแปรแรงแงนับและตัวแปรต่อเนื่อง คือ Simple logistic regression analysis ในกรณีที่ตัวแปรตามเป็นประเภทต่อเนื่อง เมื่อพิจารณาตัวแปรต้นเป็นชนิดต่อเนื่องและ/หรือแรงแงนับสถิติที่ใช้ คือ Simple linear regression analysis

นอกจากนี้ การพิจารณาการเลือกใช้สถิติในการทดสอบความสัมพันธ์และการทำนายด้วยวิธีการทดสอบพหุตัวแปร (Multivariate analysis) [23, 28-29] ซึ่งวัตถุประสงค์ของการทดสอบดังกล่าวเพื่อเป็นการทดสอบความสัมพันธ์หรือการทำนายด้วยสมการถดถอย (Regression modeling) ระหว่างตัวแปรต้น (X_s) และตัวแปรตาม (Y) โดยมีการควบคุมตัวแปรกวน (Confounding factors หรือ Multicollinearity) ในกรณีมีระดับการวัดของตัวแปรตามประเภทต่อเนื่องควรใช้สถิติ Multiple regression analysis สำหรับการทดสอบด้วยสถิติ Multiple logistic regression analysis เป็นการทดสอบที่เหมาะสมกับตัวแปรตามที่มีระดับการวัดประเภทแรงแงนับ และเป็นชนิด 2 ตัวแปร (Dichotomous หรือ binary outcome) เท่านั้น

บทสรุป

การเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัยทางวิทยาศาสตร์สุขภาพมีองค์ประกอบในการพิจารณาหลายประการ คือ 1) ข้อมูล ตัวแปรและระดับการวัดของข้อมูล 2) การทดสอบทางสถิติ 3) ค่าถามวิจัย และ 4) ประเภทของสถิติซึ่งแต่ละองค์ประกอบมีความสำคัญและเชื่อมโยงสู่การเลือกใช้สถิติในการวิเคราะห์ข้อมูลทางวิทยาศาสตร์สุขภาพอย่างถูกต้องและเหมาะสม

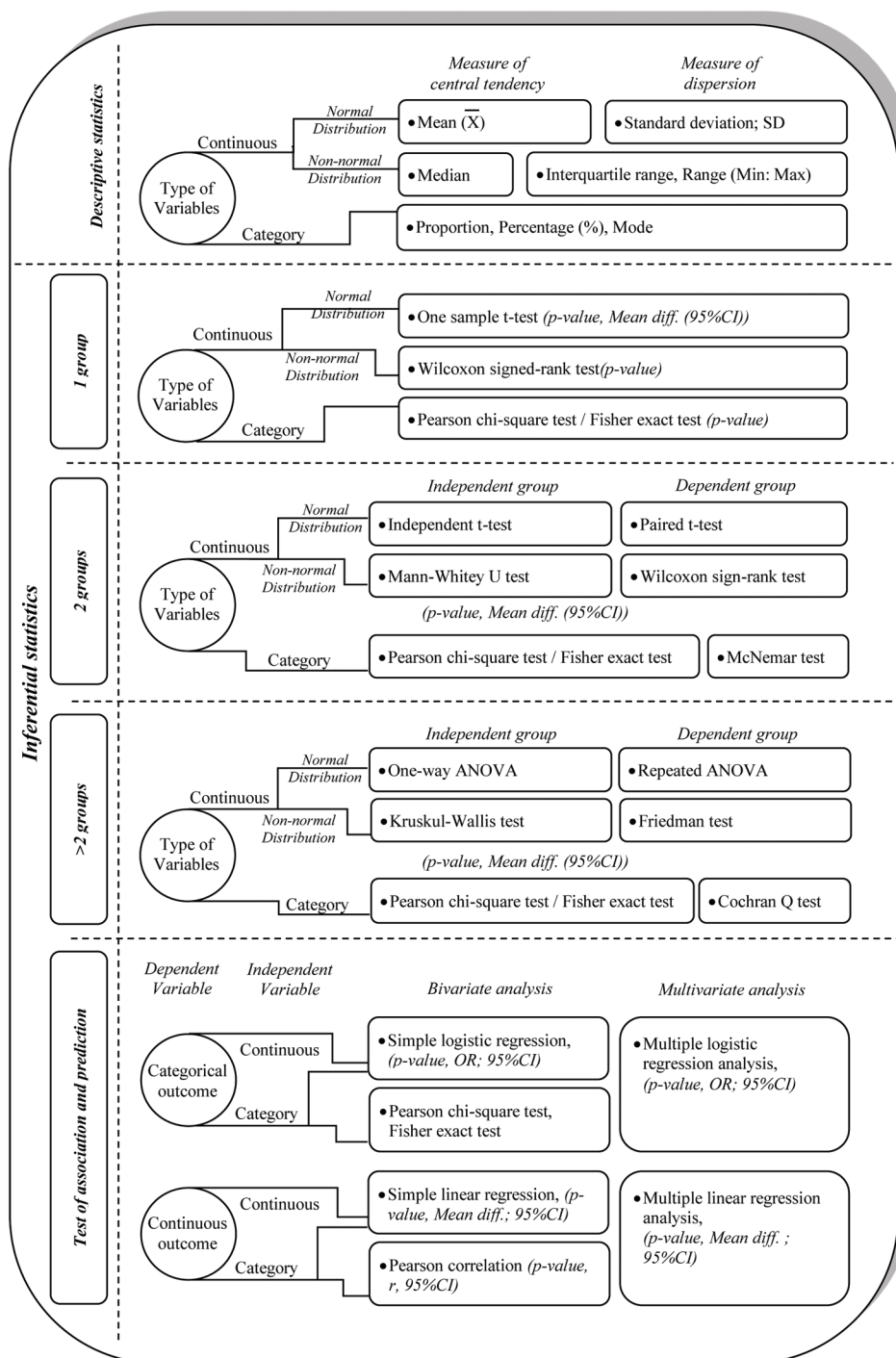
การดำเนินงานวิจัยผู้วิจัยควรพิจารณาและวางแผนในการเก็บข้อมูลอย่างละเอียด และทำความเข้าใจถึงวัตถุประสงค์การวิจัย ซึ่งเป็นประเด็นสำคัญที่ควรพิจารณาว่ามีวัตถุประสงค์เพื่อการใด เช่น วัตถุประสงค์เพื่อการพรรณนา เปรียบเทียบ ทดสอบความสัมพันธ์หรือการทำนาย นอกจากนี้ สถิติในการวิเคราะห์ข้อมูลจำแนกออกเป็น 2 ประเภท คือ สถิติพารามетริกและสถิตินอนพารามетริก ซึ่งสถิติทั้งสองประเภทมีความแตกต่างกัน โดยที่สถิติพารามетริกเป็นการทดสอบโดยใช้ข้อมูลที่มีการกระจายแบบโค้งปกติ และสถิติเป็นการทดสอบโดยใช้ข้อมูลที่ไม่มีการกระจายแบบโค้งปกติ ซึ่งแต่ละประเภทมีการเลือกใช้โดยขึ้นอยู่กับระดับการวัดของตัวแปรตามและข้อตกลงเบื้องต้นของสถิติแต่ละประเภท

ดังนั้น ผู้วิจัยควรมีการวางแผนในการเลือกใช้สถิติที่เหมาะสมในการวิเคราะห์ข้อมูล เพื่อเป็นประโยชน์ในการวิเคราะห์และนำเสนอข้อมูลที่ถูกต้องและเหมาะสม ต่อไป

เอกสารอ้างอิง

- [1] Perez-Vicente S, Exposito Ruitz M. (2009). Descriptive statistics. *Allergol Immunopathol (Madr)*. 2009; 37(6): 314-320.
- [2] เชิดศักดิ์ โฆวาสินธุ์. (2531). สถิติวิเคราะห์สำหรับการวิจัย. *วารสารการพัฒนการศึกษา*. 10(29): 7-19.
- [3] นิภา เมธาวีชัย. (2543). สถิติสำหรับการวิจัย. *วารสารเผยแพร่ความรู้ทางวิชาการและงานวิจัย*. 9: 123-150.
- [4] Song JW, Haas A, Chung KC. (2009). Application of statistical tests in hand surgery. *J Hand Surg*. 34: 1872-1881.
- [5] Cruess DF. (1989). Review of use of statistics in THE AMERICAN JOURNAL OF TROPICAL MEDICINE AND HYGIENE FOR JANUARY-DECEMBER 1988. *Am. J. Trop. Med. Hyg*. 41(6): 619-529.
- [6] Goldin J, Zhu W, Sayre JW. (1996). A review of the statistical analysis used in papers published in Clinical Radiology and British Journal of Radiology. *Clinical Radiology*. 51: 47 -50.
- [7] อรุณ จิรวัดน์กุล. (2549). การเขียนวิธีวิเคราะห์ทางสถิติในโครงร่างวิจัย. *วารสารวิชาการสาธารณสุข*. 15(5): 661-662.
- [8] Jaykaran. (2010). How to select appropriate statistical test? *J Pharm Negat Results*. 2010;1(2) :61-63.
- [9] Nam CM, Chung SY. (2012). Statistical methods for medical studies. *J Korean Med Assoc*. 55(6): 573 - 581.
- [10] du Prel JB, Röhrig B, Hommel G, Blettner M. (2010). Choosing statistical tests: part 12 of a series on evaluation of scientific publications. *Dtsch Arztebl Int*. 107(19): 343-348.
- [11] Soyemi K. (2012). Research and statistics: Choosing the right statistical test. *Pediatrics in Review*. 33(5): e38-e44.
- [12] บุญมี พันธุ์ไทย. (2553). การใช้วิธีการทางสถิติในการทำวิจัย. *วารสารรวมคำแหง*. 27(4): 37-52.
- [13] สมรัตน์ เลิศมหาฤทธิ์. (2543). สถิติในการวิจัย. ใน: ภิรมย์ กมลรัตนกุล, มนต์ชัย ขาลาประวรรตน์, ทวีสิน ต้นประยูร. *หลักการวิจัยให้สำเร็จ*. (พิมพ์ครั้งที่ 2). กรุงเทพฯ: บริษัท เท็กซ์ แอนด์ เจอร์นัล พับลิเคชั่น จำกัด.
- [14] Overholser BR, Sowinski KM. (2007). Biostatistics Primer: Part I. *Nutr Clin Pract*. 22(6): 629-635.
- [15] อนันต์ เกิดดำ. (2552). แนวทางการเลือกใช้สถิติในการวิจัย. *คอมพิวเตอร์และเทคโนโลยีขั้นสูง*. 10: 8-14.
- [16] บุญเรือง ศรีเหรียญ. (2542). การเลือกใช้สถิติในการวิจัย. *วารสารเพชรบุรีวิทยาลัยการณ.*: 28-39.
- [17] Fisher MJ, Marshall AP. (2009). Understanding descriptive statistics. *Australian Critical Care*. 22: 93-97.
- [18] Thompson CB. (2009). Descriptive data analysis. *Air Medical Journal*. 28(2): 56-59.
- [19] อรุณ จิรวัดน์กุล. (2548). การเลือกตัวอย่าง และการอนุมานด้วยสถิติ. *วารสารวิชาการสาธารณสุข*. 14(5): 739-740.
- [20] อรุณ จิรวัดน์กุล, มาลินี เหล่าไพบุลย์, จิราพร เขียวอยู่, ยุพา ถาวรพิทักษ์, จารุวรรณ โชคคณาพิทักษ์, บัณฑิต ถิ่นคำพร และคณะ. (2550). *ชีวสถิติ*. (พิมพ์ครั้งที่ 4). ขอนแก่น: โรงพิมพ์คลังน่านวิทยา.
- [21] สมเกียรติ จงประสิทธิ์พร. (2537). การนำเสนอข้อมูลด้วยสถิติ. *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*. 4(1): 10-25.
- [22] Wissing DR, Timm D. (2012). Statistics for the nonstatistician: Part I. *Southern Medical Association*. 105(3): 126-130.

- [23] Kocher MS, Zurakowski D. (2004). Clinical epidemiology and biostatistics: A primer for orthopaedic surgeons. *J Bone Joint Surg Am.* 86: 607-620.
- [24] มณีรัตน์ ชีระวิวัฒน์. (2543). เทคนิคการเลือกสถิติวิเคราะห์ข้อมูลในงานวิจัยพฤติกรรมสุขภาพ. *วารสารเพื่อนสุขภาพ.* 12(1): 54-57.
- [25] O'Brien PC. Shampo MA. Robertson JS. (1983). Statistics for nuclear medicine: introduction, descriptive statistics and graphic displays. *The Journal of Nuclear Medicine.* 24(1): 83-88.
- [26] วรณชนก จันทชุม. (2541). การใช้สถิติวิเคราะห์ข้อมูลในการวิจัยทางการแพทย์. *วารสารคณะพยาบาลศาสตร์.* 21(4): 1-8.
- [27] บุญญพัฒน์ ไชยเมธ. (2553). ข้อควรระวังในการใช้ Chi-square Test ในงานวิจัยทางด้านวิทยาศาสตร์สุขภาพ. *วารสารมหาวิทยาลัยทักษิณ.* 13(1): 55-58
- [28] Leeper JD. The University of California at Los-angles. **What statistical analysis should I use?** Retrieved 11 November 2012 from http://www.ats.ucla.edu/stat/mult_pkg/whatstat/default.htm.



ภาพที่ 1 การพิจารณาเลือกใช้สถิติในการวิเคราะห์ข้อมูลทางวิทยาศาสตร์สุขภาพ