

การหาปัจจัยเสี่ยงของการเกิดโรคหรือปัญหาสุขภาพ ด้วยเทคนิคการวิเคราะห์ถดถอยโลจิสติกเชิงพหุคูณแบบทวินาม

Determination of Risk Factors or Health Issues Utilizing Multiple Binary Logistic Regression Analysis

ถนอม นามวงศ์¹Thanom Namwong¹นริศรา อารีรักษ์²Narisara Arrirak²¹วิทยาลัยการสาธารณสุขสิรินธร จังหวัดอุบลราชธานี¹Sirindhorn College of Public Health, Ubon

คณะสาธารณสุขศาสตร์และสหเวชศาสตร์

Ratchathani Province, Faculty of Public

สถาบันพระบรมราชชนก

Health and Allied Health Sciences,

Praboromarajchanok Institute

²สำนักงานสาธารณสุขจังหวัดยโสธร²Yasothon Provincial Health Office

DOI: 10.14456/dcj.2025.47

Received: July 23, 2025 | Revised: September 30, 2025 | Accepted: October 6, 2025

บทคัดย่อ

การเกิดโรคหรือปัญหาสุขภาพขึ้นอยู่กับหลายปัจจัย การวิเคราะห์ข้อมูลที่ไม่ได้คำนึงถึงปัจจัยต่างๆ อย่างรอบด้าน อาจนำไปสู่ความรู้หรือข้อสรุปที่ผิดพลาด ส่งผลเสียต่อการวางแผนเพื่อดำเนินมาตรการอย่างมีประสิทธิภาพ การวิเคราะห์ถดถอยโลจิสติกแบบทวินาม (Binary Logistic Regression) เป็นวิธีการทางสถิติที่นิยมใช้มากสำหรับวิเคราะห์ข้อมูลที่ตัวแปรตามเป็นตัวแปรกลุ่มหรือตัวแปรแจกแจง (Categorical variable) ที่มีเพียง 2 กลุ่ม (Dichotomous) ส่วนตัวแปรต้นเป็นตัวแปรกลุ่ม (Categorical variable) ที่อาจมีมากกว่า 2 กลุ่ม หรือเป็นตัวแปรต่อเนื่อง (Continuous variable) หากวิเคราะห์ถดถอยโลจิสติกแบบทวินามที่คำนึงถึงผลกระทบจากตัวแปรอื่นเรียกว่าการวิเคราะห์ถดถอยโลจิสติกเชิงพหุคูณแบบทวินาม (Multiple Binary Logistic Regression) มีเป้าหมายในการอธิบายความสัมพันธ์ระหว่างตัวแปรต้นหลายตัวกับตัวแปรตาม ที่บ่งชี้ขนาดของความเสี่ยงหรือขนาดของความสัมพันธ์ โดยคำนึงถึงอิทธิพลของปัจจัยอื่น ๆ หรือการควบคุมตัวแปรกวน ผลจากการวิเคราะห์ข้อมูล จะทำให้ได้ผลการศึกษาที่ถูกต้อง และน่าเชื่อถือ สามารถนำผลการศึกษาไปวางแผน ส่งเสริม ป้องกัน และควบคุมโรคได้อย่างมีประสิทธิภาพ

ติดต่อผู้พิมพ์: ถนอม นามวงศ์

อีเมล: thanom@scphub.ac.th

Abstract

The occurrence of diseases or health problems is influenced by multiple factors. Analyzing data without considering all relevant variables may lead to inaccurate conclusions, which can negatively affect the planning of effective interventions. Binary logistic regression is a widely used statistical method for analyzing

data when the dependent variable is dichotomous, while independent variables may be categorical (with more than two levels) or continuous. When the dependent variable still has two categories, but the model includes two or more independent variables, it is referred to as multiple binary logistic regression analysis. The method aims to describe the relationship between independent and dependent variables, indicating the magnitude of risk or association, while controlling confounding factors. Accurate and reliable results from such analyses can provide valuable evidence to support the planning, promotion, prevention, and control of diseases.

Correspondence: Thanom Namwong

E-mail: thanom@scphub.ac.th

คำสำคัญ

ปัจจัยเสี่ยง, ปัญหาสุขภาพ, ถดถอยโลจิสติก

Keywords

Risk Factor, Health Issue, Logistic Regression

บทนำ

การวิเคราะห์หาปัจจัยเสี่ยงของการเกิดโรคหรือปัญหาสุขภาพเป็นขั้นตอนพื้นฐานที่สำคัญในการวางแผนส่งเสริม ป้องกัน และควบคุมโรคอย่างมีประสิทธิภาพ เนื่องจากจะสามารถระบุได้ว่าปัจจัยใดมีความสัมพันธ์กับการเกิดโรคหรือปัญหาสุขภาพทั้งในเชิงพฤติกรรมส่วนบุคคล ปัจจัยทางสังคม เศรษฐกิจ และสิ่งแวดล้อม การเข้าใจกลไกของปัจจัยเสี่ยงเหล่านี้จะช่วยให้ผู้ปฏิบัติงานด้านสาธารณสุขสามารถกำหนดมาตรการเชิงรุกได้อย่างตรงจุด ผลการดำเนินงานบรรลุเป้าหมายในระดับชุมชน การวิเคราะห์ปัจจัยเสี่ยงมีบทบาทในการวินิจฉัยปัญหาสุขภาพเฉพาะพื้นที่ โดยสะท้อนถึงบริบทเฉพาะด้านวัฒนธรรม วิถีชีวิตหรือปัจจัยแวดล้อมของประชาชน ทำให้สามารถพัฒนาแผนงานโครงการที่สอดคล้องกับความต้องการและสภาพปัญหาจริงของชุมชน ช่วยเสริมสร้างการมีส่วนร่วมของประชาชนอย่างยั่งยืน และลดการใช้ทรัพยากรโดยไม่จำเป็น⁽¹⁾ แม้ว่า การวิเคราะห์หาปัจจัยเสี่ยงจะเป็นที่ยอมรับและใช้กันอย่างแพร่หลาย แต่หลาย ๆ การศึกษาที่ผ่านมาก็ยังมีความจำกัด โดยเฉพาะการใช้สถิติพื้นฐานในการวิเคราะห์ข้อมูลที่ไม่สามารถควบคุมปัจจัยแทรกซ้อนหรือไม่ได้คำนึงถึงปัจจัยอื่น^(2,3) ซึ่งปัจจัยเสี่ยงของการเกิดโรคหรือปัญหาสุขภาพนั้น มีอยู่หลายปัจจัย การวิเคราะห์ข้อมูลที่ไม่ได้คำนึงถึงปัจจัยต่างๆ อย่างรอบด้าน อาจนำไปสู่ความรู้หรือข้อสรุปที่ผิดพลาด ส่งผลต่อการ

วางแผนส่งเสริมสุขภาพ ป้องกัน และควบคุมโรคอย่างมีประสิทธิภาพ⁽⁴⁾

บทความนี้จะนำเสนอหลักการและวิธีการวิเคราะห์ข้อมูล เพื่อหาปัจจัยเสี่ยงของการเกิดโรคหรือปัญหาสุขภาพด้วยเทคนิคการวิเคราะห์ถดถอยโลจิสติกแบบพหุ พร้อมทั้งยกตัวอย่างประกอบ เพื่อให้บุคลากรทางการแพทย์และสาธารณสุข นำไปประยุกต์ใช้ได้เหมาะสม

การวิเคราะห์ถดถอยโลจิสติก

การวิเคราะห์ถดถอยโลจิสติก (Logistic Regression) มี 2 ประเภท คือ การวิเคราะห์การถดถอยโลจิสติกพหุนาม (Multinomial Logistic Regression) เป็นการวิเคราะห์ที่ตัวแปรตามที่มีค่ามากกว่า 2 กลุ่ม (Polychromatic variable) และการวิเคราะห์การถดถอยโลจิสติกแบบทวินามหรือแบบไบนารี (Binary Logistic Regression; BLR) ที่ตัวแปรตามมี 2 กลุ่ม ทั้งนี้ หากตัวแปรตามมีหลายตัวและตัวแปรตามมี 1 ตัว และแบ่งออกเป็น 2 กลุ่ม จะเรียกว่าการวิเคราะห์ถดถอยโลจิสติกเชิงพหุคูณแบบทวินาม (Multiple Binary Logistic Regression; MBLR) ซึ่งเป็นวิธีการทางสถิติที่นิยมใช้มากในการวิจัยทางการแพทย์และสาธารณสุข มักพบในรูปแบบการวิจัยที่เป็นเชิงวิเคราะห์ (Analytic study) ที่มีกลุ่มเปรียบเทียบ เช่น Cohort study, Case-control study และ Cross-sectional analytic study ช่วยวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรต้นหรือตัวแปร

อิสระ (Independent Variable) หลายตัว ซึ่งอาจเป็นปัจจัยเสี่ยง (Risk factor/Exposure/Predictor) หรือปัจจัยกำหนดสุขภาพ (Health determinant) กับตัวแปรตาม (Dependent Variable) หรือผลลัพธ์ (Outcome/Response variable/event) ที่เป็นลักษณะการเกิดหรือไม่เกิดโรค หรือการเกิดหรือไม่เกิดปัญหาสุขภาพ โดยสามารถควบคุมปัจจัยแทรกซ้อน (Control confounder) และระบุความเสี่ยง (Odds ratio; OR) หรือขนาดของความเสี่ยงหรือความสัมพันธ์ (Magnitude of association) ได้อย่างมีประสิทธิภาพ ทำให้การวิเคราะห์มีความแม่นยำมากขึ้น ช่วยลดอคติ (bias) ในการแปลผล และส่งเสริมการวางแผนโดยใช้หลักฐานเชิงประจักษ์ (evidence-based planning) ได้ดียิ่งขึ้น^(5,6)

วัตถุประสงค์การวิเคราะห์ MBLR

การวิเคราะห์ MBLR มีวัตถุประสงค์ 2 ข้อ คือ 1) เพื่อหาปัจจัยเสี่ยง (Risk assessment) และ 2) เพื่อสร้างสมการทำนายเหตุการณ์ (Prediction) บทความนี้กล่าวถึงเฉพาะการหาปัจจัยเสี่ยง ประเด็นหลักของการวิเคราะห์โดยใช้ MBLR นั้นคือการอธิบายความสัมพันธ์ระหว่างตัวแปรต้นกับตัวแปรตามที่มีการควบคุมผลกระทบของปัจจัยอื่น ๆ ซึ่งแบบจำลองทางคณิตศาสตร์ที่เหมาะสม คือ Logistic Model (ความน่าจะเป็นที่จะเกิดโรค) การใช้ข้อมูลที่มีอยู่เพื่อประมาณค่าพารามิเตอร์เรียกว่าการ “Fit Model” วิธีการทางคณิตศาสตร์ที่ใช้ในการประมาณค่าพารามิเตอร์เรียกว่า Maximum Likelihood (ค่าพารามิเตอร์ เช่น ค่าคงที่ (Constant) และค่าสัมประสิทธิ์ (Coefficient))⁽⁵⁾

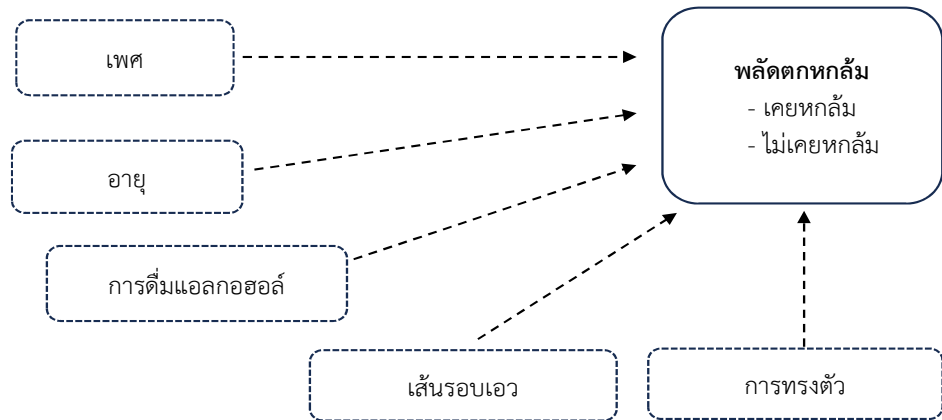
ตัวแปรต้นและตัวแปรตามของการวิเคราะห์ MBLR

การวิเคราะห์ MBLR เป็นวิธีการหนึ่งของการวิเคราะห์ตัวแปรพหุ (Multivariable analysis) ใช้สำหรับวิเคราะห์ข้อมูลที่ตัวแปรตามเป็นตัวแปรกลุ่มหรือตัวแปรแจกแจง (Categorical variable) มีค่าได้เพียง 2 กลุ่ม (Dichotomous or Binary) เช่น ป่วย-ไม่ป่วย หาย-ไม่หาย เกิด-ไม่เกิดเหตุการณ์ ส่วนตัวแปรต้นเป็นตัวแปรกลุ่ม (Categorical variable) ที่อาจมีมากกว่า 2 ค่า หรือเป็นตัวแปรต่อเนื่อง (Continuous variable) ก็ได้ กรณีตัวแปรต้นเป็นแบบต่อเนื่อง (continuous variable) จะวิเคราะห์หาความสัมพันธ์เชิงเส้น (linear) กับตัวแปรตาม เพื่อเป็นแนวทางการตัดสินใจว่าจะนำเข้าสู่การวิเคราะห์แบบตัวแปรต่อเนื่องหรือแบบตัวแปรกลุ่ม ถ้ามีความสัมพันธ์เชิงเส้นก็จะนำเข้าสู่การวิเคราะห์แบบข้อมูลต่อเนื่อง อย่างไรก็ตาม เพื่อให้ง่ายต่อการแปลความหมาย และการนำไปใช้ประโยชน์ทางการแพทย์และสาธารณสุข มักนำเข้าสู่การวิเคราะห์โดยจัดเป็นกลุ่ม เช่น การนำตัวแปรอายุที่เป็นตัวแปรต่อเนื่องไปจัดเป็นกลุ่ม (Classify data) แบ่งเป็นช่วง ๆ หรือกลุ่มอายุ

ยกตัวอย่าง เช่น การวิเคราะห์ปัจจัยเสี่ยงต่อการพลัดตกหกล้มในผู้สูงอายุ ดังนี้

ตัวแปรต้น คือ เพศ (มี 2 ค่า คือ ชาย หญิง) อายุ (ปี มี 3 ค่า/กลุ่มอายุ) การดื่มแอลกอฮอล์ (มี 2 ค่า คือ ไม่ดื่ม ดื่ม) เส้นรอบเอว (มี 2 ค่า คือ ต่ำ-ปกติ สูงกว่าปกติ) และการทรงตัว (มี 2 ค่า คือ ปกติ บกพร่อง)

ตัวแปรตาม คือ การพลัดตกหกล้ม (มี 2 ค่า คือ เคยหกล้ม และ ไม่เคยหกล้ม)



ภาพที่ 1 ตัวอย่างการวิเคราะห์ปัจจัยเสี่ยงต่อการพลัดตกหกล้มในผู้สูงอายุด้วย Logistic Regression

Figure 1 Example of risk factors analysis for falls in the elderly using logistic regression

ขั้นตอนการวิเคราะห์ข้อมูล

ขั้นตอนการวิเคราะห์ข้อมูล มี 3 ขั้นตอน ได้แก่

1) การเตรียมข้อมูล 2) การวิเคราะห์ข้อมูล และ 3) การนำเสนอและแปลผลการวิเคราะห์ข้อมูล โดยมีรายละเอียดแต่ละขั้นตอนดังนี้

1. การเตรียมข้อมูล

1.1 ตรวจสอบข้อมูล (Clean data) เช่น ค่าที่ผิดปกติ (Outlier) ค่าว่าง (Missing) ซึ่งอาจจะต้องมีการตรวจสอบกับแบบสอบถามหรือแหล่งข้อมูลเพื่อยืนยันความถูกต้อง หรือใช้เทคนิคการจัดการข้อมูลที่เหมาะสมตามหลักที่มีเหตุผลทางวิชาการรองรับ

1.2 การจัดการข้อมูล (Modify data) เช่น การยุบรวมตัวแปร การจัดกลุ่มตัวแปร ซึ่งจะต้องมีเหตุผลทางวิชาการรองรับ หรือมีเป้าหมายในการวิเคราะห์ข้อมูลเพื่อนำไปใช้ประโยชน์ ในการวิเคราะห์ด้วย Logistic Regression กรณีตัวแปรต้นเป็นแบบต่อเนื่องจะใช้การวิเคราะห์หาความสัมพันธ์เชิงเส้น (linear) กับตัวแปรตามเพื่อเป็นแนวทางการตัดสินใจว่าจะนำเข้าสู่การวิเคราะห์แบบตัวแปรต่อเนื่องหรือแบบตัวแปรกลุ่ม ดังที่กล่าวถึงข้างต้น

2. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล แบ่งออกเป็น 6 ขั้นตอนย่อย ได้แก่

- 1) การวิเคราะห์อย่างหยาบ (Crude analysis)
- 2) การเลือกตัวแปรเริ่มต้นเพื่อนำเข้าสู่การวิเคราะห์แบบหลายตัวแปร (Initial Model)
- 3) การตรวจสอบข้อตกลง (Assumption) ในการใช้ MBLR
- 4) การวิเคราะห์แบบหลายตัวแปร (Multi-variable analysis)
- 5) การตรวจสอบหรือประเมินความเหมาะสมของการวิเคราะห์ข้อมูลหรือ Model (Goodness of fit)
- 6) การประเมินประสิทธิภาพของสมการหรือ Model

โดยมีรายละเอียดดังนี้

2.1 การวิเคราะห์อย่างหยาบ

การวิเคราะห์อย่างหยาบ (Crude analysis) บางคนเรียก Bivariate analysis เพราะเป็นการทดสอบความสัมพันธ์ (test of association) ของตัวแปรเพียงสองตัว หรือตัวแปรต้นกับตัวแปรตามที่จะดู โดยไม่คำนึงถึงผลกระทบของปัจจัยอื่น เป็นขั้นตอนแรกที่ต้องทำในทุกกรณีการศึกษา การวิเคราะห์ BLR มีข้อตกลงที่สำคัญ ได้แก่ 1) การสังเกตเป็นอิสระจากกัน ซึ่งไม่ควรวัดซ้ำในบุคคลเดียวกัน หรือมีความเกี่ยวข้องกันทางใดทางหนึ่ง วิธีทดสอบโดยการสร้างกราฟค่าตกค้างเทียบกับลำดับ (residuals vs. order plot) ซึ่งรูปแบบต้องเป็น

แบบลุ่ม และ 2) ไม่มีค่าผิดปกติอย่างรุนแรง extreme outliers ในตัวแปรที่ศึกษา วิธีทดสอบ เช่น การเรียงลำดับข้อมูล การทำกราฟ box plots หรือ histogram⁽⁶⁾

การวิเคราะห์ความสัมพันธ์อย่างหยาบอาจทำได้ 2 แบบ คือ 1) วิเคราะห์ถดถอยแบบทวิอย่างง่ายด้วย Simple Binary Logistic Regression (SBLR) โดยทดสอบระดับนัยสำคัญด้วย Likelihood Ratio Test และ 2) ใช้สถิติไค-สแควร์ (χ^2 -test) เว้นแต่เมื่อมีเซลล์มีค่า Expected value น้อยกว่า 5 เกินร้อยละ 20 ของจำนวนเซลล์ทั้งหมด ให้ใช้ Fisher’s Exact test แทน ผลจากการวิเคราะห์จะได้ขนาดความสัมพันธ์ (Measure of association) คือ Relative Risk (RR) ถ้าเป็นการศึกษาแบบ Cohort study หรือใช้ Odds Ratio (OR) ถ้าเป็นการศึกษาแบบ Cross-sectional หรือ Case-control Study หรือ Cohort study ค่า Odds ratio เป็นอัตราส่วนที่บ่งบอกถึงความสัมพันธ์ หรือขนาดของความสัมพันธ์ของปัจจัยหรือตัวแปรต้นที่คาดว่าน่าจะส่งผลต่อตัวแปรตาม ซึ่งนำเสนอค่า OR จากการวิเคราะห์อย่างหยาบ ไม่ว่าจะเป็นการศึกษารูปแบบใด

จะมีข้อดีในการเปรียบเทียบขนาดของความสัมพันธ์ที่ได้จากการวิเคราะห์แบบหลายตัวแปร⁽⁵⁾ (รายละเอียดเพิ่มเติมในหัวข้อ 2.4)

การเลือกกลุ่มอ้างอิง หรือ reference group ในการวิเคราะห์ BLR ในการศึกษาแบบ Case-control หรือ Cohort study นิยมให้กลุ่มที่มีความชุกหรืออุบัติการณ์ของโรคต่ำเป็นกลุ่มอ้างอิง เช่น การศึกษาว่าเพศเป็นปัจจัยเสี่ยงของการพลัดตกหกล้มหรือไม่ ซึ่งส่วนใหญ่เพศชายจะมีความชุกหรืออุบัติการณ์ต่ำกว่าเพศหญิง เพศชายจึงควรกำหนดให้เป็นกลุ่มอ้างอิง และค่า OR ที่ได้ก็จะมากกว่า 1 และง่ายต่อการแปลผล

ทั้งนี้ เพื่อให้มีความเข้าใจเกี่ยวกับขั้นตอนการวิเคราะห์และการแปลผล ขอยกตัวอย่างการวิเคราะห์ปัจจัยที่สัมพันธ์กับการพลัดตกหกล้มในผู้สูงอายุ รูปแบบการวิจัยเป็นเชิงวิเคราะห์แบบภาคตัดขวาง (Cross-sectional analytic study) ในกลุ่มตัวอย่างที่เป็นผู้สูงอายุ 18,818 ราย ซึ่งได้มาจากฐานข้อมูลการคัดกรองผู้สูงอายุ (Secondary data) ของจังหวัดแห่งหนึ่ง ผลของการวิเคราะห์อย่างหยาบดังตารางที่ 1

ตารางที่ 1 การวิเคราะห์ปัจจัยที่สัมพันธ์กับการพลัดตกหกล้มในผู้สูงอายุด้วยการวิเคราะห์อย่างหยาบ หรือ การวิเคราะห์ทวินามอย่างง่าย
Table 1 Analysis of factors associated with falls in the elderly using crude analysis Simple binary logistic regression

ตัวแปร	N (18,818)	เคยพลัดตกหกล้ม ใน 6 เดือนที่ผ่านมา		COR	95% CI COR	p-value
		จำนวน	%			
เพศ						<0.001*
ชาย	8,326	506	6.1	ref		
หญิง	10,492	902	8.6	1.5	1.3, 1.6	
อายุ (ปี)						
60-69	10,467	652	6.2	ref		
70-79	6,037	507	8.4	1.4	1.2, 1.6	<0.001*
80 ขึ้นไป	2,314	247	10.8	1.8	1.6, 2.1	<0.001*
การดื่มแอลกอฮอล์						0.862
ไม่ดื่ม	15,149	1,131	7.5	ref		
ดื่ม	3,669	277	7.6	1.1	0.8, 1.2	
เส้นรอบเอว						0.13*
ต่ำ-ปกติ	8,545	684	8.0	ref		
สูงกว่าปกติ	10,273	724	7.1	0.8	0.8, 0.9	
การทรงตัว						<0.001*
ปกติ	13,352	759	5.7	ref		
บกพร่อง	5,466	649	11.9	2.2	2.0, 2.5	

* ตัวแปรที่จะนำเข้าสู่การวิเคราะห์แบบหลายตัวแปร ที่มีค่า p-value<0.25 ตามคำแนะนำของ Hosmer & Lemeshow⁽⁶⁾

การวิเคราะห์ความสัมพันธ์อย่างหยาบทำได้ทั้งการใช้ตารางการณักร (Contingency table) ในรูปตาราง 2x2 และการใช้ SBLR คือมีตัวแปรต้นเพียงตัวเดียวกับตัวแปรตาม อย่างไรก็ตาม การใช้ตาราง 2x2 ไม่สามารถใช้กรณีที่ตัวแปรต้นมีค่ามากกว่า 2 ค่า เช่น ในตารางที่ 1 ตัวแปรอายุ ซึ่งมี 3 ค่ากรณีนี้ต้องใช้ SBLR ในการวิเคราะห์ ค่าสถิติที่ได้จากการวิเคราะห์ความสัมพันธ์อย่างหยาบ เรียกว่า crude OR (COR) การเลือกกลุ่มอ้างอิง หรือ reference group ในการวิเคราะห์เพื่อให้ได้ค่า COR ที่ง่ายต่อการแปลผล คือ การเลือกกลุ่มอ้างอิงที่มีความเสี่ยงต่ำๆ เช่น ในตารางที่ 1 ผู้สูงอายุที่อายุน้อยจะเสี่ยงต่อการพลัดตกหกล้มน้อยกว่าคนที่อายุมาก ดังนั้น กลุ่มอายุ 60-69 ปี จึงเป็นกลุ่มอ้างอิงของกลุ่ม 70-79 ปี และ 80 ปีขึ้นไป และค่า COR ที่ได้ก็จะมากกว่า 1 และแปลผลได้ว่า คนที่อายุ 70-79 ปี มีความเสี่ยงมากกว่าคนที่ อายุ 60-69 ปี อยู่ 1.4 เท่า ทั้งนี้ นอกจากค่า COR และค่า p -value แล้ว ยังมีค่า 95% CI (confident interval) ที่มีความสำคัญในการแปลผล เป็นค่าที่ใช้ประมาณค่าไปหาประชากรแบบช่วง มี 2 ค่า คือ ค่าที่อยู่ช่วงล่าง (Lower) และช่วงบน (Upper) ค่า 95% CI มีความหมาย คือ การบ่งบอกว่าตัวแปรนั้นมีนัยสำคัญทางสถิติ คือ และเป็นช่วงที่คาดว่าค่าพารามิเตอร์จริงในประชากรจะอยู่ในช่วงนี้ 95 ครั้ง ถ้าทำการทดลองแบบเดียวกันนี้ 100 ครั้ง ค่า 95% CI ที่มีช่วงแคบ สะท้อนความแม่นยำในการทำนาย และขนาดตัวอย่างที่เพียงพอ ผลที่บอกว่ามีนัยสำคัญทางสถิติ (p -value<0.05) อาจมีค่า OR ที่ต่ำมาก ๆ ดังนั้น ค่า 95% CI ของ OR จึงควรแสดงไว้ทุกครั้งที่นำเสนอผลการวิเคราะห์ข้อมูล ค่า OR จะเป็นการประมาณค่าแบบจุด และจะอยู่ระหว่างค่า 95% CI OR เสมอ บางครั้งการนำเสนอข้อมูลในตารางอาจไม่นำเสนอค่า p -value แต่จะนำเสนอค่า 95% CI แทน เพราะค่า 95% CI OR ที่ไม่คร่อม 1 แสดงว่าตัวแปรต้นนั้นมีความสัมพันธ์กับตัวแปรตามอย่างมีนัยสำคัญทางสถิติ⁽⁵⁾

2.2 การเลือกตัวแปรเริ่มต้นเพื่อนำเข้าสู่การวิเคราะห์แบบ MBLR

เป้าหมายหนึ่งที่สำคัญของการวิเคราะห์อย่างหยาบ คือ การเลือกตัวแปรเข้าสู่การวิเคราะห์แบบหลายตัวแปร (Multivariable analysis) ในโมเดลเริ่มต้น (Initial regression model) โดยพิจารณาจากตัวแปรที่มีค่า p -value น้อยกว่า 0.25 ตามคำแนะนำของ Hosmer & Lemeshow⁽⁶⁾ ซึ่งจากตารางที่ 1 ตัวแปรต้นที่จะนำเข้าสู่การวิเคราะห์แบบหลายตัวแปร มีจำนวน 4 ตัวแปร คือ เพศ อายุ เส้นรอบเอว และการทรงตัว ส่วนตัวแปรการดื่มแอลกอฮอล์จะทำการคัดออก เพราะมีค่า p -value มากกว่า 0.25 ดังนั้น การวิเคราะห์อย่างหยาบ ต้องวิเคราะห์ตัวแปรต้นทุกตัวกับตัวแปรตาม ทั้งนี้ การเลือกตัวแปรต้นที่พิจารณาจากค่า p -value ด้วยการวิเคราะห์ความสัมพันธ์อย่างหยาบ รวมทั้ง จากการทบทวนวรรณกรรมที่เกี่ยวข้องที่ระบุว่าตัวแปรนั้นเป็นตัวแปรที่มีความสัมพันธ์กับตัวแปรตามหรือเป็นตัวแปรกวน หรืออาจเป็นตัวแปรที่เราสนใจที่มีแผนเพื่อนำผลการวิจัยไปใช้ ทั้งนี้ เพื่อให้การเลือกตัวแปรเข้าสู่การวิเคราะห์ในโมเดลเริ่มต้นมีความเหมาะสมที่สุด และเกิดประโยชน์จากการวิเคราะห์ข้อมูลสูงสุด

2.3 การตรวจสอบข้อตกลง (Assumption) ในการวิเคราะห์แบบ MBLR

ข้อตกลงเบื้องต้นการวิเคราะห์ Logistic regression ที่สำคัญ คือ 1) ตัวแปรต้นหรือตัวแปรอิสระไม่มีความสัมพันธ์กัน (multicollinearity) วิธีการของ Hosmer & Lemeshow⁽⁶⁾ แนะนำการทดสอบความสัมพันธ์กันของตัวแปรต้นโดยการหาค่า VIF (Variance Inflation Factor) ซึ่งค่า VIF ต้องไม่เกิน 10 วิธีการของ Burns and Grove⁽⁷⁾ ใช้สถิติ Pearson correlation ดูจากค่า r ไม่เกิน 0.65 และถ้าใช้เกณฑ์ ของ Stevens⁽⁸⁾ ใช้ค่า r ไม่เกิน 0.80 ซึ่งในบทความนี้เลือกวิธีโดยการหาค่า VIF และ 2) ขนาดตัวอย่างต้องเพียงพอ สูตรการคำนวณขนาดตัวอย่างที่นิยมใช้ คือ ของ Hsieh et al⁽⁹⁾ หรือใช้สูตรการคำนวณขนาดตัวอย่างตามกฎ Rule of Thumb ที่กำหนดให้ตัวแปรต้น 1 ตัวต่อ 10 เหตุการณ์

หรือผู้ป่วย ยกตัวอย่าง เช่น ถ้าตัวแปรต้นที่เราจะทำการวิเคราะห์ด้วย Multivariable analysis มีอยู่ 7 ตัวแปร กลุ่มตัวอย่างต้องมีผู้ป่วยหรือที่มีเหตุการณ์ หรือ events ที่เราสนใจ อย่างน้อย 70 ราย⁽¹⁰⁾

2.4 การวิเคราะห์แบบหลายตัวแปร

การวิเคราะห์แบบหลายตัวแปร (Multivariable analysis) เรียกว่า MBLR ในกรณีที่ไม่มี การจับคู่ (Un-Matched case control) บางคนนิยมเรียกว่า Un-conditional MBLR การวิเคราะห์โดยใช้ MBLR ที่มีการควบคุมอิทธิพลหรือผลกระทบของตัวแปรอื่น ๆ ใน Model ค่าสถิติที่ได้จึงเรียกว่าเป็น Adjusted OR (AOR) เมื่อได้ตัวแปรเริ่มต้นจากการวิเคราะห์อย่างหยาบ (Crude analysis) แล้วจะมีการทดสอบความสัมพันธ์กันของตัวแปรต้น (Multicollinearity) เพื่อหาค่า VIF (Variance inflation factor) หรือค่า r ซึ่งต้องไม่เกินที่กำหนดตามข้อตกลงเบื้องต้น จากนั้นจะทำการคัดเลือกตัวแปรเข้าสู่การวิเคราะห์หรือเรียกว่าการ fit model

การ fit model ซึ่งมีอยู่ 4 เทคนิค คือ 1) วิธีนำตัวแปรเข้าทั้งหมด (Enter regression) 2) วิธีเพิ่มตัวแปร (Forward selection) 3) วิธีลดตัวแปร (Backward elimination) และ 4) วิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise selection) ซึ่งในแต่ละวิธีมีขั้นตอนที่แตกต่างกันไป การเลือกแต่ละวิธีขึ้นอยู่กับวัตถุประสงค์หรือเป้าหมายในการวิเคราะห์ Logistic regression รวมทั้ง ความชำนาญของผู้ที่ทำการวิเคราะห์ข้อมูล⁽¹¹⁾ รายละเอียดดังนี้

1) การเลือกตัวแปรโดยวิธีนำตัวแปรเข้าทั้งหมด (Enter regression) เป็นวิธีการเอาตัวแปรต้นทุกตัวทั้งตัวแปรต้นที่มีความสัมพันธ์กับตัวแปรตามอย่างมีนัยสำคัญทางสถิติและไม่มีนัยสำคัญทางสถิติ เข้าไปวิเคราะห์ใน Model การคัดเลือกตัวแปรด้วยวิธีนี้ ตัวแปรต้นที่นำเข้าต้องพิจารณามาแล้วเป็นอย่างดีว่าสมควรที่จะอยู่ใน Model ของการวิเคราะห์ข้อมูลขั้นสุดท้าย (final model)

2) วิธีเพิ่มตัวแปร (Forward selection) วิธีการนี้จะคัดเลือกตัวแปรต้นที่มีความสัมพันธ์กับตัวแปรตาม

มากที่สุดและมีนัยสำคัญทางสถิติจะถูกคัดเลือกเข้าสู่การวิเคราะห์หรือเข้า Model ก่อน โดยพิจารณาจากค่าสัมประสิทธิ์ถดถอย (Regression coefficients หรือ ค่า B) หรือค่า OR ที่ได้จากการวิเคราะห์อย่างหยาบ จากนั้นจึงนำตัวแปรต้นที่มีความสัมพันธ์กับตัวแปรตามอันดับรองลงมาและมีนัยสำคัญทางสถิติเข้า Model ตามลำดับ การนำตัวแปรต้นเข้า Model จะทำเช่นนี้เรื่อยๆ ไปจนกระทั่งไม่มีตัวแปรต้นใดที่มีความสัมพันธ์กับตัวแปรตามอีกแล้ว

3) การเลือกตัวแปรโดยวิธีลดตัวแปร (Backward elimination) เป็นวิธีการคัดเลือกตัวแปรต้นออกจากสมการหรือออกจากการวิเคราะห์ในโปรแกรมทีละตัว โดยเริ่มจากการนำตัวแปรต้นทุกตัวเข้าสู่สมการ แล้วจึงคัดเลือกตัวแปรต้นออกทีละตัว โดยตัวแปรต้นที่มีความสัมพันธ์กับตัวแปรตามน้อยที่สุด ซึ่งพิจารณาจากค่า p -value ที่มีค่ามากที่สุด จะถูกคัดออกก่อน

4) วิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise selection) เป็นวิธีการคัดเลือกตัวแปรแบบผสมผสานระหว่างวิธีการคัดเลือกตัวแปรต้นทั้งแบบการเพิ่มตัวแปรและการลดตัวแปรเข้าด้วยกัน ซึ่งรายละเอียดการวิเคราะห์ไม่ขอนำเสนอในบทความนี้

ทั้งนี้ วิธีการคัดเลือกตัวแปรที่เหมาะสมและคิดว่าทำได้ง่ายสำหรับบุคลากรสาธารณสุขหรือนักวิจัยที่ยังไม่ค่อยมีประสบการณ์ด้านการวิเคราะห์ MBLR ที่แนะนำมี 2 วิธี คือ 1) วิธีนำตัวแปรเข้าทั้งหมด (Enter regression) และ 2) วิธีลดตัวแปร (Backward elimination)

การพิจารณาเลือกตัวแปรไว้ใน การวิเคราะห์ขั้นสุดท้าย (final model) มีอยู่ 2 แบบหลัก ได้แก่ 1) การคงไว้เฉพาะตัวแปรที่มีนัยสำคัญทางสถิติ (significant predictor) และ 2) การคงตัวแปรที่มีนัยสำคัญทางสถิติและไม่มีนัยสำคัญทางสถิติ (significant predictor+confounder) ซึ่งตัวแปรที่ไม่มีนัยสำคัญทางสถิติ นั้น มักจะเป็นตัวแปรกวนหรือตัวแปรที่ต้องการควบคุมอิทธิพล ซึ่งการคงตัวแปรกวนหรือตัวแปรที่ต้องการควบคุมอิทธิพลไว้ใน การวิเคราะห์ขั้นสุดท้ายนั้น

จะใช้เหตุผลด้านวิชาการจากการทบทวนวรรณกรรมที่ผ่านมาว่าตัวแปรนั้นมีผลต่อตัวแปรตาม และการใช้เหตุผลด้านสถิติ คือ ไม่สามารถถอดตัวแปรนั้นออกจากสมการ หรือ Model ได้เพราะถ้าถอดออกจะมีผลต่อค่าสถิติ เช่น ค่า OR ของตัวแปรหลักที่เปลี่ยนไปเกินร้อยละ 10^(12,13)

ตัวอย่างการวิเคราะห์แบบหลายตัวแปรโดยใช้ MBLR ในตารางที่ 2 เริ่มจากนำตัวแปรต้นที่ได้จากการ

ตารางที่ 2 การวิเคราะห์ปัจจัยที่สัมพันธ์กับการพลัดตกหกล้มในผู้สูงอายุด้วย Multiple binary logistic regression (MBLR)

Table 2 Analysis of factors associated with falls in the elderly using multiple logistic regression (MBLR)

ตัวแปร	จำนวน ตัวอย่าง (18,818)	เคสพลัดตกหกล้ม ใน 6 เดือนที่ผ่านมา จำนวน %	COR	AOR	95% CI AOR	p-value
เพศ						<0.001
ชาย	8,326	506 6.1	ref	ref		
หญิง	10,492	902 8.6	1.5	1.7	1.6, 2.3	
อายุ (ปี)						
60-69	10,467	652 6.2	ref	ref		
70-79	6,037	507 8.4	1.4	1.2	1.1, 1.3	0.011
80 ขึ้นไป	2,314	247 10.8	1.8	1.3	1.1, 1.6	0.001
เส้นรอบเอว						<0.001
ต่ำ-ปกติ	8,545	684 8.0	ref	ref		
สูงกว่าปกติ	10,273	724 7.1	0.8	1.5	1.3, 1.7	
การทรงตัว						<0.001
ปกติ	13,352	759 5.7	ref	ref		
บกพร่อง	5,466	649 11.9	2.2	2.0	1.8, 2.3	

Hosmer-Lemeshow goodness-of-fit ได้ค่า p -value=0.187

2.5 การตรวจสอบหรือประเมินความเหมาะสมของการวิเคราะห์ข้อมูลหรือ Model

สมการหรือ Model ที่ได้จะต้องมีการทดสอบความเหมาะสมโดยพิจารณาความเป็นสารรูปสนธิหรือ Goodness of fit สามารถกระทำได้ด้วยเทคนิควิธีต่าง ๆ บทความนี้เสนอแนะเทคนิคการประเมินด้วย Hosmer and Lemeshow test เป็นการทดสอบโดยใช้สถิติ Chi-Square เพื่อดูว่าสมการหรือ Model ที่สร้างขึ้นนั้นมีความเหมาะสมหรือสามารถให้ค่าทำนายความน่าจะเป็นของการเกิดเหตุการณ์สอดคล้องกับการเกิดเหตุการณ์จริงที่วัดได้จากข้อมูลที่เก็บมาหรือไม่ ถ้าผลการทดสอบสถิติ Chi-Square ไม่มีนัยสำคัญ

วิเคราะห์อย่างหยาบ 4 ตัวแปร คือ เพศ อายุ เส้นรอบเอว และการทรงตัว ทำการวิเคราะห์ได้ค่า VIF เท่ากับ 1.1-2.9 ซึ่งน้อยกว่า 10 จึงไม่เกิด Multicollinearity จากนั้นทำการวิเคราะห์แบบหลายตัวแปร ผลการวิเคราะห์สุดท้าย พบว่าทั้ง 4 ตัวแปรมีความสัมพันธ์กับการพลัดตกหกล้มอย่างมีนัยสำคัญทางสถิติ (p -value<0.05) จึงไม่ได้คัดตัวแปรใดออกจากสมการหรือโมเดล

แสดงว่ายอมรับได้หรือมีความเหมาะสม ตัวอย่างที่ตารางที่ 2 การทดสอบ Hosmer-Lemeshow goodness-of-fit ได้ค่า p -value 0.187 แสดงว่า ทั้ง 4 ปัจจัยที่อยู่ใน model สุดท้ายมีความเหมาะสม หรือโมเดลเข้ากับข้อมูลได้ดี (fit)

2.6 การประเมินประสิทธิภาพของสมการหรือ Model

การประเมินประสิทธิภาพของสมการในการทำนายเป็นสิ่งจำเป็นถ้าเราต้องการนำแบบจำลองที่ได้นั้นไปทำนายเหตุการณ์ เช่น การประเมินโมเดลด้วยการหาค่าความแม่นยำในการทำนายและค่าจุดตัด (Cut-off) ที่เหมาะสมในการจำแนกกลุ่มที่มีโอกาส

พลัดตกหกล้มในผู้สูงอายุ โดยการหาพื้นที่ใต้โค้ง ROC (Receiver operating characteristics [ROC] curve) ในทำนองเดียวกันจุดตัด (cut off) เพื่อแยกความแตกต่างกันระหว่างคนที่มีความเสี่ยงและไม่มีความเสี่ยงพลัดตกหกล้ม สามารถนำมาคำนวณหาค่าความไว (sensitivity) ความจำเพาะ (specificity) ค่าความถูกต้อง (Accuracy) ได้ ซึ่งขออนุญาตไม่กล่าวถึงในการบทความนี้⁽¹⁴⁻¹⁶⁾

3. การนำเสนอและแปลผลการวิเคราะห์ข้อมูล

ในการนำเสนอผลการวิเคราะห์ข้อมูลจาก MBLR มักมีการนำเสนอค่า COR เปรียบเทียบกับค่า AOR เพื่อให้เห็นความแตกต่างระหว่างผลการวิเคราะห์อย่างหยาบและการวิเคราะห์แบบหลายตัวแปร ส่วนการนำเสนอผลการศึกษาที่จะสรุปว่ามีปัจจัยใดบ้างสัมพันธ์กับการพลัดตกหกล้ม จะพิจารณาจากค่า AOR ค่า 95% CI AOR และค่า *p*-value รวมทั้ง การนำเสนอจำนวน และร้อยละ ของแต่ละตัวแปรที่มีประโยชน์ต่อการนำเสนอผลการวิเคราะห์ข้อมูลไปใช้ประโยชน์ เช่นเดียวกัน ทั้งนี้ การแปลผลการวิเคราะห์ข้อมูลจากตารางที่ 2 รายละเอียดดังนี้

“การวิเคราะห์ปัจจัยที่สัมพันธ์กับการพลัดตกหกล้มในผู้สูงอายุด้วยการวิเคราะห์แบบหลายตัวแปรพบว่าปัจจัยที่สัมพันธ์กับการพลัดตกหกล้มในผู้สูงอายุอย่างมีนัยสำคัญทางสถิติ (*p*-value<0.05) จำนวน 4 ปัจจัย ได้แก่ เพศ อายุ เส้นรอบเอว และการทรงตัวบกพร่อง กล่าวคือ ผู้สูงอายุเพศหญิงมีโอกาสพลัดตกหกล้มอยู่ 1.7 เท่า เมื่อเปรียบเทียบกับผู้สูงอายุเพศชาย (95% CI AOR

1.6, 2.3) ผู้สูงอายุ 70-79 ปี มีความเสี่ยงต่อการพลัดตกหกล้มเป็น 1.2 เท่า (95% CI AOR 1.1, 1.3) และ ผู้สูงอายุ 80 ปีขึ้นไป มีความเสี่ยงต่อการพลัดตกหกล้มเป็น 1.3 เท่า (95% CI AOR 1.1, 1.6) เมื่อเปรียบเทียบกับคนที่อายุ 60-69 ปี คนที่เส้นรอบเอวสูงกว่าปกติ มีความเสี่ยงต่อการพลัดตกหกล้มเป็น 1.5 เท่า (95% CI AOR 1.3, 1.7) เมื่อเปรียบเทียบกับคนที่เส้นรอบเอวต่ำ-ปกติ ผู้สูงอายุทรงตัวบกพร่อง มีความเสี่ยงต่อการพลัดตกหกล้มเป็น 2.0 เท่า (95% CI AOR 1.8, 2.3) เมื่อเปรียบเทียบกับคนที่มีการทรงตัวปกติ การประเมินความเหมาะสมของ Model ด้วยวิธี Hosmer-Lemeshow goodness-of-fit ได้ค่า *p*-value 0.187 สะท้อนว่าโมเดลเข้ากับข้อมูลได้ดี (Fit)” อย่างไรก็ตาม การนำเสนอ และการแปลผลการวิเคราะห์ข้อมูล ขึ้นอยู่กับสไตล์ ประสิทธิภาพของผู้วิเคราะห์ข้อมูล หรือเป้าหมายการนำเสนอผลการวิเคราะห์ข้อมูลไปใช้ หากเป็นบทความวิจัยที่ต้องเผยแพร่ในวารสารอาจขึ้นอยู่กับรูปแบบที่วารสารกำหนด

ค่า OR เป็นค่าที่สะท้อนความหนักแน่นของความสัมพันธ์ระหว่างตัวแปรต้นกับตัวแปรตาม ตัวแปรต้นใดที่มีค่า OR มากกว่า แสดงว่ามีความสัมพันธ์กับตัวแปรตามมากกว่า อาจนำไปสู่การให้ข้อเสนอแนะในการเลือกปัจจัยนั้นเป็นอันดับต้น ๆ (priority) เพื่อวางแผนหรือออกแบบมาตรการในการป้องกันหรือแก้ไขปัญหาระดับความหนักแน่นของความสัมพันธ์^(8,17) รายละเอียดดังตารางที่ 3

ตารางที่ 3 เกณฑ์ระดับความหนักแน่นของความสัมพันธ์ (Strength of Association)

Table 3 Criteria for the strength of association

ค่า OR	ระดับความสัมพันธ์
1.00-1.49	อ่อน (Weak)
1.50-2.99	ปานกลาง (Moderate)
≥3.00	สูง (Strong)

ข้อสังเกต จากตารางที่ 2 ตัวแปร “เส้นรอบเอว” จากการวิเคราะห์อย่างหยาบ พบว่าไม่มีความสัมพันธ์กับการพลัดตกหกล้ม แต่ในการวิเคราะห์แบบหลายตัวแปร หรือการวิเคราะห์ที่คำนึงถึงอิทธิพล

ของปัจจัยอื่น พบว่ามีความสัมพันธ์กับการพลัดตกหกล้ม นั่นสะท้อนว่าผู้สูงอายุที่เส้นรอบเอวเกินจะมีความเสี่ยงต่อการพลัดตกหกล้มหรือไม่ อาจขึ้นอยู่กับปัจจัยอื่นๆ เช่น เพศ อายุ และการทรงตัว ผลจากการวิเคราะห์นี้

ชี้ให้เห็นถึงความสำคัญของการวิเคราะห์หาปัจจัยเสี่ยงหรือการวิเคราะห์หาความสัมพันธ์แบบหลายตัวแปรด้วย MBLR ซึ่งจะทำให้ได้ผลการศึกษาหรือข้อสรุปที่น่าเชื่อถือ

วิจารณ์

การวิเคราะห์หาปัจจัยเสี่ยงของการเกิดโรคหรือปัญหาสุขภาพด้วยเทคนิคการวิเคราะห์ถดถอยโลจิสติกแบบทวี (MBLR) จะทำให้ได้ผลการศึกษาหรือข้อสรุปที่ถูกต้อง และน่าเชื่อถือยิ่งขึ้น ถึงแม้ว่าจะมีหลายขั้นตอนและข้อตกงที่ต้องคำนึงถึง แต่เป็นวิธีการที่ทำได้ไม่ยากนัก โดยใช้โปรแกรมสำเร็จรูปช่วยวิเคราะห์ เช่น SPSS, STATA หรือ R ขนาดตัวอย่างต้องเพียงพอ กับจำนวนตัวแปรหรือปัจจัยเสี่ยงที่จะทำการวิเคราะห์ หากมีจำนวนน้อยอาจเกิด Over estimate ได้ค่า OR มาก และอาจมีค่า 95% CI ที่กว้าง การวางแผนเก็บข้อมูลหรือออกแบบการวิเคราะห์ ควรคำนึงถึงข้อตกง ทั้งขนาดตัวอย่าง รวมถึง Multicollinearity การอธิบายความสัมพันธ์ระหว่างปัจจัยเสี่ยงและการเกิดโรค อาจคำนึงถึงข้อจำกัดของรูปแบบการวิจัย โดยเฉพาะการวิจัยแบบภาคตัดขวาง Cross-sectional study ที่เป็นการวัดปัจจัยเสี่ยงและการเกิดโรคพร้อมๆ กัน ไม่สามารถระบุความเป็นเหตุเป็นผลได้อย่างชัดเจน การพิจารณาปัจจัยเสี่ยงด้วยน้ำหนักของค่า OR ควบคู่กับค่า 95% CI รวมทั้งจำนวนและร้อยละของปัจจัยต่างๆ จำแนกตามกลุ่มป่วย-ไม่ป่วย จะช่วยให้ตัดสินใจเกี่ยวกับการนำตัวแปรหรือปัจจัยเสี่ยงนั้นไปออกแบบการวางแผนดำเนินงานได้อย่างมั่นใจยิ่งขึ้น

สรุป

การวิเคราะห์หาปัจจัยเสี่ยงของการเกิดโรคหรือปัญหาสุขภาพด้วยเทคนิคการวิเคราะห์ข้อมูลอย่างเหมาะสมมีความสำคัญอย่างยิ่ง การวิเคราะห์แบบหลายตัวแปรด้วยเทคนิคการวิเคราะห์ MBLR เป็นวิธีการทางสถิติที่นิยมใช้มากในการวิจัยทางด้านการแพทย์และสาธารณสุข สำหรับวิเคราะห์ข้อมูลที่ตัวแปรตาม

เป็นตัวแปรแจกนับ (Categorical variable) มีค่าได้เพียง 2 ค่า (Dichotomous) ส่วนตัวแปรต้นเป็นตัวแปรแจกนับ (Categorical variable) ที่อาจมีมากกว่า 2 ค่า หรือเป็นตัวแปรต่อเนื่อง (Continuous data) ก็ได้ วัตถุประสงค์การวิเคราะห์ คือ เพื่อหาปัจจัยเสี่ยง และเพื่อสร้างสมการทำนายเหตุการณ์ ประเด็นหลักของการวิเคราะห์โดยใช้ MBLR คือการอธิบายความสัมพันธ์ระหว่างตัวแปรต้นกับตัวแปรตามที่มีการควบคุมอิทธิพลจากปัจจัยอื่น ๆ จะทำให้ได้ผลการศึกษาหรือข้อสรุปที่ถูกต้อง และน่าเชื่อถือ สามารถนำผลการศึกษาไปวางแผน ส่งเสริม ป้องกัน และควบคุมโรคอย่างมีประสิทธิภาพ

ข้อเสนอแนะ

เทคนิคการวิเคราะห์ MBLR เพื่อปัจจัยเสี่ยงของการเกิดโรคหรือปัญหาสุขภาพ มีข้อเสนอแนะเพื่อการนำไปใช้ดังนี้

- 1) ควรพิจารณานำเทคนิคการวิเคราะห์ถดถอยโลจิสติกแบบทวี ไปวิเคราะห์หาปัจจัยเสี่ยงของการเกิดโรคหรือความสัมพันธ์ของปัจจัยต่าง ๆ จะทำให้ได้ผลการศึกษาหรือข้อสรุปที่ถูกต้อง และน่าเชื่อถือยิ่งขึ้น
- 2) การวิเคราะห์ถดถอยโลจิสติกแบบพหุมีข้อตกงที่สำคัญที่ต้องคำนึงถึง คือ ขนาดตัวอย่างต้องเพียงพอ กับจำนวนตัวแปรหรือปัจจัยเสี่ยงที่จะทำการวิเคราะห์ และระมัดระวังความสัมพันธ์กันสูงของตัวแปรต้น (Multicollinearity)

เอกสารอ้างอิง

1. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. Health Promotion International. 2000; 15(3):259-67.
2. Khowhit D, Kumvan W, Dana K, Kumsom C. Factors related to fall prevention behavior among older adults in rural communities, Yang Sisurat District, Maha Sarakham Province. Nursing

- Journal of the Ministry of Public Health. 2023;33(1):131–46. (in Thai)
3. Pothipim M, Ketbunroong V, Norkaew J, Udomporn T, Chumklang T. Factors associated with falls among older adults in Nakhon Ratchasima Province. *Journal of Nursing and Health Care*. 2025;43(2):1–13. (in Thai)
 4. Baum F. *The New Public Health*. 2nd ed. Australia: Oxford University Press Australia; 2008.
 5. Tinkhamrob B. *Analysis of health science research data using logistic regression*. Khon Kaen: Faculty of Public Health, Khon Kaen University; 2000. (in Thai)
 6. Hosmer, DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: McGraw-Hill; 2000.
 7. Burns N, Grove SK. *The Practice of Nursing Research: Conduct, Critique and Utilization*. Philadelphia: W.B. Saunders Company; 1993.
 8. Stevens J. *Applied multivariate statistics for the social science*. Mahwah, New Jersey: Lawrence Erlbaum Associate; 1996.
 9. Hsieh FY, Block DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. *Stat Med*. 1998;
 10. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49(12):1373–9.
 11. Montgomery DC, Peck EA, G.G. V. *Introduction to Linear Regression Analysis*. 5th ed. New Jersey: John Wiley & Sons; 2012.
 12. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic research: Principles and quantitative methods*. Belmont, CA: Lifetime Learning Publications; 1982.
 13. Greenland-s. Generalized mantel-haenszel estimators for k 2xj tables. *Biometrics*. 1989;45(1):183–91.
 14. Fu Y, Li KB, Yang XC. A risk score model for predicting cardiac rupture after acute myocardial infarction. *Chin Med J (Engl)* [Internet]. 2010 [cited 2022 Apr 2];132(9):1039–44. Available from: <https://mednexus.org/doi/full/10.1097/CM9.0000000000000175>
 15. Sarakarn P, Mulpolsri P. The appropriate cut-off points for Receiver Operating Characteristic (ROC) curve analysis in developing health innovation tools: A case study using STATA program. *Thai Bull Pharm Sci*. 2021;16(1):93–108. (in Thai)
 16. Kaiyawan Y. *Principles and using logistic regression analysis for research*. Rajamangala University of Technology Srivijaya Research Journal. 2012;4(1):1–12. (in Thai)
 17. Zhang J, Yu KF, What's the Relative Risk? A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes. *Am Med Assoc*. 1998;17(14):1623–34.