

การรายงานการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติก ในวารสารทางการแพทย์และสาธารณสุขในประเทศไทย

ศักดิ์สิน สิมสินธุ์¹, จิราพร เขียวอยู่¹

¹สาขาวิชาวิทยาการระบาดและชีวสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

A Report of the Management of Continuous Independent Variables in Logistic Regression Analysis in Medical and Public Health Journals in Thailand

Saksin Simsin¹, Jiraporn Khiewyoo¹

¹Department of Epidemiology and Biostatistics, Faculty of Public Health, Khon Kaen University

หลักการและวัตถุประสงค์: วิธีการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์การถดถอยลอจิสติกมีผลต่อตัวแปรอิสระที่คงอยู่ในตัวแบบสุดท้าย การจัดกลุ่มตัวแปรอิสระต่อเนื่องที่ไม่เหมาะสมนอกจากทำให้สูญเสียความสามารถในการสื่อความหมายของข้อมูล ยังทำให้ประสิทธิภาพในการทำนายของตัวแบบลดลงและเพิ่มความคลาดเคลื่อนชนิดที่ 1 จากเหตุผลที่กล่าวผู้วิจัยจึงมีความสนใจศึกษาวิธีการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติกที่รายงานในบทความวิจัยทางการแพทย์และสาธารณสุขในประเทศไทย

วิธีการศึกษา: การศึกษาเชิงพรรณนา โดยทำการทบทวนวิธีการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติกที่รายงานในบทความวิจัยที่ตีพิมพ์ในวารสารทางการแพทย์และสาธารณสุขในประเทศไทย การคัดเลือกบทความวิจัยเข้าในการศึกษา เริ่มจากจัดวารสารทางการแพทย์และสาธารณสุขที่อยู่ในฐานข้อมูลศูนย์ดัชนีการอ้างอิงวารสารไทยออกเป็น 4 กลุ่มและคัดเลือกบทความวิจัยจาก 4 กลุ่มวารสาร ที่ใช้การวิเคราะห์การถดถอยลอจิสติกที่ตัวแปรตามเป็นข้อมูลกลุ่ม 2 กลุ่มและมีตัวแปรอิสระต่อเนื่องอย่างน้อย 1 ตัวแปร ในการวิเคราะห์เข้าในการศึกษา ตัวอย่างมี 119 บทความจากบทความวิจัยตีพิมพ์เผยแพร่ระหว่างวันที่ 1 มกราคม พ.ศ. 2556 ถึง 31 ธันวาคม พ.ศ. 2558 ทั้งหมดจำนวน 1,997 บทความ

ผลการศึกษา: จาก 119 บทความที่คัดเข้าในการศึกษา มี 4 บทความที่กล่าวถึงการตรวจสอบความสัมพันธ์เชิง

Background and Objective: The management of continuous independent variables before being inputted in logistic regression analysis can affect the independent variables in the final model. Improper grouping of continuous independent variables can cause losing the informational meaning, decreasing the predictive ability and increasing type 1 error. From these reasons, we were interested in studying the handling of these continuous independent variables in the logistic regression analysis in Thai medical and public health journals.

Methods: This descriptive study reviewed the management of the continuous independent variables in logistic regression analysis which reported in medical and public health journals in Thailand. The journal article selection was initiated by classifying the medical and public health journals that published in the database of Thai-Journal Citation Index Center into 4 groups. We chose the journal articles from those which used binary logistic regression analysis with at least 1 continuous independent variable into this study. There were 119 from a total of 1997 research articles, published between January 1, 2013 and December 31, 2015.

Results: Of 119 articles, 4 articles were reported an investigation of the linear relationship between the continuous independent variables and the logit but only 3 articles were reported the results of the test. There were 115 articles which did not mention investigating

เส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิต แต่มีเพียง 3 บทความที่รายงานผลการทดสอบ สำหรับ 115 บทความที่ไม่ได้กล่าวถึงการตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิต มีตัวแปรอิสระต่อเนื่องในการวิเคราะห์ 341 ตัวแปร นักวิจัยนำเข้าในการวิเคราะห์แบบตัวแปรกลุ่มร้อยละ 71.8 เป็นการจัดกลุ่มโดยในบทความไม่ได้กล่าวถึงหลักเกณฑ์ที่ใช้อ้างอิงในการจัดกลุ่ม ร้อยละ 59.6

สรุป: มีการนำเสนอข้อมูลเกี่ยวกับการตรวจความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิต และรายละเอียดวิธีการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์น้อยมาก ตัวแปรอิสระต่อเนื่องส่วนมากถูกแปลงเป็นตัวแปรกลุ่ม โดยที่ในบทความวิจัยไม่ได้กล่าวถึงหลักเกณฑ์ที่ใช้อ้างอิงในการจัดกลุ่ม

คำสำคัญ: การวิเคราะห์การถดถอยลอจิสติก, การจัดการตัวแปรอิสระต่อเนื่อง

ศรีนครินทร์เวชสาร 2560; 32(1): 38-44. • Srinagarind Med J 2017; 32(1): 38-44.

บทนำ

ตัวแปรอิสระในการวิเคราะห์การถดถอยลอจิสติกเป็นได้ทั้งตัวแปรกลุ่มและตัวแปรต่อเนื่อง กรณีเป็นตัวแปรต่อเนื่องอาจนำเข้ามาในตัวแปรต่อเนื่องหรือตัวแปรกลุ่มที่มีเกณฑ์การจัดกลุ่มอ้างอิง ขึ้นอยู่กับเงื่อนไขความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องนั้นกับตัวแปรตามในรูปลอจิต จากการศึกษาที่ผ่านมาพบว่าในบทความวิจัยทางด้านระบาดวิทยาและการแพทย์ทั่วไปมีการจัดกลุ่มตัวแปรอิสระต่อเนื่องในการวิเคราะห์ข้อมูลสูงถึงร้อยละ 86¹ ซึ่งในกรณีที่ตัวแปรอิสระต่อเนื่องมีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิตการจัดกลุ่มหรือการแปลงค่าตัวแปรอิสระต่อเนื่องเป็น 2 กลุ่มหรือมากกว่าก่อนนำเข้าในการวิเคราะห์นั้นไม่มีความจำเป็นในทางสถิติและยังส่งผลให้เกิดผลเสีย² มีการศึกษาถึงผลของการนำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์การถดถอยลอจิสติกที่มีวิธีการจัดการที่แตกต่างกัน 3 วิธี โดยวิธีแรกตัวแปรอิสระต่อเนื่องทั้งหมดในการศึกษาถูกนำเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่องวิธีที่สองนำเข้าโดยจัดกลุ่มตัวแปรอิสระต่อเนื่องเป็นกลุ่มมากกว่า 2 กลุ่มและวิธีสุดท้ายจัดกลุ่มตัวแปรอิสระต่อเนื่องเป็น 2 กลุ่ม โดยใช้ค่ากลางของข้อมูลเป็นจุดแบ่ง ผลการศึกษาพบว่าการจัดการตัวแปรอิสระต่อเนื่องที่แตกต่างก่อนนำเข้าในการวิเคราะห์มีผลต่อตัวแปรอิสระที่คงอยู่ในตัวแบบสุดท้ายที่แตกต่าง (Model content) ทั้งยังพบว่าการจัดกลุ่มตัวแปร

อิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์ที่ไม่เหมาะสมนอกจากทำให้สูญเสียความสามารถในการสื่อความหมายของข้อมูลแล้ว ยังทำให้ประสิทธิภาพในการทำนายของตัวแบบลดลงและเพิ่มความคลาดเคลื่อนชนิดที่ 1³ จากเหตุผลที่กล่าวข้างต้นประกอบกับการใช้การวิเคราะห์การถดถอยลอจิสติกในการวิเคราะห์ข้อมูลทางการแพทย์และสาธารณสุขมีแนวโน้มเพิ่มขึ้น⁴ แต่ยังไม่มีการศึกษาถึงรายละเอียดวิธีการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติกที่รายงานในบทความวิจัยที่ได้รับการตีพิมพ์ในวารสารทางการแพทย์และสาธารณสุขในประเทศไทย ผู้วิจัยจึงมีความสนใจที่จะทำการศึกษาในประเด็นดังกล่าว โดยคาดหวังว่าจะได้ข้อเสนอแนะที่เป็นข้อมูลเชิงประจักษ์สำหรับนักวิจัยหรือผู้ที่มีส่วนเกี่ยวข้องให้ตระหนักถึงความสำคัญในการจัดการตัวแปรอิสระต่อเนื่องที่เหมาะสมก่อนนำเข้าในการวิเคราะห์ข้อมูล

วิธีการศึกษา

รูปแบบการศึกษา

เป็นการศึกษาเชิงพรรณนา (descriptive study) โดยทำการทบทวนวิธีการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติกที่รายงานในบทความวิจัยที่ตีพิมพ์ในวารสารทางการแพทย์และสาธารณสุขในประเทศไทย ประชากรศึกษา คือ บทความวิจัยที่ตีพิมพ์ในวารสารทางการแพทย์

แพทย์และสาธารณสุขในประเทศไทย ที่ใช้การวิเคราะห์การถดถอยลอจิสติกที่ตัวแปรตามเป็นข้อมูลกลุ่ม 2 กลุ่มในการวิเคราะห์ข้อมูล และมีตัวแปรอิสระต่อเนื่องในการวิเคราะห์อย่างน้อย 1 ตัวแปร ที่เผยแพร่ระหว่างวันที่ 1 มกราคม พ.ศ. 2556 ถึง 31 ธันวาคม พ.ศ. 2558 การคัดเลือกตัวอย่างบทความวิจัยเข้าในการศึกษา เริ่มจากผู้วิจัยจัดวารสารทางการแพทย์และสาธารณสุขที่อยู่ในฐานข้อมูลศูนย์ดัชนีการอ้างอิงวารสารไทยออกเป็น 4 กลุ่ม ดังนี้

1. กลุ่มแพทยศาสตร์ และทันตแพทยศาสตร์
 2. กลุ่มเภสัชศาสตร์ กายภาพบำบัด เทคนิคการแพทย์ วิทยาศาสตร์การแพทย์และการกีฬา
 3. กลุ่มพยาบาลศาสตร์ สุขภาพจิตและจิตวิทยา
 4. กลุ่มสาธารณสุขศาสตร์ การแพทย์แผนไทยและอื่นๆ
- การคัดเลือกวารสารเข้าในแต่ละกลุ่มด้วยค่า Thai-journal Impact Factors เฉลี่ย 3 ปี สูงสุด 5 ลำดับแรกในแต่ละกลุ่ม และคัดเลือกบทความวิจัยจาก 4 กลุ่มวารสารที่เผยแพร่ระหว่างวันที่ 1 มกราคม พ.ศ. 2556 ถึง 31 ธันวาคม พ.ศ. 2558 ที่ใช้การวิเคราะห์การถดถอยลอจิสติกที่ตัวแปรตามเป็นข้อมูลกลุ่ม 2 กลุ่มและมีตัวแปรอิสระต่อเนื่องอย่างน้อย 1 ตัวแปรในการวิเคราะห์เข้าในการศึกษา

เครื่องมือที่ใช้ในการศึกษา

เครื่องมือที่ใช้ในการศึกษาเป็นแบบบันทึกที่ผู้วิจัยพัฒนาขึ้นจากการศึกษาทบทวนวรรณกรรมที่เกี่ยวข้องและให้ปฏิบัติตามกรอบแนวคิดการวิจัย เพื่อตอบวัตถุประสงค์การวิจัย

แบบบันทึกนี้ประกอบไปด้วย 2 ส่วน

ส่วนที่ 1 แบบบันทึกเพื่อการคัดกรองบทความวิจัย

ส่วนที่ 2 แบบบันทึกวิธีการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติก

การเก็บรวบรวมข้อมูล

ผู้วิจัยเก็บรวบรวมข้อมูล โดยการอ่านทบทวนบทความวิจัยที่ผ่านการคัดเลือกเข้าในการศึกษา บันทึกรายละเอียดวิธีการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติกและประเมินความเหมาะสมในการจัดการตัวแปรอิสระต่อเนื่องตามเกณฑ์ที่ผู้วิจัยสร้างขึ้น โดยผู้วิจัยอ่านบทความที่ผ่านการคัดเลือกเข้าในการศึกษาอย่างน้อย 2 รอบ รอบแรกอ่านเพื่อให้เข้าใจในรายละเอียดสาระของงานวิจัยและรอบที่สองอ่านเพื่อเก็บรายละเอียดการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าตัวแบบการถดถอยลอจิสติก

การวิเคราะห์ข้อมูล

บรรยายลักษณะทั่วไปของบทความวิจัยและรายละเอียดวิธีการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์การถดถอยลอจิสติกด้วยสถิติพรรณนา นำเสนอค่าความถี่

และร้อยละวิเคราะห์ข้อมูลโดยใช้โปรแกรม STATA Version 10 (stataCorp, College Station, texas)

ผลการศึกษา

จากการทบทวนวารสารทั้ง 4 กลุ่ม พบว่าระหว่างวันที่ 1 มกราคม พ.ศ. 2556 ถึง 31 ธันวาคม พ.ศ. 2558 มีบทความวิจัยตีพิมพ์เผยแพร่ทั้งหมดจำนวน 1,997 บทความ เป็นบทความวิจัยที่ใช้การวิเคราะห์การถดถอยลอจิสติกในการวิเคราะห์ข้อมูลจำนวน 158 บทความ โดยในจำนวนนี้เป็นบทความวิจัยที่ใช้การวิเคราะห์การถดถอยลอจิสติกที่ตัวแปรตามเป็นข้อมูลกลุ่ม 2 กลุ่มและมีตัวแปรอิสระต่อเนื่องในการวิเคราะห์อย่างน้อย 1 ตัวแปร จำนวน 119 บทความ ข้อมูลทั่วไปของตัวอย่างบทความวิจัย (ตารางที่ 1)

จากการทบทวนบทความวิจัย 119 บทความ มีบทความวิจัยที่กล่าวถึงการตรวจความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิต 4 บทความ แต่มีเพียง 3 บทความที่รายงานผลการทดสอบโดยพบว่านักวิจัยมีการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์ดังนี้มี 1 บทความ ที่ตัวแปรอิสระต่อเนื่องมีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิต ซึ่งในบทความนี้มีตัวแปรอิสระต่อเนื่องในการวิเคราะห์ 1 ตัวแปรและนักวิจัยได้นำตัวแปรเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่อง ส่วนอีก 2 บทความ ผลการทดสอบพบว่าตัวแปรอิสระต่อเนื่องไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิตซึ่งมีตัวแปรอิสระต่อเนื่องในการวิเคราะห์ 6 ตัวแปร นักวิจัยได้นำตัวแปรเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่อง 1 ตัวแปร และนำเข้าในการวิเคราะห์ในรูปแบบตัวแปรกลุ่ม 5 ตัวแปร สำหรับเกณฑ์ที่นักวิจัยใช้ในการจัดกลุ่มตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์นั้น พบว่าเป็นการจัดกลุ่มด้วยค่ากลางของข้อมูล 2 ตัวแปร จัดกลุ่มโดยที่ไม่ได้กล่าวถึงหลักเกณฑ์ที่ใช้อ้างอิง 2 ตัวแปรและจัดกลุ่มโดยมีหลักเกณฑ์ในการอ้างอิง 1 ตัวแปร ทั้ง 5 ตัวแปรนักวิจัยได้จัดเป็น 2 กลุ่ม ส่วน 1 บทความที่รายงานการตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิตแต่ไม่ได้รายงานผลการทดสอบนั้น มีตัวแปรอิสระต่อเนื่องในการวิเคราะห์ 9 ตัวแปร นักวิจัยได้นำตัวแปรเข้าในการวิเคราะห์ในรูปแบบตัวแปรกลุ่มทั้ง 9 ตัวแปร เป็นการจัดกลุ่มโดยไม่มีหลักเกณฑ์ที่ใช้อ้างอิง 5 ตัวแปร ทั้ง 5 ตัวแปรนี้จัดเป็น 2 กลุ่ม และจัดกลุ่มโดยมีหลักเกณฑ์ที่ใช้ในการอ้างอิง 4 ตัวแปร ใน 4 ตัวแปรนี้มี 3 ตัวแปรที่นักวิจัยได้จัดเป็น 2 กลุ่มและอีก 1 ตัวแปรที่จัดเป็น 3 กลุ่ม (ตารางที่ 2)

ตารางที่ 1 ข้อมูลทั่วไปของตัวอย่างบทความวิจัย

ข้อมูลทั่วไปของบทความวิจัย	จำนวนบทความ (n=119)	
	จำนวน	ร้อยละ
กลุ่มวารสาร		
กลุ่มที่ 1 กลุ่มแพทยศาสตร์และทันตแพทยศาสตร์	43	36.1
กลุ่มที่ 2 กลุ่มเภสัชศาสตร์ เทคนิคการแพทย์ วิทยาศาสตร์การแพทย์	12	10.1
กลุ่มที่ 3 กลุ่มพยาบาลศาสตร์ สุขภาพจิตและจิตวิทยา	20	16.8
กลุ่มที่ 4 กลุ่มสาธารณสุขศาสตร์ การแพทย์แผนไทยและอื่นๆ	44	37.0
วัตถุประสงค์ของงานวิจัยที่ใช้ Logistic regression		
จุดประสงค์หลักของการวิจัย	73	61.3
ไม่ใช่จุดประสงค์หลักของการวิจัย	46	38.7
เป้าหมายการใช้ Logistic regression		
เพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม	110	92.4
เพื่อสร้างสมการทำนาย	9	7.6
จำนวนตัวแปรอิสระทั้งหมดในการวิเคราะห์		
1-5 ตัวแปร	44	37.0
6-10 ตัวแปร	50	42
11-16 ตัวแปร	22	18.5
>16 ตัวแปร	3	2.5
จำนวนตัวแปรอิสระต่อเนื่องในการวิเคราะห์		
1 ตัวแปร	28	23.5
2 ตัวแปร	29	24.4
3 ตัวแปร	23	19.3
4 ตัวแปร	15	12.6
5 ตัวแปร	12	10.1
> 5 ตัวแปร	12	10.1

ตารางที่ 2 การรายงานการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์ กรณีที่มีการรายงานการตรวจสอบความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิต

การรายงานการตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิตของบทความวิจัย	บทความ	
	จำนวน	ร้อยละ
การตรวจสอบความสัมพันธ์เชิงเส้น (n= 119)		
- ไม่รายงานการตรวจสอบความสัมพันธ์เชิงเส้น	115	96.6
- รายงานการตรวจสอบความสัมพันธ์เชิงเส้น	4	3.4
(ไม่รายงานผลการทดสอบ 1 บทความ)		
(รายงานผลการทดสอบ 3 บทความ)		
การจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์ กรณีที่มีการรายงานผลการทดสอบความสัมพันธ์เชิงเส้น (3 บทความ)	ตัวแปร	
	จำนวน	ร้อยละ
ผลการทดสอบพบตัวแปรอิสระต่อเนื่องมีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิต (1 บทความ 1 ตัวแปร)		
- นำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์โดยตรงในรูปแบบตัวแปรต่อเนื่อง	1	100
ผลการทดสอบพบตัวแปรอิสระต่อเนื่องไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิต (2 บทความ 6 ตัวแปร)		
- นำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์โดยตรงในรูปแบบตัวแปรต่อเนื่อง	1	16.7
- นำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์ในรูปแบบตัวแปรกลุ่มเกณฑ์ที่ใช้ในการจัดกลุ่มตัวแปรอิสระต่อเนื่อง (5 ตัวแปร)	5	83.3
- จัดกลุ่มโดยไม่มีอ้างอิงจากทฤษฎี เนื้อหาหรืองานวิจัยก่อนหน้า	2	40
- อ้างอิงจากทฤษฎี เนื้อหาหรืองานวิจัยก่อนหน้า ที่มีความหมายในการแปลผล	1	20
- แบ่งเป็น 2 กลุ่มตามค่ากลางของข้อมูล จำนวนกลุ่มที่จัด (5 ตัวแปร)	2	40
- จัดเป็น 2 กลุ่ม	5	100
การจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์ กรณีที่ไม่รายงานผลการทดสอบความสัมพันธ์เชิงเส้น (1 บทความ 9 ตัวแปร)		
- นำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์ในรูปแบบตัวแปรกลุ่มเกณฑ์ที่ใช้ในการจัดกลุ่มตัวแปรอิสระต่อเนื่อง (9 ตัวแปร)	9	100
- จัดกลุ่มโดยไม่ได้กล่าวถึงเกณฑ์ในการอ้างอิง	5	55.6
- จัดกลุ่มโดยอ้างอิงจากทฤษฎี เนื้อหาหรืองานวิจัยก่อนหน้า จำนวนกลุ่มที่จัด (9 ตัวแปร)	4	44.4
- จัดเป็น 2 กลุ่ม	8	88.9
- จัดเป็น 3 กลุ่ม	1	11.1

สำหรับ 115 บทความ ที่ไม่รายงานตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปโลจิสติก นักวิจัยมีการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติก ดังนี้

1. นักวิจัยนำตัวแปรอิสระต่อเนื่องทุกตัวแปรเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่อง จำนวน 27 บทความ คิดเป็นร้อยละ 23.5

2. นักวิจัยนำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่องและตัวแปรกลุ่ม จำนวน 15 บทความ คิดเป็นร้อยละ 13.0

3. นักวิจัยนำตัวแปรอิสระต่อเนื่องทุกตัวแปรเข้าในการวิเคราะห์ในรูปแบบตัวแปรกลุ่ม จำนวน 73 บทความ คิดเป็นร้อยละ 63.5

บทความวิจัยทั้ง 115 บทความ มีตัวแปรอิสระต่อเนื่องในการวิเคราะห์ 341 ตัวแปร นักวิจัยนำเข้าในการวิเคราะห์ในรูปแบบตัวแปรกลุ่ม 245 ตัวแปร คิดเป็นร้อยละ 71.8 (245/341)

นำเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่อง 94 ตัวแปร คิดเป็นร้อยละ 27.6 (94/341) และนำเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่องและตัวแปรกลุ่มที่ได้จากการจัดกลุ่มตัวแปรอิสระต่อเนื่อง 2 ตัวแปร คิดเป็นร้อยละ 0.6 (2/341)

สำหรับหลักเกณฑ์ที่นักวิจัยใช้ในการจัดกลุ่มตัวแปรอิสระต่อเนื่องในการวิเคราะห์นั้น พบว่าเป็นการจัดกลุ่มโดยไม่ได้กล่าวถึงเกณฑ์ที่ใช้อ้างอิง ร้อยละ 59.6 (146/245) จัดกลุ่มจากการอ้างอิงทฤษฎี เนื้อหา ร้อยละ 35.9 (88/245) จัดกลุ่มด้วยการใช้ค่ากลางของชุดข้อมูล ร้อยละ 2.9 (7/245) ใช้เกณฑ์แบ่งเป็นช่วงรอบเท่าๆ กัน ร้อยละ 1.2 (3/245) และใช้เกณฑ์แบ่งจุดตัดที่ให้ค่าพินัยที่น้อยที่สุด ร้อยละ 0.4 (1/245)

จำนวนกลุ่มที่จัด พบว่าการจัดกลุ่มตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์ มีการจัดกลุ่มตัวแปรอิสระต่อเนื่องเป็น 2 กลุ่ม ร้อยละ 86.1 (211/245) จัดเป็น 3 กลุ่ม ร้อยละ 11.0 (27/245) จัดเป็น 4 กลุ่ม ร้อยละ 2.4 (6/245) และจัดเป็นกลุ่ม 6 กลุ่ม ร้อยละ 0.4 (1/245) (ตารางที่ 3)

ตารางที่ 3 การรายงานการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์ กรณีไม่รายงานการตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปโลจิสติก

การรายงานการจัดการตัวแปรอิสระต่อเนื่องก่อนการนำเข้าในการวิเคราะห์ (115 บทความ 341 ตัวแปร)	จำนวนบทความ (n = 115)	
	จำนวน	ร้อยละ
การจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลอจิสติกในบทความวิจัย		
- นำตัวแปรอิสระต่อเนื่องทุกตัวแปรเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่อง	27	23.5
- นำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์ในรูปแบบตัวแปรต่อเนื่องและตัวแปรกลุ่ม	15	13.0
- นำตัวแปรอิสระต่อเนื่องทุกตัวแปรเข้าในการวิเคราะห์ในรูปแบบตัวแปรกลุ่ม	73	63.5
การจัดการตัวแปรอิสระต่อเนื่องก่อนการนำเข้าในการวิเคราะห์ กรณีไม่รายงานผลการทดสอบความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปโลจิสติก	จำนวนตัวแปร	ร้อยละ
วิธีการจัดการตัวแปรอิสระต่อเนื่อง (n= 341)		
- นำเข้าในการวิเคราะห์แบบตัวแปรกลุ่ม	245	71.8
- นำเข้าในการวิเคราะห์แบบตัวแปรต่อเนื่อง	94	27.6
- นำเข้าในการวิเคราะห์แบบตัวแปรต่อเนื่องและตัวแปรกลุ่ม	2	0.6
เกณฑ์ในการจัดกลุ่มตัวแปรอิสระต่อเนื่อง (n= 245)		
- จัดกลุ่มโดยไม่ได้กล่าวถึงเกณฑ์การอ้างอิงจากทฤษฎีหรือเนื้อหา	146	59.6
- อ้างอิงเนื้อหาหรืองานวิจัยที่มีการทำมาก่อน ที่มีความหมายในการแปลผล	88	35.9
- ใช้เกณฑ์ค่ากลางในการแบ่ง (mean, median)	7	2.9
- ใช้เกณฑ์แบ่งเป็นช่วงรอบเท่าๆ กัน	3	1.2
- ใช้เกณฑ์แบ่งจุดตัดที่มีนัยสำคัญทางสถิติ ที่ให้ค่าพินัยที่น้อยที่สุด	1	0.4
จำนวนกลุ่มการจัดกลุ่มตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์ (n= 245)		
- จัดเป็น 2 กลุ่ม	211	86.2
- จัดเป็น 3 กลุ่ม	27	11.0
- จัดเป็น 4 กลุ่ม	6	2.4
- จัดเป็น 6 กลุ่ม	1	0.4

ตารางที่ 4 ความเหมาะสมในการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์; พิจารณาเฉพาะบทความที่มีการรายงานผลการทดสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิสต์

ความเหมาะสมในการจัดการตัวแปรอิสระต่อเนื่อง	จำนวนตัวแปร (n = 7)
นำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์ที่ไม่เหมาะสม	5
การจัดการตัวแปรอิสระต่อเนื่องเหมาะสม กรณีแรก กรณีตัวแปรอิสระต่อเนื่องมีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิสต์และมีการนำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์โดยตรงในรูปแบบของตัวแปรต่อเนื่อง	1
การจัดการตัวแปรอิสระต่อเนื่องเหมาะสม กรณีที่สอง กรณีตัวแปรอิสระต่อเนื่องไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิสต์และมีการนำตัวแปรอิสระต่อเนื่องเข้าในการวิเคราะห์รูปแบบตัวแปรกลุ่มที่มีการจัดกลุ่มโดยการอ้างอิงเนื้อหา ทฤษฎีหรือจากการศึกษาวิจัยก่อนหน้า	1
ความเหมาะสมในภาพรวม	2

ความเหมาะสมในการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์

จากการทบทวนบทความความวิจัย 119 บทความ มีบทความที่กล่าวถึงการตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิสต์ 4 บทความ แต่สามารถประเมินความเหมาะสมในการจัดการตัวแปรอิสระต่อเนื่องได้เพียง 3 บทความ เนื่องจากมีการรายงานผลการทดสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิสต์ โดยจาก 3 บทความมีตัวแปรอิสระต่อเนื่องในการวิเคราะห์ 7 ตัวแปร ในจำนวนนี้มีการจัดการตัวแปรอิสระต่อเนื่องเหมาะสม 2 ตัวแปร โดยเป็นการจัดการตัวแปรอิสระต่อเนื่องอย่างเหมาะสมในสองกรณี กรณีแรกคือผลการทดสอบพบว่าตัวแปรอิสระต่อเนื่องมีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิสต์และผู้วิจัยนำตัวแปรเข้าในการวิเคราะห์โดยตรงในรูปแบบของตัวแปรต่อเนื่อง ตัวแปรที่จัดการเหมาะสมกรณีนี้มี 1 ตัวแปร การจัดการตัวแปรอิสระต่อเนื่องอย่างเหมาะสม อีกกรณีหนึ่งคือผลการทดสอบพบว่าตัวแปรอิสระต่อเนื่องไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิสต์และผู้วิจัยนำตัวแปรเข้าในการวิเคราะห์ในรูปแบบของตัวแปรกลุ่มที่มีการจัดกลุ่มโดยการอ้างอิงจากทฤษฎีหรือการศึกษาวิจัยก่อนหน้า ตัวแปรที่จัดการเหมาะสมในกรณีนี้มี 1 ตัวแปรเช่นกัน (ตารางที่ 4)

อย่างไรก็ตามการจัดการตัวแปรอิสระต่อเนื่องที่เหมาะสมยังเป็นไปได้ในอีก 2 กรณี คือ กรณีผลการทดสอบพบว่าตัวแปรอิสระต่อเนื่องมีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิสต์ แต่ผู้วิจัยตัดสินใจนำตัวแปรเข้าในการวิเคราะห์ในรูปแบบตัวแปรกลุ่มที่มีการจัดกลุ่มโดยการอ้างอิงจากทฤษฎีหรือการศึกษาวิจัยก่อนหน้า หรือกรณีที่ผลการทดสอบพบว่าตัวแปรอิสระต่อเนื่องไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิสต์ แต่ผู้วิจัยสร้างตัวแปรต่อเนื่องใหม่ที่มีความ

สัมพันธ์เชิงเส้นกับตัวแปรตามในรูปลอจิสต์ แล้วผู้วิจัยนำตัวแปรที่สร้างใหม่นี้เข้าในการวิเคราะห์ ความเหมาะสมในสองกรณีหลังที่กล่าวนี้ไม่พบในบทความที่ทบทวน

วิจารณ์

การวิเคราะห์การถดถอยลอจิสต์เป็นวิธีการทางสถิติที่นิยมใช้ในการวิเคราะห์ข้อมูลทางด้านการแพทย์และสาธารณสุข ในกรณีที่ตัวแปรอิสระเป็นตัวแปรต่อเนื่องจำเป็นต้องตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิสต์ตั้งแต่ขั้นตอนของการวิเคราะห์หรืออย่างหยาบ เพื่อการจัดการตัวแปรที่เหมาะสมก่อนนำเข้าในการวิเคราะห์ ผลการศึกษาในครั้งนี้แสดงให้เห็นได้ว่าบทความวิจัยที่ตีพิมพ์ในวารสารทางด้านทางการแพทย์และสาธารณสุข มีการนำเสนอข้อมูลในประเด็นที่เกี่ยวข้องกับการตรวจสอบความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปลอจิสต์และรายละเอียดในการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าในการวิเคราะห์การถดถอยลอจิสต์ค่อนข้างน้อย ยังพบว่าตัวแปรอิสระต่อเนื่องมักถูกแปลงเป็นตัวแปรกลุ่มด้วยการจัดกลุ่มเป็น 2 กลุ่มหรือมากกว่า สาเหตุอาจเกิดจากข้อจำกัดของเนื้อหาที่วารสารทำให้การนำเสนอในประเด็นสำคัญขาดหายไปหรืออาจเกิดจากนักวิจัยบางส่วนเข้าใจว่าการจัดกลุ่มตัวแปรอิสระต่อเนื่องเป็นวิธีการที่ควรปฏิบัติไม่มีประสพการณ์ในการวิเคราะห์ตัวแปรอิสระต่อเนื่อง ไม่ทราบความสำคัญและวิธีการที่เหมาะสมในการวิเคราะห์ ซึ่งการจัดการตัวแปรอิสระต่อเนื่องที่ไม่เหมาะสมนั้นมีผลต่อเนื้อหาของตัวแปร³และการจัดกลุ่มตัวแปรอิสระต่อเนื่องนอกจากจะทำให้สูญเสียความสามารถในการสื่อความหมายของข้อมูล ยังทำให้ตัวแบบสูญเสียประสิทธิภาพในการทำนาย^{2,3,6}

สรุป

ผลการศึกษาในครั้งนี้แสดงให้เห็นได้ว่าบทความวิจัยที่ตีพิมพ์ในวารสารทางด้านกายภาพและสาธารณสุข มีการนำเสนอข้อมูลในประเด็นที่เกี่ยวข้องกับการตรวจสอบความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปแบบจุดและรายละเอียดในการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าไปในการวิเคราะห์การถดถอยลจีสติค น้อยมาก ยังพบว่าตัวแปรอิสระต่อเนื่องมักถูกแปลงเป็นตัวแปรกลุ่มด้วยการจัดกลุ่มเป็น 2 กลุ่มหรือมากกว่า โดยที่ในบทความวิจัยไม่ได้กล่าวถึงหลักเกณฑ์ที่อ้างอิงในการจัดกลุ่มตัวแปรอิสระต่อเนื่อง

ข้อจำกัด

การศึกษานี้เป็นการทบทวนรายละเอียดและประเมินการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลจีสติคจากข้อความที่ปรากฏอยู่ในบทความวิจัยตามเกณฑ์ที่นักวิจัยสร้างขึ้น ผลการศึกษาเพียงแสดงให้เห็นภาพของการรายงานการจัดการตัวแปรอิสระต่อเนื่องในการวิเคราะห์การถดถอยลจีสติคการตระหนักและการให้ความสำคัญในการจัดการตัวแปรอิสระต่อเนื่องที่เหมาะสมในการใช้การวิเคราะห์การถดถอยลจีสติคในการวิเคราะห์ข้อมูลที่ปรากฏอยู่ในบทความที่ทบทวนเท่านั้น

ข้อเสนอแนะในการวิจัย

นักวิชาการ นักวิจัยหรือนักสถิติควรตระหนักและให้ความสำคัญในการรายงานผลการทดสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปแบบจุด

นักวิชาการ นักวิจัยหรือนักสถิติควรตระหนักและให้ความสำคัญในการจัดการตัวแปรอิสระต่อเนื่องที่เหมาะสมก่อนนำเข้าไปในการวิเคราะห์การถดถอยลจีสติค

ผู้ทรงคุณวุฒิหรือผู้เชี่ยวชาญที่ทำหน้าที่ตรวจสอบบทความวิจัยก่อนการตีพิมพ์ของแต่ละวารสาร (Peer review) ควรพิจารณาในเรื่องการตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระต่อเนื่องกับตัวแปรตามในรูปแบบจุดและการรายงานการจัดการตัวแปรอิสระต่อเนื่องก่อนนำเข้าไปในการวิเคราะห์การถดถอยลจีสติคเพื่อให้ผู้ใช้งานวิจัยสามารถประเมินความเหมาะสม ความน่าเชื่อถือของบทความวิจัยได้

กิตติกรรมประกาศ

ขอขอบพระคุณ รองศาสตราจารย์ ดร.นพ. สมพงษ์ ศรีแสนปาง ผู้ช่วยศาสตราจารย์ ดร.จุฬาลักษณ์ โกมลตรี ผู้ช่วยศาสตราจารย์ ดร. ศิริพร คำสะอาด ผู้ช่วยศาสตราจารย์ ดร.พอใจ พิทยานิตย์ธรรม และ ดร.กรรณิการ์ ตฤณวุฒิมพงษ์ ที่กรุณาตรวจสอบความสมบูรณ์ของเครื่องมือในการเก็บรวบรวมข้อมูลวิจัย รวมทั้งให้ข้อเสนอแนะที่ดียิ่งในการศึกษาครั้งนี้

เอกสารอ้างอิง

1. Turner EL, Dobson JE, Pocock SJ. Categorization of continuous risk factors in epidemiological publications: a survey of current practice. *Epidemiol Perspect Innov* 2010; 7: 9.
2. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: A bad idea. *Stat Med* 2006; 25: 127-41.
3. Schellingerrout JM, Heymans MW, De Vet HCW, Koes BW, Verhagen AP. Categorizing continuous variables resulted in different predictors in a prognostic model for nonspecific neck pain. *J Clin Epidemiol* 2009; 68: 868-74.
4. กรรณิการ์ ตฤณวุฒิมพงษ์, จิราพร เขียวอยู่. การประเมินการใช้การวิเคราะห์ถดถอยพหุลจีสติคในวารสารทางการแพทย์และสาธารณสุขในประเทศไทยสองวารสาร. *วารสารวิชาการสาธารณสุข* 2552; 18: 665-73.
5. บัณฑิต ถิ่นคำรพ. การวิเคราะห์ข้อมูลการวิจัยทางวิทยาศาสตร์สุขภาพโดยใช้การถดถอยลจีสติค. ขอนแก่น: ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น, 2543.
6. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using "Optimal" cutpoint in the evaluation of prognostic factor. *J Natl Cancer Inst* 1994; 86: 829-35.

