

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของการวิจัย

วิธีการวิเคราะห์ข้อมูลทางสถิติส่วนใหญ่มีข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระภายในตัวแปรตาม โดยตัวแปรตามต้องไม่มีความสัมพันธ์กันภายในตัวเองแต่อย่างไรก็ตามมีข้อมูลบางประเภทที่มีลักษณะไม่สอดคล้องกับข้อตกลงข้างต้น เช่น ข้อมูลจากการวิจัยระยะยาว (Longitudinal Study) และข้อมูลที่มีการเก็บรวบรวมแบบวัดซ้ำ (Repeated Measurement) เป็นต้น ถ้าใช้วิธีการวิเคราะห์ข้อมูลแบบที่ไม่คำนึงถึงความสัมพันธ์ภายในตัวแปรตาม อาจส่งผลกระทบต่อคุณสมบัติของตัวประมาณค่าที่ได้ โดยเฉพาะเมื่อความสัมพันธ์มีค่ามาก ค่าที่ประมาณได้จะมีประสิทธิภาพลดลง ดังนั้น เพื่อเพิ่มความเชื่อมั่นของค่าที่ประมาณได้ จึงควรใช้วิธีการวิเคราะห์ข้อมูลที่ทำให้ได้มาซึ่งตัวประมาณที่ดีที่สุดมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ที่ดีที่สุดและต้องไม่มีความเอนเอียงด้วย (McCullagh and Nelder, 1989; Pindyck and Rubinfeld, 1998) ซึ่งเทคนิคในการวิเคราะห์ข้อมูลที่มีความสัมพันธ์กันของตัวแปรตามนั้น ขึ้นอยู่กับประเภทและความซับซ้อนของข้อมูล

Generalized Estimating Equations (GEE) เป็นวิธีการหนึ่งที่กำลังได้รับความสัมพันธ์ภายในตัวแปรตาม ซึ่งวิธีการนี้ได้พัฒนามาจาก Generalized Linear Model (GLM) โดย Liang และ Zeger (1986) และ GEE มีวัตถุประสงค์เพื่อหาความสัมพันธ์ระหว่างตัวแปรอิสระต่างๆ กับตัวแปรตาม ซึ่งมีการนำเสนอในรูปแบบของวิธีการวิเคราะห์การถดถอย และวิธีการนี้สามารถเพิ่มประสิทธิภาพ และไม่มีความเอนเอียงในการประมาณค่าพารามิเตอร์

Harrison และ Hulin (1989) ได้แสดงให้เห็นว่า GEE เป็นเทคนิคในการวิเคราะห์ข้อมูลที่ตัวแปรตามมีความสัมพันธ์ในตัวเอง และยังสามารถอนุโลมใช้ได้กับข้อมูลที่ตัวแปรตามไม่มีการแจกแจงแบบปกติ ปัจจุบันได้มีการนำวิธีการนี้ไปใช้อย่างแพร่หลายทางด้าน ระบาดวิทยา การแพทย์ และชีววิทยา เป็นต้น

การศึกษาครั้งนี้มีจุดมุ่งหมาย เพื่ออธิบายวิธีการวิเคราะห์ข้อมูลโดยใช้ GEE และวิธีการเลือกตัวแบบที่เหมาะสมในการวิเคราะห์ GEE โดยใช้เกณฑ์ Quasi Information Criterion

(QIC) รวมถึงขั้นตอนวิธีการในการนำ GEE ไปประยุกต์ใช้ในการวิเคราะห์ข้อมูล “โครงการวิจัยระยะยาวในเด็กไทย” (The Prospective Cohort study of Thai Children: PCTC) เพื่อศึกษาผลของระยะเวลาการเลี้ยงลูกด้วยนมมารดา และอายุที่เริ่มให้อาหารเสริมที่มีต่อการเจริญเติบโตด้านน้ำหนักและภาวะน้ำหนักน้อยกว่าเกณฑ์

1.2 วัตถุประสงค์

1.2.1 วัตถุประสงค์ทั่วไป

ศึกษาวิธีการวิเคราะห์ข้อมูลการวิจัยระยะยาว และวิธีการเลือกตัวแปรที่เหมาะสม

1.2.2 วัตถุประสงค์เฉพาะ

1. ศึกษาวิธีการวิเคราะห์ข้อมูลการวิจัยระยะยาวด้วยวิธี GEE
2. ศึกษาวิธีการเลือกตัวแปรที่เหมาะสม ในการวิเคราะห์ GEE โดยใช้เกณฑ์ QIC
3. ศึกษาผลของระยะเวลาการเลี้ยงลูกด้วยนมมารดาและอายุที่เริ่มให้อาหารเสริมต่อการเจริญเติบโตด้านน้ำหนัก และภาวะน้ำหนักน้อยกว่าเกณฑ์ในเด็กแรกเกิดถึง 2 ปี
4. ศึกษาผลของภาวะน้ำหนักน้อยกว่าเกณฑ์ที่มีต่อระยะเวลาการเลี้ยงลูกด้วยนมมารดา

1.3 ประโยชน์ที่คาดว่าจะได้รับ

1. ทราบแนวคิด ทฤษฎี และแนวทางในการนำวิธี GEE ไปประยุกต์ใช้
2. นำความรู้เกี่ยวกับความสัมพันธ์ระหว่างระยะเวลาเลี้ยงลูกด้วยนมมารดา และอายุที่เริ่มให้อาหารเสริม กับการเจริญเติบโตด้านน้ำหนัก และภาวะน้ำหนักน้อยกว่าเกณฑ์ของเด็กแรกเกิดถึง 2 ปี ที่ได้ไปพัฒนากลยุทธ์ในการแนะนำวิธีการปฏิบัติเกี่ยวกับการให้อาหารทารกและเด็กที่เหมาะสมและมีศักยภาพยิ่งขึ้น
3. นำวิธีการศึกษามาใช้เป็นแนวทางในการศึกษาวิจัยในประเด็นที่เกี่ยวข้อง

1.4 ขอบเขตการศึกษาและวิธีการดำเนินการวิจัย

1.4.1 ขอบเขตการศึกษา

การวิจัยครั้งนี้เป็นการศึกษาวิธีการ GEE และนำวิธีการ GEE มาใช้ในการวิเคราะห์ข้อมูลการวิจัยระยะยาว โดยสร้างตัวแบบความสัมพันธ์ระหว่างภาวะน้ำหนักน้อยกว่าเกณฑ์ในเด็กแรกเกิดถึง 2 ปี กับ ระยะเวลาการเลี้ยงทารกด้วยนมมารดา อาหารเสริม พื้นฐานของมารดาและครอบครัว และข้อมูลทั่วไปของเด็ก และทำการคัดเลือกตัวแบบที่เหมาะสมโดยใช้เกณฑ์ QIC รวมทั้งสร้างตัวแบบของการเจริญเติบโตด้านน้ำหนัก และตัวแบบของการเลี้ยงทารกด้วยนมมารดา

1.4.2 แหล่งข้อมูล

ข้อมูลได้มาจาก “โครงการวิจัยระยะยาวในเด็กไทย” ซึ่งโครงการดังกล่าวได้เริ่มเก็บข้อมูลหญิงตั้งครรภ์ทั้งหมดที่มีกำหนดคลอดใน 1 ปี ปฏิทินที่กำหนดไว้ในแต่ละพื้นที่ ในช่วงเวลาตั้งแต่ปลายปี พ.ศ. 2543 เป็นต้นมา ใน 5 พื้นที่ของประเทศไทย ประกอบด้วย 1) อำเภอนมทวน จังหวัดกาญจนบุรี 2) อำเภเทพา จังหวัดสงขลา 3) อำเภอกระนวน จังหวัดขอนแก่น 4) อำเภอเมือง จังหวัดน่าน 5) กรุงเทพมหานคร ภายใต้การสนับสนุนของสำนักงานกองทุนสนับสนุนการวิจัย สถาบันวิจัยระบบสาธารณสุข กระทรวงสาธารณสุข และองค์การอนามัยโลก

1.4.3 การเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการขออนุญาตใช้ข้อมูลประกอบด้วย น้ำหนักตามเกณฑ์อายุ ข้อมูลพื้นฐานครอบครัวและมารดา (อายุ ระดับการศึกษา อาชีพ จำนวนครั้งของการตั้งครรภ์ สภาพสังคมเศรษฐกิจ ศาสนา พื้นที่อยู่อาศัย) ข้อมูลทั่วไปของเด็ก (อายุครรภ์ เพศ น้ำหนักแรกเกิด การเจ็บป่วยในช่วง 2 ปี) การให้นม และอาหารเสริม (ระยะเวลาการให้นมมารดา อายุที่เริ่มให้นมผสม อายุที่เริ่มให้อาหารเสริมชนิดต่างๆ) หลังจากนั้นผู้วิจัยได้นำข้อมูลมาตรวจสอบความครบถ้วนสมบูรณ์ และจัดการข้อมูลให้เหมาะสมกับวิธีการวิเคราะห์ โดยคัดเลือกข้อมูลในช่วงเวลาที่มีการเก็บรวบรวมข้อมูลมีความสมบูรณ์ที่สุดใน 5 ช่วงเวลา คือ แรกเกิด อายุ 6 เดือน 1 ปี 1 ปีครึ่ง และ 2 ปีตามลำดับ

1.4.4 การวิเคราะห์ข้อมูล

1. Descriptive Statistics: วิเคราะห์ข้อมูลเบื้องต้นโดยใช้ ความถี่ และร้อยละ เมื่อนำเสนอข้อมูลเชิงคุณภาพ ใช้ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน เมื่อนำเสนอข้อมูลเชิงปริมาณ
2. Univariate Analysis: ศึกษาความสัมพันธ์ระหว่างภาวะน้ำหนักน้อยกว่าเกณฑ์ กับตัวแปรอิสระเชิงคุณภาพทีละตัวแปรใช้ Pearson's Chi-Squared test และศึกษาความแตกต่างของค่าเฉลี่ยการเจริญเติบโตด้านน้ำหนัก กับตัวแปรอิสระเชิงคุณภาพทีละตัวแปรใช้ t-test และ ANOVA
3. Multivariate Analysis: สร้างตัวแบบความสัมพันธ์กรณีตัวแปรตามเป็นตัวแปรเชิงปริมาณ โดย Multiple Regression และกรณีตัวแปรตามเป็นตัวแปรเชิงคุณภาพโดย Logistic Regression
4. Analysis of Longitudinal Data: สร้างตัวแบบความสัมพันธ์โดยใช้ GEE เมื่อนำถึงถึงความสัมพันธ์ภายในตัวแปรตาม และใช้เกณฑ์ QIC ในการเลือกตัวแบบที่เหมาะสม วิเคราะห์ข้อมูลโดยใช้โปรแกรม STATA Version 7

1.5 นิยามศัพท์

การเจริญเติบโตด้านน้ำหนัก	หมายถึง	ค่ามาตรฐานน้ำหนักตามเกณฑ์อายุ (z-score of weight for age)
ภาวะน้ำหนักน้อยกว่าเกณฑ์	หมายถึง	ค่ามาตรฐานน้ำหนักตามเกณฑ์อายุน้อยกว่า -2SD
การวิจัยระยะยาว	หมายถึง	การศึกษาข้อมูลที่มีการเก็บรวบรวมข้อมูลจากหน่วยตัวอย่างเดียวกันในช่วงเวลาต่างๆ กันอย่างต่อเนื่อง
การวิจัยแบบภาคตัดขวาง	หมายถึง	การศึกษาข้อมูลที่มีการเก็บรวบรวมข้อมูลเพียงครั้งเดียวจากหน่วยตัวอย่าง
GLM	หมายถึง	Generalized Linear Model
GEE	หมายถึง	Generalized Estimating Equation
QIC	หมายถึง	Quasi Information Criterion