

ตัวแบบพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone ผู้ป่วยในเด็กโรคโลหิตจางธาลัสซีเมีย

พัชรี มณีรัตน์^{1*} และ พิศิษฐ์ นาคใจ¹

^{1,2}คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏอุดรดิตถ์

เลขที่ 27 ถนนอินใจมี ตำบลท่าอิฐ อำเภอเมือง จังหวัดอุดรดิตถ์ 53000, M.patcharee@uru.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบการพยากรณ์ที่เหมาะสมกับปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone ในผู้ป่วยเด็กโรคโลหิตจางธาลัสซีเมีย ณ คลินิกโรคเลือดเด็ก โรงพยาบาลอุดรดิตถ์ ทำการสำรวจข้อมูลย้อนหลัง จากเวชระเบียนผู้ป่วยตั้งแต่ 1 มกราคม 2554 ถึง 31 ธันวาคม 2558 จำนวน 15 รายที่รักษาด้วยยา Deferasirox และจำนวน 28 รายที่รักษาด้วยยา Deferiprone โดยศึกษาความสัมพันธ์ระหว่างคุณลักษณะทั่วไปของผู้ป่วย, ระดับธาตุเหล็กในเลือดก่อนการรักษา, ผลข้างเคียง, ขนาดยา กับปริมาณธาตุเหล็กในเลือดหลังรักษาจากยาทั้ง 2 ชนิด สำหรับระเบียบวิธีทางสถิติที่ใช้คือ การวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ พร้อมทั้งการตรวจสอบความผิดพลาดและความเชื่อถือมั่นพิจารณาจากสัมประสิทธิ์ตัวกำหนดแบบพหุที่มีการปรับค่าแล้ว (R_{adj}^2) และความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ผลการศึกษาพบว่า ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox มีความสัมพันธ์กับปริมาณธาตุเหล็กก่อนรักษา และระดับเอนไซม์ตับ ALT ที่เพิ่มขึ้น โดยที่ $R_{adj}^2 = 0.7550$ และ $MSE = 0.011$ และปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferiprone มีความสัมพันธ์กับการวินิจฉัยของโรค, อายุ, สถานะของม้าม, ปริมาณธาตุเหล็กในเลือดก่อนการรักษา และระดับเอนไซม์ตับ AST ที่เพิ่มขึ้น โดยที่ $R_{adj}^2 = 0.506$ และ $MSE = 0.014$

คำสำคัญ: Deferasirox, Deferiprone, การวิเคราะห์การถดถอยเชิงพหุ

Model of Serum Ferritin level by Deferasirox and Deferiprone In Pediatric Thalassemic Patients

Patcharee Maneerat^{1*} and Pisit Nakjai¹

^{1,2}Faculty of Science and Technology, Uttaradit Rajabhat University

No.27 In-jai-mee Road, Ta-it Sub-district, Muang District, Uttaradit, 53000. M.patcharee@uru.ac.th

Abstract

The objectives of this research are to construct model approximation using Serum ferritin level by Deferasirox and Deferiprone in pediatric Thalassemic patients at clinical trial, Uttaradit hospital. This study is the retrospective cohort study from January 2011 to December 2014. There are 15 patients treated with Deferasirox and 28 patients treated with Deferiprone. This model is the relation between general performances, pre- Serum Ferritin, the sided effect of a drug, doses and post - Serum Ferritin by drug cases. The statistical methodology is used multiple linear regression. The error investigation and reliability was considered by Adjusted R square and Mean Square Error. The results of this research are as follows: The post - Serum Ferritin of Deferasirox is involved with pre- Serum Ferritin and elevated ALT while $R_{adj}^2 = 0.755$, $MSE = 0.011$. The post - Serum Ferritin of Deferiprone is associated with Medical diagnosis, age, splenectomy, pre- Serum Ferritin and elevated AST while $R_{adj}^2 = 0.506$, $MSE = 0.014$

Keywords: Deferasirox, Deferiprone, multiple linear regression

บทนำ

ในปัจจุบันบุคลากรทางการแพทย์และประชาชนทั่วไปคงได้ยินคำว่า “โรคโลหิตจางธาลัสซีเมีย” บ่อยครั้ง กระทรวงสาธารณสุขได้ให้ความสำคัญกับโรคนี้เป็นอย่างมาก เพราะว่าเป็นโรคที่มีการถ่ายทอดทางพันธุกรรมมากที่สุดในประเทศไทย เนื่องจากมีคนไทยเป็นพาหะของโรคนี้ ร้อยละ 30 ของประชากรหรือประมาณ 18 - 20 ล้านคน ซึ่งในส่วนของโรคเลือดในเด็กสาเหตุส่วนหนึ่งเกิดจากความผิดปกติของยีนในการควบคุมการสร้างฮีโมโกลบินซึ่งเป็นสารสำคัญในเม็ดเลือดแดง ผู้ป่วยบางคนตัวซีดมาก ตับม้ามโตมาก [1] สำหรับสาเหตุการเสียชีวิตในผู้ป่วยเป็นผลมาจากวิธีการรักษาคือ ผู้ป่วยทุกคนจำเป็นต้องได้รับการให้เลือดเป็นประจำ แต่อาจพบภาวะแทรกซ้อนที่เกิดขึ้นได้บ่อยคือ ภาวะเหล็กเกิน อันเกิดจากการสะสมธาตุเหล็กในร่างกายมากเกินไป จนก่อให้เกิดผลเสียต่อการทำงานของอวัยวะต่างๆ ในร่างกาย [2]

การรักษาด้วยยาขับธาตุเหล็กในผู้ป่วยที่มีภาวะเหล็กเกินนั้นมีอัตราการรอดชีวิตเพิ่มขึ้น เนื่องจากมียาที่ใช้รักษาคือยา desferrioxamine (Desferal®) เป็นยาขับเหล็กที่มีประสิทธิภาพดี แต่ต้องบริหารยาด้วยการฉีดเข้าใต้ผิวหนัง ซึ่งผู้ป่วยเด็กส่วนใหญ่ มักไม่ให้ความร่วมมือ ปัจจุบันมียาขับเหล็กชนิดรับประทานได้แก่ Deferipron (GPO-L-ONE®) และ Deferasirox (Exjade®) สามารถบริหารยาได้ง่าย ลดระดับธาตุเหล็กในเลือดและหัวใจได้ดี และที่สำคัญสามารถใช้ได้อย่างปลอดภัยสำหรับผู้ป่วยเด็ก [2]

การให้ยาทั้ง 2 ชนิดมีวัตถุประสงค์ที่สำคัญเพื่อลดปริมาณธาตุเหล็กในผู้ป่วยเด็กโรคโลหิตจางธาลัสซีเมีย ซึ่งปัญหาที่เป็นประเด็นสำคัญคือ สิ่งใดบ้างที่มีอิทธิพลต่อปริมาณธาตุเหล็กในเลือดของผู้ป่วยเด็กดังกล่าว อันนำมาสู่ประเด็นในการศึกษาครั้งนี้คือการสร้างตัวแบบพยากรณ์ที่เหมาะสมกับปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone ในผู้ป่วยเด็กโรคโลหิตจางธาลัสซีเมีย ณ คลินิกโรคเลือดเด็ก โรงพยาบาลอุดรดิตถ์ ซึ่งพิจารณาจากตัวแปรที่มีอิทธิพลต่อปริมาณธาตุเหล็กคือ อายุ (Age), การวินิจฉัยโรค (Medical diagnosis), สถานะของม้าม (splenectomy), ปริมาณธาตุเหล็กก่อนรักษา (Pre-serum ferritin), ขนาดของยาที่ได้รับ, ระดับเอนไซม์ตับ Aspartate Aminotransferase ที่เพิ่มขึ้น (AST), ระดับเอนไซม์ตับ Alanine Aminotransferase ที่เพิ่มขึ้น (ALT), การทำงานของไตผิดปกติ (Rising Creatinine), โปรตีนรั่วในปัสสาวะ (Proteinuria), ระดับเม็ดเลือดขาวนิวโทรฟิล (absolute neutrophil count: ANC) โดยการใช้ระเบียบวิธีทางสถิติคือ การวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ (Multiple Linear Regression พร้อมทั้งการตรวจสอบความผิดพลาดและความเชื่อถือได้อันพิจารณาจากสัมประสิทธิ์ตัวกำหนดแบบพหุที่มีการปรับค่าแล้ว (Adjusted R square: R_{adj}^2), ความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE)

ซึ่งหากได้ตัวแบบพยากรณ์ที่เหมาะสมแล้ว จะสามารถทราบถึงปัจจัยที่สำคัญอันมีอิทธิพลต่อปริมาณธาตุเหล็กในเลือดของผู้ป่วยเด็กโรคโลหิตจางธาลัสซีเมียจากการรักษาด้วยยาทั้ง 2 ชนิด ท้ายสุดแล้วข้อสรุปที่ได้รับมีความน่าเชื่อถือสะท้อนให้เห็นถึงสถานการณ์จริงที่เกิดขึ้นนำไปสู่แนวทางการพัฒนา และแก้ไขปรับปรุงในการรักษาผู้ป่วยโรคดังกล่าวต่อไป

วิธีดำเนินการวิจัย

1. ขอบเขตประชากร/ขอบเขตตัวอย่าง

1.1 ขอบเขตประชากร

ผู้ป่วยเด็กอายุ 2-15 ปีที่ได้รับการวินิจฉัยว่าเป็นโรคโลหิตจางธาลัสซีเมียที่มีภาวะเหล็กเกิน และได้รับการรักษาด้วยการใช้ยาขับธาตุเหล็ก Deferasirox หรือ Deferipron ที่คลินิกโรคเลือดเด็ก โรงพยาบาลอุดรดิตถ์

1.2 ขอบเขตตัวอย่าง

ผู้ป่วยเด็กที่ได้รับการวินิจฉัยว่าเป็นโรคโลหิตจางธาลัสซีเมียที่มีภาวะเหล็กเกิน และได้รับการรักษาด้วยการใช้ยาขับธาตุเหล็ก Deferasirox จำนวน 15 ราย และยาขับธาตุเหล็ก Deferiprone จำนวน 28 ราย ณ คลินิกโรคเลือดเด็ก โรงพยาบาลอุดรดิตถ์ ตั้งแต่วันที่ 1 มกราคม 2554 ถึง 31 ธันวาคม 2558

2. ขอบด้านตัวแปร ประกอบด้วยดังนี้

2.1 ตัวแปรตาม แสดงดังนี้

Y_1 = ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox : หน่วยเป็น ng/m

Y_2 = ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferiprone : หน่วยเป็น ng/m

2.2 ตัวแปรอิสระ แสดงดังนี้

X_1 = อายุ : หน่วยเป็นเดือน , X_2 = การวินิจฉัยโรค
 X_3 = สถานะของมะเร็ง , X_4 = ปริมาณธาตุเหล็กก่อนรักษา : หน่วยเป็น ng/m
 X_5 = ขนาดของยาที่ได้รับ : หน่วยเป็น $mg/kg/day$
 X_6 = ระดับเอนไซม์ตับ AST ที่เพิ่มขึ้น , X_7 = ระดับเอนไซม์ตับ ALT ที่เพิ่มขึ้น
 X_8 = การทำงานของไตผิดปกติ , X_9 = โปรตีนรั่วในปัสสาวะ
 X_{10} = ระดับเม็ดเลือดขาวนิวโทรฟิล : หน่วยเป็น $cell/mm^3$

3. ระเบียบวิธีสถิติที่ใช้

ทำการสร้างตัวแบบการพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone โดยใช้ระเบียบวิธีทางสถิติคือ การวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ ซึ่งมีรายละเอียดดังต่อไปนี้

3.1 การวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ

ในทางปฏิบัติการสร้างสมการการถดถอยแสดงความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระเพียง 1 ตัวเกิดขึ้นได้ไม่บ่อยนัก เนื่องจากย่อมมีหลายปัจจัย หรืออิทธิพลที่มีต่อตัวแปรตาม อันเป็นที่มาของการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ เพื่อรองรับสถานการณ์ที่เกิดขึ้นดังกล่าว

3.1.1 ตัวแบบถดถอยเชิงเส้นตรงแบบพหุ

ตัวแบบดังกล่าวแสดงความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ k ตัว มีรูปแบบดังนี้ [3]

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad ; i = 1, 2, 3, \dots, n \quad \dots\dots\dots(1)$$

กำหนดให้

- Y_i แทนตัวแปรตามของค่าสังเกตที่ i
- X_{ij} แทนตัวแปรอิสระของค่าสังเกตที่ i ลำดับที่ j โดยที่ $j = 0, 1, 2, \dots, k$ และ $X_{i0} = 1$
- β_j แทนค่าสัมประสิทธิ์การถดถอย (Regression Coefficients) โดยที่ $j = 0, 1, 2, \dots, k$
- ε_i แทนความคลาดเคลื่อนเชิงสุ่ม (Random error)

จากตัวแบบ (1) ค่าสัมประสิทธิ์การถดถอย β_j ถือเป็นพารามิเตอร์ที่ไม่ทราบค่า ซึ่งใช้แทนสำหรับการเปลี่ยนแปลงของค่าเฉลี่ยของ Y ต่อการเปลี่ยนแปลงของ X_j 1 หน่วย เมื่อตัวแปรอิสระอื่นคงที่ ด้วยเหตุนี้จึงนิยมเรียก β_j ว่าสัมประสิทธิ์การถดถอยบางส่วน (Partial regression coefficients)

3.1.2 ข้อตกลงเบื้องต้น (Assumptions) ของการวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ

1. ε_i เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน และมีรูปแบบการแจกแจงเดียวกัน คืออยู่ภายใต้รูปแบบการแจกแจงปกติด้วยค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ^2 แทนเป็นสัญลักษณ์ $\varepsilon_i \sim NID(0, \sigma^2)$
2. Y_i เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน และมีรูปแบบการแจกแจงเดียวกัน คืออยู่ภายใต้รูปแบบการแจกแจงปกติด้วยค่าเฉลี่ยเท่ากับ $\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}$ และความแปรปรวนเท่ากับ σ^2 แทนเป็นสัญลักษณ์ $Y_i \sim NID(\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}, \sigma^2)$
3. X_j เป็นตัวแปรที่ทราบค่า (Known value) เมื่อ $j = 1, 2, \dots, k$
4. X_j และ X_l เป็นอิสระต่อกัน เมื่อ $j, l = 1, 2, \dots, k$

3.1.3 การประมาณค่าพารามิเตอร์โดยวิธีกำลังสองน้อยที่สุด

การประมาณค่าพารามิเตอร์ในแบบการวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ สามารถทำได้เช่นเดียวกับการประมาณค่าพารามิเตอร์ในแบบการถดถอยอย่างง่าย เมื่อเก็บข้อมูลตัวอย่างประกอบด้วยค่าสังเกต n ค่า โดยที่ $n > k$

การสร้างตัวประมาณพารามิเตอร์ $\beta_0, \beta_1, \dots, \beta_k$ โดยวิธีกำลังสองน้อยที่สุดมีหลักการแนวคิดที่ว่า ตัวประมาณค่าพารามิเตอร์ดังกล่าวต้องสามารถทำให้ผลรวมกำลังสองของความคลาดเคลื่อน (Error sum of square : SSE) มีค่าต่ำสุด โดยที่ความคลาดเคลื่อนเกิดจากความแตกต่างระหว่างค่าสังเกต Y_i กับค่าพยากรณ์ที่ได้จากสมการถดถอย \hat{y} แสดงดังนี้

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik})^2 \quad \dots\dots\dots (2)$$

ดังนั้นจึงทำให้

$$SSE = \sum_{i=1}^n \left(Y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij} \right) \quad \dots\dots\dots (3)$$

จากนั้นการสร้างตัวประมาณพารามิเตอร์ β_j โดยที่ $j = 0, 1, 2, \dots, k$ ทำได้โดยการหาอนุพันธ์ของ SSE เทียบกับพารามิเตอร์แต่ละตัวดังนี้

$$\frac{\partial SSE}{\partial \beta_0} = -2 \sum_{i=1}^n \left(Y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij} \right) \quad \dots\dots\dots (4)$$

$$\vdots$$

$$\frac{\partial SSE}{\partial \beta_k} = -2 \sum_{i=1}^n X_{ik} \left(Y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij} \right) \quad \dots\dots\dots (5)$$

แล้วให้สมการ (4) และ (5) เท่ากับ 0 และ แทนค่า $\beta_0, \beta_1, \dots, \beta_k$ ด้วย b_0, b_1, \dots, b_k หลังจากนั้นจัดรูปแบบสมการข้างต้นใหม่ จะได้สมการปกติ $p = k + 1$ สมการ สามารถสร้างตัวประมาณกำลังสองน้อยที่สุดของ $\beta_0, \beta_1, \dots, \beta_k$ แต่การแก้สมการทั้งหมดด้วยหลักการพีชคณิตค่อนข้างยุ่งยาก โดยเฉพาะอย่างยิ่งเมื่อมีตัวแปรอิสระหลายตัว ดังนั้นเพื่อสะดวกต่อการสร้างสมการถดถอยจากสมการ (1) สามารถเขียนอยู่ในรูปของเมทริกซ์ได้ดังนี้

$$Y = X\beta + \varepsilon \quad \dots\dots\dots (6)$$

จากสมการ (2) สามารถเขียนผลรวมกำลังสองให้อยู่ในรูปของเมทริกซ์ได้ดังนี้

$$SSE = \sum_{i=1}^n e^2 = \varepsilon' \varepsilon = (Y - X\beta)'(Y - X\beta) = Y'Y - 2\beta'X'Y + \beta'X'X\beta \quad \dots\dots\dots (7)$$

จาก (7) หาอนุพันธ์เทียบกับพารามิเตอร์ β แล้วให้สมการเท่ากับ 0 และแทนค่า β ด้วยตัวประมาณ b จะได้สมการปกติในรูปของเมทริกซ์เป็น

$$X'Xb = X'Y \quad \dots\dots\dots (8)$$

แก้สมการ(7) โดยการคูณ $(X'X)^{-1}$ ในทั้ง 2 ข้างของสมการ ดังนั้นตัวประมาณ b ด้วยวิธีกำลังสองน้อยที่สุดของ β คือ

$$b = (X'X)^{-1} X'Y \quad \dots\dots\dots (9)$$

เมทริกซ์ $(X'X)^{-1}$ หาได้เสมอ ถ้าหากตัวแปรอิสระไม่มีความสัมพันธ์กัน (Linearly independent) หมายความว่าไม่มีคอลัมน์ใดของ X มีความสัมพันธ์เชิงเส้นกับคอลัมน์อื่น จากข้างต้นเมทริกซ์ $(X'X)^{-1}$ เป็นเมทริกซ์สมมาตร (Symmetric matrix) ที่มีรูปแบบเฉพาะค่าที่อยู่บนเส้นทแยงมุม (Diagonal element) เป็นผลรวมกำลังของ X ใดๆ และค่าที่อยู่นอกเส้นทแยงมุม (Off-diagonal element) เป็นผลรวมของผลคูณ (Cross product) ระหว่าง X ใดๆ ส่วน $X'Y$ เป็นเวกเตอร์ขนาด $p \times 1$ ที่เกิดจากผลรวมของผลคูณระหว่าง X และค่าสังเกต Y ดังนั้นสมการถดถอยในรูปของเมทริกซ์ คือ $\hat{Y} = Xb$

3.1.4 เกณฑ์ในการคัดเลือกเลือกตัวแบบ

การเลือกตัวแปรอิสระที่อยู่ในตัวแบบมี 2 ขั้นตอนคือ การสร้างชุดของตัวแปรอิสระที่เป็นไปได้ทั้งหมด และพิจารณาเลือกชุดของสมการที่ดีกว่าชุดอื่น สำหรับเกณฑ์ที่ใช้ประเมินและเปรียบเทียบเพื่อการตรวจสอบความผิดพลาดและความเชื่อถือได้ของสมการถดถอยพิจารณาจากสัมประสิทธิ์ตัวกำหนดแบบพหุที่มีการปรับค่าแล้ว และความคลาดเคลื่อนกำลังสองเฉลี่ย แสดงรายละเอียดดังนี้

สัมประสิทธิ์ตัวกำหนดแบบพหุที่มีการปรับค่าแล้ว

สัมประสิทธิ์ตัวกำหนดแบบพหุ (R^2) มีค่าเพิ่มขึ้นเมื่อเพิ่มตัวแปรอิสระเข้าไปในตัวแบบ จึงเป็นที่มาของการพิจารณาสัมประสิทธิ์ตัวกำหนดแบบพหุที่มีการปรับค่าแล้ว สำหรับสับเซตของตัวแบบที่มี p เทอม ($R^2_{adj(p)}$) เพื่อช่วยป้องกันการเพิ่มตัวแปรอิสระที่ไม่จำเป็นเข้าไปในตัวแบบ (Overfitting) แสดงรูปแบบความสัมพันธ์ดังนี้

$$R^2_{adj(p)} = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} = 1 - \left(\frac{n-1}{n-p} \right) (1 - R^2) \quad \text{โดยที่ } 0 \leq R^2_{adj} \leq 1, p = k + 1$$

ความคลาดเคลื่อนกำลังสองเฉลี่ย

ความคลาดเคลื่อนกำลังสองเฉลี่ยสำหรับสับเซตของตัวแบบที่มี p เทอมแทนด้วย “ MSE_p ” แสดงความสัมพันธ์ดังนี้

$$MSE_p = \frac{SSE_p}{n - p}$$

สำหรับเกณฑ์ในการเลือกจำนวนตัวแปรอิสระ เมื่อพิจารณาจาก MSE_p คือ เลือก p ที่ทำให้ MSE_p ต่ำสุด โดยที่มีค่าใกล้เคียงกับ MSE ของ full model ดังนั้น การเลือกตัวแปรอิสระเข้าไปในตัวแบบโดยใช้เกณฑ์ที่ทำให้ MSE_p มีค่าต่ำสุดเทียบเท่ากับเกณฑ์ที่ทำให้ $R^2_{adj(p)}$ มีค่าสูงสุด

3.2 การตรวจสอบความเหมาะสมของสมการถดถอย

การอนุมานทางสถิติทั้งในส่วนการทดสอบสมมติฐาน และการประมาณค่าแบบช่วง ต้องอาศัยข้อกำหนดเกี่ยวกับการแจกแจงปกติของความคลาดเคลื่อน อันเป็นหนึ่งในข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอย ดังนั้นก่อนที่จะนำตัวแบบการถดถอยมาใช้ในการพิจารณาไปใช้ประโยชน์ จำเป็นอย่างยิ่งที่ต้องมีการตรวจสอบความเหมาะสมของตัวแบบก่อน ซึ่งการตรวจสอบดังกล่าวไม่สามารถใช้สถิติได้โดยทั่วไปจึงมีการตรวจสอบประกอบด้วยดังนี้

3.2.1 การวิเคราะห์ความคลาดเคลื่อนของตัวอย่าง

ความแตกต่างระหว่างค่าสังเกตจริง Y_i กับค่าที่อยู่บนเส้นสมการพยากรณ์ \hat{y}_i ถูกเรียกว่า “ความคลาดเคลื่อน” แสดงอยู่ในรูปแบบดังนี้

$$e_i = Y_i - \hat{y}_i \quad \text{โดยที่ } i = 1, 2, \dots, n$$

ความคลาดเคลื่อนในการพยากรณ์ เป็นค่าความผันแปรที่ไม่สามารถอธิบายได้ด้วยเส้นสมการถดถอย หากความคลาดเคลื่อนที่ได้จากตัวแบบการถดถอยมีลักษณะเบี่ยงเบนไปจากข้อตกลงเบื้องต้น การเบี่ยงเบนดังกล่าวจะปรากฏในความคลาดเคลื่อนของตัวอย่างด้วยเช่นกัน ดังนั้นความคลาดเคลื่อนตัวอย่างจึงถือว่าเป็นวิธีหนึ่งที่วัดประสิทธิภาพของสมการถดถอย ซึ่งความคลาดเคลื่อนตัวอย่างมีค่าเฉลี่ยเท่ากับ 0 ($\bar{e} = 0$) และค่าประมาณความแปรปรวนคงที่เท่ากับ

$$S^2 = MSE = \frac{SSE}{n - (k + 1)} = \frac{\sum_{i=1}^n e_i^2}{n - (k + 1)} = \frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n - (k + 1)}$$

3.2.2 การตรวจสอบความเท่ากันของความแปรปรวน (Test for Constancy of Error variance)

การตรวจสอบความแปรปรวนของของความคลาดเคลื่อนคงที่หรือไม่นั้น สามารถพิจารณาได้จากการสร้างกราฟระหว่าง e_i กับ \hat{y} หรือ X_i ซึ่งหากความคลาดเคลื่อนมีการกระจายสม่ำเสมออย่างสม่ำเสมอค่า 0 นั้นหมายความว่าความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่ ซึ่งเรียกว่า “Homoscedasticity” แต่ในทางกลับกันหากความคลาดเคลื่อนมีการกระจายที่มีรูปแบบไม่สม่ำเสมอ แสดงว่าสมการถดถอยเกิดปัญหาซึ่งเรียกว่า “Heteroscedasticity”

ดังนั้นวิธีการตรวจสอบความแปรปรวนของความคลาดเคลื่อนคงที่นั้นสามารถใช้ในการทดสอบสมมติฐานทางสถิติโดยใช้สถิติทดสอบของ Levene ที่มีการปรับค่า (Modified Levene Test) พร้อมทั้งการแก้ไขปัญหาเมื่อกรณีเกิดความแปรปรวนของความคลาดเคลื่อนไม่คงที่ แสดงดังต่อไปนี้

สถิติทดสอบของ Levene ที่มีการปรับค่า

สถิติทดสอบที่ถูกปรับ หรือพัฒนามาจากสถิติทดสอบของ Levene ซึ่งเป็นสถิติทดสอบที่ไม่ใช้พารามิเตอร์ (Non parametric Test) สำหรับสมมติฐานที่ต้องการทดสอบคือ $H_0 : Var(e_i) = \sigma^2$ กำหนดให้ e_{i1}, e_{i2} แทนความคลาดเคลื่อนที่ได้จากข้อมูลกลุ่มที่ 1 และ 2 ตามลำดับ, n_1, n_2 แทนขนาดตัวอย่างของข้อมูลในกลุ่มที่ 1 และ 2 ตามลำดับโดยที่ $n = n_1 + n_2$, \tilde{e}_1, \tilde{e}_2 แทนมัธยฐานของความคลาดเคลื่อนในกลุ่มที่ 1 และ 2 ตามลำดับ โดยที่ $d_{i1} = |e_{i1} - \tilde{e}_1|$, $d_{i2} = |e_{i2} - \tilde{e}_2|$ เมื่อ d_{i1}, d_{i2} แทนส่วนเบี่ยงเบนสัมบูรณ์ของความคลาดเคลื่อนรอบมัธยฐานของข้อมูลในกลุ่มที่ 1 และ 2 ตามลำดับ, \bar{d}_1, \bar{d}_2 แทนค่าเฉลี่ยของ d_{i1}, d_{i2} ตามลำดับ ดังนั้นสถิติทดสอบของ Levene ที่มีการปรับค่า คือ

$$t_L^* = \frac{\bar{d}_1 - \bar{d}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad \text{โดยที่} \quad S_p^2 = \frac{\sum_{i=1}^{n_1} (d_{i1} - \bar{d}_1)^2 + \sum_{i=1}^{n_2} (d_{i2} - \bar{d}_2)^2}{n - 2}$$

ณ ระดับนัยสำคัญ α ; หาก $|t_L^*| > t_{1-\alpha, n-2}$ หมายความว่าความแปรปรวนของค่าความคลาดเคลื่อนไม่เท่ากันหรือไม่คงที่ ทำให้เกิดปัญหา “Heteroscedasticity”

การแก้ไขปัญหาเมื่อความแปรปรวนของความคลาดเคลื่อนไม่คงที่

เมื่อความแปรปรวนของความคลาดเคลื่อนไม่คงที่ ถือเป็นปัญหาสำคัญหากไม่ได้รับการแก้ไขแล้ว แม้ว่าตัวประมาณกำลังสองเป็นตัวประมาณที่ไม่เอนเอียง แต่ขาดคุณสมบัติความแปรปรวนต่ำสุด มีผลให้ค่าสัมประสิทธิ์การถดถอยมีความคลาดเคลื่อนมาตรฐานสูงกว่าที่ควรจะเป็น ซึ่งการแปลงข้อมูลเป็นอีกวิธีหนึ่งที่ทำให้ค่าการประมาณค่าพารามิเตอร์ที่แม่นยำและมีความไวต่อการทดสอบสมมติฐานทางสถิติมากขึ้น ดังนั้นการแก้ไขปัญหาเมื่อเกิดความแปรปรวนของความคลาดเคลื่อนไม่คงที่ในการศึกษาครั้งนี้โดยการแปลงค่าตัวแปรตามด้วยการแปลงเป็นลอการิทึม (Logarithm)

3.2.4 การตรวจสอบความสัมพันธ์ของค่าความคลาดเคลื่อน

ในบางสถานการณ์ได้นำสมการถดถอยไปประยุกต์ใช้แล้วพบว่าตัวแปรอิสระ และตัวแปรตามขึ้นอยู่กับเวลาที่ผ่านไป ซึ่งเรียกข้อมูลลักษณะเช่นนี้ว่า ข้อมูลอนุกรมเวลา (Time Series Data) หากข้อมูลเกิดขึ้นในลักษณะเช่นนี้ มักส่งผลให้ความคลาดเคลื่อนมีความสัมพันธ์กัน (Autocorrelation) นั่นคือ

$$E(\varepsilon_i, \varepsilon_j) \neq 0 \text{ เมื่อ } i \neq j$$

สำหรับสาเหตุของปัญหาดังกล่าวคือ สมการถดถอยสร้างจากข้อมูลอนุกรมเวลา และที่สำคัญคือ การขาดตัวแปรอิสระที่สำคัญบางตัวในแบบการถดถอยนั่นเอง ที่ผ่านมามีนักสถิติหลายท่านได้คิดค้นวิธีเพื่อตรวจสอบปัญหา Autocorrelation ซึ่งสถิติทดสอบของ Durbin และ Watson ใช้สำหรับตรวจสอบปัญหาดังกล่าว โดยมีข้อกำหนดที่สำคัญคือข้อมูลมีระยะห่างของช่วงเวลาเท่าๆกัน และความคลาดเคลื่อนในตัวแบบการถดถอยอยู่ในกระบวนการ Autoregressive อันดับที่ 1 หรือ AR(1) พิจารณาตัวแบบของความคลาดเคลื่อน AR(1)

$$\varepsilon_i = \rho \varepsilon_{i-1} + a_i \quad \dots\dots\dots(10)$$

กำหนดให้ ε_t แทนความคลาดเคลื่อนในตัวแบบ ณ เวลา t ; $t = 1, 2, \dots, n$ สำหรับ ρ แทนค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง ε_t กับ ε_{t-1} โดยที่ $|\rho| < 1$ หาก $|\rho| > 1$ เกิดปัญหา Positive Autocorrelation และหาก $\rho < 0$ เกิดปัญหา Negative Autocorrelation และ a_t แทนตัวแปรปรวณเชิงสุ่ม (Random Disturbance) โดยที่ $a_t \sim NID(0, \sigma_a^2)$ จากสมการ (10) เมื่อแทนค่า $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ ต่อเนื่องกัน เข้าไปในสมการ (10) จะได้ว่า $\varepsilon_t = \sum_{u=0}^{\infty} \rho^u a_{t-u}$

จากข้างต้นความคลาดเคลื่อน ณ เวลา t ใดๆ เป็นฟังก์ชันเชิงเส้นของตัวแปรสุ่ม a_t ที่เกิดขึ้นก่อนหน้านี้ โดยที่ตัวแปรสุ่ม $a_t \sim NID(0, \sigma_a^2)$ และ $|\rho| > 1$ หมายถึงเมื่อระยะเวลาที่ยาวนานขึ้น $|\rho|$ จะมีค่าลดลง ทำให้ a_{t-u} มีค่าลดลงตามไปด้วย และหาก $|\rho| < 0$ จะเห็นได้ว่าช่วงเวลาที่ยาวนานมากขึ้นความคลาดเคลื่อนจะมีความสัมพันธ์กันน้อยลงจนกระทั่งไม่มีความสัมพันธ์เชิงเส้นตรงต่อกันเมื่อ $\rho = 0$ แล้วเนื่องจาก $\varepsilon_t = a_t$ และตัวแปรปรวณเชิงสุ่ม a_t เป็นอิสระต่อกัน ดังนั้นวิธีการตรวจสอบปัญหาความคลาดเคลื่อนมีความสัมพันธ์กัน โดยใช้สถิติทดสอบของ Durbin และ Watson เมื่อสมมติฐานคือ $H_0 : \rho = 0$ และ $H_1 : \rho > 0$

$$\text{สถิติทดสอบที่ใช้คือ} \quad d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad \dots\dots\dots(11)$$

เมื่อ e_t แทนความคลาดเคลื่อนของสมการถดถอยที่ได้จากวิธีกำลังสองน้อยที่สุด โดยที่ $e_t = Y_t - \hat{y}_t$; $t = 1, 2, \dots, n$
 สำหรับเกณฑ์ในการตัดสินใจ โดยคำนวณค่าวิกฤตที่แท้จริงของสถิติทดสอบ Durbin-Watson ทำได้ยากจึงได้กำหนดขอบเขตล่าง d_L และ ขอบเขตบน d_U ของสถิติทดสอบดังกล่าว หากอยู่นอกอาณาเขตที่กำหนดนำไปสู่การตัดสินใจคือ หาก $d > d_U$ จะยอมรับ H_0 เมื่อ $d < d_L$ จะปฏิเสธ H_0 และนอกจากขอบเขตนี้ถือว่าไม่สามารถสรุปไม่ได้

3.2.5 การตรวจสอบ Multicollinearity

หนึ่งในข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอยเชิงเส้นแบบพหุคือ ตัวแปรอิสระต้องไม่มีความสัมพันธ์กัน “Multicollinearity” ถือเป็นปัญหาที่ตัวแปรอิสระมีความสัมพันธ์กันที่ใกล้เคียงกับเส้นตรง (Near-linear dependence) ระหว่างตัวแปรอิสระ ส่งผลให้ $X'X$ เป็นเมทริกซ์ซิงกูลาร์ อันจะกระทบต่อการประมาณค่าสัมประสิทธิ์การถดถอย สำหรับการตรวจสอบปัญหา Multicollinearity ทำได้ด้วยการคำนวณ Variance Inflation Factor (VIF) แสดงดังต่อไปนี้

$$VIF_j = \frac{1}{1 - R_j^2} \quad \dots\dots\dots(12)$$

เมื่อ VIF_j แทนค่า Variance inflation Factor ของตัวแปรอิสระที่ j และ R_j^2 แทนค่าสัมประสิทธิ์ตัวกำหนดระหว่างตัวแปรอิสระ X_j กับตัวแปรอิสระตัวอื่น หาก $VIF_j > 5$ ถือว่าเกิดปัญหา “Multicollinearity” และ $VIF_j > 10$ ถือว่าเกิดปัญหา “Multicollinearity” ขั้นรุนแรง

ในการวิจัยครั้งนี้ทำการวิเคราะห์ และประมวลผลโดยใช้โปรแกรมทางสถิติ R project

วิธีดำเนินการรวบรวมข้อมูล

การศึกษานี้สำรวจข้อมูลย้อนหลัง (Retrospective cohort study) จากเวชระเบียนผู้ป่วยตั้งแต่ 1 มกราคม 2554 ถึง 31 ธันวาคม 2558 จำนวน 43 รายโดยรวบรวมข้อมูลจากเวชระเบียนผู้ป่วยเด็กที่ได้รับการวินิจฉัยว่าเป็นโรคโลหิตจางธาลัสซีเมียที่มีภาวะเหล็กเกิน และได้รับการรักษาด้วยการใช้ยาขับธาตุเหล็ก Deferrasirox จำนวน 15 ราย และ Derferiprone จำนวน 28 ราย ณ คลินิกโรคเลือดเด็ก โรงพยาบาลอุดรดิตถ์

ผลการวิจัย

ในขั้นตอนนี้เป็นการตรวจสอบข้อมูลเพื่อการได้มาซึ่งตัวแบบการพยากรณ์พยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone ในผู้ป่วยเด็กโรคโลหิตจางธาลัสซีเมีย ประกอบด้วยดังนี้

1. การตรวจสอบข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนในรูปแบบการแจกแจงของตัวแปรตาม

จากข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ เพื่อทำการตรวจสอบปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone อยู่ภายใต้รูปแบบการแจกแจงปกติหรือไม่ ทำการตรวจสอบแสดงดังนี้ ตารางที่ 2 แสดงค่าสถิติทดสอบของ Shapiro-Wilk [5] ในการตรวจสอบรูปแบบการแจกแจงความน่าจะเป็นของปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone

ข้อสมมุติเบื้องต้น (Assumption)	ยา Deferasirox		ยา Deferiprone	
	ค่าสถิติ	p-value	ค่าสถิติ	p-value
รูปแบบการแจกแจงปกติของปริมาณธาตุเหล็กในเลือดหลังการรักษา	0.739	0.001*	0.734	0.000*

* ณ ระดับนัยสำคัญ 0.05 การทดสอบไม่มีนัยสำคัญทางสถิติ

จากข้างต้นผลการวิเคราะห์พบว่า ณ ระดับนัยสำคัญ 0.05; รูปแบบการแจกแจงของปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone มีรูปแบบการแจกแจงไม่ปกติ คือเป็นลักษณะเบ้ขวา ดังนั้นจึงทำการแปลงข้อมูล(Transformation) ปริมาณธาตุเหล็กในเลือดหลังรักษาของยาทั้ง 2 ชนิดให้มีรูปแบบใหม่คือ

$$\hat{y}_1 = \log_{10} \text{ข้อมูลปริมาณธาตุเหล็กในเลือดหลังรักษา Deferasirox}$$

$$\hat{y}_2 = \log_{10} \text{ข้อมูลปริมาณธาตุเหล็กในเลือดหลังรักษา Deferiprone}$$

จากนั้นทำการตรวจสอบรูปแบบการแจกแจงผลการวิเคราะห์ แสดงดังนี้

ตารางที่ 3 แสดงค่าสถิติทดสอบของ Shapiro-Wilk ในการตรวจสอบรูปแบบการแจกแจงความน่าจะเป็นของปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone เมื่อการแปลงค่าตัวแปรตาม

ข้อสมมุติเบื้องต้น (Assumption)	ยา Deferasirox		ยา Deferiprone	
	ค่าสถิติ	p-value	ค่าสถิติ	p-value
การตรวจสอบรูปแบบการแจกแจงปกติของปริมาณธาตุเหล็กในเลือดหลังรักษา	0.923	0.217**	0.918	0.030*

* ณ ระดับนัยสำคัญ 0.01 การทดสอบไม่มีนัยสำคัญทางสถิติ **ณ ระดับนัยสำคัญ 0.05 การทดสอบไม่มีนัยสำคัญทางสถิติ

จากข้างต้นผลการวิเคราะห์พบว่า ณ ระดับนัยสำคัญ 0.05; รูปแบบการแจกแจงของปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ ณ ระดับนัยสำคัญ 0.01 รูปแบบการแจกแจงของปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferiprone มีรูปแบบการแจกแจงเป็นปกติเช่นกัน ดังนั้นรูปแบบการแจกแจงของตัวแปรตามเป็นไปตามข้อสมมุติเบื้องต้นและนำไปสู่การสร้างตัวแบบพยากรณ์ในหัวข้อต่อไป

2. การสร้างตัวแบบพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone

การวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุเพื่อใช้ในการพยากรณ์ หรือประมาณปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox ได้ผลการวิเคราะห์ดังนี้

2.1 ตัวแบบพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา โดยมีวิธีการคัดเลือกตัวแปรอิสระด้วยวิธี Stepwise Regression คือ

$$\log \hat{y} = 2.959 + 7.875E - 5x_4 - 0.005x_7$$

ณ ระดับนัยสำคัญ 0.05 ตัวแปรที่มีความสัมพันธ์กับปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox คือ ปริมาณธาตุเหล็กก่อนรักษา และ ALT คิดเป็นร้อยละ 88.90 (R = 0.889) และตัวแปรอิสระดังกล่าวสามารถอธิบายความผัน

แปรของปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox คิดเป็นร้อยละ 75.50 ($R^2_{adj} = 0.755$) และมีค่าประมาณความคลาดเคลื่อนมาตรฐานเท่ากับ 0.011 ($MSE = 0.011$)

2.2 ตัวแบบพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferiprone โดยมีวิธีการคัดเลือกตัวแปรอิสระด้วยวิธี Backward Elimination คือ

$$\hat{y} = 2.391 + 0.003x_1 - 0.092x_2 - 1.360x_3 + 1.142E^{-4}x_4 + 0.008x_6$$

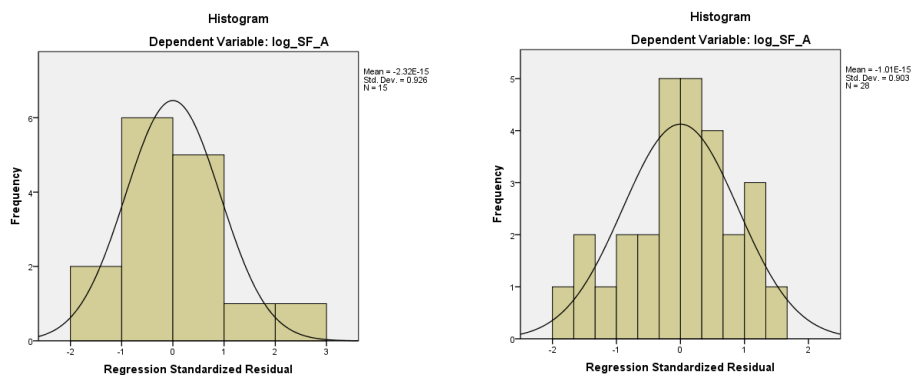
ณ ระดับนัยสำคัญ 0.10 ตัวแปรที่มีความสัมพันธ์กับปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferiprone ในผู้ป่วยเด็กโรคโลหิตจางธาลัสซีเมียคือ อายุ, การวินิจฉัยโรค, สถานะของม้าม, ปริมาณธาตุเหล็กก่อนรักษา และ AST คิดเป็นร้อยละ 77.30 ($R = 0.773$) และตัวแปรอิสระดังกล่าวสามารถอธิบายความผันแปรของปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferiprone คิดเป็นร้อยละ 50.60 และมีค่าประมาณความคลาดเคลื่อนมาตรฐานเท่ากับ 0.014

3. การตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับค่าความคลาดเคลื่อน

พิจารณาจากค่าความคลาดเคลื่อนว่าเป็นไปตามข้อสมมติเบื้องต้นหรือไม่ แสดงผลการทดสอบดังนี้

3.1 การตรวจสอบรูปแบบการแจกแจงของค่าความคลาดเคลื่อน

การตรวจสอบดังกล่าว เพื่อต้องการทราบถึงค่าความคลาดเคลื่อน (e_i) เป็นอิสระต่อกัน และมีการแจกแจงปกติด้วยค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนทุก ๆ ค่าของ x_i เท่ากับ σ^2 แสดงดังนี้



ภาพที่ 3 คะแนนมาตรฐานของความคลาดเคลื่อนของปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone

ตารางที่ 11 แสดงผลการตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความคลาดเคลื่อน ของตัวแบบพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone

ข้อสมมติเบื้องต้น (Assumption)	ยา Deferasirox		ยา Deferiprone	
	ค่าสถิติ	p-value	ค่าสถิติ	p-value
1.ค่าเฉลี่ยของค่าความคลาดเคลื่อน	0	0.500*	0	0.500*
2.ค่าความคลาดเคลื่อนอยู่ภายใต้รูปแบบการแจกแจงปกติ โดยใช้สถิติทดสอบของ Shapiro-Wilk	0.923	0.213*	0.965	0.444*
3.ความเป็นอิสระของค่าความคลาดเคลื่อนโดยใช้หลักสถิติของ Durbin - Watson	2.310	มากกว่า 1.5 เข้าใกล้ 2	2.569	มากกว่า 1.5 เข้าใกล้ 2

4.ความแปรปรวนของค่าความคลาดเคลื่อนคงที่โดยใช้สถิติทดสอบของ Levene ที่มีการปรับค่า				
• อายุ	-	-	1.008	0.325*
• การวินิจฉัยโรค	-	-	0.043	0.838*
• สถานะมะเร็ง	-	-	0.604	0.444*
• ปริมาณธาตุเหล็กก่อนรักษา	2.567	0.133	1.008	0.325*
• AST	-	-	1.008	0.325*
• ALT	2.567	0.133	-	-

* ณ ระดับนัยสำคัญ 0.05 การทดสอบไม่มีนัยสำคัญทางสถิติ

จากตารางข้างต้นข้อสมมติเบื้องต้นของค่าความคลาดเคลื่อนเป็นไปตามเงื่อนไขของการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ คือในการตรวจสอบค่าเฉลี่ยของค่าความคลาดเคลื่อน จากข้อมูลยา Deferasirox และ Deferiprone พบว่าค่า ณ ระดับนัยสำคัญ 0.05; ค่าเฉลี่ยของค่าความคลาดเคลื่อนของยาทั้ง 2 ชนิดเท่ากับ 0 ; $E(\varepsilon_i) = 0$ การตรวจสอบรูปแบบการแจกแจงค่าความคลาดเคลื่อน จากการใช้สถิติทดสอบของ Shapiro-Wilk พบว่า ณ ระดับนัยสำคัญเท่ากับ 0.05; ค่าความคลาดเคลื่อนอยู่ภายใต้รูปแบบการแจกแจงปกติ และค่าความคลาดเคลื่อนเป็นอิสระต่อกันเนื่องจากค่าสถิติ Durbin – Watson มีค่ามากกว่า 1.5 และเข้าใกล้ 2 และการตรวจสอบความแปรปรวนของค่าความคลาดเคลื่อนคงที่โดยใช้สถิติทดสอบของ Levene ที่มีการปรับค่าพบว่า ณ ระดับนัยสำคัญเท่ากับ 0.05; ค่าความแปรปรวนของค่าความคลาดเคลื่อนคงที่

3.2 การตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระของตัวแปรอิสระที่อยู่ในสมการถดถอย

เป็นการตรวจสอบว่าสมการถดถอยที่ได้เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอย แล้วจึงจะสามารถใช้สมการถดถอยที่ได้จากกลุ่มตัวอย่าง สำหรับการทำนายค่าของตัวแปรตามได้อย่างมีความน่าเชื่อถือ

ตารางที่ 9 แสดงผลการตรวจสอบความเป็นอิสระของตัวแปรอิสระที่อยู่ในสมการถดถอย สำหรับพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox

ตัวแปรอิสระในสมการถดถอย	VIF
X_4 : ปริมาณธาตุเหล็กก่อนรักษา	1.530
X_7 : ALT	1.138

จากตารางข้างต้นเป็นการตรวจสอบความเป็นอิสระของตัวแปรอิสระในสมการถดถอย พบว่าตัวแปรอิสระ X_4 มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระ X_7 น้อยมากนั้นทำให้ค่าสัมประสิทธิ์ตัวกำหนดระหว่างตัวแปรอิสระ มีค่าเข้าใกล้ 0 และทำให้ค่า VIF มีค่าเข้าใกล้ 1 ซึ่งยังไม่ถือว่าเกิดปัญหาตัวแปรอิสระมีความสัมพันธ์ต่อกัน

ตารางที่ 10 แสดงผลการตรวจสอบความเป็นอิสระของตัวแปรอิสระที่อยู่ในสมการถดถอย สำหรับพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferiprone

ตัวแปรอิสระในสมการถดถอย	VIF
X_1 : อายุ	1.756
X_2 : การวินิจฉัยโรค	1.100
X_3 : สถานะของมะเร็ง	1.908
X_4 : ปริมาณธาตุเหล็กก่อนรักษา	1.049
X_6 : AST	1.027

จากตารางข้างต้นเป็นการตรวจสอบความเป็นอิสระของตัวแปรอิสระในสมการถดถอย พบว่าตัวแปรอิสระในสมการถดถอยมีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระอื่นๆน้อยมาก ซึ่งยังไม่ถือว่าเกิดปัญหาตัวแปรอิสระมีความสัมพันธ์ต่อกัน

อภิปรายผล

จากการสร้างตัวแบบการถดถอยเชิงเส้นแบบพหุ เพื่อการพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone ในผู้ป่วยเด็กโรคโลหิตจางธาลัสซีเมีย ณ โรงพยาบาลอุตรดิตถ์ ซึ่งผลการศึกษาที่ได้มีความสอดคล้องกับการศึกษาของ Cappellini MD. และคณะ, Piga A. และคณะ [6,7] ทำการศึกษาในผู้ป่วยระยะที่ 2-3 พบว่าหนึ่งในตัวแปรที่มีอิทธิพลต่อปริมาณธาตุเหล็กในเลือด คือ ระดับเอนไซม์ AST และ ALT ที่เพิ่มขึ้น

จะเห็นได้ว่าตัวแบบการถดถอยเชิงเส้นตรงแบบพหุที่ได้ยังเป็นตัวแบบที่ไม่น่าเชื่อถือมากนัก เนื่องจากสัมประสิทธิ์ตัวกำหนดแบบพหุที่มีการปรับค่าแล้วมีค่าค่อนข้างต่ำ และมีค่า MSE ค่อนข้างสูงอาจมีสาเหตุเนื่องจากตัวแบบการถดถอยดังกล่าวขาดตัวแปรอิสระที่สำคัญบางตัว และกรณีที่ใช้ศึกษาค่อนข้างน้อย อันส่งผลให้ตัวแบบที่ได้รับยังขาดความแม่นยำในการพยากรณ์

สำหรับข้อเสนอแนะในการศึกษาครั้งต่อไป คือ ตัวแบบพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox ควรมีตัวแปรอิสระอื่นที่อาจมีความสัมพันธ์ คือ ภาวะผิดปกติของระบบทางเดินอาหาร, การมีผื่นที่ผิวหนัง, การวัดค่าการทำงานของไต ภาวะแทรกซ้อนทางตา และภาวะการได้ยิน และสำหรับยา Deferiprone ควรมีตัวแปรอิสระอื่นที่อาจมีความสัมพันธ์ คือ ภาวะผิดปกติของระบบทางเดินอาหาร และการปวดข้อ เพื่อทำให้ตัวแบบมีความน่าเชื่อถือมากยิ่งขึ้น

สรุปผลการวิจัย

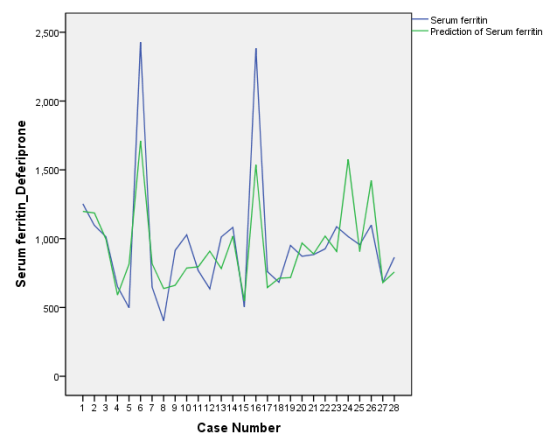
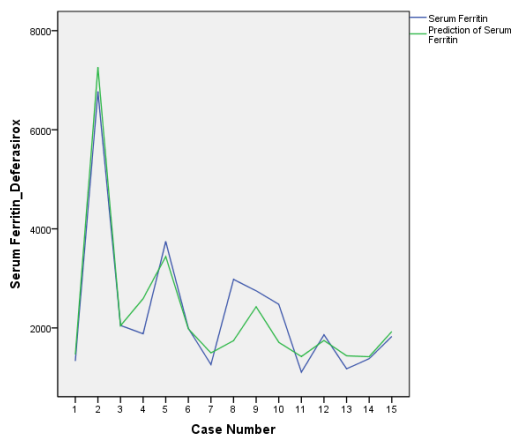
ตัวแบบการพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox มีปัจจัยที่มีความสัมพันธ์คือ ปริมาณธาตุเหล็กก่อนรักษา และ ALT โดยที่ $R^2_{adj} = 0.755$ และ $MSE = 0.011$ และสำหรับตัวแบบการพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferiprone มีปัจจัยที่มีความสัมพันธ์คือ อายุ, การวินิจฉัยโรค, สถานะของม้าม, ปริมาณธาตุเหล็กก่อนรักษา และ AST โดยที่ $R^2_{adj} = 0.506$ และ $MSE = 0.014$ ซึ่งทั้ง 2 สมการถดถอยเป็นไปตามเงื่อนไขข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ

กิตติกรรมประกาศ

งานวิจัยครั้งนี้สำเร็จลงได้ ด้วยความอนุเคราะห์จากทุกฝ่ายที่เกี่ยวข้อง ผู้วิจัยขอขอบพระคุณมหาวิทยาลัยราชภัฏอุตรดิตถ์ที่ได้อนุมัติทุนอุดหนุนการวิจัย และ ผศ.ดร. มานะชัย รอดชื่น, ผศ.ดร. เกตุจันทร์ จำปาไชยศรี ที่ได้กรุณาให้คำปรึกษาและข้อเสนอแนะต่างๆ ที่ เป็นประโยชน์อย่างยิ่งที่ทำงานวิจัยนี้มีความสมบูรณ์มากยิ่งขึ้น

เอกสารอ้างอิง

- [1] อนุวัฒน์ สุตตันทวิบูลย์. เรื่องน่ารู้เกี่ยวกับโรคธาลัสซีเมีย. ภาควิชาสูติศาสตร์ – นรีเวชวิทยา คณะแพทยศาสตร์ศิริราชพยาบาล
- [2] พัทธนา จงอัจฉริยกุล. ดารินทร์ ขอโสทธิกุล.การรักษาภาวะเหล็กเกินด้วยยา Deferrasirox เปรียบเทียบกับ Deferiprone, วารสารกุมารเวชศาสตร์ 2557; 53:32-42
- [3] เกตุจันทร์ จำปาไชยศรี. 2549. การวิเคราะห์การถดถอย. พิมพ์ครั้งที่ 1. สำนักหอสมุด มหาวิทยาลัยนเรศวร.
- [4] Garson G.D. 2012. Testing Statistical Assumptions. Statistical Associates Publishing, .20, 36.
- [5] Nomadih MR, Yap BW. 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and [4] Anderson-Darling tests. Journal of Statistical Modeling and Analytics 2: 31-33.
- [6] Cappellini MD, Cohen A, Piga A, et al. 2006. A phase 3 study of deferiasirox (ICL670), a once-daily oral iron chelator, in patients with beta-thalassemia. Blood 107:3455-3462.
- [7] Piga A, Galanello R, Forni GL, et al. 2006. Randomized phase II trial of deferiasirox (Exjade, ICL670), a once-daily, orally-administered iron chelator, in comparison to deferoxamine in thalassemia patients with transfusional iron overload. Haematologica 91:873-880.



ภาพที่ 4 เปรียบเทียบค่าจริงกับค่าพยากรณ์ปริมาณธาตุเหล็กในเลือดหลังรักษาด้วยยา Deferasirox และ Deferiprone ตามลำดับ