

การพยากรณ์ความน่าจะเป็นระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิด้วย ตัวแบบลอจิตสะสม และตัวแบบลอจิตสองกลุ่ม

Probability Forecasting of the Primary Biliary Cirrhosis Disease Stages Using Cumulative Logit and Dichotomous Logit Models

ปทุมวดี กันเนตร สารินี ชุ่มสิงห์ และ วีรานันท์ พงศาภักดิ์*

Pathumwadee Kunnat, Sarinee Chumsing, and Veeranun Pongsapukdee*

ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร วิทยาเขตพระราชวังสนามจันทร์

บทคัดย่อ

การพยากรณ์ความน่าจะเป็นของระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ อาศัยวิธีการวิเคราะห์ข้อมูลด้วยตัวแบบลอจิตสะสม 4 กลุ่ม เมื่อตัวแปรตอบสนองเป็นแบบจำแนกประเภท 4 กลุ่ม คือ กลุ่มที่ 1: โรคตับแข็งจากท่อน้ำดีปฐมภูมิระยะที่ 1 (porital stage) กลุ่มที่ 2: โรคตับแข็งจากท่อน้ำดีปฐมภูมิระยะที่ 2 (periportal stage) กลุ่มที่ 3: โรคตับแข็งจากท่อน้ำดีปฐมภูมิระยะที่ 3 (septal stage) กลุ่มที่ 4: โรคตับแข็งจากท่อน้ำดีปฐมภูมิระยะที่ 4 (biliary stage) และเพื่อให้เห็นผลลัพธ์ในภาพรวมระหว่างการเป็นโรคในระยะเวลาที่ไม่รุนแรงและในระยะที่รุนแรง เราจึงวิเคราะห์ด้วยตัวแบบลอจิต 2 กลุ่ม โดยการยุบกลุ่มของตัวแปรตอบสนองที่ลำดับอยู่ติดกันให้เป็น 2 กลุ่ม คือ กลุ่มที่ 1: โรคตับแข็งจากท่อน้ำดีปฐมภูมิระยะที่ 1-2 (porital stage-periportal stage) ซึ่งยังไม่รุนแรง และกลุ่มที่ 2: โรคตับแข็งจากท่อน้ำดีปฐมภูมิระยะที่ 3-4 (septal stage-biliary stage) ซึ่งเป็นระยะรุนแรง ส่วนตัวแปรอธิบายที่ใช้มีทั้งหมด 11 ตัว ได้แก่ เพศ (sex), การตรวจพบภาวะท้องมาน (ascites), การตรวจพบภาวะตับโต (hepatomegaly), การตรวจพบไฟแดง คล้ายแมงมุม (spiders), การตรวจพบอาการบวมที่ขา (edema), ค่าซีรั่มบิลิรูบิน (bili), ค่าซีรั่มคอเลสเตอรอล (cholesterol), ค่าอัลบูมิน (albumin), ค่าปริมาณสารทองแดงในปัสสาวะ (urine copper), ค่า SGOT, จำนวนเกล็ดเลือด (platelets) เป็นข้อมูลจริงของผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิจำนวน 278 คน ในฐานข้อมูลจริงด้านสาธารณสุขเว็บไซต์ <http://lib.stat.cmu.edu> และประมวลผลด้วยโปรแกรม SAS version 9.1 ผลการวิจัยพบว่า ตัวแบบลอจิตสะสม [1] 4 กลุ่มมีความเหมาะสมอย่างมีนัยสำคัญทางสถิติที่ $P < 0.001$ และพบว่า ตัวแปรอธิบายที่ส่งผลต่อระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ ได้แก่ การตรวจพบภาวะท้องมาน (ascites), การตรวจพบภาวะตับโต (hepatomegaly), การตรวจพบไฟแดงที่มี

ลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) การพยากรณ์ความน่าจะเป็น สามารถใช้ตัวแบบความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มโรคตับแข็งจากท่อน้ำดีปฐุมภูมิใน ระยะที่ 1 (Porital stage)

$$\text{คือ } P(Y=1) = P(Y \leq 1) = \frac{e^{-5.4755+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}{1+e^{-5.4755+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}} \text{ ความน่าจะเป็นที่}$$

ผู้ป่วยจะอยู่ในกลุ่มโรคตับแข็งจากท่อน้ำดีปฐุมภูมิในระยะที่ 2 (Periportal stage) คือ $P(Y=2)$

$$= P(Y \leq 2) - P(Y \leq 1)$$

$$= \frac{e^{-3.1847+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}{1+e^{-3.1847+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}$$

$$- \frac{e^{-5.4755+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}{1+e^{-5.4755+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}$$

ความน่าจะเป็นที่จะอยู่ในกลุ่มโรคตับแข็งจากท่อน้ำดีปฐุมภูมิในระยะที่ 3 (Septal stage) คือ $P(Y=3) =$

$$P(Y \leq 3) - P(Y \leq 2)$$

$$= \frac{e^{-0.7774+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}{1+e^{-0.7774+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}$$

$$- \frac{e^{-3.1847+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}{1+e^{-3.1847+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}$$

ความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มโรคตับแข็งจากท่อน้ำดีปฐุมภูมิในระยะที่ 4 (Biliary stage) คือ

$$P(Y=4) = 1 - P(Y \leq 3) = 1 - \frac{e^{-0.7774+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}{1+e^{-0.7774+1.8276\text{ascites}+0.8424\text{hepatom}+0.39961\text{spiders}+0.3617\text{bili}}}$$

สำหรับตัวแบบลอจิสต์ 2 กลุ่ม พบว่า มีความเหมาะสม อย่างมีนัยสำคัญทางสถิติ ที่ $P < 0.02$ และตัวแปรอธิบายที่ส่งผลต่อระยะของโรคตับแข็งจากท่อน้ำดีปฐุมภูมิ คือ การตรวจพบ ภาวะตับโต (hepatom) , การตรวจพบไฟแดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) การพยากรณ์ความน่าจะเป็น สามารถใช้ตัวแบบความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐุม

$$\text{ภูมิในระยะรุนแรง คือ } P(x) = \frac{e^{1.4486-0.7024\text{hepatom}-0.5224\text{spiders}-0.4464\text{bili}}}{1+e^{1.4486-0.7024\text{hepatom}-0.5224\text{spiders}-0.4464\text{bili}}}$$

ดังนั้น ความน่าจะเป็นของระยะโรคตับแข็งจากท่อน้ำดีปฐุมภูมิที่ยังไม่รุนแรง เท่ากับ $1 - P(x)$

คำสำคัญ: Categorical Variables, cumulative logit models for ordinal responses, Dichotomous logit models, Primary Biliary Cirrhosis (PBC) disease stages.

Abstract

The main objective of this research is to forecast probabilities of Primary Biliary Cirrhosis (PBC) disease stages. The analytical models are the cumulative logit models for ordinal responses under the tetrachotomous response categories: PBC disease patients stage 1 (porital stage), stage 2 (periportal stage), stage 3 (septal stage) and stage 4 (biliary stage). Furthermore, we also perform another model using the logit model for dichotomous response categories in the case from collapsing the adjacent stage of PBC disease: group 1: patients in stage 1 – 2, and group 2: patients in stage 3 - 4. All 11 explanatory variables consist of sex, ascites, hepatomegaly (hepatom), spiders, edema, serum bilirubin (bili), serum cholesterol, serum albumin, urine copper, SGOT and platelet. A random sample of 278 patients is from the website <http://lib.stat.cmu.edu> which are processed by the statistical package SAS version 9.1.

The research results reveal that the statistical significant factors effecting to stage of PBC disease are ascites, hepatomegaly (hepatom), spiders, edema, and serum bilirubin at $\alpha = 0.01$ ($P < 0.001$). The forecasting probabilities can be obtained from each model. When the patient of PBC disease in stage 1 (porital stage) it's probability is $P(Y=1) =$

$$P(Y \leq 1) = \frac{e^{-5.4755+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}}{1 + e^{-5.4755+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}} .$$

The probability when the patient of PBC disease in stage 2 (Periportal stage) is $P(Y=2) = P(Y \leq 2) - P(Y \leq 1)$

$$= \frac{e^{-3.1847+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}}{1 + e^{-3.1847+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}}$$

$$- \frac{e^{-5.4755+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}}{1 + e^{-5.4755+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}} .$$

The probability when the patient of PBC disease in stage 3 (Septal stage) is $P(Y=3) = P(Y \leq 3) - P(Y \leq 2)$

$$= \frac{e^{-0.7774+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}}{1+e^{-0.7774+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}}$$

$$= \frac{e^{-3.1847+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}}{1+e^{-3.1847+1.8276ascites+0.8424hepatom+0.39961spiders+0.3617bili}} \cdot$$

The probability when the patient of PBC disease in stage 4 (Biliary stage) is $P(Y=4) = 1 - P(Y \leq 3)$

The estimated dichotomous logit model indicates the factors that effecting to stage of PBC disease are hepatomegaly (hepatom), spiders, edema, and serum bilirubin at $\alpha = 0.05$ ($P < 0.02$). The forecasting probabilities can be obtained from the model when the patients of PBC

disease in the severe stage is $P(x) = \frac{e^{1.4486-0.7024hepatom-0.5224spiders-0.4464bili}}{1+e^{1.4486-0.7024hepatom-0.5224spiders-0.4464bili}}$. Therefore, the probability when the patient of PBC disease is in the stage of not severe is $1 - P(x)$.

Keywords : Categorical variables, cumulative logit models for ordinal responses, Dichotomous logit models, Primary Biliary Cirrhosis (PBC) disease stages.

1. บทนำ

โรคตับแข็งจากท่อน้ำดีปฐมภูมิ (Primary Biliary Cirrhosis : PBC) คือโรคตับเรื้อรังที่ทำลายท่อน้ำดีภายในตับ อย่างช้า ๆ ทำให้ท่อน้ำดีของตับเกิดการอักเสบเรื้อรัง (ท่อน้ำดี intrahepatic) การอักเสบของท่อน้ำดีของตับเป็นระยะเวลานาน อาจจะเป็นสาเหตุทำให้เกิดแผลเป็น ซึ่งเป็น สาเหตุหนึ่งของโรคตับแข็งที่พบมากในประเทศแถบยุโรปและสหรัฐอเมริกา จากการศึกษาพบว่าโรคตับแข็งจากท่อน้ำดีปฐมภูมิเกิดขึ้นได้กับกลุ่มคนทุกเพศทุกวัย โดยส่วนใหญ่พบในกลุ่มผู้ป่วยที่มี อายุ ระหว่าง 40 – 60 ปี ผู้ป่วยที่พบส่วนใหญ่เป็นเพศหญิงร้อยละ 90 [2]

ปัจจุบันไม่ทราบสาเหตุของการเกิดโรคตับแข็งจากท่อน้ำดีปฐมภูมิอย่างชัดเจน แต่เป็นที่ทราบกันว่าไม่ได้เกิดจากแอลกอฮอล์หรือยา การศึกษาในปัจจุบันแนะนำว่า อาจเกี่ยวข้องกับ ระบบภูมิคุ้มกัน การติดเชื้อหรือข้อบกพร่องทางพันธุกรรม ที่ถูกส่งผ่านจากพ่อแม่สู่ลูก

ผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิมีการแสดงอาการอย่างช้า ๆ โดยในระยะแรกจะไม่แสดงอาการ อาการเฉพาะที่เกิดขึ้นส่วนใหญ่ ร้อยละ 70- 80 คือ ความเมื่อยล้า แต่แพทย์ยังไม่พบการเจ็บป่วยใด ๆ ร้อยละ 50-60 คือ อาการคัน ภาวะแทรกซ้อนของโรคตับแข็งจากท่อน้ำดีปฐมภูมิที่มี โอกาสเกิดขึ้นสูง คือ น้ำในช่องท้อง มีเลือดออกทางเดินอาหาร อาการตัวเหลือง ผิวน้ำหรือตามีสีเหลือง [3-4]

ระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ สามารถแบ่งออกได้เป็น 4 ระยะ คือ ระยะที่ 1 Portal Stage, ระยะที่ 2 Periportal Stage, ระยะที่ 3 Septal Stage และระยะที่ 4 Biliary Stage ในแต่ละระยะของโรค จะแสดงถึงความก้าวหน้าของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ และระดับความเสียหายของท่อน้ำดีภายในเซลล์ตับ โดยโรคตับแข็งและภาวะแทรกซ้อนจะเกิดขึ้นในระยะสุดท้าย

เนื่องจากผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิ ไม่มีการค้นพบอาการที่ผ่านผลการตรวจเลือดผิดปกติในตับตามปกติ แพทย์จำเป็นต้องทำแบบทดสอบหลายอย่างเพื่อยืนยันการวินิจฉัยโรคตับแข็งจากท่อน้ำดีปฐมภูมิ โดยแพทย์มักจะตระหนักถึงปัญหาในระหว่างการทดสอบผลเลือด หรือ การประเมินผลตรวจสำหรับเงื่อนไขอื่น หากแพทย์สงสัยโรคตับแข็งจากท่อน้ำดีปฐมภูมิ การทดสอบ หลาย ๆ อย่างสามารถช่วยในการวินิจฉัย เช่น การทำงานของตับ ภาพอัลตราซาวด์ Anti - mitochondrial แอนติบอดี (AMAs) และการตรวจชิ้นเนื้อตับ นอกจากนี้ปัจจัยทางพันธุกรรมก็มีส่วนร่วมในการวินิจฉัยโรค [5] แต่ในบางครั้งอาจทำให้เป็นเรื่องยากสำหรับแพทย์ที่จะวินิจฉัยโรคตับแข็งจากท่อน้ำดีปฐมภูมิ ได้ ตามความเป็นจริง เพราะผู้ป่วยส่วนใหญ่ยังคงไม่มี อาการเป็นเวลาหลายปี และบางส่วนไม่อาจสังเกตเห็นอาการใด ๆ จากอาการเริ่มต้นแตกต่างกันไปในผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิ [4]

ในการวินิจฉัยโรคของผู้ป่วย ถ้าแพทย์สามารถทราบปัจจัยที่มีส่วนสำคัญหรือเกี่ยวข้องกับระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ จะทำให้แพทย์สามารถวินิจฉัยและเลือกวิธีการรักษาได้ ทันทีทั้งที่และตรงกับระยะอาการของโรค และเพื่อให้เกิดผลดีต่อผู้ป่วยมากที่สุด หลักเล็งการได้รับยาที่ไม่จำเป็น และป้องกันการเกิดภาวะแทรกซ้อนที่อาจเกิดจากการได้รับยา โดยอาศัยปัจจัยหรือประวัติทางระบาดวิทยา และผลตรวจทางห้องปฏิบัติการเพื่อยืนยันเช่น ผลตรวจซีรั่ม รวมถึงอาศัยตัวแบบในการช่วยวินิจฉัยระยะของโรค เช่นนี้ในงานวิจัยจึงสนใจศึกษาปัจจัยต่าง ๆ ที่มีผลต่อระยะของผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิ

ปัจจุบันนักวิจัยให้ความสนใจวิธีวิเคราะห์ข้อมูลและการพยากรณ์ที่ตัวแปรตอบสนองเป็นจำแนกประเภทเพิ่มขึ้น เช่น ตัวแปรตอบสนองจำแนกประเภทแบบไม่มีลำดับ (เป็นโรค, ไม่เป็นโรค) และตัวแปรตอบสนองแบบอันดับ (ระดับของโรครุนแรง ปานกลาง น้อย ปกติ) เช่น ตัวแบบต่างๆสำหรับการวิเคราะห์ข้อมูลที่มีตัวแปรตอบสนองจำแนกประเภทแบบมีลำดับมีอยู่หลายตัวแบบ ได้แก่ cumulative logit model เช่น proportional odds model [6-7], continuation ratio model [8], constrained and unconstrained partial proportional odds model [9] และ adjacent – category logistic model [10] โดยงานวิจัยนี้สนใจที่จะประยุกต์ใช้ตัวแบบ proportional odds

ในงานวิจัยนี้ผู้วิจัยทำการศึกษาปัจจัยที่มีผลต่อระยะอาการของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ [11] และศึกษาวิธีการพยากรณ์ความน่าจะเป็นของแต่ละกลุ่มในตัวแปรตอบสนอง รวมทั้งเพื่อเปรียบเทียบความน่าจะเป็นต่าง ๆ โดยใช้ตัวสถิติ Odds ratio ซึ่งในการวิเคราะห์ข้อมูล ผู้วิจัยทำการแบ่งตัวแปรตอบสนองออกเป็น 2 กรณี คือ กรณีที่ตัวแปรตอบสนองเป็นแบบจำแนกประเภทแบบมีลำดับ 4 กลุ่มและกรณีขุบกลุ่มตัวแปรตอบสนองเป็น 2 กลุ่ม ในการวิเคราะห์ข้อมูล ผู้วิจัยจะใช้ตัวแบบลอจิสติกสะสม (cumulative logit model) เมื่อตัวแปรตอบสนองเป็นแบบจำแนกประเภทแบบมีลำดับ 4 กลุ่ม และใช้ตัวแบบลอจิสติก 2 กลุ่ม (logit model) เมื่อมีการขุบกลุ่มของตัวแปรตอบสนองแบบมีลำดับ 4 กลุ่ม เป็น 2 กลุ่ม ซึ่งตัวแบบที่ได้จากทั้ง 2 กรณีจะทำการตรวจสอบภาวะสารูปดี (Goodness-of-fit test) และทำการตรวจสอบค่าสถิติ Likelihood Ratio, Score และ Wald นอกจากนี้ ผู้วิจัยทำการหาตัวแบบที่เหมาะสมมากที่สุดกับข้อมูลชุดนี้ โดยการเปรียบเทียบค่าของตัวสถิติอัตราส่วนภาวะน่าจะเป็น หรือ $-2 \log L$, SC (Schwarz Criterion) และการวิเคราะห์ส่วนเหลือ โดยข้อมูลที่ใช้ในการวิจัย เป็นข้อมูลจริงจากเว็บไซต์ <http://lib.stat.cmu.edu/> เรื่อง Primary Biliary Cirrhosis ซึ่งเก็บข้อมูลหุติยภูมิจากผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิ ที่เข้ารับการรักษาที่ Mayo Clinic จำนวน 278 คน

1.1 วัตถุประสงค์ของการวิจัย

1. เพื่อใช้ตัวแบบลอจิสติกสะสมมีลำดับ 4 กลุ่มในการวิเคราะห์ปัจจัยที่มีผลต่อระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ
2. เพื่อใช้ตัวแบบลอจิสติกสองกลุ่มในการวิเคราะห์ปัจจัยที่มีผลต่อระยะ โรคตับแข็งจากท่อน้ำดีปฐมภูมิ
3. เพื่อพยากรณ์ความน่าจะเป็นแต่ละกลุ่มในตัวแปรตอบสนอง ด้วยตัวแบบลอจิสติกสะสมและตัวแบบลอจิสติก 2 กลุ่ม
4. เพื่อเปรียบเทียบความน่าจะเป็นต่าง ๆ โดยใช้ตัวสถิติ Odds ratio

1.2 ขอบเขตของการวิจัย

1. ข้อมูลที่ใช้ในการทำวิจัยในครั้งนี้เป็นข้อมูล Primary Biliary Cirrhosis ซึ่งเก็บข้อมูลตัวอย่างมาจากผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิ ที่เข้ารับการรักษาที่ Mayo Clinic ในช่วงปี 1974 – 1984 จำนวน 278 คน
2. ตัวแปรตอบสนอง ประกอบด้วย 2 กรณี คือ
 - กรณีตัวแปรตอบสนองจำแนกประเภท 4 กลุ่ม จำแนกระยะโรคตับแข็งจากท่อน้ำดีปฐมภูมิ 4 ระดับ
 - ผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิระยะที่ 1 (Portal stage) = 1
 - ผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิระยะที่ 2 (Periportal stage) = 2

- ผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐุมภูมิมะเยที่ 3 (Septal stage) = 3
- ผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐุมภูมิมะเยที่ 4 (Biliary stage) = 4
- กรณีตัวแปรตอบสนองจำแนกประเภท 2 กลุ่ม จำแนกมะเย โรคตับแข็งจากท่อน้ำดีปฐุมภูมิ 2 มะเย
- ผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐุมภูมิมะเยที่ 1 และมะเยที่ 2 = 0
- ผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐุมภูมิมะเยที่ 3 และมะเยที่ 4 = 1

3. ตัวแปรอธิบาย มีลักษณะจำแนกประเภท เป็นปัจจัยหรือประวัติทางระบาดวิทยา และผลตรวจทางห้องปฏิบัติการที่แพทย์ระบุให้ตรวจเพื่อยืนยันโรคตับแข็งจากท่อน้ำดีปฐุมภูมิ ทั้งหมด 11 ตัว ได้แก่

- เพศ (sex): 0 แทน เพศหญิง, 1 แทน เพศชาย
 - การตรวจพบภาวะท้องมาน (ascites) 0 แทน no, 1 แทน yes
 - การตรวจพบภาวะตับโต (hepatomegaly) 0 แทน no, 1 แทน yes
 - การตรวจพบไฟแดงลักษณะคล้ายแมงมุม (spiders) 0 แทน no, 1 แทน yes
 - การตรวจพบอาการบวมหน้า (edema) 0 แทน ไม่มีอาการบวมหน้าและไม่มีการบำบัดโรคสำหรับ-อาการบวมหน้า, 0.5 แทน การมีอาการบวมหน้าปรากฏ ปราศจากยาขับปัสสาวะ, 1 แทน การมีอาการบวมหน้า แม้จะมีการบำบัดโรคด้วยยาขับปัสสาวะ
 - ค่าซีรั่มบิลิรูบิน (bili) (มิลลิกรัม/เดซิลิตร) 0 แทน 0.3-1.0 มิลลิกรัม/เดซิลิตร (ค่าปกติ), 1 แทน อื่นๆ
 - ค่าซีรั่มคอเลสเตอรอล (cholesterol) (มิลลิกรัม/เดซิลิตร) 0 แทน 150 – 200 มิลลิกรัม/เดซิลิตร (ค่าปกติ), 1 แทน อื่น ๆ
 - ค่าอัลบูมิน (albumin) (กรัม/เดซิลิตร) 0 แทน 3.8 - 5.0 กรัม/เดซิลิตร(ค่าปกติ), 1 แทน อื่นๆ
 - ค่าปริมาณสารทองแดงในปัสสาวะ (urine copper) (ไมโครกรัม/วัน) 0 แทน 30-50 ไมโครกรัม/วัน (ค่าปกติ), 1 แทน อื่น ๆ
 - ค่า SGOT (U/ml) 0 แทน น้อยกว่าหรือเท่ากับ 38.4 U/ml (ค่าปกติ), 1 แทน 38.5 -171.5 U/ml, 2 แทน 171.6 - 346.5 U/ml, 3 แทน มากกว่า 346.5 U/ml
 - จำนวนเกล็ดเลือด (platelets) (มิลลิลิตร/1000) 0 แทน 14,000 – 450,000 มิลลิลิตร/1000 (ค่าปกติ) และ 1 แทน อื่นๆ
4. การเลือกตัวแปรอธิบายเข้าในแบบเพื่อศึกษาอิทธิพลของตัวแปรอธิบายที่มีผลต่อตัวแปรตอบสนอง ในกรณีที่ข้อมูลเป็นจำแนกประเภท จะใช้วิธีการคัดเลือกตัวแปรแบบขั้นตอน (Stepwise) ส่วนการประมวลผลโดยใช้โปรแกรมทางสถิติ SAS (Statistic Analysis System) version 9.1

1.3 ศัพท์เฉพาะในการวิจัย

1. โรคตับแข็งจากท่อน้ำดีปฐมภูมิ หมายถึง โรคตับเรื้อรังที่จะทำให้ท่อน้ำดีภายในตับอย่างช้า ๆ (ท่อน้ำดี intrahepatic)
2. ภาวะท้องมาน (ascites) หมายถึง ผู้ป่วยจะมีท้องโตเนื่องจากมีน้ำคั่งภายในช่องท้องมักมีขาบวมกดบวม 2 ข้างเกิดจากตับสร้างโปรตีนที่เรียกว่าอัลบูมินในกระแสเลือดลดลง
3. ภาวะตับโต (hepatomegaly) หมายถึง อาการที่ตับมีขนาดบวมโตกว่าปกติจนสามารถคลำได้ (ตับขนาดปกติจะซ่อนอยู่บริเวณชายโครง คลำไม่พบ แต่ถ้าโตขึ้นจะเล็ดลงมาอยู่ใต้ชายโครงจนคลำพบได้)
4. ไฟแดงคล้ายแมงมุม (spiders) หมายถึง พบไฟแดงมีลักษณะคล้ายแมงมุม เป็นการเปลี่ยนแปลงของเส้นเลือดของผิวหนังที่พบได้บ่อยมากในผู้ป่วยตับเรื้อรัง
5. อาการบวมน้ำ (edema) หมายถึง ภาวะที่มีของเหลวปริมาณมากกว่าปกติคั่งค้างสะสมอยู่ในระหว่างเซลล์ เนื้อเยื่อและส่วนต่างๆของร่างกาย
6. ระดับ bilirubin หมายถึง เป็นผลผลิตจากการสลายของ hemoglobin ซึ่งจะถูกขับถ่ายออกทางน้ำดี ค่าปกติ คือ 0.3-1.0 mg/dl
7. ระดับ cholesterol หมายถึง ระดับของ cholesterol เป็นสิ่งสะท้อนถึง การเผาผลาญของไขมันส่วนหนึ่ง ซึ่งมีผลมาจาก พันธุกรรม, อาหาร, ตับ, ไต, ฮอร์โมน และการทำหน้าที่ของอวัยวะหลายๆส่วน ค่าปกติ คือ 150-200 mg/dl
8. ระดับ albumin หมายถึง ระดับอัลบูมินในเลือดเป็นตัวชี้ถึงความรุนแรงของโรคตับเรื้อรัง ค่าปกติคือ 3.8-5.0 g/dl
9. ระดับ urine copper หมายถึง ปริมาณสารทองแดงในปัสสาวะ ช่วงปกติ คือ 30-50 ไมโครกรัมต่อวัน ช่วงค่าปกติอาจแตกต่างกันเล็กน้อยในห้องปฏิบัติการที่แตกต่างกัน
10. ค่า SGOT หมายถึง ค่าผลตรวจทางห้องปฏิบัติการ SGOT ผลตรวจเหล่านี้บ่งบอกสภาพการทำงานของตับ ค่าปกติของ SGOT คือ 1-38.4 U/ml
11. platelets หมายถึง จำนวนเกล็ดเลือด ค่าปกติ คือ 140,000-450,000 ml/1000
12. ตัวแบบ (model) หมายถึง ตัวแบบเชิงสถิติ (Statistical model) เช่น ตัวแบบลอจิสติกสะสม (Cumulative logit mode [1])
13. ตัวแบบลอจิสติกสะสม ในงานวิจัยนี้ ใช้ตัวแบบลอจิสติกสะสมของกรณี Proportional odds model หมายถึง ตัวแบบเชิงเส้นทั่วไป (Generalized Linear Models) ชนิดหนึ่งใช้สำหรับวิเคราะห์ข้อมูลจำแนกประเภทแบบมีลำดับ [7] ซึ่งมีรูปแบบ คือ

$$\log \left[\frac{P(Y \leq j | X_i)}{P(Y > j | X_i)} \right] = \beta_{0j} + \beta' X_i \quad i = 1, \dots, p \quad j = 1, \dots, k-1.$$

โดยที่ $X_i = (X_1, X_2, \dots, X_p)$ แทนเวกเตอร์ขนาด p ของค่าสังเกตของตัวแปรอธิบาย $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ แทนเวกเตอร์ของพารามิเตอร์ขนาด p จุดตัด $\beta_{01} \leq \beta_{02} \leq \dots \leq \beta_{0(k-1)}$ คือ log odds ของ Y ที่น้อยกว่าหรือเท่ากับ j เมื่อ X มีค่าเป็นศูนย์ ซึ่งข้อสมมติของตัวแบบคือ พารามิเตอร์ของตัวแบบ (log odds ratio) เท่ากัน ($\beta_k = \beta$) ในทุกจุดแบ่ง j

14. อัตราส่วน Odds (Odds ratio) คืออัตราส่วนความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจกับความน่าจะเป็นที่จะไม่เกิดเหตุการณ์ที่สนใจ มีค่าเป็นบวกใดโดยความน่าจะเป็นในทุกเซลล์เป็นบวก ถ้าตัวแปรอิสระต่อกันจะมีค่าเท่ากับ 1

15. ตัวแปรตอบสนองจำแนกประเภทแบบมีลำดับ (Ordinal response variable) หมายถึง ตัวแปรตอบสนองที่มีหลายระดับและมีการจัดเรียงลำดับของตัวแปรตอบสนองนั้นๆ อาจเรียงลำดับจากน้อยไปมาก เช่น ในงานวิจัยนี้ คือ ระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ(มีการตรวจชิ้นเนื้อตับ) โดยกำหนดให้เป็น ระยะที่ 1 (Portal Stage), ระยะที่ 2 (Periportal Stage), ระยะที่ 3 (Septal Stage) และระยะที่ 4 (Biliary Cirrhosis) ตามลำดับอาการของโรคที่ไม่รุนแรงจนถึงขั้นรุนแรงมาก

16. ตัวแปรอธิบาย (Explanatory variables) หมายถึง ตัวแปรที่คาดว่าจะมีอิทธิพลต่อตัวแปรตอบสนอง ในงานวิจัยนี้ คือ เพศ (sex) มี 2 กลุ่ม, การตรวจพบภาวะท้องมาน (ascites) มี 2 กลุ่ม, การตรวจพบภาวะตับโต (hepatomegaly) มี 2 กลุ่ม, การตรวจพบไฟแดงคล้ายแมงมุม (spiders) มี 2 กลุ่ม, การตรวจพบอาการบวมน้ำ (edema) มี 3 กลุ่ม, ระดับ Bilirubin (bili)

(mg/dl) มี 2 กลุ่ม, ระดับ cholesterol (mg/dl) มี 2 กลุ่ม, ระดับ albumin (gm/dl) มี 2 กลุ่ม, ระดับ urine copper (ug/day) มี 3 กลุ่ม, ค่า SGOT (U/ml) มี 4 กลุ่ม, platelets (ml /1000) มี 3 กลุ่ม

ข้อมูลของตัวแปรอธิบายเป็นแบบจำแนกประเภท (Categorical explanatory variable) หมายถึง ข้อมูลที่จำแนกตามระดับหรือกลุ่มของตัวแปรนั้นๆ เช่น กลุ่ม 1, 2, ..., k เมื่อมี k กลุ่ม

17. การแจกแจงของตัวแปรจำแนกประเภท หมายถึง การแจกแจงของตัวแปรสุ่มแบบไม่ต่อเนื่อง เช่น การแจกแจงเบอร์นูลลี ทวินาม และพหุนาม

18. ตัวแปรจำแนกประเภท (Categorical variable) หมายถึง ตัวแปรที่มีลักษณะจำแนกประเภท โดยอาจจัดหน่วยทดลองเป็นประเภทที่สนใจ เช่น ไม่รุนแรง ปานกลาง รุนแรงมาก

2. วิธีการทดลอง

2.1 การวิเคราะห์ข้อมูลของตัวแปรตอบสนองภายใต้ตัวแบบลอจิสติก 4 กลุ่ม

เป็นการสร้างตัวแบบมาศึกษาปัจจัยที่มีผลต่อระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ และพยากรณ์ความน่าจะเป็นของแต่ละกลุ่มในตัวแปรตอบสนอง 4 กลุ่ม

2.2 การวิเคราะห์ข้อมูลของตัวแปรตอบสนองภายใต้ตัวแบบลอจิสติก 2 กลุ่ม

เป็นการสร้างตัวแบบมาศึกษาปัจจัยที่มีผลต่อระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ และพยากรณ์ความน่าจะเป็นของแต่ละกลุ่มในตัวแปรตอบสนอง 2 กลุ่ม

2.3 การเลือกตัวแบบที่เหมาะสมกับข้อมูล

การหาตัวแบบที่เหมาะสมกับข้อมูลมากที่สุดในแต่ละกรณีของตัวแปรตอบสนองจำแนกประเภท 4 กลุ่ม หรือ 2 กลุ่ม คือในแต่ละกรณีนั้นทำการเลือกตัวแบบที่เหมาะสม จากตัวแบบต่าง ๆ โดยเปรียบเทียบค่าของตัวสถิติอัตราส่วนภาวะน่าจะเป็นหรือ $-2 \log L$, ตัวสถิติ SC (Schwarz Criterion), ตัวสถิติ AIC และการวิเคราะห์ส่วนเหลือ ส่วนการเลือกตัวแบบที่เหมาะสมระหว่างกรณี 4 กลุ่ม และ 2 กลุ่ม ไม่จำเป็นต้องเลือก ถ้าต่างก็เป็นตัวแบบที่มีภาวะสารูปดีทั้งคู่ สามารถเลือกใช้ตัวแบบที่เหมาะสมได้ตามวัตถุประสงค์ของผู้วิจัยทั้งสองกรณี

3. ผลการทดลองและวิจารณ์

3.1 การวิเคราะห์ข้อมูลของตัวแปรตอบสนองภายใต้ตัวแบบลอจิสติก 4 กลุ่ม

ตารางที่ 1. ตัวแปรอธิบายโดยใช้ตัวแบบลอจิสติก กรณีตัวแปรตอบสนอง 4 กลุ่ม

Effect	DF	Wald - ChiSquare	Pr > ChiSq (P-value)
ascites	1	11.3941	0.0007
hepatom	1	37.5632	< 0.0001
spiders	1	7.5847	0.0059
bili	1	7.7274	0.0054

จากตารางที่ 1 การตรวจพบภาวะท้องมาน (ascites), การตรวจพบภาวะตับโต (hepatom), การตรวจพบไฟแดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) เป็นตัวแปรอธิบายมีผลกระทบที่มีความเกี่ยวข้องกับระยะของโรคอย่างมีนัยสำคัญทางสถิติที่ $\alpha = 0.05$

ตารางที่ 2. ค่าประมาณพารามิเตอร์ของตัวแบบลอจิสติก 4 กลุ่ม ด้วยวิธีภาวะน่าจะเป็นสูงสุด

parameter	DF	Estimate	Standard Error	Wald Chi - Square	Pr > ChiSq (P-value)
Intercept 1	1	-5.4755	0.6346	74.4502	< 0.0001
Intercept 2	1	-3.1847	0.5727	30.9212	< 0.0001
Intercept 3	1	-0.7774	0.5505	1.9945	0.1579
ascites 0	1	1.8276	0.5414	11.3941	0.0007
hepatom 0	1	0.8424	0.1374	37.5632	< 0.0001
spiders 0	1	0.3961	0.1438	7.5847	0.0059
bili 0	1	0.3617	0.1301	7.7274	0.0054

จากตารางที่ 2 ค่าประมาณสัมประสิทธิ์ของตัวแบบลอจิสติกที่ได้จากการประมาณค่าโดยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood) พบว่าค่าประมาณสัมประสิทธิ์ของตัวแปร การตรวจพบภาวะท้องมาน (ascites), การตรวจพบภาวะตับโต (hepatom), การตรวจพบไฟแดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) เท่ากับ 1.8276, 0.8424, 0.3961 และ 0.3617 ตามลำดับ เนื่องจากค่า P-value น้อยกว่า 0.05 แสดงว่าสัมประสิทธิ์ลอจิสติกสำหรับตัวแปรการตรวจพบภาวะท้องมาน (ascites), การตรวจพบภาวะตับโต (hepatom), การตรวจพบไฟแดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) มีนัยสำคัญทางสถิติที่ระดับ $\alpha = 0.05$ สามารถเขียนสมการลอจิสติกได้ 3 สมการคือ

$$\log \left[\frac{P(Y \leq 1 | x)}{P(Y > 1 | x)} \right] = -5.4755 - 1.8276 \text{ ascites}(0) - 0.8424 \text{ hepatom}(0) - 0.3961 \text{ spiders}(0) + 0.3617 \text{ bili}(0) \quad \text{_____}(1)$$

$$\log \left[\frac{P(Y \leq 2 | x)}{P(Y > 2 | x)} \right] = -3.1847 - 1.8276 \text{ ascites}(0) - 0.8424 \text{ hepatom}(0) - 0.3961 \text{ spiders}(0) + 0.3617 \text{ bili}(0) \quad \text{_____}(2)$$

$$\log \left[\frac{P(Y \leq 3 | x)}{P(Y > 3 | x)} \right] = -0.7774 - 1.8276 \text{ ascites}(0) - 0.8424 \text{ hepatom}(0) - 0.3961 \text{ spiders}(0) + 0.3617 \text{ bili}(0). \quad \text{_____}(3)$$

การพยากรณ์ค่าความน่าจะเป็นจากสมการทั้ง 3 สมการของกรณีตัวแปรตอบสนอง 4 กลุ่ม

สามารถหาค่าความน่าจะเป็นของตัวแปรตอบสนอง ได้ดังนี้

ค่าความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะที่ 1 (Porital stage) สามารถคำนวณจากสมการ $P(Y=1) = P(Y \leq 1)$

ค่าความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะที่ 2 (Periportal stage) สามารถคำนวณจากสมการ $P(Y=2) = P(Y \leq 2) - P(Y \leq 1)$

ค่าความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะที่ 3 (Septal stage) สามารถคำนวณจากสมการ $P(Y=3) = P(Y \leq 3) - P(Y \leq 2)$

ค่าความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะที่ 4 (Biliary stage) สามารถคำนวณจากสมการ $P(Y=4) = 1 - P(Y \leq 3)$

การวิเคราะห์ข้อมูล และการตีความหมายของผลลัพธ์จากตัวแบบลอจิสติกส์สม ชกตัวอย่าง ดังนี้

$$P(Y=1) = P(Y < 1) = \frac{e^{-5.4755+1.8276ascites(1)+0.8424hepatom(1)+0.39961spiders(1)+0.3617bili(1)}}{1 + e^{-5.4755+1.8276ascites(1)+0.8424hepatom(1)+0.39961spiders(1)+0.3617bili(1)}}$$

$$= \frac{e^{-2.0477}}{1 + e^{-2.0477}} = 0.1143$$

นั่นคือ ความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยที่ป่วยในระยะที่ 1 (Porital stage) เท่ากับ 0.1143 เมื่อผู้ป่วยตรวจพบภาวะท้องมาน, ภาวะตับโต, ไข่แดงมีลักษณะคล้ายแมงมุมและค่าซีรั่มบิลิรูบิน อยู่ในเกณฑ์ผิดปกติ

$$P(Y=2) = P(Y \leq 2) - P(Y \leq 1)$$

$$= \frac{e^{-3.1847+1.8276ascites(1)+0.8424hepatom(1)+0.39961spiders(1)+0.3617bili(1)}}{1 + e^{-3.1847+1.8276ascites(1)+0.8424hepatom(1)+0.39961spiders(1)+0.3617bili(1)}} - 0.1143 = 0.4462$$

นั่นคือ ความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยที่ป่วยในระยะที่ 2 (Periportal stage) เท่ากับ 0.4462 เมื่อผู้ป่วยตรวจพบภาวะท้องมาน, ภาวะตับโต, ไข่แดงมีลักษณะคล้ายแมงมุมและค่าซีรั่มบิลิรูบิน อยู่ในเกณฑ์ผิดปกติ

$$P(Y=3) = P(Y \leq 3) - P(Y \leq 2)$$

$$= \frac{e^{-0.7774+1.8276ascites(1)+0.8424hepatom(1)+0.39961spiders(1)+0.3617bili(1)}}{1 + e^{-0.7774+1.8276ascites(1)+0.8424hepatom(1)+0.39961spiders(1)+0.3617bili(1)}} - 0.5605 = 0.3736$$

นั่นคือ ความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยที่ป่วยในระยะที่ 3 (Septal stage) เท่ากับ 0.3736 เมื่อผู้ป่วยตรวจพบภาวะท้องมาน, ภาวะตับโต, ไข่แดงมีลักษณะคล้ายแมงมุมและค่าซีรั่มบิลิรูบิน อยู่ในเกณฑ์ผิดปกติ และ $P(Y=4) = 1 - P(Y \leq 3) = 0.0660$

นั่นคือ ความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยที่ป่วยในระยะที่ 4 (Biliary stage) เท่ากับ 0.0660 เมื่อผู้ป่วยตรวจพบภาวะท้องมาน, ภาวะตับโต, ไข่แดงมีลักษณะคล้ายแมงมุมและค่าซีรั่มบิลิรูบิน อยู่ในเกณฑ์ผิดปกติ

3.2 การวิเคราะห์ข้อมูลของตัวแปรตอบสนองภายใต้ตัวแบบลอจิสต์ 2 กลุ่ม ดังตารางที่ 3

ตารางที่ 3. ตัวแปรอธิบายโดยใช้ตัวแบบลอจิสต์สองกลุ่มกรณีตัวแปรตอบสนอง 2 กลุ่ม

Effect	DF	Wald Chi-Square	Pr > ChiSq (P-value)
hepatom	1	16.9919	< 0.0001
spiders	1	5.3991	0.0201
bili	1	8.2032	0.0042

จากตารางที่ 3 สรุปได้ว่าตัวแปรการตรวจพบภาวะตับโต (hepatom) การตรวจพบไข่แดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) มีความสัมพันธ์กับระยะของโรคอย่างมีนัยสำคัญทางสถิติที่ $\alpha = 0.05$ ค่าประมาณพารามิเตอร์ของตัวแบบลอจิสต์สองกลุ่ม ดังตารางที่ 4

ตารางที่ 4. ค่าประมาณพารามิเตอร์ของตัวแบบลอจิสต์สองกลุ่ม

parameter	DF	Estimate	Standard Error	Wald Chi - Square	Pr > ChiSq (P-value)	Exp(Est)
Intercept	1	1.4486	0.2267	40.8295	< 0.0001	0.235
hepatom 0	1	-0.7024	0.1704	16.9919	< 0.0001	2.019
spiders 0	1	-0.5224	0.2248	5.3991	0.0201	1.686
bili 0	1	-0.4464	0.1559	8.2032	0.0042	1.563

จากตารางที่ 4 ค่าประมาณสัมประสิทธิ์ของตัวแบบลอจิสต์สองกลุ่มที่ได้จากการประมาณค่าโดยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood) พบว่า ค่าประมาณสัมประสิทธิ์ของตัวแปร การตรวจพบภาวะตับโต (hepatom), การตรวจพบไข่แดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) เท่ากับ 0.7024, 0.5224 และ 0.4464 ตามลำดับ เนื่องจากค่า P-value น้อยกว่า 0.05 แสดงว่าสัมประสิทธิ์

ลอจิตสองกลุ่มสำหรับตัวแปร การตรวจพบภาวะตับโต (hepatom), การตรวจพบไฟแดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) อย่างมีนัยสำคัญทางสถิติที่ระดับ $\alpha=0.05$ ดังนั้นสามารถเขียนสมการลอจิต 2 กลุ่มได้คือ

$$\text{logit}[P(Y = 1|x)] = 1.4486 - 0.7024 \text{ hepatom}(0) - 0.5224 \text{ spiders}(0) - 0.4464 \text{ bili}(0) \quad (4)$$

การพยากรณ์ค่าความน่าจะเป็นจากสมการ (4) ของกรณีตัวแปรตอบสนอง 2 กลุ่ม

การพยากรณ์ค่าความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิใน ระยะเวลารุนแรง กรณีที่ตรวจพบภาวะตับโต (hepatom), ไฟแดงมีลักษณะคล้ายแมงมุม (spiders) และค่าบิลิรูบิน (bili) ในเกณฑ์ผิดปกติคือมีค่าอยู่นอกช่วง 0.3 ถึง 1.0 มิลลิกรัม/เดซิลิตร

$$\text{คือ} \quad P(x) = \frac{e^{1.4486-0.7024\text{hepatom}(1)-0.5224\text{spiders}(1)-0.4464\text{bili}(1)}}{1+e^{1.4486-0.7024\text{hepatom}(1)-0.5224\text{spiders}(1)-0.4464\text{bili}(1)}} = \frac{e^{-0.2226}}{1+e^{-0.2226}} = 0.4446$$

นั่นคือ ความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยในระยะรุนแรงเท่ากับ 0.4446 เมื่อผู้ป่วยตรวจพบภาวะท้องมาน, ภาวะตับโต, ไฟแดงมีลักษณะคล้ายแมงมุมและค่าซีรั่มบิลิรูบินอยู่ในเกณฑ์ผิดปกติ

3.3 การเลือกตัวแบบที่เหมาะสมกับข้อมูลสำหรับแต่ละกรณี

แสดงตัวแบบที่เลือกได้แล้วในแต่ละกรณี โดยทั้งสองตัวแบบมีภาวะสารูปดี ดังตารางที่ 5

ตารางที่ 5. การเปรียบเทียบความเหมาะสมของตัวแบบ ด้วยตัวสถิติ AIC, SC และ -2LogL

ตัวแบบ	AIC	SC	-2LogL
ลอจิตสะสม 4 กลุ่ม	559.927	585.321	545.927
ลอจิต 2 กลุ่ม	269.819	284.330	261.819

จากตารางที่ 5 พบว่า ตัวแบบลอจิตสะสม 4 กลุ่ม (Proportional Odds model) ให้ค่าของตัวสถิติ AIC, SC และ -2LogL เท่ากับ 559.927, 585.321 และ 545.927 ตามลำดับ ตัวแบบลอจิต 2 กลุ่มให้ค่า 269.819, 284.330 และ 261.819 ตามลำดับ โดยตัวแบบทั้งคู่ประกอบด้วยตัวแปรอธิบายที่แตกต่างกันบ้าง

และทั้งสองตัวแบบต่างมีภาวะรูปคืออย่างมีนัยสำคัญทางสถิติที่ระดับ $p < 0.001$ และ $p < 0.02$ ตามลำดับ โดยสามารถเลือกใช้ตัวแบบที่เหมาะสมได้ตามวัตถุประสงค์ของผู้วิจัยทั้งสองกรณี

4. สรุปผลการทดลอง

4.1 การวิเคราะห์ข้อมูลกรณีที่ตัวแปรตอบสนองจำแนกประเภทมีลำดับ 4 กลุ่ม

1. ปัจจัยที่มีผลต่อระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ คือ การตรวจพบภาวะท้องมาน (ascites), การตรวจพบภาวะตับโต (hepatom), การตรวจพบไฟแดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) โดยตัวแบบแสดงใน 3 สมการ ดังนี้

$$\log \left[\frac{P(Y \leq 1 | x)}{P(Y > 1 | x)} \right] = -5.4755 - 1.8276 \text{ ascites}(0) - 0.8424 \text{ hepatom}(0) - 0.3961 \text{ spiders}(0) + 0.3617 \text{ bili}(0)$$

$$\log \left[\frac{P(Y \leq 2 | x)}{P(Y > 2 | x)} \right] = -3.1847 - 1.8276 \text{ ascites}(0) - 0.8424 \text{ hepatom}(0) - 0.3961 \text{ spiders}(0) + 0.3617 \text{ bili}(0)$$

$$\log \left[\frac{P(Y \leq 3 | x)}{P(Y > 3 | x)} \right] = -0.7774 - 1.8276 \text{ ascites}(0) - 0.8424 \text{ hepatom}(0) - 0.3961 \text{ spiders}(0) + 0.3617 \text{ bili}(0)$$

2. การประมาณค่าความน่าจะเป็นของตัวแปรตอบสนอง พบว่าความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะที่ 1 (Portal stage) กรณีที่ตรวจพบภาวะท้องมาน, ภาวะตับโต, ไฟแดงมีลักษณะคล้ายแมงมุมและค่าซีรั่มบิลิรูบินอยู่ในเกณฑ์ผิดปกติคือมีค่าอยู่นอกช่วง 0.3 ถึง 1.0 มิลลิกรัม/เดซิลิตร เท่ากับ 0.1143 ความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะที่ 2 (Periportal stage) เท่ากับ 0.4462 ความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะที่ 3 (Septal stage) เท่ากับ 0.3736 และความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะที่ 4 (Biliary stage) เท่ากับ 0.0660

4.2 การวิเคราะห์ข้อมูล กรณีที่ตัวแปรตอบสนองจำแนกประเภทสองกลุ่ม

1. ปัจจัยที่มีผลต่อระยะของโรคตับแข็งจากท่อน้ำดีปฐมภูมิ คือ การตรวจพบภาวะตับโต (hepatom), การตรวจพบไฟแดงมีลักษณะคล้ายแมงมุม (spiders) และค่าซีรั่มบิลิรูบิน (bili) โดยตัวแบบที่ได้ คือ

$$\text{logit}[P(Y = 1|x)] = 1.4486 - 0.7024 \text{ hepatom}(0) - 0.5224 \text{ spiders}(0) - 0.4464 \text{ bili}(0).$$

2. การประมาณค่าความน่าจะเป็น พบว่าความน่าจะเป็นที่ผู้ป่วยจะอยู่ในกลุ่มผู้ป่วยโรคตับแข็งจากท่อน้ำดีปฐมภูมิในระยะรุนแรง เมื่อตรวจพบภาวะตับโต (hepatom), ไขมันแดงลักษณะคล้ายแมงมุม (spiders) และค่าบิลิรูบิน (bili) ในเกณฑ์ผิดปกติออกช่วง 0.3 ถึง 1.0 มก/เดซิลิตร เท่ากับ 0.4446

4.3 การเลือกตัวแบบที่เหมาะสมกับข้อมูล

ตัวแบบลอจิสติกและตัวแบบลอจิสติกสองกลุ่มโดยพิจารณาจากค่าตัวสถิติ AIC, SC และอัตราส่วนภาวะน่าจะเป็น (G^2) หรือ -2LogL และการตรวจสอบส่วนเหลือ พบว่าตัวแบบลอจิสติกสองกลุ่มมีค่าตัวสถิติ AIC, SC และอัตราส่วนภาวะน่าจะเป็น (G^2) หรือ -2LogL เท่ากับ 269.819 284.33 และ 261.819 ตามลำดับ ซึ่งมีค่าน้อยกว่าค่าตัวสถิติของตัวแบบลอ- จิสติกเท่ากับ 559.927, 585.321 และ 545.927 ตามลำดับ และจากการตรวจสอบส่วนเหลือของตัวแบบลอจิสติกสองกลุ่มและตัวแบบลอจิสติก พบว่ามีการกระจายของส่วนเหลือโดยแต่ละค่าสังเกตมีการกระจายที่ดี (ไม่ได้แสดงไว้ เนื่องจากเนื้อที่จำกัด) และเมื่อวิเคราะห์ความไวด้วย ROC Curve ของตัวแบบลอจิสติกสองกลุ่มแล้วพบว่ามีความไวในการพยากรณ์กลุ่มที่สนใจได้ถูกต้องสูงทั้งสองตัวแบบ (ไม่ได้แสดงไว้ เนื่องจากเนื้อที่จำกัด)

4.4 ข้อเสนอแนะเพื่อการวิจัย

ในการศึกษาตัวแบบลอจิสติกเมื่อตัวแปรตอบสนอง 4 กลุ่มและตัวแบบลอจิสติก 2 กลุ่ม เมื่อตัวแปรตอบสนองจำแนกประเภท 2 กลุ่ม นี้เป็นกรณีศึกษาโดยใช้ข้อมูลจริงจากเว็บไซต์ <http://lib.stat.cmu.edu/> เรื่อง Primary Biliary Cirrhosis ดังนั้นผลลัพธ์ที่ได้จากการวิเคราะห์สามารถสรุปได้จากตัวแบบที่นำมาศึกษาเท่านั้น หากต้องการศึกษาจากตัวแบบอื่น ๆ อาจทำเพิ่มเติมและตรวจสอบความเหมาะสมของตัวแบบ ซึ่งอาจพบความแตกต่างกันภายใต้เงื่อนไขที่นำไปประยุกต์เป็นกรณี ๆ ไป ตัวแบบอื่น ๆ เช่นตัวแบบ Continuation ratio model [8], Constrained and unconstrained partial proportional odds model [9], และ Adjacent – category logistic model [10]

5. เอกสารอ้างอิง

- [1] Agresti, A., 2000. Categorical Data Analysis. Second edition, New York : John Wiley & Sons.
- [2] Clavien, P. A. and Killenberg, P. G., 2006. Medical Care of the Liver Transplant Patient: TotalPre-, Intra- and Post-Operative Management, 3rd Edition. Blackwell: John Wiley & Sons.
- [3] PBCers, 2009. [online] Available at: < <http://pbcers.org/> > [Accessed 12 March 2011].

- [4] Primary-biliary-cirrhosis, 2009. [online] Available at:
<<http://emedicine.medscape.com/gastroenterology/>> [Accessed 8 March 2011].
- [5] Mayo Clinic staff, 2009. [online] Available at:
<<http://www.mayoclinic.com/health/primary-biliary-cirrhosis/>> [Accessed 8 March 2011].
- [6] Walker, S.H. and Duncan D.B., 1967. Estimation of the probability of an event as a function of several independent variables. *Biometrika* 54, 167-179.
- [7] McCullagh, P., 1980. Regression Models for Ordinal Data. *J. Royal .Statist. Soc.Ser. B*42, 109-142.
- [8] Fienberg, S.E., 1980. The Analysis of Cross-Classified Categorical Data, 2nd ed., MIT Press. Cambridge.
- [9] Peterson, B.L. and Harrell, F.E., 1990. Partial Proportional Odds Models for Ordinal Response Variables. *Appl. Stat.*, 39, 205-217.
- [10] Agresti, A., 1984. Analysis of Ordinal Categorical Data. New York: John Wiley & Sons.
- [11] สมาคม PBC, 2009. *Canadian PBC Society*. [online] Available at:
<<http://www.pbc-society.ca/>> [Accessed 8 March 2011].