

ตัวแบบพยากรณ์และปัจจัยเสี่ยงที่จะเกิดเบาหวานจอประสาทตาขั้นคุกคาม ต่อการมองเห็น : กรณีศึกษาโรงพยาบาลนครปฐม

Predictive models and risk factors for development of vision threatening diabetic retinopathy: a case of Nakhonpathom Hospital

เอกจิตต์ จิ่งเจริญ ปร.ด.

*Ekachidd Chungcharoen Ph.D.**

บุญเพ็ญ จงเสรีจิตต์ พ.บ.,(จักษุวิทยา)

*Boonphen Jongsareejit M.D.(Ophthalmology)***

*คณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์

Thammasat Business School, Thammasat university

**โรงพยาบาลนครปฐม

Nakhonpathom Hospital

Received: August 22,2022 Revised: September 28,2022 Accepted: October 3,2022

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาตัวแบบพยากรณ์ภาวะการเกิดโรคเบาหวานที่จอประสาทตาในขั้นคุกคามต่อการมองเห็นของผู้ป่วยเบาหวาน ซึ่งผลลัพธ์ของการพยากรณ์จะถูกใช้เป็นข้อมูลเพิ่มเติมในกระบวนการคัดกรองและป้องกันในผู้ป่วยเบาหวานทั้งหมดเพื่อที่จะสามารถทำการเฝ้าระวัง แทรกแซงปรับเปลี่ยนพฤติกรรมของผู้ป่วยและให้การรักษอย่างเข้มข้นทันทั่วถึง เพื่อให้ผู้ป่วยมีคุณภาพการมองเห็นและมีคุณภาพชีวิตที่ดีอยู่ได้ในระยะยาว ตัวแบบพยากรณ์จะใช้เทคนิคหลักสองเทคนิคได้แก่ Logistic Regression และ Probabilistic Neural Network เนื่องจากถูกใช้อย่างแพร่หลายในการพยากรณ์ตัวแปรตามในลักษณะของการแบ่งกลุ่ม ข้อมูลที่ใช้ในการพัฒนาตัวแบบรวบรวมจากข้อมูลผู้ป่วยโรงพยาบาลนครปฐมประกอบไปด้วยตัวแปรอิสระ 18 ตัวแปร ในระยะเวลา 6 เดือนล่าสุด โดยตัดข้อมูลที่ไมครบถ้วนและผิดพลาดออก มีจำนวนรวมทั้งสิ้น 355 คน ประกอบไปด้วยผู้ป่วยที่ยังไม่เป็นโรคเบาหวานจอประสาทตาในขั้นคุกคามต่อการมองเห็นจำนวน 124 คน และเป็นโรคเบาหวานจอประสาทตาในขั้นคุกคามต่อการมองเห็นจำนวน 231 คน โดยมีการแบ่งข้อมูลออกเป็นสองชุดได้แก่ชุดที่พัฒนาตัวแบบ (in sample test) จำนวน 90% และชุดสำหรับทดสอบความแม่นยำของตัวแบบ (out of sample test) จำนวน 10% โดยมีสัดส่วนของการเกิดโรคใกล้เคียงกัน ผลการศึกษาพบว่าตัวแบบ logistic regression มีความแม่นยำโดยรวม 69.01% ในขณะที่ตัวแบบ probabilistic neural network มีความแม่นยำโดยรวม 96.90% โดยทั้งสองตัวแบบมีความสามารถจะพยากรณ์การเกิดภาวะ VTDR ในผู้ป่วยได้แม่นยำกว่าการพยากรณ์การไม่เกิดภาวะ VTDR ส่วนในด้านปัจจัยสำคัญที่เป็นสาเหตุได้แก่ เพศ อายุ cholesterol hematocrit creatinine HDL-cholesterol และ body mass index

คำสำคัญ : โรคเบาหวานจอประสาทตาขั้นคุกคามต่อการมองเห็นของผู้ป่วย, การวิเคราะห์การถดถอยโลจิสติกส์, การวิเคราะห์เครือข่ายประสาทเทียมความน่าจะเป็น, ตัวแบบพยากรณ์

Abstract

This research aimed to develop the model to predict vision threatening diabetic retinopathy (VTDR) condition in diabetic patients so that preventive medical treatments could be started at early stage for prolonging patients' good quality of life. Two popular predicting techniques for categorical dependent variable were used in comparison. These were logistic regression and probabilistic neural network. Data for the last six months of DR patients at Nakhonpathom hospital were collected based on 18 variables from the literature. Data cleansing were implemented. Eventually, there was a total of 355 data used with 124 patients without VTDR condition and 231 patients with VTDR condition. These data were also separated into two groups. The first group (in sample test) with 90% of the total data was used for developing the models. The second group (out of sample test) with 10% of data was used for checking the models' accuracy. The research results showed the overall accuracy of logistic regression model was 69.01% while the overall accuracy of probabilistic neural network model was 96.90%. Both logistic regression and probabilistic neural network models were more accurate in predicting patients with VTDR condition than without VTDR condition. The study showed significant risk factors related to VTDR condition were sex, age, cholesterol, hematocrit, creatinine, HDL-cholesterol, and body mass index.

Keywords : vision threatening diabetic retinopathy, logistic regression, probabilistic neural network, forecasting model.

บทนำ

โรคเบาหวาน (diabetic mellitus, DM) เป็นโรคไม่ติดต่อที่พบบ่อย มากกว่าร้อยละ 5 ของประชากรไทยวัย 30-60 ปี ภาวะแทรกซ้อนที่สำคัญที่สุดอันหนึ่งคือเบาหวานจอประสาทตา (diabetic retinopathy, DR) ซึ่งพบได้ประมาณ 1 ใน 4 ของผู้ป่วยเบาหวาน และเป็นสาเหตุที่ทำให้เกิดการสูญเสียการมองเห็นอันดับสองของประเทศไทย ในประชากรทั่วโลกพบว่าความชุกของ DR ในผู้ป่วยเบาหวานมีอัตราเฉลี่ย 34.6% โดยเป็นเบาหวานจอประสาทตาขั้นคุกคามต่อการมองเห็น (vision-threatening diabetic retinopathy, VTDR) 10.2% การศึกษาในไทยพบว่ามีความชุกของ DR ในผู้ป่วยเบาหวานประมาณ 24-31% โดยเป็นระยะที่ยังไม่มีหลอดเลือดฝอยงอกใหม่ (non proliferative diabetic retinopathy, NPDR) ประมาณ 21-23% และระยะหลอดเลือดฝอยงอกใหม่ (proliferative diabetic retinopathy, PDR) 2.3-9.4% ในขั้น severe NPDR

และ PDR ให้การรักษาโดยยิงเลเซอร์จอตา (panretinal photocoagulation, PRP) ส่วน DME, PDR, เลือดออกในน้ำวุ้นตา (vitreous hemorrhage) รักษาโดยฉีดยายับยั้งการเจริญของเส้นเลือด (anti-vascular endothelial growth factor, anti-VEGF) เข้าวุ้นตาพร้อมด้วย ซึ่งผู้ป่วยกลุ่มที่ได้รับการรักษา Laser PRP หรือฉีดยา anti-VEGF เข้าวุ้นตานี้คือผู้ป่วยเบาหวานจอประสาทตาขั้นคุกคามต่อการมองเห็น ถ้าไม่ได้รับการรักษาก็จะทำให้สูญเสียการมองเห็น พบว่าปัจจัยเสี่ยงหลักของ DR ได้แก่ ระยะเวลาที่เป็นเบาหวานระดับน้ำตาลในเลือดที่สูงอย่างไรก็ตามยังมีความแตกต่างของปัจจัยเสี่ยง รูปแบบ ความรุนแรง และความแปรปรวนในการให้ระดับความผิดปกติของ DR (stage of DR) นอกจากนั้นยังมีอิทธิพลจากพฤติกรรมการบริโภคอาหาร ภาวะเศรษฐกิจและสังคม ในส่วนของงานวิจัยตัวแบบพยากรณ์โรคเบาหวานจอประสาทตา ซึ่งสร้างจากปัจจัยเสี่ยงที่มีนัยสำคัญหลายตัวแปร ทำให้มี

ความแม่นยำมากกว่าการดูปัจจัยเสี่ยงตัวเดียว และเป็นแนวโน้มนงานวิจัยในต่างประเทศ พบว่าสามารถใช้ในการพยากรณ์ผู้ป่วย DR ได้อย่างแม่นยำและง่ายต่อการนำไปใช้ทำให้วางแผนนัดตรวจคัดกรองตามความเหมาะสมในผู้ป่วยแต่ละคน เกิดประสิทธิภาพและความคุ้มค่าในการดูแลผู้ป่วย⁽¹⁻²⁾ ซึ่งการศึกษาประเภทนี้ในประเทศไทยมีน้อยมาก สถิติที่ใช้สำหรับงานวิจัยนี้ได้แก่ การวิเคราะห์ logistic regression และโครงข่ายประสาทเทียมประเภท probabilistic neural network (PNN) เพื่อพัฒนาตัวแบบพยากรณ์ผู้ป่วยที่มีลักษณะตามตัวแปรต่าง ๆ ที่เป็นปัจจัยเสี่ยงที่จะเกิด VTDR

ตัวแบบ logistic regression เป็นรูปแบบของ regression model ที่ตัวแปรตามมีลักษณะเป็นข้อมูลเชิงกลุ่มมีได้สองค่าเช่น 0 เท่ากับ “ไม่เกิดเหตุการณ์” และ 1 เท่ากับ “เกิดเหตุการณ์” โดยที่ตัวแปรอิสระที่เป็นได้ทั้งตัวแปรเชิงกลุ่ม (categorical variable) หรือตัวแปรแบบต่อเนื่อง (continuous variable) ถูกนำมาใช้ในการสร้างแบบจำลองเพื่อพยากรณ์ในหลายสถานการณ์ไม่ว่าจะเป็นทางธุรกิจหรือทางการแพทย์หรืออื่น ๆ โดยในงานวิจัยนี้มีวัตถุประสงค์เพื่อหาความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามที่มีสองสถานะคือ 0 เท่ากับผู้ป่วยไม่เกิดภาวะ VTDR และ 1 เท่ากับผู้ป่วยเกิดภาวะ VTDR โดยสามารถเขียนฟังก์ชันความน่าจะเป็นของการเกิดเหตุการณ์ผู้ป่วยเกิดภาวะ VTDR หรือ $\Pr(y=1)$ ได้ดังนี้⁽³⁾

$$\Pr(y = 1) = \frac{1}{1 + e^{-x}} \quad (1)$$

กรณีมีตัวแปรอิสระจำนวน n ตัว เมื่อ X เป็นผลรวมเชิงเส้นของตัวแปรต่าง ๆ ดังนั้น

$$X = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (2)$$

โดยที่ $b_i; i = 1, 2, \dots, n$ คือ สัมประสิทธิ์ของตัวแปรอิสระ

x_1, x_2, \dots, x_n คือ ตัวแปรอิสระทั้งหมด n ตัวแปร

ดังนั้นจึงขยายสมการ (1) ได้ดังนี้

$$\Pr(y = 1) = \frac{1}{1 + e^{-(b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n)}} \quad (3)$$

สำหรับความน่าจะเป็นที่จะเกิดเหตุการณ์ผู้ป่วยไม่มีภาวะ VTDR หรือ $\Pr(y = 0)$ จะเขียนได้จากกฎของ mutually exclusive and collectively exhaustive $1 - \Pr(y = 1)$ ดังนี้

$$\Pr(y = 0) = 1 - \left(\frac{1}{1 + e^{-X}} \right) = \frac{e^{-X}}{1 + e^{-X}} \quad (4)$$

จากสมการที่ (3) และ (4) สามารถแปลงค่าให้อยู่ในรูปของอัตราส่วนของความน่าจะเป็นของการเกิดเหตุการณ์ผู้ป่วยเกิดภาวะ VTDR ต่อความน่าจะเป็นที่ผู้ป่วยไม่เกิดภาวะ VTDR หรือเรียกว่า Odds Ratio ดังนี้

$$\frac{\Pr(y = 1)}{\Pr(y = 0)} = \frac{\left(\frac{1}{1 + e^{-X}} \right)}{\frac{e^{-X}}{1 + e^{-X}}} = e^X \quad (5)$$

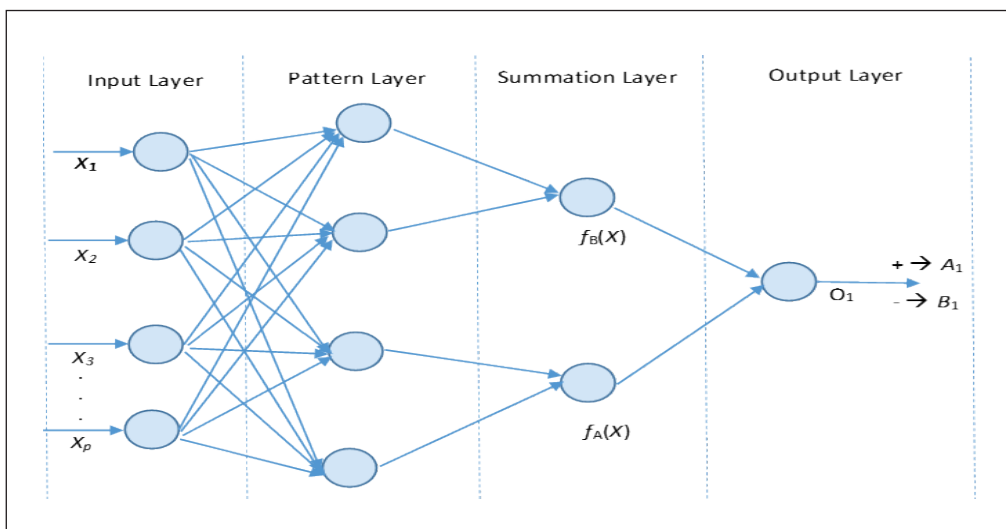
แปลงสมการที่ (5) ให้อยู่ในรูปของสมการเชิงเส้นโดยการใส่ \ln เข้าไปทั้งสองข้างเป็นการทำ logit transformation เพื่อประมาณค่าสัมประสิทธิ์ b_i ด้วยวิธี ordinary least square (OLS) ดังนี้

$$\ln \left[\frac{\Pr(y=1)}{\Pr(y=0)} \right] = x = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (6)$$

ซึ่งการอธิบายทิศทางความสัมพันธ์ของตัวแปรอิสระกับความน่าจะเป็นที่จะเกิดเหตุการณ์ภาวะ VTDR ในผู้ป่วยสามารถอธิบายได้โดยตรงจากค่า logistic coefficients เช่นเดียวกับกรณีของ multiple regression หรือดูจากค่า exponential coefficients ซึ่งในกรณีที่ค่าสัมประสิทธิ์มีค่ามากกว่า 1.0 จะสะท้อนความสัมพันธ์ที่เป็นบวกแต่ถ้าน้อยกว่า 1.0 จะสะท้อนความสัมพันธ์ในทิศทางตรงกันข้าม

ในส่วนของตัวแบบโครงข่ายประสาทเทียม (artificial neural network) ได้ถูกพัฒนาขึ้นในช่วงหลายปีที่ผ่านมาโดยเป็นการประยุกต์การจำลองการทำงานของสมองมนุษย์ด้วยโปรแกรมคอมพิวเตอร์ให้มีความสามารถในการเรียนรู้และจดจำผ่านตัวอย่างนำไปสู่การตัดสินใจเช่น การจดจำรูปแบบที่มีความไม่แน่นอน การประมาณความสัมพันธ์ของตัวแปรอิสระและตัวแปรตามที่ไม่ทราบรูปแบบของความสัมพันธ์ที่แน่นอน การจัดกลุ่มของข้อมูลและการพยากรณ์ทางธุรกิจ ฯลฯ ซึ่งในการพัฒนาตัวแบบนั้นตัวแปรอิสระที่นำมาใช้นั้นต้องมีความสัมพันธ์กับตัวแปรตามและจะต้องไม่มีปัญหาที่มีความสัมพันธ์กันเองสูง (multicollinearity) และต้องไม่มีปัญหาความแปรปรวนของค่าความคลาดเคลื่อนไม่เท่า

กัน (heteroscedasticity) มิฉะนั้นจะส่งผลให้เกิดความผิดพลาดในการตัดสินใจตัวแบบประสาทเทียมได้ถูกพัฒนาขึ้นโดยมีหลายประเภทเหมาะกับงานที่มีลักษณะแตกต่างกัน DF. Specht⁽⁴⁾ พัฒนาแบบจำลองโครงข่ายประสาทความน่าจะเป็น (probabilistic neural network, PNN) ซึ่งเป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมแบบหลายชั้น (multi-layers feedforward neural network) และอาศัยแนวคิดจากหลักการของ Bayesian classification ในส่วนของการจัดหมวดหมู่และแยกแยะข้อมูล PNN มีกระบวนการเรียนรู้แบบไม่มีการสอน (unsupervised learning) มีการจัดเรียงโครงสร้างด้วยตัวเองตามลักษณะของข้อมูลผลลัพธ์ที่ได้ ซึ่งจะจัดเป็นหมวดหมู่ของผลลัพธ์ที่ได้เป็นเครือข่ายแบบ feedforward ที่มี 4 ชั้นได้แก่ ชั้นป้อนเข้า (input layer) ชั้นการจำลองแบบ (pattern layer) ชั้นค่าผลรวม (summation layer) ชั้นของผลลัพธ์ (output layer) ตามภาพที่ 1 โดย PNN ถูกนำมาใช้กันอย่างแพร่หลายในปัญหาการจำแนกประเภท (classification) และการจดจำรูปแบบ (pattern recognition) รวมถึงการพยากรณ์ความล้มเหลวของธุรกิจ การพยากรณ์ทางการแพทย์ และด้านอื่น ๆ⁽⁵⁻⁸⁾



ภาพที่ 1 แสดงกรอบแนวคิดของ probabilistic neural network ดัดแปลงจาก Specht (1990)

จากภาพที่ 1 เมื่อมีปัจจัยนำเข้า เลเยอร์แรกจะคำนวณระยะทางจากเวกเตอร์อินพุตไปยังเวกเตอร์อินพุตการฝึก ในเลเยอร์ที่สองหรือ pattern layer เพื่อสร้างเวกเตอร์องค์ประกอบระบุว่าอินพุตอยู่ใกล้กับอินพุตการฝึกมากเพียงใด เลเยอร์ที่สามรวมผลรวมของอินพุตแต่ละคลาส และสร้างเอาต์พุตสุทธิเป็นเวกเตอร์ของความน่าจะเป็น สุดท้ายฟังก์ชันการถ่ายโอนที่แข่งขันกันบนเอาต์พุตของเลเยอร์ที่สามจะเลือกความน่าจะเป็นสูงสุด และสร้าง 1 (การระบุเชิงบวก) สำหรับคลาสนั้น และ 0 (การระบุ

เชิงลบ) สำหรับคลาสที่ไม่ได้กำหนดเป้าหมาย โดยทั่วไป การหาความน่าจะเป็นของแต่ละผลลัพธ์สามารถอธิบายจากหลักการของ Bayesian classification ได้คือเมื่อพิจารณากรณีสถานะของตัวแปรตาม θ ที่มีสองสถานะคือ θ_A คือผู้ป่วยไม่อยู่ในภาวะ VTDR และ θ_B เท่ากับผู้ป่วยอยู่ในภาวะ VTDR โดยที่การตัดสินใจ $\theta = \theta_A$ หรือ $\theta = \theta_B$ นั้นจะขึ้นอยู่กับเวกเตอร์ขนาด p คือ $X^t = [X_1 \dots X_i \dots X_p]$ กฎการตัดสินใจตามทฤษฎี Bayes สามารถเขียนได้ดังนี้⁽⁵⁾

$$d(X) = \theta_A \text{ ถ้า } h_A l_A f_A(X) > h_B l_B f_B(X)$$

และ

$$d(X) = \theta_B \text{ ถ้า } h_A l_A f_A(X) < h_B l_B f_B(X) \quad (7)$$

โดยที่ X คือเวกเตอร์ในชั้นปัจจัยนำเข้า (input layer)

ค่า h_A และ $h_B = 1 - h_A$ คือค่า prior probability ก่อนการจัดให้อยู่ในสถานะ A และสถานะ B ตามลำดับ ส่วน l_A และ l_B คือ ต้นทุนของการจำแนกผิดพลาด (loss function) ของ A และ B ได้แก่

$$l_A: d(X) = \theta_B \text{ เมื่อ } \theta = \theta_A$$

$$l_B: d(X) = \theta_A \text{ เมื่อ } \theta = \theta_B$$

โดยที่ต้นทุนของการจำแนกผิดพลาดในกรณีที่จำแนกสถานะถูกต้องมีค่าเท่ากับ 0 ส่วน $f_A(X)$ และ $f_B(X)$ คือความน่าจะเป็นของ X ซึ่งถูกกำหนดจากฟังก์ชันความหนาแน่นของค่าความน่าจะเป็น (probability density function, PDF) ของสถานะ A และ B ตามลำดับ

ดังนั้นขอบเขตการตัดสินใจระหว่างพื้นที่ซึ่ง $d(X) = \theta_A$ และ $d(X) = \theta_B$ จะอธิบายได้จากสมการดังต่อไปนี้

$$f_A(X) = K f_B(X) \quad (8)$$

โดยที่

$$K = h_B l_B / h_A l_A \quad (9)$$

ซึ่งโดยทั่วไปการจำแนกสถานะของข้อมูลในสมการ (8) จะมีความซับซ้อนเนื่องจากไม่ได้มีข้อจำกัดที่เกี่ยวข้องกับ PDF อย่างชัดเจนนอกจากต้องไม่เป็นค่าลบและพื้นที่รวมแล้วต้องมีค่าเท่ากับ 1 ดังนั้นการหาค่าตอบของสมการที่ (8) จะอยู่ที่ความสามารถในการประมาณค่า PDF ของสถานะต่าง ๆ จาก training samples ซึ่งค่า prior probability และ loss function นั้นสามารถประมาณได้ล่วงหน้า

ดังนี้

$$f_A(X) = \frac{1}{(2\pi)^{p/2} \sigma^p n_A} \sum_{j=1}^{n_A} e^{-\frac{(X-Y_{Aj})'(X-Y_{Aj})}{2\sigma^2}} \quad (10)$$

โดยที่ p คือจำนวนตัวแปรในเวกเตอร์ของปัจจัยนำเข้า X ส่วน n_A คือ จำนวนตัวอย่างในกลุ่ม A ค่า Y_{Aj} คือค่ากลางของตัวอย่างที่ j^{th} ในกลุ่ม A และค่า σ คือค่า Smoothing Parameter

ในประเทศไทยการคัดกรอง DR มีการพัฒนาปัญญาประดิษฐ์เพื่อแปลผลภาพจอประสาทตาอย่างรวดเร็วและแม่นยำ ทำให้เกิดความครอบคลุม แต่ถึงกระนั้นยังมีผู้ป่วยในชั้น VTDR มากถึง 9% ของผู้ป่วยเบาหวาน จึงเป็นที่มาของงานวิจัยนี้เพื่อพัฒนาตัวแบบพยากรณ์ภาวะการเกิดโรคเบาหวานจอประสาทตาใน

ดังนั้นค่า PDF ของแต่ละสถานะที่จำแนกจะเป็นค่าที่ไม่ทราบและต้องมีการประมาณ โดย Parzen (1962) ได้พัฒนาการประมาณฟังก์ชัน PDF ของแต่ละสถานะโดยวิธีการแบบ non-parametric และ Cacoulas (1966) ได้ต่อยอดงานของ Parzen (1962) ให้ครอบคลุมถึงสถานะที่มีหลายระดับชั้น ทั้งนี้ฟังก์ชัน PDF ของแต่ละสถานะจะประกอบด้วยปัจจัยนำเข้า p ตัว ดังแสดงในสมการที่ (10) ดังนี้

ชั้นคุณภาพต่อการมองเห็นของผู้ป่วยซึ่งผลลัพธ์ของการพยากรณ์จะถูกใช้เป็นข้อมูลเพิ่มเติมในกระบวนการคัดกรองและป้องกันในผู้ป่วยเบาหวานทั้งหมดเพื่อที่จะสามารถทำการเฝ้าระวังแทรกแซงปรับเปลี่ยนพฤติกรรมเพื่อให้ผู้ป่วยมีคุณภาพการมองเห็นและมีคุณภาพชีวิตที่ดีอยู่ได้ในระยะยาว

วัสดุและวิธีการศึกษา

จากงานวิจัยที่ผ่านมาหลายปัจจัยเสี่ยงต่อการเกิด DR และการดำเนินของโรคที่แย่ง ในการวิจัยครั้งนี้จะนำปัจจัยเหล่านี้มาวิเคราะห์โดยใช้ข้อมูลผู้ป่วยโรงพยาบาลนครปฐม ที่เข้ารับการตรวจรักษาตาอ่อนหลัง 6 เดือนล่าสุด ตั้งแต่ 1 มกราคม-31 พฤษภาคม 2565 มีข้อมูลเบื้องต้น 621 ข้อมูลผ่านกระบวนการ data-cleansing สำหรับข้อมูลที่ไม่สมบูรณ์จะถูกตัดออกไป ใช้เฉพาะข้อมูลที่บันทึกครบถ้วนจำนวน 355 ข้อมูล

แบ่งข้อมูลเป็น 2 ส่วน ส่วนแรกจำนวน 325 ข้อมูลหรือ 90% เพื่อพัฒนาตัวแบบและทดสอบความเหมาะสมของตัวแบบ (in sample test) แบ่งเป็นผู้ป่วยที่ไม่มีภาวะ VTDR จำนวน 112 คนและมีภาวะ VTDR จำนวน 213 คน และส่วนที่สองจำนวน 30 คน สำหรับทดสอบความแม่นยำของตัวแบบ (out of sample test) แบ่งเป็นผู้ป่วยที่ไม่มีภาวะ VTDR จำนวน 12 คนและมีภาวะ VTDR จำนวน 18 คน โดยมีตัวแปรอิสระดังตาราง

ประเภทของตัวแปรอิสระ	ตัวแปร	ความหมาย
ลักษณะทั่วไปของผู้ป่วย	Age	อายุ
	Sex	เพศผู้ป่วย
	BMI	ค่าดัชนีมวลกาย body mass index
ข้อมูลทางคลินิก	SBP	ค่าความดันตัวบน systolic blood pressure
	DBP	ค่าความดันตัวล่าง diastolic blood pressure
ผลทางพยาธิวิทยา	HbA1c	ค่าระดับน้ำตาลในเลือดสะสม
	FBS	ระดับน้ำตาลในเลือดหลังอดอาหาร 6 ชั่วโมง
	Chol	ค่า total cholesterol
	Trig	ค่า triglyceride
	HDL	ค่า high density lipoprotein-cholesterol
	LDL	ค่า low density lipoprotein-cholesterol
	Cr	ค่า creatinine
	Bun	ค่า blood urea nitrogen
	PLT	ค่าเกร็ดเลือด platelet count
	HCT	ค่าปริมาณความเข้มข้นของเม็ดเลือดแดงต่อปริมาตรของเลือดทั้งหมด hematocrit
	mAlb	ค่า microalbumin urine
โรคร่วม	HT	ความดันสูง hypertension
	DLD	

ตัวแปรตาม มีสองค่าคือ 0 หมายถึง ผู้ป่วยที่ไม่อยู่ในภาวะ VTDR และ 1 หมายถึงเป็นผู้ป่วยที่อยู่ในภาวะ VTDR ที่ได้รับการรักษาด้วย laser PRP หรือ รักษาโดยฉีดยา anti-VEGF เข้าวันตา

โดยภาวะ VTDR ในการศึกษาี้ หมายถึงกลุ่ม severe NPDR และ PDR ให้การรักษาโดยยิงเลเซอร์จอตา (panretinal photocoagulation, PRP) และกลุ่ม DME, PDR รักษาโดยฉีดยายับยั้งการเจริญของเส้นเลือด (anti-vascular endothelial growth factor, anti-VEGF) เข้าวันตาพร้อมด้วย

การวิเคราะห์ข้อมูลด้วยวิธี logistic regression การสร้างแบบจำลองด้วยเทคนิค logistic regression เริ่มจากการพิจารณาเลือกตัวแปรอิสระทั้งหมดที่คาดว่าจะมีอิทธิพลต่อความเป็นไปได้ที่ผู้ป่วยจะมีความเสี่ยงที่จะ

เกิดภาวะ VTDR ข้างต้น โดยมีการแบ่งข้อมูลออกเป็น 2 ชุดเพื่อพัฒนาตัวแบบ (in sample test) และเพื่อทดสอบความน่าเชื่อถือของตัวแบบ (out of sample test) โดยข้อมูลชุดแรกนำตัวแปรอิสระทั้งหมดเข้าสู่กระบวนการคัดเลือกเข้าสู่การวิเคราะห์ logistic regression ด้วยวิธี stepwise selection เพื่อขจัดปัญหา multicollinearity จากนั้นนำตัวแปรอิสระที่ผ่านการคัดเลือกมาสร้างตัวแบบพยากรณ์ภาวะ VTDR และตรวจสอบความแม่นยำของตัวแบบที่สร้างขึ้นมาโดยพิจารณาค่าสถิติได้แก่ การลดลงของค่า $-2 \text{ Log Likelihood } (-2LL)$ ของตัวแบบ logistic regression สุดท้ายเทียบจาก base model ค่า pseudo R^2 ได้แก่ Cox and Snell R^2 ค่า Nagelkerke R^2 ซึ่งแสดงถึงความแปรปรวนของตัวแปรตามคือภาวะ VTDR ที่ถูกอธิบายได้โดยตัวแปรอิสระ และค่าสถิติ Hosmer and

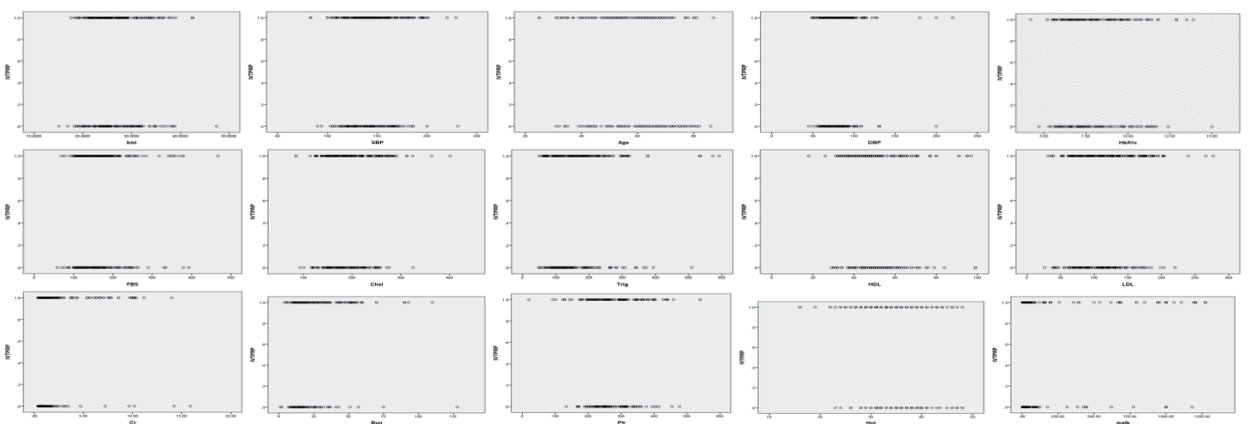
Lemeshow ซึ่งสะท้อนความแม่นยำในการแยกกลุ่มของผู้ป่วยที่มีความเสี่ยงที่จะเกิดภาวะ VTDR นั้นเอง ตัวแบบนี้จะถูกนำไปใช้ในการพยากรณ์ข้อมูลในชุดที่สอง โดยรายงานผลการพยากรณ์ในรูปแบบของ classification matrices ซึ่งเป็นการทดสอบความแม่นยำตัวแบบ

การวิเคราะห์ข้อมูลด้วยวิธี PNN การวิเคราะห์ข้อมูลด้วยเครือข่ายแบบ PNN นั้นขั้นป้จจัยนำเข้าจะประกอบด้วย Nodes จากป้จจัยต่าง ๆ ของผู้ป่วยที่มีความเสี่ยงที่จะเกิดภาวะ VTDR มีการแบ่งกลุ่มข้อมูลสำหรับเป็น training cases และ testing cases จากข้อมูลผู้ป่วยที่นำมา train ในชั้นค่าผลรวมจะประกอบไปด้วย 2 nodes จากสถานะไม่มีภาวะการเกิด VTDR และมีภาวะ VTDR และชั้นของผลลัพธ์ เมื่อมีป้จจัยนำเข้าแต่ละ node ในชั้นการจำลองแบบจะคำนวณระยะทางจาก training case ที่แทนด้วย node กับ input case ผลที่ได้จะถูกผ่านไปยังชั้นค่าผลรวมในรูปแบบของฟังก์ชันระยะทางและ smoothing factors ซึ่งเป็นตัวกำหนดถึงความเร็วที่แต่ละ training case ที่แปรผันกับระยะทางในชั้นค่าผลรวม ซึ่งแต่ละ node แสดงถึงสถานะของผู้ป่วย จะทำหน้าที่รวม

ผลลัพธ์สำหรับ nodes ต่าง ๆ ของ training cases สำหรับผลลัพธ์สถานะของการเกิดอาการของผู้ป่วยในชั้นนี้ของแต่ละกลุ่มจะอยู่ในรูปของ probability density function (PDF) แล้ว output node จะทำหน้าที่เลือกสถานะของผู้ป่วยที่มีค่า PDF สูงสุดเป็นการสรุปอาการที่ควรจะเป็นว่ามีภาวะ VTDR หรือไม่ การตรวจสอบความแม่นยำของตัวแบบ PNN สามารถกระทำได้โดยผ่าน classification matrices เช่นเดียวกัน

ผลการศึกษา

ตัวแบบ logistic regression ผลจากแผนภาพการกระจายระหว่างตัวแปรอิสระทั้ง 15 ตัวที่เป็น continuous variable กับตัวแปรตาม VTDR พบว่า logistic curve ไม่สามารถฟิตกับข้อมูลได้ดีนักเนื่องจากข้อมูลของตัวแปรอิสระส่วนใหญ่ไม่สามารถแยกแยะการไม่เกิดภาวะ VTDR (VTDR=0) และการเกิด VTDR (VTDR=1) ได้อย่างชัดเจนตามภาพที่ 2 ซึ่งกรณีนี้จะส่งผลต่อการคัดเลือกตัวแปรที่ถูกคัดเลือกเข้าสู่ตัวแบบ logistic regression ในลำดับถัดไป



ภาพที่ 2 แผนภาพการกระจายระหว่างตัวแปรอิสระ 15 ตัวแปรกับตัวแปรตาม VTDR

ผลจากการพัฒนาตัวแบบ logistic regression ด้วยวิธี stepwise selection เพื่อลดปัญหา multicollinearity พบว่าค่า $-2LL$ ลดลงอย่างมีนัยสำคัญจาก base model ที่ 405.33 เหลือ 374.99 โดยมีค่า Cox and Snell R^2 และ Nagelkerke R^2 อยู่ที่ 0.13 และ 0.18 ตามลำดับ ค่า Hosmer and Lemeshow χ^2 test อยู่ที่ 10.36 ที่ Sig. เท่ากับ 0.24 แสดงว่าตัวแบบ logistic regression มี

ความเหมาะสมในการนำมาใช้ในการพยากรณ์ ตารางที่ 1 แสดงผลการวิเคราะห์ด้วยตัวแบบ logistic regression จะเห็นว่าค่าสัมประสิทธิ์ของตัวแปรอิสระทุกตัวที่ถูกเลือกเข้ามาในตัวแบบ logistic regression มีค่า Wald statistics สูงอย่างมีนัยสำคัญ เนื่องจาก Sig. < 0.05 แสดงถึงความเหมาะสมในการถูกเลือกเข้ามาไว้ในตัวแบบ

ตารางที่ 1 ผลการวิเคราะห์ด้วยตัวแบบ logistic regression

	B	S.E.	Wald	df	Sig.	Exp(B)
Age	-0.030	0.011	7.374	1	0.007	0.971
Sex	-1.008	0.262	14.814	1	0.000	0.365
Chol	0.007	0.003	7.097	1	0.008	1.007
HCT	-0.099	0.027	13.691	1	0.000	0.906
Constant	5.043	1.379	13.364	1	0.000	154.922

จากผลการวิเคราะห์พบว่าปัจจัยที่ส่งผลต่อการเกิด VTDR อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ถูกคัดเลือกเข้ามาทั้งสิ้น 4 ตัวแปร ประกอบด้วย อายุ และเพศ มีสัมประสิทธิ์ logistic เป็นลบ อายุมากขึ้นส่งผลให้ภาวะ VTDR ลดลง และเพศหญิงมีโอกาสเกิดภาวะ VTDR น้อยกว่าเพศชาย (sex=0 ผู้ชาย sex=1 ผู้หญิง)

ส่วนตัวแปรผลทางพยาธิวิทยาได้แก่ คอเลสเตอรอลรวม (Chol) โดยมีความสัมพันธ์ทางบวก ค่าคอเรสเตอรอลที่สูงขึ้นส่งผลให้โอกาสของการเกิดภาวะ VTDR ที่สูงขึ้น และ HCT มีค่าสัมประสิทธิ์เป็นลบแสดงถึงค่า HCT ที่ต่ำมีโอกาสในการเกิดภาวะ VTDR ที่มากขึ้น ดังแสดงใน ตารางที่ 2

ตารางที่ 2 แสดงความแม่นยำในการพยากรณ์ของตัวแบบ logistic regression in sample test และ out of sample test

ภาวะผู้ป่วย	การพยากรณ์ logistic regression (in sample test)		
	ไม่เป็น VTDR	เป็น VTDR	ความถูกต้อง (%)
ไม่เป็น VTDR	39	73	34.80
เป็น VTDR	29	184	86.40
รวม			68.60
ภาวะผู้ป่วย	การพยากรณ์ logistic regression (out of sample test)		
	ไม่เป็น VTDR	เป็น VTDR	ความถูกต้อง (%)
ไม่เป็น VTDR 5 7 41.70	5	7	41.70
เป็น VTDR 1 17 94.40	1	17	94.40
รวม			73.30

จากตารางที่ 2 พบว่าในส่วนของ in sample test ตัวแบบ logistic regression สามารถพยากรณ์ภาวะของผู้ป่วยที่ไม่เป็น VTDR ได้ถูกต้องเพียง 34.80% และภาวะของผู้ป่วยที่เป็น VTDR ถูกต้อง 86.40% โดยมีความแม่นยำของการพยากรณ์โดยรวม 68.60% ในส่วน out of sample test สามารถพยากรณ์ภาวะของผู้ป่วยที่ไม่เป็น VTDR ได้ถูกต้องเพียง 41.70% และภาวะของผู้ป่วยที่เป็น VTDR

ถูกต้อง 94.40% โดยมีความแม่นยำของการพยากรณ์โดยรวม 69.01%

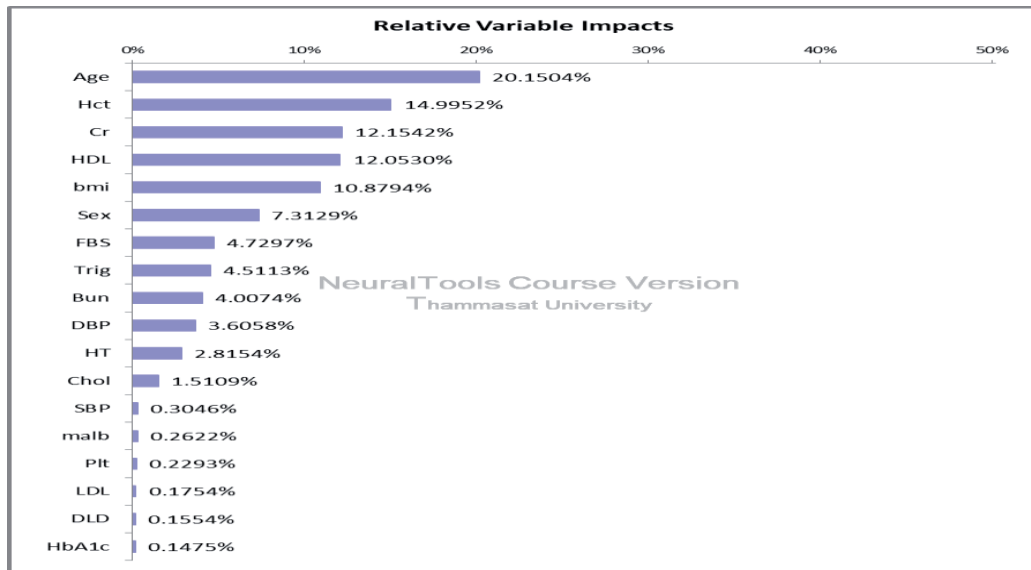
ตัวแบบ PNN ผลการ train โครงข่าย PNN และทดสอบการพยากรณ์ ได้ผลลัพธ์ในรูปแบบของ classification matrix ซึ่งแสดงความแม่นยำในการพยากรณ์ของตัวแบบ PNN ในกรณีของ in sample test และ out of sample test ดังแสดงในตารางที่ 3

ตารางที่ 3 แสดงความแม่นยำในการพยากรณ์ของตัวแบบ PNN ในกรณีของ in sample test และ out of sample test

ภาวะผู้ป่วย	การพยากรณ์ PNN (training)		
	ไม่เป็น VTDR	เป็น VTDR	ความถูกต้อง (%)
ไม่เป็น VTDR	109	0	100.00
เป็น VTDR	0	210	100.00
รวม			100.00
ภาวะผู้ป่วย	การพยากรณ์ PNN (testing)		
	ไม่เป็น VTDR	เป็น VTDR	ความถูกต้อง (%)
ไม่เป็น VTDR	8	7	53.33
เป็น VTDR	4	17	80.95
รวม			69.44

จากตารางที่ 3 พบว่าในส่วนของ in sample test ตัวแบบ PNN สามารถพยากรณ์ผู้ป่วยที่ไม่มีภาวะ VTDR ได้ถูกต้อง 100 % และผู้ป่วยที่มีภาวะ VTDR ได้ 100% โดยมีความแม่นยำของการพยากรณ์โดยรวมสูงถึง 100% ในส่วน out of sample test ตัวแบบมีความแม่นยำในการจำแนกผู้ป่วยที่ไม่มีภาวะ VTDR ได้ถูกต้อง 53.33%

และผู้ป่วยที่มีภาวะ VTDR ได้ 80.95% โดยมีความแม่นยำของการพยากรณ์โดยรวม 69.44% ซึ่งถ้ารวมทั้ง 2 กรณีส่งผลให้ตัวแบบ PNN มีความแม่นยำโดยรวมอยู่ที่ 96.90% ภาพที่ 3 แสดงลำดับของความสัมพันธ์ของตัวแปรอิสระต่อการจำแนกผู้ป่วยที่มีภาวะ VTDR ได้โดยตัวแบบ PNN



ภาพที่ 3 แสดงลำดับของความสัมพันธ์ของตัวแปรอิสระต่อการจำแนกผู้ป่วยที่มีภาวะ VTDR จากตัวแบบ

จากภาพที่ 3 พบว่าตัวแปรที่มีผลต่อการจำแนกภาวะโดยรวมตามความสำคัญโดยใช้หลัก ABC analysis ที่ 70:20:10 ได้แก่ กลุ่ม A ประกอบไปด้วยตัวแปร Age HCT Cr HDL และ BMI มีผลกระทบ 70.22% กลุ่ม B ประกอบไปด้วย Sex FBS Trig และ Bun มีผลกระทบ

รวม 20.56% และกลุ่ม C ประกอบไปด้วย DBP HT Chol SBP mAlb PLT LDL DLD และ HbA1c มีผลกระทบรวม 9.21% ตามลำดับ ทั้งนี้ ผลลัพธ์จาก relative variable impact ไม่ได้สะท้อนทิศทางของความสัมพันธ์ของตัวแปรอิสระต่อโอกาสที่ผู้ป่วยจะมีภาวะ VTDR ได้

วิจารณ์

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยต่าง ๆ ที่มีความสัมพันธ์ต่อการเกิดภาวะ VTDR ในผู้ป่วยและพัฒนาตัวแบบเพื่อพยากรณ์ภาวะ VTDR โดยใช้เทคนิค logistic regression และเทคนิค PNN ที่นิยมใช้กันอย่างแพร่หลายทั้งในประเทศและต่างประเทศสำหรับกรณีที่ตัวแปรตามมีลักษณะแบ่งกลุ่ม จากการวิเคราะห์ที่ตัวแปรที่มีความสัมพันธ์กับภาวะ VTDR ในผู้ป่วยของทั้ง 2 ตัวแบบ ปรากฏว่าตัวแบบ logistic regression มีตัวแปรสำคัญที่ถูกคัดเลือกเข้ามาไว้ในตัวแบบได้แก่ อายุ เพศ chol HCT โดยอายุมีสัมประสิทธิ์ logistic เป็นลบ อายุมากขึ้นส่งผลให้ภาวะ VTDR ลดลง การศึกษาส่วนใหญ่

พบว่าอายุผู้ป่วยที่มากขึ้นเป็นปัจจัยเสี่ยงต่อการเกิด DR แต่ก็มีงานวิจัยที่พบว่าอายุกับโรคเบาหวานจอประสาทตาเป็นไปในทิศทางกลับกันซึ่งอาจเกี่ยวข้องกับกลุ่มผู้ป่วยที่ศึกษาและในเชื้อชาติที่ต่างกัน มีงานวิจัยพบว่าผู้ป่วยอายุมากกว่า 65 ปี ความเสี่ยงที่จะเกิด DR ลดลง นอกจากนี้ผู้ป่วยเบาหวานชนิด 2 หากเริ่มเป็นเบาหวานในอายุน้อยจะมีความเสี่ยงที่จะเกิด DR และมีความรุนแรงของโรคมากด้วย ในการศึกษาที่พบว่าอายุผู้ป่วยกับการเกิด VTDR มีความสัมพันธ์ในทิศทางกลับกัน ซึ่งอาจจะเป็นเพราะปัจจัยช่วงอายุหรืออายุเริ่มเป็นโรคในกลุ่มผู้ป่วยที่ศึกษาหรือเชื้อชาติ ซึ่งน่าจะได้มีการศึกษาเพิ่มเติมในระยะต่อไป

เพศหญิงมีความเสี่ยงเกิดภาวะ DR น้อยกว่าเพศชาย ซึ่งสอดคล้องกับงานวิจัยส่วนใหญ่ กลไกที่เกี่ยวข้องไม่ชัดเจน น่าจะเป็นบทบาทของฮอร์โมนโมนทางเพศ (sex-hormone-related pathway)

ส่วนตัวแปรคอเลสเตอรอล (Chol) ที่สูงขึ้นส่งผลให้โอกาสเกิดภาวะ VTDR สูงขึ้น การศึกษาเกี่ยวกับค่าไขมันในเลือดที่ผิดปกติกับการเป็น DR ยังไม่มีความสอดคล้องในทางเดียวกันทั้งหมด อย่างไรก็ตามในปัจจุบันพบว่าความผิดปกติของการเผาผลาญไขมัน (lipid metabolism abnormalities) มีบทบาทสำคัญในพยาธิกำเนิด DR และพบวาระดับน้ำตาลในเลือดสูงร่วมกับระดับไขมันในเลือดที่ผิดปกติมีผลต่อการเกิด DR และต่อความรุนแรงของโรค เพราะทำให้เกิดการอักเสบระดับต่ำเรื้อรัง (chronic low-grade inflammation) ที่เกี่ยวข้องกับDR ไขมันมีกลไกความผิดปกติที่หลากหลาย เช่น complex molecular signaling event ความดื้อต่ออินซูลิน (insulin resistance) ทำให้มีการเผาผลาญไขมันผิดปกติด้วย ในการศึกษาพบว่า chol และ HDL เป็นปัจจัยเสี่ยงของ VTDR ซึ่งเป็นประเด็นที่น่าสนใจเพื่อให้ความสำคัญกับปัจจัยนี้ในการดูแลผู้ป่วย ค่าHCT มีค่าสัมพันธ์เป็นลบ ซึ่งสอดคล้องกับการวิจัยอื่น ๆ และการรักษาภาวะซีดช่วยทำให้ชะลอความรุนแรงของ DR ด้วยโดยสมมติฐานว่าน่าจะเป็นสาเหตุจากทำลายอวัยวะโดยตรง และความหนืดของเลือดที่ลดลงทำให้ลดแรงที่กระทำต่อเซลล์ผนังเส้นเลือด (anti-atherogenic functions of endothelial cells) ยับยั้งการสร้างสารต่อต้านการทําหน้าที่แข็งตัวของเส้นเลือด

สำหรับความแม่นยำในการพยากรณ์พบว่าตัวแบบมีความแม่นยำในการพยากรณ์ภาวะการเกิด VTDR ในผู้ป่วยได้ถูกต้องแม่นยำสูงเท่ากับ 86.40% และ 94.40% ในกลุ่มของข้อมูล in sample test และ out of sample test ตามลำดับ แต่ตัวแบบมีการพยากรณ์ผิดพลาดในกรณีผู้ป่วยไม่มีภาวะ VTDR ว่ามีภาวะ VTDR ค่อนข้างสูง โดยสามารถพยากรณ์ภาวะของผู้ป่วยที่ไม่

เป็น VTDR ได้ถูกต้องเพียง 34.80% และ 41.70% ในกลุ่มของข้อมูล in sample test และ out of sample test ตามลำดับ อย่างไรก็ตามในมุมมองของการรักษาเชิงป้องกันการวินิจฉัยผู้ป่วยที่ไม่มีภาวะ VTDR ว่ามีอาการทำให้บุคลากรทางการแพทย์สามารถมีความระมัดระวังและให้การรักษาได้อย่างทันท่วงที ทำให้โอกาสเกิดภาวะ VTDR ในอนาคตของผู้ป่วยลดลง สำหรับตัวแบบ PNN ตัวแปรสำคัญที่มีผลต่อการเกิด VTDR ซึ่งประเมินจากรelative variable impact คิดเป็น 70.22% ของ impact ทั้งหมด ได้แก่อายุผู้ป่วย ค่า HCT ค่า creatinine ค่า HDL และ BMI ตามลำดับ ซึ่งผลจากการวิเคราะห์ทั้ง 2 แบบพบว่าปัจจัยสำคัญตรงกัน 2 ตัวแปร คืออายุผู้ป่วยและค่า HCT

ปัจจัยเสี่ยงที่มีความสำคัญในตัวแบบพยากรณ์ PNN อายุ HCT creatinine HDL และ BMI สอดคล้องกับการศึกษาอื่น ๆ ค่า creatinine ที่สูง มีความสัมพันธ์กับการเกิด PDR และ DME และการรักษาภาวะไตเสื่อมป้องกันไม่ให้ DR ลุกลามมากขึ้น BMI เป็นปัจจัยเสี่ยงต่อ DR ยังไม่มีข้อสรุปที่ชัดเจน แต่เป็นปัจจัยเสี่ยงต่อการลุกลามของโรคเบาหวาน นอกจากนั้นการเพิ่มขึ้นของ BMI มีความสัมพันธ์อย่างมีนัยสำคัญกับ HbA1c cholesterol และความดันโลหิตสูงที่แย่ง ซึ่งบ่งบอกถึง DR ที่รุนแรงมากขึ้นด้วย โรคเบาหวานมีปัจจัยเกี่ยวข้องด้วยหลายอย่าง (multi-factorial) กลไกการเกิด DR เป็น microvascular complication ซึ่งซับซ้อน ปัจจัยเสี่ยงที่สำคัญได้แก่ค่า HbA1c แต่งานวิจัยนี้ไม่ปรากฏว่า HbA1c เป็นปัจจัยเสี่ยงที่ชัดเจน ซึ่งอาจมีเหตุผลจากลักษณะกลุ่มผู้ป่วยที่ศึกษาควรน่าจะได้มีการศึกษาประเด็นนี้ให้ชัดเจนขึ้นในอนาคตต่อไป

งานวิจัยนี้นับเป็นจุดเริ่มต้นของการนำตัวแบบพยากรณ์มาช่วยในกระบวนการวินิจฉัยภาวะ VTDR ที่จะเกิดในผู้ป่วยเพื่อได้วางแผนขั้นตอนการรักษาที่เหมาะสมตั้งแต่ช่วงต้นสำหรับโรงพยาบาลนครปฐม ดังนั้นงานวิจัยนี้จึงมีข้อจำกัดในหลายส่วน โดยส่วนแรกจากขนาด

ของกลุ่มตัวอย่างที่ไม่ใหญ่มาก ทั้งนี้เนื่องมาจากฐานข้อมูลของผู้ป่วยที่ไม่ครบถ้วนในแต่ละตัวแปรทำให้เกิด missing values เป็นจำนวนมากส่งผลให้มีการตัดข้อมูลเหล่านี้่ออกจากการพิจารณาไป ดังนั้นเพื่อการวิจัยในอนาคตที่มีความน่าเชื่อถือมากขึ้นควรมีฐานข้อมูลของผู้ป่วยที่มีความสมบูรณ์นอกจากนี้อาจมีการขยายขอบเขตของงานวิจัยให้ครอบคลุมข้อมูลของผู้ป่วยในภูมิภาคอื่น ๆ โดยอาจมีการวิจัยในเชิงเปรียบเทียบหรือวิจัยในภาพรวมของประเทศ ในส่วนของตัวแบบพยากรณ์นั้น อาจมีการใช้เทคนิคอื่น ๆ เพิ่มเติมเช่นตัวแบบทางด้าน machine learning หรืออื่น ๆ ก็อาจจะทำให้ได้ตัวแบบมีความแม่นยำในการพยากรณ์มากยิ่งขึ้น นอกจากนี้การที่ผู้วิจัยไม่นำตัวแปรอิสระจากการทบทวนวรรณกรรมและงานวิจัยก่อนหน้ามาพิจารณาทั้งหมด จึงอาจมีการเพิ่มตัวแปรอิสระที่เกี่ยวข้อง เพื่อวิเคราะห์ตัวแปรที่ส่งผลต่อการเกิด VTDR ได้ชัดเจนมากขึ้นทำให้เกิดความแม่นยำของตัวแบบพยากรณ์และทำให้งานวิจัยมีความสมบูรณ์มากยิ่งขึ้น

เอกสารอ้างอิง

1. Wanyue Li, Yanan Song, Kang Chen, Jun Ying, Zhong Zheng, Shen Qiao, et al. Predictive model and risk analysis for diabetic retinopathy using machine learning: A retrospective cohort study in China. *BMJ Open* 2021; 11: e050989.
2. Ashish Bora, Siva Balasubramanian, Boris Babenko, Sunny Virmani, Subhashini Venugopalan, Akinori Mitani, et al. Predicting the risk of developing diabetic retinopathy using deep learning. *The Lancet Digit Health* 2021; 3: e10-19.
3. Hair et al (2006). *Multivariate Data Analysis*. Sixth Edition. Prentice Hall.
4. Specht, D. F. Probabilistic neural networks. *Neural Networks*. 1990; 3: 109-18
5. Berno E, Brambilla L, Canaparo R, Casale F, Costa M, Della Pepa C, et al. Application of probabilistic neural networks to population pharmacokinetics. *Proceedings of the International Joint Conference on Neural Networks 2003*: 2637-42.
6. Fernandez-Gamez M.A, Cisneros-Ruiz A.J., Callejon-Gil A. Apply a Probabilistic Neural Network to Hotel Bankruptcy Prediction. *Tourism & Management Studies*. 2016; 12(1): 40-52.
7. Huang Chenn-Jung, Liao Wei-Chen. Application of Probabilistic Neural Networks to the Class Prediction of Leukemia and Embryonal Tumor of Central Nervous System. *Neural Processing Letters*. 2004; 19(3): 211-26.
8. 17. Li, Q. B., Li, X., Zhang, G. J., Xu, Y. Z., Wu, J. G., Sun, X. J. Application of Probabilistic Neural Networks Method to Gastric Endoscope Samples Diagnosis Based on FTIR Spectroscopy. *Spectroscopy and Spectral Analysis* 2009; 29: 1553-7