

Trong phương trình hồi quy đa biến, một số yếu tố được xác định có liên quan chặt chẽ với bí tiểu. Nam giới có nguy cơ bí tiểu cao hơn nữ giới (OR = 36,47; 95% CI: 11,38–116,96). Tiểu đường làm tăng đáng kể nguy cơ bí tiểu với OR = 12,06; 95% CI: 2,76–52,69. Bệnh lý tiết niệu làm tăng nguy cơ bí tiểu 4,25 lần (OR = 4,25; 95% CI: 1,02–17,72). Phẫu thuật chấn thương làm tăng nguy cơ 5,02 lần (OR = 5,02; 95% CI: 1,65–15,24). Những người bệnh được gây tê tùy sống có nguy cơ bí tiểu cao hơn những người được gây mê 3,02 lần (OR = 3,02; 95% CI: 1,01–8,08). Các yếu tố khác như nhóm tuổi, tăng huyết áp và phẫu thuật tiết niệu nam học không liên quan chặt chẽ với nguy cơ bí tiểu.

IV. BÀN LUẬN

Nghiên cứu của chúng tôi thực hiện trên 384 người bệnh phẫu thuật ngoại khoa tại Trung tâm Y tế khu vực Cẩm Khê ghi nhận tỷ lệ người bệnh có triệu chứng bí tiểu sau phẫu thuật (POUR) là 16,9%. Trong đó, tỷ lệ người bệnh được chẩn đoán bí tiểu cần can thiệp đặt thông tiểu là 8,9%.

Tỷ lệ 8,9% của chúng tôi cao hơn so với một số nghiên cứu quốc tế trước đây, như nghiên cứu của Stallar và Prescott (6%) [4] hay Tammela T. (3,8%) [3]. Tuy nhiên, kết quả này lại thấp hơn đáng kể so với báo cáo của Lamonerie và cộng sự (44%) [5]. Sự khác biệt lớn này có thể được giải thích chủ yếu là do phương pháp chẩn đoán. Nghiên cứu của Lamonerie sử dụng siêu âm bàng quang, một phương pháp có độ chính xác cao, phát hiện tới 54% người bệnh bí tiểu không có triệu chứng lâm sàng [5]. Trong khi đó, tại Trung tâm Y tế khu vực Cẩm Khê, do chưa được trang bị máy siêu âm bàng quang di động thường quy tại phòng hậu phẫu, nghiên cứu của chúng tôi chẩn đoán POUR chủ yếu dựa trên các triệu chứng lâm sàng (như đau tức vùng trên khớp mu, buồn tiểu không tiểu được) và quyết định can thiệp đặt thông tiểu. Phương pháp này có thể đã bỏ sót các trường hợp người bệnh không xuất hiện triệu chứng hoặc các triệu chứng bị che lấp bởi thuốc gây tê, thuốc giảm đau.

Tuy nhiên, so với các nghiên cứu trong nước về tình trạng bí tiểu sau phẫu thuật không sử dụng phương pháp siêu âm bàng quang để chẩn đoán POUR thì tỷ lệ 8,9% người bệnh phải đặt thông tiểu của chúng tôi là cao hơn với nghiên cứu tại Bệnh viện Việt Đức (5,8%) [8]. Tỷ lệ này là phù hợp với đặc thù của một trung tâm y tế tuyến khu vực, nơi tiếp nhận đa dạng các loại phẫu thuật và có tỷ lệ người bệnh nhóm tuổi

≥40 cao (79,4%), tỷ lệ nam giới cao (61,2%) và nhiều bệnh lý nền (70,9%), vốn là các yếu tố làm tăng nguy cơ bí tiểu.

Phân tích các yếu tố liên quan cho thấy, giới tính nam là yếu tố nguy cơ mạnh nhất đối với bí tiểu sau phẫu thuật ngoại khoa (OR = 36,47). Điều này phù hợp với các nghiên cứu trước đây, cho thấy nam giới dễ gặp bí tiểu do cấu trúc giải phẫu đường tiết niệu và tỷ lệ phì đại tuyến tiền liệt cao hơn [6]. Bệnh lý nền cũng đóng vai trò quan trọng. Tiểu đường làm tăng nguy cơ xuất hiện bí tiểu 12,6 lần (OR = 12,6; 95% CI: 2,76–52,69), trong khi các bệnh lý tiết niệu làm tăng nguy cơ bí tiểu 4,25 lần (OR = 4,25; 95% CI: 1,02–17,72). Các kết quả này phù hợp với cơ chế sinh lý, khi bệnh lý nền làm giảm trương lực bàng quang, ảnh hưởng dẫn truyền thần kinh và làm tăng nguy cơ bí tiểu sau phẫu thuật [6, 7]. Phẫu thuật chấn thương làm tăng nguy cơ bí tiểu 5,02 lần (OR = 5,02; 95% CI: 1,65–15,24). Những người bệnh được gây tê tùy sống có nguy cơ bí tiểu cao hơn những người được gây mê 3,02 lần (OR = 3,02; 95% CI: 1,01–8,08). Phẫu thuật chấn thương cũng được xác định là yếu tố nguy cơ trong nhiều nghiên cứu, nguyên nhân là do tình trạng bất động trước và sau phẫu thuật [8]. Do đó, "bàng quang chứa nước tiểu ở tư thế nằm nghiêng, cản trở việc làm trống bàng quang do áp lực dẫn tới ứ nước ở bàng quang, bể thận" [7]. Ngược lại, phương pháp vô cảm gây mê được xác định có tác dụng bảo vệ, với nguy cơ bí tiểu thấp hơn đáng kể so với gây tê tùy sống, phản ánh ảnh hưởng của tê tùy sống lên thần kinh vận động và cảm giác bàng quang [7].

V. KẾT LUẬN

Bí tiểu sau phẫu thuật ngoại khoa là một biến chứng phổ biến (16,9% có triệu chứng bí tiểu, 8,9% cần can thiệp đặt thông tiểu) tại Trung tâm Y tế khu vực Cẩm Khê. Giới tính nam, bệnh lý nền (tiết niệu, tiểu đường), phẫu thuật chấn thương và gây tê tùy sống là các yếu tố nguy cơ chính làm gia tăng tỷ lệ bí tiểu sau phẫu thuật. Cần sàng lọc người bệnh có nguy cơ cao, theo dõi sớm sau mổ và đầu tư trang thiết bị siêu âm bàng quang di động để phát hiện bí tiểu không triệu chứng nhằm nâng cao chất lượng chăm sóc người bệnh.

TÀI LIỆU THAM KHẢO

1. Abram P, Cardozo L, Fall M, et al. The standardisation of terminology of lower urinary tract function: Report from the Standardisation Sub-committee of the International Continence Society. *Neurourol Urodyn*. 2002;21(2):167-78.

2. Agrawal K, Majhi S, Garg R. Post-operative urinary retention: Review of literature. World J Anesthesiol. 2019;8(1):1-12.
3. Tammela T, Kontturi M, Lukkarinen O. Postoperative urinary retention. I. Incidence and predisposing factors. Scand J Urol Nephrol. 1986;20(3):197-201.
4. Stallard S, Prescott S. Postoperative urinary retention in general surgical patients. Br J Surg. 1988;75(11):1141-3.
5. Lamonerie L, Chartier-Kastler E, MOMS, et al. Prevalence of postoperative bladder distension and urinary retention detected by ultrasound measurement. Br J Anaesth. 2004;92(4):544-6.
6. Darrah DM, Griebing TL, Silverstein JH. Postoperative urinary retention. Anesthesiol Clin. 2009;27(3):465-84.
7. Keita H, Diouf E, Fssatou S, et al. Predictive factors of early postoperative urinary retention in the postanesthesia care unit. Anesth Analg. 2005;101(2):592-6.
8. Nguyễn Bá Anh, và cộng sự. Đánh giá tình trạng bí tiểu của bệnh nhân sau phẫu thuật và một số yếu tố liên quan tại khoa Điều trị theo yêu cầu, Bệnh viện Việt Đức. Điều dưỡng Việt Nam. 2019.
9. Nguyễn Duy Hưng, Nguyễn Xuân Hùng, Đỗ Huy Tùng, & Lương Quang Tuyên. (2024). Kết quả phẫu thuật nội soi điều trị tăng sinh lành tính tuyến tiền liệt bằng dao điện lưỡng cực tại Bệnh viện Trung ương Thái Nguyên. Tạp chí Y học Việt Nam, 536(1B).

GIÁ TRỊ TIÊN LƯỢNG TỬ VONG CỦA MÔ HÌNH MÁY HỌC RỪNG NGẪU NHIÊN Ở BỆNH NHÂN NHỒI MÁU CƠ TIM CẤP

Trần Công Duy*, Lý Quang Sang*, Mai Quốc Thành**, Nguyễn Đình Quốc Anh*, Thái Thanh Trúc*, Hoàng Văn Sỹ*

TÓM TẮT

Mục tiêu: Đánh giá giá trị tiên lượng tử vong của mô hình máy học rừng ngẫu nhiên ở bệnh nhân nhồi máu cơ tim (NMCT) cấp. **Đối tượng và phương pháp:** Nghiên cứu đoàn hệ hồi cứu trên 540 bệnh nhân NMCT cấp điều trị tại Bệnh viện Chợ Rẫy từ tháng 01/2020 đến 09/2021. Các dữ liệu lâm sàng, cận lâm sàng, tổn thương động mạch vành và điều trị được thu thập để huấn luyện và kiểm tra mô hình rừng ngẫu nhiên bằng Python 3.12. Dữ liệu được chuẩn hóa, cân bằng mẫu bằng phương pháp SMOTE và phân tích thành phần chính nhằm xác định các yếu tố quan trọng. Hiệu năng mô hình được đánh giá bằng diện tích dưới đường cong ROC (AUC) và độ chính xác. **Kết quả:** Tuổi trung bình của dân số nghiên cứu là $64,0 \pm 11,6$; nam giới chiếm 71,3%. Tỷ lệ tử vong trong vòng 12 tháng là 10,7%. Mười yếu tố quan trọng nhất trong mô hình tiên lượng tử vong gồm phân độ Killip, thể lâm sàng NMCT, tổn thương động mạch vành phải và động mạch xuống trước trái, hút thuốc lá, sử dụng thuốc ức chế men chuyển/chen thụ thể angiotensin II (ACEi/ARB), thiếu máu, nồng độ troponin I lúc nhập viện, tuổi và điểm Gensini. Trên bộ dữ liệu ban đầu, mô hình đạt AUC từ 0,653 đến 0,730 và độ chính xác từ 0,889 đến 0,907. Sau khi chuẩn hóa và cân bằng mẫu, AUC tăng lên 0,709–0,730. **Kết luận:** Mô hình máy học rừng ngẫu nhiên có khả năng

tiên lượng tử vong ở bệnh nhân NMCT cấp với hiệu năng ở mức khá. **Từ khóa:** nhồi máu cơ tim cấp, tiên lượng tử vong, mô hình máy học, rừng ngẫu nhiên

SUMMARY

PROGNOSTIC VALUE OF A RANDOM FOREST MACHINE LEARNING MODEL FOR MORTALITY IN PATIENTS WITH ACUTE MYOCARDIAL INFARCTION

Objectives: To evaluate the prognostic value of a random forest machine learning model for predicting mortality in patients with acute myocardial infarction (AMI). **Subjects and methods:** This retrospective cohort study included 540 patients with AMI treated at Cho Ray Hospital from January 2020 to September 2021. Clinical, laboratory, coronary angiographic, and treatment data were used to train a random forest model using Python 3.12. Data were standardized, class imbalance was addressed with the SMOTE method, and principal component analysis was applied to identify key predictors. Model performance was assessed using the area under the receiver operating characteristic curve (AUC) and accuracy. **Results:** The mean age was 64.0 ± 11.6 years, and 71.3% were male. The 12-month mortality rate was 10.7%. The most important predictors of mortality included Killip class, clinical type of AMI, lesions of the right coronary artery and the left anterior descending artery, smoking, use of angiotensin-converting enzyme inhibitors/angiotensin II receptor blockers, anemia, admission troponin I level, age, and Gensini score. In the original dataset, the model achieved AUC values ranging from 0.653 to 0.730, with accuracy between 0.889 and 0.907. After data standardization and class balancing, AUC improved to 0.709–0.730.

*Bộ môn Nội Tổng Quát, Trường Y, Đại học Y Dược Thành Phố Hồ Chí Minh

**Trung tâm Tiêm Chủng VNNC Cần Giuộc, Tây Ninh

Chịu trách nhiệm chính: Trần Công Duy

Email: dr.trancongduy@ump.edu.vn

Ngày nhận bài: 9.1.2026

Ngày phản biện khoa học: 13.2.2026

Ngày duyệt bài: 5.3.2026

Conclusion: The random forest machine learning model demonstrated fair prognostic performance for mortality in patients with AMI. **Keywords:** acute myocardial infarction, prognostic value for mortality, machine learning, random forest

I. ĐẶT VẤN ĐỀ

Nhồi máu cơ tim (NMCT) cấp là nguyên nhân hàng đầu gây tử vong và tàn tật trên toàn cầu mặc dù đã có nhiều tiến bộ trong chẩn đoán sớm, tái tưới máu và điều trị nội khoa. Việc phân tầng nguy cơ chính xác ngay từ giai đoạn sớm có vai trò quan trọng trong định hướng chiến lược điều trị, theo dõi và tiên lượng cho người bệnh. Trong thực hành lâm sàng, các thang điểm nguy cơ truyền thống như GRACE và TIMI được sử dụng rộng rãi, tuy nhiên các mô hình này được xây dựng dựa trên phương pháp thống kê cổ điển với số lượng biến số hạn chế và giả định tuyến tính, do đó chưa phản ánh đầy đủ tính phức tạp và tương tác đa chiều giữa các yếu tố lâm sàng trong bối cảnh điều trị hiện đại [1].

Trong những năm gần đây, sự phát triển của trí tuệ nhân tạo đã mở ra hướng tiếp cận mới trong tiên lượng bệnh tim mạch. Các mô hình máy học cho phép xử lý khối lượng lớn dữ liệu, khai thác mối quan hệ phi tuyến tính giữa các biến số và đã được chứng minh có khả năng dự đoán tử vong sau NMCT cấp tốt hơn hoặc tương đương các thang điểm nguy cơ truyền thống trong nhiều nghiên cứu [2–4]. Bên cạnh đó, các tổng quan hệ thống và phân tích gộp cũng ghi nhận hiệu năng dự báo tốt của các mô hình máy học ở nhiều mốc thời gian theo dõi khác nhau trong NMCT cấp [5].

Trong số các thuật toán máy học, rừng ngẫu nhiên là phương pháp được sử dụng phổ biến nhờ khả năng kháng nhiễu, hạn chế nguy cơ quá khớp và cho phép đánh giá tầm quan trọng của các biến số. Các nghiên cứu ứng dụng rừng ngẫu nhiên trong tiên lượng tử vong sau NMCT cấp cho thấy hiệu năng dự báo tốt và có thể áp dụng sớm ngay từ thời điểm nhập viện [6-7]. Tuy nhiên, hiệu năng mô hình phụ thuộc đáng kể vào đặc điểm quần thể và chất lượng dữ liệu. Tại Việt Nam, nghiên cứu ứng dụng máy học trong tiên lượng tử vong NMCT cấp chưa được thực hiện. Do đó, mục tiêu của nghiên cứu này là huấn luyện và kiểm tra mô hình máy học rừng ngẫu nhiên dự báo tử vong do mọi nguyên nhân trong 12 tháng ở bệnh nhân NMCT cấp, từ đó đánh giá giá trị tiên lượng của mô hình máy học này.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Thiết kế nghiên cứu: đoàn hệ hồi cứu

Đối tượng nghiên cứu: Những bệnh nhân NMCT cấp nhập Khoa Nội Tim Mạch và Khoa Tim Mạch Can Thiệp, Bệnh viện Chợ Rẫy từ tháng 01/2020 đến 09/2021 được chọn vào nếu thỏa tiêu chuẩn: đủ 18 tuổi trở lên, được chẩn đoán xác định NMCT cấp theo định nghĩa toàn cầu lần thứ 4 [8], có đầy đủ dữ liệu để thu thập và theo dõi được tình trạng sống còn hay tử vong trong 12 tháng. Tiêu chuẩn loại trừ bao gồm tiền sử NMCT, can thiệp mạch vành qua da và phẫu thuật bắc cầu mạch vành và động mạch vành bình thường trên chụp động mạch vành cản quang.

Phương pháp nghiên cứu: Các bệnh nhân NMCT cấp thỏa tiêu chuẩn được thu thập dữ liệu, ghi nhận vào phiếu thu thập soạn sẵn thông tin về các đặc điểm nhân khẩu học, yếu tố nguy cơ bệnh mạch vành, các thông số lâm sàng, cận lâm sàng, đặc điểm tổn thương động mạch vành, phương pháp điều trị và tình trạng tử vong trong 12 tháng từ ngày nhập viện. Chúng tôi lựa chọn ngôn ngữ lập trình Python để nhập dữ liệu đầu vào cho mô hình, phân chia, huấn luyện và đánh giá kết quả các mô hình máy học rừng ngẫu nhiên. Bộ dữ liệu huấn luyện (70% dữ liệu) và kiểm tra (30% dữ liệu) được chuyển đổi dạng dữ liệu phù hợp với mô hình máy học bằng Python 3.12, các thư viện Scikit-learn, Imbalanced-learn và các công cụ của Pandas, Numpy, Matplotlib. Dữ liệu được chuẩn hóa và cân bằng mẫu bằng phương pháp SMOTE (Synthetic Minority Oversampling Technique). Phương pháp phân tích thành phần chính PCA (Principal Component Analysis) được sử dụng để tìm ra 10 yếu tố chính ảnh hưởng đến toàn bộ dữ liệu liên quan đến bệnh nhân có tử vong hoặc không tử vong. Các thuật toán được xây dựng gồm hàm phân nhánh Entropy, hàm Gini, hàm Log-loss, tính AUC và độ chính xác của mô hình rừng ngẫu nhiên.

III. KẾT QUẢ NGHIÊN CỨU

Nghiên cứu được thực hiện trên 540 bệnh nhân NMCT cấp. Tuổi trung bình của dân số nghiên cứu là $64,0 \pm 11,6$ năm, trong đó nam giới chiếm tỷ lệ ưu thế với 71,3%. Các yếu tố nguy cơ tim mạch thường gặp là tăng huyết áp (81,8%), rối loạn lipid máu (89,3%), hút thuốc lá (42,0%) và đái tháo đường (24,4%). Thể lâm sàng NMCT cấp ST chênh lên chiếm đa số với 63,7%, trong khi NMCT không ST chênh lên chiếm 36,3%. Phần lớn bệnh nhân nhập viện ở

phân độ Killip I (76,3%), tuy nhiên vẫn ghi nhận tỷ lệ đáng kể bệnh nhân ở các phân độ nặng hơn, bao gồm Killip III và IV lần lượt là 6,1% và 9,3% (Bảng 1).

Về đặc điểm cận lâm sàng, nồng độ troponin I lúc nhập viện tăng cao với giá trị trung vị 15,2 ng/mL. Thiếu máu được ghi nhận ở 38,1% bệnh nhân. Phân suất tổng máu thất trái trung vị là 46,5%, cho thấy tỷ lệ không nhỏ bệnh nhân có rối loạn chức năng tâm thu. Tổn thương động mạch vành thường gặp nhất là động mạch xuống trước trái (LAD) (89,3%), tiếp theo là động mạch vành phải (RCA) (71,7%) và động mạch mũ (55,7%). Tổn thương thân chung động mạch vành trái chiếm 10,0%. Mức độ tổn thương động mạch vành được phản ánh qua điểm Gensini trung vị là 34,3. Tỷ lệ tái thông mạch vành rất cao, đạt 94,4%. Hầu hết bệnh nhân được điều trị nội khoa theo khuyến cáo với aspirin (99,8%), thuốc ức chế P2Y12 (100,0%) và statin (98,5%), trong khi ACEi/ARB và chẹn beta lần lượt được sử dụng ở 90,2% và 75,6% bệnh nhân (Bảng 2).

Nghiên cứu ghi nhận có 58 bệnh nhân tử vong trong 12 tháng, chiếm tỉ lệ 10,7%. Phân tích thành phần chính tử bộ dữ liệu ban đầu cho thấy 10 yếu tố có mức độ đóng góp quan trọng nhất vào mô hình tiên lượng tử vong theo thứ tự giảm dần bao gồm phân độ Killip, thể lâm sàng NMCT, tổn thương RCA, tổn thương LAD, hút thuốc lá, sử dụng ACEi/ARB, tình trạng thiếu máu, nồng độ troponin I lúc nhập viện, tuổi và điểm Gensini. Trong đó, phân độ Killip là yếu tố có trọng số cao nhất, phản ánh vai trò trung tâm của mức độ suy tim cấp trong tiên lượng tử vong ở bệnh nhân NMCT cấp (Bảng 3).

Khi áp dụng mô hình rừng ngẫu nhiên trên bộ dữ liệu ban đầu, khả năng tiên lượng tử vong đạt mức trung bình với AUC lần lượt là 0,691 với hàm Entropy, 0,653 với hàm Gini và 0,730 với hàm Log-loss. Độ chính xác của mô hình ở bộ dữ liệu này dao động từ 0,889 đến 0,907, cao nhất ở mô hình sử dụng Log-loss (Bảng 4). Sau khi dữ liệu được chuẩn hóa và hiệu chỉnh mất cân bằng mẫu bằng phương pháp SMOTE, hiệu năng của mô hình rừng ngẫu nhiên được cải thiện rõ rệt. Giá trị AUC tăng lên 0,712 đối với hàm Entropy và 0,709 đối với hàm Gini, trong khi hàm Log-loss duy trì AUC ở mức 0,730. Mặc dù độ chính xác giảm nhẹ so với bộ dữ liệu ban đầu, các giá trị này vẫn duy trì ở mức cao, dao động từ 0,833 đến 0,858 (Bảng 5).

Bảng 1. Đặc điểm nhân khẩu và lâm sàng của dân số nghiên cứu

Biến số	Giá trị (n = 540)
Tuổi (năm)	64,0 ± 11,6
Giới tính	
Nam	385 (71,3%)
Nữ	155 (28,7%)
Tăng huyết áp	442 (81,9%)
Đái tháo đường	132 (24,4%)
Rối loạn lipid máu	482 (89,3%)
Béo phì	110 (20,4%)
Hút thuốc lá	227 (42,0%)
Tiền sử gia đình bệnh mạch vành sớm	36 (6,7%)
Thể lâm sàng NMCT cấp	
NMCT cấp ST chênh lên	344 (63,7%)
NMCT cấp không ST chênh lên	196 (36,3%)
Tần số tim (lần/phút)	80,0 (68,3 – 92,0)
Huyết áp tâm thu (mmHg)	120 (100 – 130)
Huyết áp tâm trương (mmHg)	70 (60 – 80)
Phân độ Killip	
Độ I	412 (76,3%)
Độ II	45 (8,3%)
Độ III	33 (6,1%)
Độ IV	50 (9,3%)

Bảng 2. Đặc điểm cận lâm sàng và điều trị của dân số nghiên cứu

Biến số	Giá trị (n = 540)
Glucose (mg/dL)	115 (97 – 149)
eGFR (ml/ph/1,73 m ²)	83,0 (64,0 – 94,5)
Troponin I lúc nhập viện (ng/mL)	15,2 (3,0 – 50,0)
Thiếu máu	206 (38,1%)
Phân suất tổng máu thất trái (%)	46,5 (39,0 – 53,0)
Vị trí tổn thương động mạch vành	
Thân chung động mạch vành trái	54 (10,0%)
Động mạch xuống trước trái	482 (89,3%)
Động mạch mũ	301 (55,7%)
Động mạch vành phải	387 (71,7%)
Điểm Gensini	34,3 (17,0 – 58,0)
Tái thông mạch vành	510 (94,4%)
Aspirin	539 (99,8%)
Ức chế P2Y12	540 (100,0%)
Statin	532 (98,5%)
ACEi/ARB	487 (90,2%)
Chẹn beta	408 (75,6%)
Kháng aldosterone	189 (35,0%)

Bảng 3. Mười yếu tố quan trọng nhất trong phân tích thành phần chính

Yếu tố	Trọng số
Phân độ Killip	4,531
Thể lâm sàng NMCT cấp	4,471
RCA	4,471
LAD	4,466
Hút thuốc lá	4,464
ACEi/ARB	4,452
Thiếu máu	4,407
Troponin I lúc nhập viện	4,406
Tuổi	4,390
Điểm Gensini	4,338

Bảng 4. Giá trị của mô hình tiên lượng tử vong của bệnh nhân NMCT cấp ở bộ dữ liệu ban đầu

Mô hình	AUC	Độ chính xác
Entropy	0,691	0,901
Gini	0,653	0,889
Log-loss	0,730	0,907

Bảng 5. Giá trị của mô hình tiên lượng tử vong của bệnh nhân NMCT cấp ở bộ dữ liệu chuẩn hóa và hiệu chỉnh cân bằng mẫu

Mô hình	AUC	Độ chính xác
Entropy	0,712	0,858
Gini	0,709	0,833
Log-loss	0,730	0,846

IV. BÀN LUẬN

Chúng tôi xây dựng và đánh giá mô hình máy học rừng ngẫu nhiên nhằm tiên lượng tử vong ở bệnh nhân NMCT cấp dựa trên dữ liệu lâm sàng, cận lâm sàng và điều trị thu thập trong thực hành thường quy. Kết quả cho thấy mô hình rừng ngẫu nhiên đạt khả năng phân biệt ở mức khá, với giá trị AUC dao động từ 0,653 đến 0,730 tùy theo thuật toán và phương pháp xử lý dữ liệu. Đáng chú ý, sau khi chuẩn hóa dữ liệu và hiệu chỉnh mất cân bằng mẫu bằng phương pháp SMOTE, hiệu năng mô hình được cải thiện rõ rệt, đặc biệt về khả năng phân biệt (AUC tăng lên khoảng 0,71–0,73), cho thấy vai trò quan trọng của xử lý dữ liệu tiền mô hình trong các nghiên cứu máy học y sinh.

Trong nhiều thập kỷ qua, các thang điểm nguy cơ như GRACE và TIMI được sử dụng rộng rãi để tiên lượng tử vong ở bệnh nhân NMCT cấp. Tuy nhiên, các mô hình này được xây dựng dựa trên phương pháp thống kê truyền thống với

giả định tuyến tính và số lượng biến đầu vào hạn chế, do đó khó phản ánh đầy đủ tính phức tạp và tương tác phi tuyến giữa các yếu tố lâm sàng. Tổng quan hệ thống của Cho SM và cộng sự (cs) cho thấy các mô hình máy học nhìn chung có khả năng dự báo tử vong và tái nhập viện sau NMCT tốt hơn hoặc ít nhất tương đương các mô hình thống kê cổ điển, đặc biệt trong các bộ dữ liệu lớn và đa chiều [1]. Kết quả nghiên cứu của chúng tôi phù hợp với xu hướng này khi mô hình rừng ngẫu nhiên khai thác đồng thời nhiều biến lâm sàng, cận lâm sàng và điều trị để xây dựng mô hình tiên lượng linh hoạt hơn so với các thang điểm truyền thống.

Các nghiên cứu quốc tế quy mô lớn đã chứng minh giá trị vượt trội của mô hình máy học trong tiên lượng tử vong NMCT. Nghiên cứu của Lee HC và cs trên cơ sở dữ liệu KAMIR ở Hàn Quốc với hơn 22.000 bệnh nhân cho thấy mô hình máy học dự báo tử vong sau 1 năm đạt AUC lên tới 0,918, cao hơn đáng kể so với các mô hình dựa trên thang điểm nguy cơ truyền thống [2]. Tương tự, nghiên cứu của Yang J và cs tại Trung Quốc sử dụng mô hình XGBoost để dự đoán tử vong nội viện ở bệnh nhân NMCT cấp ST chênh lên ghi nhận AUC khoảng 0,90 và vượt trội rõ rệt so với GRACE và TIMI [3]. So với các nghiên cứu trên, giá trị AUC trong nghiên cứu của chúng tôi thấp hơn. Sự khác biệt này có thể được lý giải bởi nhiều yếu tố bao gồm cỡ mẫu nhỏ hơn và tính không đồng nhất của dân số nghiên cứu (bao gồm cả NMCT cấp ST chênh lên và NMCT cấp không ST chênh lên).

Rừng ngẫu nhiên là một trong những mô hình máy học phổ biến trong lĩnh vực tim mạch nhờ khả năng xử lý dữ liệu phi tuyến, kháng nhiễu tốt và hạn chế quá khớp. Nghiên cứu của Hadanny A và cs đã chứng minh mô hình rừng ngẫu nhiên có thể dự đoán tử vong 30 ngày sau NMCT cấp ST chênh lên với AUC khoảng 0,80 và vượt trội so với thang điểm GRACE trong bộ dữ liệu độc lập [6]. Trong nghiên cứu của chúng tôi, phân tích thành phần chính cho thấy các yếu tố đóng góp quan trọng nhất vào mô hình bao gồm phân độ Killip, thể lâm sàng NMCT cấp, tổn thương động mạch vành (LAD, RCA), thiếu máu, nồng độ troponin I lúc nhập viện, tuổi và điểm Gensini. Các yếu tố này phù hợp với y văn khi nhiều nghiên cứu quốc tế và phân tích gộp đã xác nhận phân độ Killip, tuổi, rối loạn huyết động, suy chức năng thất trái và gánh nặng tổn thương mạch vành là những yếu tố tiên lượng tử vong mạnh nhất ở bệnh nhân NMCT cấp [5].