



Original Article

Modeling the Number of Obstructed Coronary Arteries in Adults Aged 50 Years and Older: A Comparison of Poisson, Negative Binomial, Zero-Inflated, and Hurdle Models

Sedighe Rastaghi¹, Mohammad Tajfard^{2,3}, Ali Taghipour², Azadeh Saki^{2*} 

¹ Department of Acupuncture, School of Persian and Complementary Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

² Social Determinants of Health Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

³ Department of Health Education and Health Promotion, Faculty of Health, Mashhad University of Medical Sciences, Mashhad, Iran

Corresponding Author

Azadeh Saki, Ph.D.

Social Determinants of Health
Research Center, Mashhad University
of Medical Sciences, Mashhad, Iran

✉ SakiA@mums.ac.ir

☎ +985138471464

 0000-000 2-4144-5810

Article History

Received: Jul. 27, 2026

Accepted: Aug. 10, 2026

Published: Sep. 30, 2026

Keywords

Coronary artery disease; Coronary angiography; Count regression; Zero-inflated Poisson; Negative binomial regression

Publisher

Neyshabur University of Medical Sciences

ABSTRACT

Background: Coronary artery disease (CAD) is a leading cause of morbidity and mortality worldwide. Modeling the number of obstructed coronary arteries as count data may better reflect disease severity than conventional binary approaches, particularly when excess zeros are present. To identify determinants of coronary artery obstruction and compare count regression models for analyzing disease severity.

Methods: This retrospective cross-sectional study included 902 adults aged ≥ 50 years who underwent coronary angiography at Imam Reza Hospital, Iran. The outcome was the number of obstructed coronary arteries (0–3). Poisson, Negative Binomial (NB), Zero-Inflated Poisson (ZIP), Zero-Inflated Negative Binomial (ZINB), and Hurdle Negative Binomial (HNB) models were compared using AIC, BIC, log-likelihood, RMSE, MAE, and the Vuong test.

Results: Among participants, 28.9% had no coronary obstruction and 34.5% had three-vessel disease. No significant overdispersion was detected. Zero-modified models outperformed conventional Poisson and NB models. Although the HNB model showed the best fit, the ZIP model was selected because of its comparable performance and greater parsimony. Female sex was associated with fewer obstructed arteries (IRR = 0.82, $P = 0.016$), whereas younger age, female sex, and absence of diabetes increased the probability of having no coronary obstruction.

Conclusion: The ZIP model provided a simple and clinically meaningful approach for modeling coronary artery obstruction counts. Age, male sex, diabetes, and smoking were the main determinants of CAD severity.

Introduction

Coronary artery disease (CAD) remains one of the leading causes of morbidity and mortality worldwide despite major advances in prevention, diagnosis, and treatment. Cardiovascular diseases account for approximately one-third of global deaths, with ischemic heart disease representing the largest proportion of this burden (1-4). Increasing prevalence of obesity, diabetes mellitus, hypertension, dyslipidemia, and population aging has contributed to the persistent burden

of CAD, particularly among adults aged 50 years and older, who have a higher risk of obstructive coronary atherosclerosis and adverse cardiovascular outcomes (5-8). Among patients undergoing invasive coronary angiography, the number of obstructed coronary arteries is widely recognized as an important indicator of anatomical disease severity (7, 9). Multivessel coronary artery disease is consistently associated with a substantially higher risk of myocardial infarction, heart failure, repeat revascularization, and long-term cardiovascular mortality than single-vessel



disease or angiographically normal coronary arteries (10-14). Therefore, accurate characterization of the extent of coronary involvement is essential for risk stratification and therapeutic decision-making.

Despite the clinical importance of angiographic disease severity, many previous studies have simplified coronary findings into binary outcomes, such as presence versus absence of obstructive CAD or multivessel versus non-multivessel disease, and analyzed these outcomes using logistic regression models (7, 11, 15, 16). Although clinically interpretable, dichotomization results in loss of information regarding the actual number of affected vessels and may reduce statistical efficiency. Since the number of obstructed coronary arteries is a discrete count outcome ranging from 0 to 3, count regression models may provide a more appropriate framework for analyzing angiographic severity (16-19).

The conventional Poisson regression model is widely used for count outcomes; however, its assumption of equality between mean and variance is often violated in biomedical data because of patient heterogeneity and excess variability (18-21). In addition, angiographic datasets frequently contain a considerable proportion of patients without significant coronary obstruction, resulting in excess zero observations. Failure to account for overdispersion and excess zeros may lead to biased estimates and inappropriate statistical inference (20-24). Therefore, alternative approaches, including Negative Binomial regression for handling overdispersion and Zero-Inflated and Hurdle models for addressing excess zeros, have been developed to provide greater flexibility for modeling complex count distributions (18, 21-26).

Although these models have been increasingly applied in biomedical research, their application for modeling the number of obstructed coronary arteries remains limited (18, 19, 21, 27). Most cardiovascular studies have focused on binary angiographic outcomes or other measures of atherosclerotic burden, such as coronary artery calcium scores and imaging-derived indices, while few studies have directly compared competing count regression approaches for coronary vessel involvement (7, 12, 28-30). Consequently, the most appropriate statistical framework for analyzing the number of obstructed coronary arteries has not yet been fully established.

Adults aged 50 years and older represent a clinically relevant population for this investigation because cumulative exposure to cardiovascular risk factors and age-related vascular changes contribute to increased heterogeneity in coronary atherosclerosis (5, 8, 27, 29, 31). Selecting an appropriate statistical model in this population may improve the estimation of risk factor associations and provide more reliable evidence regarding determinants of coronary disease severity.

Therefore, this study aimed to compare the performance of Poisson, Negative Binomial, Zero-Inflated Poisson, Zero-Inflated Negative Binomial, and Hurdle regression models for modeling the number of obstructed coronary arteries among adults aged 50 years and older. In addition, we aimed to identify demographic, clinical, and laboratory factors associated with increasing coronary artery involvement using the most appropriate count regression approach.

Methods

Study design and population

This retrospective cross-sectional study was conducted using data extracted from the Hospital Information System (HIS) of Imam Reza Hospital, a tertiary referral center affiliated with Mashhad University of Medical Sciences, Mashhad, Iran. The dataset included patients who underwent coronary angiography during 2015 (Iranian calendar year 1394).

During the study period, 1,187 patients were screened using convenience sampling. The inclusion criterion was age ≥ 50 years. Exclusion criteria included pregnancy or lactation; rheumatic disease; chronic liver disease; chronic kidney disease; any type of cancer; infectious diseases within the previous 3 months; history of surgery during the previous 3 months; previous coronary angioplasty (PCI); immunological diseases; inflammatory bowel disease; and the use of steroids, penicillin, oral contraceptives, or hormone replacement therapy. After applying these eligibility criteria, 285 patients were excluded, and 902 patients were included in the final analysis.

The outcome variable was the number of obstructed coronary arteries with significant stenosis, recorded as a count variable ranging from 0 to 3, with higher values indicating more extensive coronary artery involvement. Among the study population, 261 patients (28.9%) had no obstructed coronary vessels, whereas 641 patients (71.1%) had at least one obstructed vessel.

The primary objective of the study was to evaluate and compare the performance of alternative count regression models for identifying demographic, clinical, and laboratory factors associated with the extent of coronary artery involvement.

Variables

Outcome variable

The dependent variable was the number of obstructed coronary arteries with significant stenosis (Y), recorded as a count variable ranging from 0 to 3, where higher values indicated more extensive coronary artery involvement.

Potential predictors were selected according to previous literature and clinical relevance. The following demographic, behavioral, medical history, and laboratory variables were included in the models: age, sex, Body Mass Index (BMI), marital status, education level, occupation, smoking history, hypertension, hyperlipidemia, previous cardiovascular disease, Fasting Blood Sugar (FBS), High-sensitivity C-Reactive Protein (hs-CRP), High-Density Lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), total cholesterol, triglyceride, diabetes mellitus. All categorical variables were entered into the regression models as indicator variables.

Statistical analysis

Descriptive statistics were first calculated for all study variables. Continuous variables were summarized using means and standard deviations, whereas categorical variables



were presented as frequencies and percentages.

Before model selection, the distributional characteristics of the response variable were evaluated by calculating the proportion of zero observations, mean, variance, and dispersion index. Overdispersion was assessed using the Cameron–Trivedi dispersion test, and residual deviance divided by residual degrees of freedom was calculated as an additional diagnostic measure. Although the outcome variable was bounded (0–3), it represented the actual number of obstructed coronary arteries rather than an ordinal severity category. Therefore, count regression models were considered appropriate because the primary objective was to model the expected number of diseased vessels and estimate incidence rate ratios. Furthermore, zero-modified count models allowed separate modeling of patients with no angiographic evidence of coronary obstruction, a common characteristic of angiographic datasets (18, 20–22, 25, 26).

Five alternative count regression models were fitted and compared, including Poisson regression, Negative Binomial (NB), Zero-Inflated Poisson (ZIP), Zero-Inflated Negative Binomial (ZINB), and Hurdle Negative Binomial (HNB) models. The Poisson model was considered the baseline model, assuming equality between the conditional mean and variance of the outcome. The NB model was fitted to account for additional variability through an extra dispersion parameter.

The ZIP model was applied to account for potential excess zeros by separating the zero-generating process from the count process among individuals at risk. The selection of zero-inflated models was motivated by the potential presence of distinct mechanisms generating zero observations. Specifically, patients with no obstructed coronary arteries may represent individuals without angiographic evidence of coronary disease, which differs clinically from patients contributing to the count-generating process. Therefore, ZIP modeling was considered to account for the excess-zero structure rather than overdispersion alone.

The ZINB model was additionally considered to simultaneously account for excess zeros and overdispersion. The HNB model was fitted as an alternative zero-modified approach, assuming that zero counts arise from a separate binary process and positive counts follow a zero-truncated distribution. (18, 20–22, 25, 26, 32, 33)

Model performance was evaluated using both goodness-of-fit and predictive criteria. Predictive performance was assessed using 10-fold cross-validation, in which the dataset was randomly divided into ten approximately equal subsets. In each iteration, nine subsets were used for model fitting and the remaining subset for validation. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were calculated from the validation predictions and averaged across the ten folds. Goodness-of-fit was assessed using Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and log-likelihood. Lower values of AIC, BIC, RMSE, and MAE indicated better model performance. Pairwise comparisons of non-nested models were performed using the Vuong test to compare the Poisson and ZIP, ZIP and ZINB, and ZIP and HNB models.

Multicollinearity among predictors was assessed using the Variance Inflation Factor (VIF), with values below 10 considered acceptable. The final model coefficients were reported as Incidence Rate Ratios (IRRs) for the count

component and Odds Ratios (ORs) for the zero-inflation component, along with 95% confidence intervals. The dataset was assessed for missing values before analysis. No missing data were identified; therefore, complete-case analysis was performed.

Software

All statistical analyses were performed using R software (version 4.5.0). Poisson and Negative Binomial regression models were estimated using the MASS package, whereas Zero-Inflated and Hurdle models were fitted using the pscl package. Multicollinearity diagnostics were conducted using the car package, and predictive performance measures were calculated using the Metrics package. A two-sided P-value less than 0.05 was considered statistically significant.

Results

Table 1 presents the distribution of demographic and clinical categorical variables according to the number of obstructed coronary vessels identified by coronary angiography. A significant association was observed between sex and the extent of coronary artery involvement ($P < 0.001$), with men accounting for a progressively larger proportion of patients as the number of diseased vessels increased. Similarly, the prevalence of diabetes mellitus increased significantly with greater coronary artery involvement ($P < 0.001$), indicating that diabetic patients were more likely to present with multivessel disease.

A significant association was also found for occupation ($P < 0.001$), suggesting that the distribution of employed and unemployed participants differed across angiographic severity groups. In addition, patients with a history of smoking were significantly more likely to have a greater number of obstructed vessels than non-smokers ($P = 0.016$). In contrast, marital status, educational level, history of hypertension, hyperlipidemia, and cardiovascular disease showed no statistically significant differences across the four angiographic groups (all $P > 0.05$).

Table 2 compares continuous clinical characteristics among patients according to the number of obstructed coronary vessels. Significant differences were observed for age, fasting blood sugar (FBS), and LDL-C cholesterol. Patients with three-vessel disease were significantly older and had higher fasting blood glucose than those with less severe or no coronary artery disease. LDL-C levels also differed across groups, with patients with one-vessel disease showing lower concentrations than those with two- and three-vessel disease. No significant differences were observed for BMI, hs-CRP, HDL-C, triglycerides, or total cholesterol. Post hoc (LSD) analyses confirmed that three-vessel disease was associated with older age and higher fasting blood glucose compared with the other angiographic groups. In addition, LDL-C was significantly lower in patients with one-vessel disease than in those with two- and three-vessel disease. No other pairwise comparisons were statistically significant.



Table 1. Comparison of categorical variables according to the number of obstructed coronary vessels

Variable	Category	0 Vessel n (%)	1 Vessel n (%)	2 Vessels n (%)	3 Vessels n (%)	P-value
Sex	Male	78 (29.9)	87 (56.1)	105 (60.0)	198 (63.7)	<0.001
	Female	183 (70.1)	68 (43.9)	70 (40.0)	113 (36.3)	
Marital status	Married	235 (90.0)	135 (87.1)	162 (92.6)	284 (91.3)	0.354
	Single/Divorced/Widowed	26 (10.0)	20 (12.9)	13 (7.4)	27 (8.7)	
Education	Below Bachelor	227 (87.0)	137 (88.4)	147 (84.0)	259 (83.3)	0.389
	Bachelor and above	34 (13.0)	18 (11.6)	28 (16.0)	52 (16.7)	
Occupation	Employed	93 (35.6)	86 (55.5)	93 (53.1)	175 (56.3)	<0.001
	Unemployed	168 (64.4)	69 (44.5)	82 (46.9)	136 (43.7)	
Smoking history	Yes	74 (28.4)	64 (41.3)	70 (40.0)	119 (38.3)	0.016
	No	187 (71.6)	91 (58.7)	105 (60.0)	192 (61.7)	
Hypertension history	Yes	124 (47.5)	72 (46.5)	82 (46.9)	165 (53.1)	0.392
	No	137 (52.5)	83 (53.5)	93 (53.1)	146 (46.9)	
Hyperlipidemia history	Yes	98 (37.5)	58 (37.4)	77 (44.0)	123 (39.5)	0.536
	No	163 (62.5)	97 (62.6)	98 (56.0)	188 (60.5)	
History of CVD	Yes	129 (49.4)	74 (47.7)	69 (39.4)	141 (45.3)	0.212
	No	132 (50.6)	81 (52.3)	106 (60.6)	170 (54.7)	
Diabetes	Yes	49 (18.8)	67 (43.2)	72 (41.1)	161 (51.8)	<0.001
	No	212 (81.2)	88 (56.8)	103 (58.9)	150 (48.2)	

Table 2. Comparison of continuous variables according to the number of obstructed coronary vessels

Variable	0 Vessel	1 Vessel	2 Vessels	3 Vessels	P-value
Age (years)	59.91±8.17	61.05±7.95	61.96±8.04	63.14±8.58	<0.001
BMI (kg/m ²)	26.90±5.36	26.93±4.40	27.53±5.85	26.89±5.37	0.581
FBS (mg/dL)	120.04±52.75	128.34±48.87	124.78±52.65	144.35±75.99	<0.001
hs-CRP (mg/L)	5.47±6.69	7.07±10.00	7.51±10.17	6.49±7.06	0.058
HDL-C (mg/dL)	42.43±11.06	42.17±18.70	40.08±9.52	40.80±17.83	0.319
TG (mg/dL)	141.33±75.13	143.35±69.66	151.12±76.39	151.20±72.81	0.324
LDL-C (mg/dL)	97.23±35.39	90.95±31.99	102.34±42.84	98.85±31.33	0.028
TC (mg/dL)	167.98±44.19	163.06±43.15	170.07±40.54	169.74±46.37	0.423

Table3: Comparison of odds ratios (ORs) for diabetes in the case group versus the control group based on different models

Number of obstructed coronary arteries	Frequency	Percentage (%)
0	261	28.9
1	155	17.2
2	175	19.4
3	311	34.5
Total	902	100.0

Distribution of the outcome variable

A total of 902 participants were included in the study. The outcome variable was defined as the number of obstructed coronary arteries detected by coronary angiography, with possible values ranging from 0 to 3. As shown in Table 3, 261 participants (28.9%) had no obstructed coronary arteries, whereas 641 participants (71.1%) had at least one obstructed vessel. Three-vessel disease was the most frequently observed category (34.5%) (Figure 1).

Distributional characteristics of the response variable

The descriptive characteristics of the response variable are presented in Table 4. The mean number of obstructed coronary arteries was 1.59, while the variance was 1.51. Approximately 28.9% of the observations were zeros. The residual deviance divided by the residual degrees of freedom

was 1.18, and the Cameron–Trivedi dispersion test did not indicate significant overdispersion (dispersion = 0.91, $P = 0.997$). These findings suggested that the response variable did not exhibit meaningful overdispersion despite the presence of a moderate proportion of zero observations.

Table 4. Distributional characteristics of the response variable

Statistic	Value
Sample size	902
Mean	1.594
Variance	1.511
Zero counts (%)	28.9
Dispersion index	0.918
Residual deviance / df	1.180
Cameron–Trivedi dispersion test	0.907
Dispersion test P-value	0.997



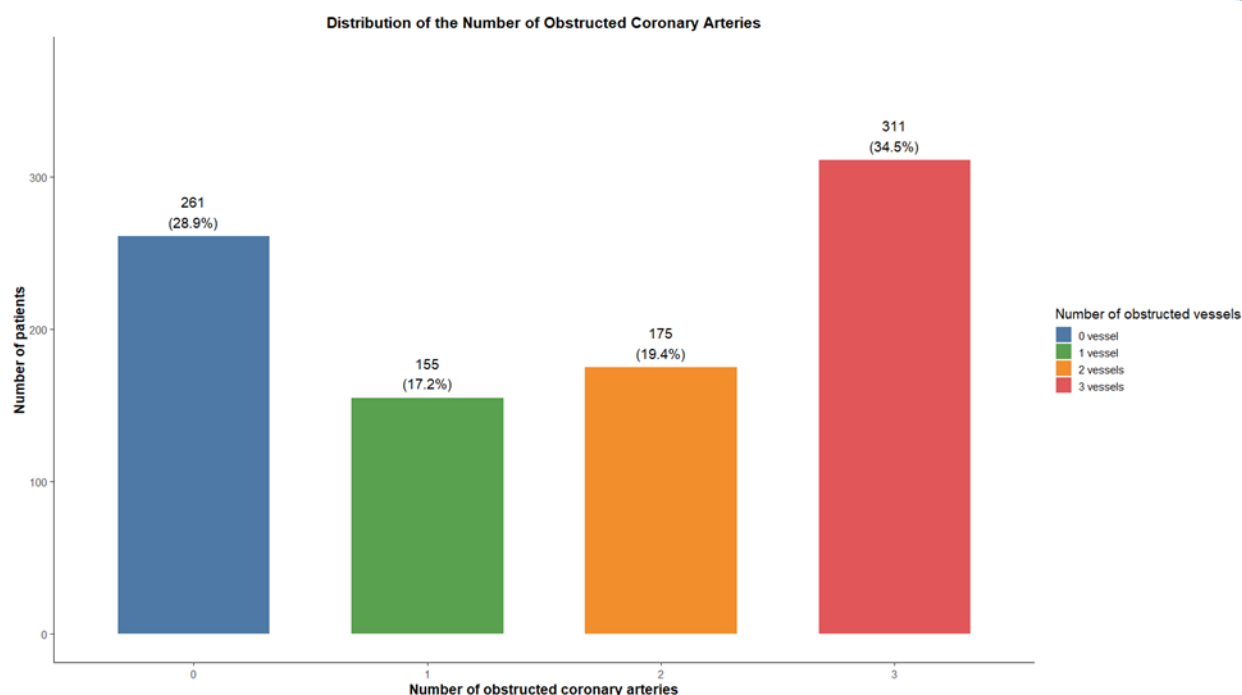


Figure 1. Distribution of the number of obstructed coronary arteries among study participants

Comparison of count regression models

To identify the most appropriate model for predicting the number of obstructed coronary arteries, five count regression models were evaluated, including Poisson, Negative Binomial (NB), Zero-Inflated Poisson (ZIP), Zero-Inflated Negative Binomial (ZINB), and Hurdle Negative Binomial (HNB). The goodness-of-fit statistics for all competing models are summarized in Table 5.

Compared with the conventional Poisson and NB models, all zero-modified models demonstrated substantially improved model fit. The HNB model achieved the smallest Akaike Information Criterion (AIC = 2695.34), the highest log-likelihood (-1310.67), and the lowest prediction error based on root mean squared error (RMSE = 1.131). The ZIP model showed nearly identical performance (AIC = 2699.17; RMSE = 1.137), whereas the ZINB model did not provide any additional improvement over the ZIP model.

Model comparisons

Pairwise comparisons between non-nested models were performed using the Vuong test, and the results are presented in Table 6. The ZIP model significantly outperformed the conventional Poisson model ($P < 0.001$). However, no statistically significant differences were observed between the ZIP and ZINB models ($P = 0.497$) or between the ZIP and HNB models ($P = 0.268$), indicating that the predictive performances of these models were statistically comparable. Considering the absence of overdispersion, the comparable predictive performance of the ZIP and HNB models, and the greater parsimony of the ZIP model, the Zero-Inflated Poisson model was selected as the final model for subsequent analyses. Importantly, the selection of the ZIP model was not based on overdispersion, which was not observed in this dataset, but rather on its ability to account for the excess-zero structure of the outcome and its superior goodness-of-fit compared with the standard Poisson model.

Table 5. Comparison of competing count regression models

Model	Parameters	Log-likelihood	AIC	BIC	RMSE	MAE
Poisson	18	-1370.31	2776.62	2863.11	1.147	1.005
Negative Binomial (NB)	19	-1370.32	2778.64	2869.92	1.147	1.005
Zero-Inflated Poisson (ZIP)	36	-1313.59	2699.17	2872.14	1.137	0.983
Zero-Inflated Negative Binomial (ZINB)	37	-1313.59	2701.17	2878.94	1.137	0.983
Hurdle Negative Binomial (HNB)	37	-1310.67	2695.34	2873.11	1.131	0.986



Table 6. Vuong test comparing non-nested models

Model comparison	Raw z	P-value	Interpretation
Poisson vs ZIP	-6.30	<0.001	ZIP superior
ZIP vs ZINB	0.008	0.497	No significant difference
ZIP vs HNB	-0.619	0.268	No significant difference

Count component of the ZIP model

The incidence rate ratios (IRRs) estimated from the count component of the final ZIP model are presented in Table 7. After adjustment for demographic, clinical, and laboratory variables, female sex was significantly associated with a lower expected number of obstructed coronary arteries than male sex (IRR = 0.82, 95% CI: 0.69–0.96, $P = 0.016$), corresponding to an approximately 18% reduction in the expected number of diseased coronary vessels. Non-diabetic participants also showed a lower expected number of obstructed coronary arteries than diabetic patients; however, this association did not reach conventional statistical significance (IRR = 0.86, 95% CI: 0.74–1.01, $P = 0.059$). None of the remaining demographic, behavioral, or biochemical variables were significantly associated with the number of obstructed coronary arteries.

Table 7. Count component of the final Zero-Inflated Poisson model

Variable	IRR	95% CI	P-value
Age	1.005	0.998–1.013	0.149
Sex (Female vs Male)	0.817	0.692–0.963	0.016
BMI	1.003	0.992–1.014	0.647
Marital status	0.982	0.794–1.213	0.864
Education	1.112	0.948–1.305	0.193
Occupation	1.089	0.936–1.268	0.271
Smoking history	1.020	0.904–1.151	0.750
Hypertension	0.930	0.821–1.054	0.258
Hyperlipidemia	0.989	0.871–1.123	0.869
Cardiovascular disease	1.023	0.908–1.154	0.705
FBS	1.000	0.999–1.001	0.469
hs-CRP	0.998	0.991–1.004	0.504
HDL-C	0.998	0.993–1.003	0.476
TG	1.000	0.999–1.001	0.703
LDL-C	1.001	0.998–1.004	0.696
TC	1.001	0.998–1.004	0.586
Diabetes (No vs Yes)	0.863	0.741–1.005	0.059

Zero-inflation component of the ZIP model

The odds ratios estimated from the zero-inflation component are presented in Table 8. Increasing age was associated with a lower probability of belonging to the structural-zero group (OR = 0.91, 95% CI: 0.86–0.97, $P = 0.005$). Female participants had substantially greater odds of having completely normal coronary arteries than males (OR = 15.75, 95% CI: 5.07–48.88, $P < 0.001$). Likewise, non-diabetic individuals exhibited markedly greater odds of having no angiographic evidence of coronary artery obstruction than diabetic patients (OR = 15.23, 95% CI: 3.98–58.32, $P < 0.001$). Participants with a history of smoking also showed higher odds of belonging to the structural-zero group (OR = 2.24, 95% CI: 1.06–4.72, $P = 0.035$).

Table 8. Zero-inflation component of the final Zero-Inflated Poisson model

Variable	OR	95% CI	P-value
Age	0.913	0.857–0.973	0.005
Sex (Female vs Male)	15.75	5.07–48.88	<0.001
BMI	0.994	0.944–1.047	0.821
Marital status	1.194	0.441–3.231	0.727
Education	1.480	0.616–3.560	0.381
Occupation	0.799	0.324–1.971	0.627
Smoking history	2.237	1.059–4.724	0.035
Hypertension	1.344	0.672–2.688	0.404
Hyperlipidemia	1.090	0.572–2.078	0.794
Cardiovascular disease	0.768	0.395–1.494	0.437
FBS	1.007	1.000–1.013	0.052
hs-CRP	0.910	0.817–1.015	0.089
HDL-C	1.005	0.986–1.024	0.624
TG	0.996	0.986–1.006	0.476
LDL-C	1.001	0.988–1.015	0.856
TC	1.003	0.989–1.018	0.655
Diabetes (No vs Yes)	15.23	3.98–58.32	<0.001

Predictive performance

The predictive performance of the competing models evaluated using 10-fold cross-validation is summarized in Table 9. Among the evaluated models, the HNB model achieved the smallest RMSE (1.131), whereas the ZIP model produced the lowest MAE (0.983). Because the differences in predictive accuracy between the ZIP and HNB models were minimal and statistically indistinguishable according to the Vuong test, and because no evidence of overdispersion was observed, the ZIP model was considered the most appropriate and parsimonious model for predicting the number of obstructed coronary arteries in the present study.

Table 9. Predictive performance of the selected model

Model	RMSE	MAE
Poisson	1.147	1.005
NB	1.147	1.005
ZIP	1.137	0.983
ZINB	1.137	0.983
HNB	1.131	0.986

Multicollinearity assessment

Multicollinearity among explanatory variables was evaluated using the Variance Inflation Factor (VIF). Most predictors exhibited VIF values below 5, suggesting a low degree of collinearity. The highest VIF values were observed for total cholesterol (VIF = 7.41) and LDL cholesterol (VIF = 4.89). Although these variables showed moderate collinearity, the values remained below the commonly accepted threshold of 10; therefore, all predictors were retained in the final analysis.

Discussion

The present study investigated the determinants of coronary artery disease (CAD) severity, defined as the number of obstructed coronary vessels identified by coronary angiography, among adults aged 50 years and older. Five count regression models, including Poisson, Negative Binomial (NB), Zero-Inflated Poisson (ZIP), Zero-Inflated Negative Binomial (ZINB), and Hurdle Negative Binomial (HNB), were compared to identify the model that best



characterized the distribution of obstructed coronary vessels. Although the HNB model achieved the best predictive performance, the ZIP model demonstrated comparable goodness-of-fit and was selected as the final model because it provided a more clinically interpretable framework for distinguishing patients without angiographic obstruction from those with established obstructive disease. Overall, sex was independently associated with the number of obstructed vessels, whereas age, diabetes mellitus, and smoking history were important determinants of belonging to the zero-obstruction subgroup.

Age was one of the strongest determinants of coronary artery disease (CAD) severity in the present study. Patients with three-vessel disease were significantly older than those without angiographic obstruction, and increasing age was associated with a lower probability of belonging to the non-obstructive group. These findings suggest that aging contributes not only to the presence of CAD but also to its anatomical extent.

This association is biologically plausible because aging promotes endothelial dysfunction, arterial stiffness, oxidative stress, chronic low-grade inflammation, and cumulative exposure to cardiovascular risk factors, all of which accelerate atherosclerotic plaque progression and increase the likelihood of diffuse multivessel coronary involvement (31, 34, 35).

Our findings are consistent with previous angiographic studies. Data from the CONFIRM registry demonstrated that advancing age was independently associated with greater coronary plaque burden and more extensive obstructive CAD (36). Similarly, analyses from the PROMISE trial showed that older patients had a higher prevalence of multivessel coronary disease than younger individuals (37). More recently, Song et al. reported that older age remained an important independent predictor of multivessel coronary artery disease after adjustment for conventional cardiovascular risk factors (38). Collectively, these findings highlight age as a major determinant of CAD severity and emphasize the importance of early cardiovascular risk assessment in adults aged 50 years and older.

These findings are further supported by the ARIC Study, which demonstrated that coronary artery calcification becomes highly prevalent in adults aged 75 years and older, reflecting the cumulative burden of atherosclerosis with advancing age. Although coronary artery calcification and angiographic stenosis represent different manifestations of coronary atherosclerosis, both indicate progressive disease burden and support the observed association between aging and more extensive coronary artery involvement in the present study (39).

Sex differences were also observed in the present study. Men had a significantly higher incidence of obstructive coronary artery disease than women, whereas women were more likely to have no angiographic evidence of coronary obstruction. These findings are consistent with the recognized cardioprotective effects of endogenous estrogen before menopause, which delay the development of obstructive atherosclerosis in women. In contrast, men tend to develop coronary atherosclerosis earlier and exhibit a greater burden of diffuse and multivessel disease (40, 41). Similar findings have been reported in large contemporary registries, including the CONFIRM study, where women undergoing

coronary angiography were more likely to have non-obstructive coronary arteries, whereas men more frequently presented with obstructive multivessel disease (42).

Diabetes mellitus was another important determinant of coronary artery disease severity in the present study. Patients with diabetes were significantly less likely to remain free of coronary obstruction, while higher fasting blood glucose levels were observed among patients with three-vessel disease. These findings support the established role of hyperglycemia in accelerating atherosclerosis through endothelial dysfunction, oxidative stress, chronic inflammation, and the formation of advanced glycation end-products, resulting in diffuse coronary plaque progression (43, 44). Consistent with our findings, previous studies have shown that diabetes is independently associated with greater coronary plaque burden, multivessel disease, and adverse cardiovascular outcomes (45). These results emphasize the importance of optimal glycemic control for preventing the progression of coronary artery disease, particularly in adults aged 50 years and older.

A history of smoking was significantly associated with a lower probability of having angiographically normal coronary arteries, indicating its important role in the development of obstructive coronary artery disease. Cigarette smoking promotes endothelial dysfunction, oxidative stress, inflammation, platelet activation, and thrombus formation, all of which accelerate atherosclerotic plaque initiation and progression (46, 47). Although smoking was not associated with the number of obstructed vessels among patients with established CAD, its significant effect on the transition from normal to obstructive coronary arteries suggests that smoking primarily contributes to the early development of coronary atherosclerosis. These findings are consistent with previous angiographic and population-based studies identifying smoking as an independent predictor of obstructive coronary artery disease (7).

In contrast to age, sex, and diabetes, lipid parameters, including LDL-C, HDL-C, triglycerides, and total cholesterol, were not independently associated with the severity of coronary artery disease in the multivariable count models. Although some lipid markers differed across angiographic groups in the univariate analyses, these associations disappeared after adjustment for potential confounders. This finding may reflect the widespread use of lipid-lowering therapy before angiography, whereby measured lipid concentrations no longer represent cumulative lifetime lipid exposure. Moreover, coronary atherosclerosis is influenced by multiple interrelated mechanisms, including chronic inflammation, glycemic control, endothelial dysfunction, and long-term exposure to cardiovascular risk factors, which may attenuate the independent contribution of circulating lipid levels. Similar findings have been reported in previous angiographic studies, where the predictive value of conventional lipid markers was substantially reduced after multivariable adjustment (48).

This observation is also consistent with the concept of residual cardiovascular risk, whereby substantial cardiovascular risk persists despite optimal lipid control because atherosclerosis is driven by multiple interrelated pathways, including chronic inflammation, metabolic dysfunction, endothelial injury, and vascular aging. Consequently, conventional lipid markers may have limited



independent predictive value once these mechanisms and other clinical risk factors are considered (49).

An important methodological aspect of this study was the comparison of alternative count regression models for the number of obstructed coronary vessels. Although approximately 29% of participants had no angiographic evidence of coronary obstruction, dispersion statistics suggested no substantial overdispersion, indicating that the Poisson model provided an acceptable baseline fit. Nevertheless, cross-validation showed slightly lower prediction errors for hurdle and zero-inflated models than for conventional Poisson and negative binomial regression. Because coronary angiographic data may arise from two distinct clinical processes—the probability of having no obstructive disease and the severity of disease among affected individuals—zero-inflated and hurdle models offer a clinically meaningful framework for separating these mechanisms. While their overall statistical fit was comparable to that of the standard count models, these approaches may provide additional interpretability when analyzing angiographic outcomes with a substantial proportion of zero counts (19, 21).

Our findings are broadly consistent with those of Shahsavari et al., who used a truncated zero-inflated double Poisson model to investigate factors associated with the number of stenotic coronary arteries. They reported that older age, male sex, diabetes mellitus, hypertension, hyperlipidemia, smoking, and higher body mass index were significant determinants of coronary artery stenosis. Similarly, our study identified age, male sex, diabetes, and smoking as important predictors of coronary artery disease severity, further supporting the role of these established cardiovascular risk factors while highlighting the value of flexible count regression models for analyzing angiographic count outcomes. (50)

Model selection was not based solely on information criteria but also incorporated cross-validation to evaluate predictive performance. While AIC and BIC quantify the trade-off between model fit and complexity using the observed data, they do not directly assess the ability of a model to generalize to independent samples. Cross-validation provides an estimate of out-of-sample prediction error and has therefore become an essential component of modern predictive modeling. In the present study, the hurdle negative binomial model yielded the lowest prediction error (RMSE), whereas ZIP and ZINB models also demonstrated superior predictive performance compared with conventional Poisson and negative binomial regression. These findings emphasize that predictive validation should complement likelihood-based model selection when identifying the most appropriate regression model for coronary angiographic count outcomes (51-53).

Our findings are also supported by recent methodological research comparing robust zero-inflated and hurdle regression models under varying degrees of zero inflation, overdispersion, and outlier contamination. In that study, model performance was shown to depend strongly on the underlying characteristics of the count data rather than on the modeling framework alone. Specifically, robust Poisson-based zero-inflated and hurdle models performed best when overdispersion and outlier contamination were minimal, whereas robust negative binomial-based models became

superior as data complexity increased. These findings reinforce the principle that no single count regression model is universally optimal and that model selection should be driven by empirical data characteristics. Consistent with this perspective, our study showed that although the Hurdle Negative Binomial model achieved the lowest prediction error, the ZIP model offered a clinically meaningful representation of coronary angiographic findings by distinguishing patients with structurally normal coronary arteries from those with established obstructive disease.

Furthermore, the observation that approximately 29% of participants had no angiographic obstruction did not automatically justify the selection of a zero-inflated model, emphasizing that the presence of excess zeros alone is insufficient to determine the optimal analytical approach. Instead, comprehensive model comparison using goodness-of-fit indices together with predictive validation provides a more robust basis for selecting an appropriate count model (54).

Our findings are also supported by recent methodological studies evaluating zero-inflated and hurdle regression models under varying degrees of zero inflation, overdispersion, and data heterogeneity (55, 56). Collectively, these studies demonstrate that no single count regression model is universally optimal and that model performance depends primarily on the underlying characteristics of the data rather than on the modeling framework alone. They further emphasize that model selection should be guided by the degree of zero inflation, overdispersion, and predictive performance instead of relying solely on the proportion of zero counts or likelihood-based goodness-of-fit criteria. Consistent with these methodological recommendations, our study showed that although the HNB model achieved the lowest prediction error, the ZIP model provided a clinically meaningful representation of coronary angiographic findings by distinguishing patients with structurally normal coronary arteries from those with established obstructive disease. These findings highlight the importance of combining predictive validation with clinical interpretability when selecting count regression models for angiographic outcomes.

Overall, the present study demonstrates that coronary artery disease severity in adults aged 50 years and older is influenced by both traditional cardiovascular risk factors and the underlying distributional characteristics of angiographic count data. Beyond identifying age, male sex, diabetes, and smoking as key determinants of disease severity, our findings underscore the importance of selecting count regression models according to empirical data characteristics and predictive performance rather than relying on a single statistical approach.

This study has several strengths. It included a relatively large cohort of adults aged 50 years and older undergoing coronary angiography and compared multiple advanced count regression approaches rather than relying solely on conventional regression methods. Moreover, coronary artery disease severity was quantified using the number of angiographically obstructed vessels, an outcome with direct clinical relevance.

Although the ZIP model provided clinically meaningful estimates, some odds ratios from the zero-inflation component had relatively wide confidence intervals,



indicating limited precision for certain predictors. These findings should therefore be interpreted cautiously, and future studies with larger sample sizes are recommended to provide more precise estimates.

Strengths and Limitations

Several limitations should also be acknowledged. First, the retrospective cross-sectional design precludes establishing causal relationships between the identified risk factors and coronary artery disease severity. Second, some estimates from the zero-inflation component were associated with relatively wide confidence intervals, indicating limited precision for certain predictors. Third, information on medication use, duration of diabetes, lipid-lowering therapy, inflammatory biomarkers other than hs-CRP, and lifestyle-related factors was unavailable. Fourth, the dataset did not contain the exact percentage of coronary artery stenosis; therefore, the threshold used to define significant stenosis could not be independently verified. Finally, the study was conducted at a single tertiary referral center, which may limit the generalizability of the findings to other populations and healthcare settings. Future multi center prospective studies with larger sample sizes and more comprehensive clinical, angiographic, and biomarker data are warranted to validate these findings.

This study has several strengths. First, it included a relatively large cohort of 902 adults aged 50 years and older who underwent coronary angiography, providing a substantial sample for evaluating determinants of coronary artery disease severity. Second, coronary artery disease severity was assessed using the number of angiographically obstructed coronary vessels, a clinically relevant count outcome that provides more information on disease extent than a simple binary classification. Third, the study comprehensively compared five count regression approaches, including Poisson, Negative Binomial, Zero-Inflated Poisson, Zero-Inflated Negative Binomial, and Hurdle Negative Binomial models. Model selection was based not only on likelihood-based measures such as AIC and BIC but also on predictive performance using cross-validation and prediction error measures, providing a more robust basis for model selection. Finally, the selected ZIP model allowed separate interpretation of the count and zero-inflation components, thereby providing clinically meaningful information regarding both the extent of coronary involvement and the probability of having no angiographic obstruction.

Conclusion

This study demonstrated that the number of obstructed coronary arteries can be effectively modeled as a count outcome using flexible regression approaches that account for the presence of zero observations. Among the evaluated models, the Zero-Inflated Poisson model provided a favorable balance between goodness of fit, predictive performance, parsimony, and clinical interpretability, despite the Hurdle Negative Binomial model showing slightly better predictive performance. The findings identified age, male sex, diabetes mellitus, and smoking history as important determinants of coronary artery disease severity. In

particular, the ZIP framework provided clinically meaningful information by distinguishing factors associated with the number of obstructed vessels from those associated with having no angiographic obstruction. These results highlight the potential value of count-based modeling for characterizing the extent of coronary artery disease and emphasize that model selection should be guided by empirical data characteristics, predictive validation, and clinical interpretability rather than by the proportion of zero observations alone. Further multicenter prospective studies are needed to validate these findings and assess their generalizability across different clinical populations.

Acknowledgment

The authors would like to thank the staff of Imam Reza Hospital for their assistance in data collection and the patients whose clinical data contributed to this study.

Funding

This study was financially supported by the Research and Technology Deputy of Mashhad University of Medical Sciences, Mashhad, Iran.

Ethical Approval

This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of Mashhad University of Medical Sciences, Iran [Ethics Approval Code: IR.MUMS.REC.1398.178]. As this was a retrospective study based on previously collected clinical data, the requirement for informed consent was waived by the ethics committee.

Conflict of Interests

The authors declare that they have no conflicts of interest related to this study.

References

1. Tsao CW, Aday AW, Almarazooq ZI, Anderson CA, Arora P, Avery CL, et al. Heart disease and stroke statistics—2023 update: a report from the American Heart Association. *Circulation*. 2023;147(8):e93. <https://doi.org/10.1161/cir.0000000000001123>.
2. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, et al. Heart disease and stroke statistics—2018 update: a report from the American Heart Association. *Circulation*. 2018;137(12):e67-e492. <https://doi.org/10.1161/cir.0000000000000558>.
3. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. *Journal of the American college of cardiology*. 2020;76(25):2982-3021. <https://doi.org/10.1016/j.jacc.2020.11.010>.
4. Collaborators GCoD. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* (London, England). 2024;403(10440):2100. [https://doi.org/10.1016/s0140-6736\(24\)00367-2](https://doi.org/10.1016/s0140-6736(24)00367-2).
5. Visseren FL, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice: Developed by the Task Force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies With the special contribution of the



- European Association of Preventive Cardiology (EAPC). European heart journal. 2021;42(34):3227-337. <https://doi.org/10.1093/eurheartj/ehab484>.
6. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Journal of the American College of cardiology. 2019;74(10):e177-e232. <https://doi.org/10.1161/cir.0000000000000678>.
7. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes: The Task Force for the diagnosis and management of chronic coronary syndromes of the European Society of Cardiology (ESC). European heart journal. 2020;41(3):407-77. <https://doi.org/10.1093/eurheartj/ehz425>.
8. Libby P. The changing landscape of atherosclerosis. Nature. 2021;592(7855):524-33. <https://doi.org/10.1038/s41586-021-03392-8>.
9. Patel MR, Calhoon JH, Dehmer GJ, Grantham JA, Maddox TM, Maron DJ, et al. ACC/AATS/AHA/ASE/ASNC/SCAI/SCCT/STS 2017 appropriate use criteria for coronary revascularization in patients with stable ischemic heart disease: a report of the American College of Cardiology appropriate use criteria task force, American Association for Thoracic Surgery, American Heart Association, American Society of Echocardiography, American Society of Nuclear Cardiology, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, and Society of Thoracic Surgeons. Journal of the American College of Cardiology. 2017;69(17):2212-41. doi: <https://doi.org/10.1016/j.jacc.2017.02.001>.
10. Head SJ, Milojevic M, Daemen J, Ahn J-M, Boersma E, Christiansen EH, et al. Mortality after coronary artery bypass grafting versus percutaneous coronary intervention with stenting for coronary artery disease: a pooled analysis of individual patient data. The Lancet. 2018;391(10124):939-48. [https://doi.org/10.1016/s0140-6736\(18\)30423-9](https://doi.org/10.1016/s0140-6736(18)30423-9).
11. Maron DJ, Hochman JS, Reynolds HR, Bangalore S, O'Brien SM, Boden WE, et al. Initial invasive or conservative strategy for stable coronary disease. New England Journal of Medicine. 2020;382(15):1395-407. <https://doi.org/10.1056/nejmoa1915922>.
12. Neumann F-J, Sousa-Uva M, Ahlsson A, Alfonso F, Banning AP, Benedetto U, et al. 2018 ESC/EACTS Guidelines on myocardial revascularization. European heart journal. 2019;40(2):87-165. <https://doi.org/10.1093/eurheartj/ehy394>.
13. Boden WE, O'Rourke RA, Teo KK, Hartigan PM, Maron DJ, Kostuk WJ, et al. Optimal medical therapy with or without PCI for stable coronary disease. New England journal of medicine. 2007;356(15):1503-16. <https://doi.org/10.1056/nejmoa070829>.
14. Andreotti F, Koskinas KC, Rossello X, Adamo M, Ainslie J, Banning A, et al. 2024 ESC guidelines for the management of chronic coronary syndromes. Eur Heart J. 2024;45:3415-537. <https://doi.org/10.1093/eurheartj/ehae177>.
15. Patel MR, Peterson ED, Dai D, Brennan JM, Redberg RF, Anderson HV, et al. Low diagnostic yield of elective coronary angiography. New England Journal of Medicine. 2010;362(10):886-95. <https://doi.org/10.1056/nejmoa0907272>.
16. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med. 2006;25(1):127-41. <https://doi.org/10.1002/sim.2331>. PMID: 16217841.
17. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332(7549):1080. <https://doi.org/10.1136/bmj.332.7549.1080>.
18. Cameron AC, Trivedi PK. Regression analysis of count data: Cambridge university press; 2013. <https://doi.org/10.1017/CBO9781139013567>.
19. Hilbe JM. Modeling count data. International encyclopedia of statistical science: Springer; 2025. p. 1504-7.
20. Hilbe JM. Negative binomial regression: Cambridge University Press; 2011.
21. Zeileis A, Kleiber C, Jackman S. Regression models for count data in R. Journal of statistical software. 2008;27:1-25. <https://doi.org/10.18637/jss.v027.i08>.
22. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics. 1992;34(1):1-14. <https://doi.org/10.2307/1269547>.
23. Ridout M, Demétrio CG, Hinde J, editors. Models for count data with many zeros. Proceedings of the XIXth international biometric conference; 1998: International Biometric Society Invited Papers Cape Town, SA.
24. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Mixed effects models and extensions in ecology with R: Springer; 2009.
25. Hall DB. Zero-inflated Poisson and binomial regression with random effects: a case study. Biometrics. 2000;56(4):1030-9. <https://doi.org/10.1111/j.0006-341x.2000.01030.x>.
26. Mullahy J. Specification and testing of some modified count data models. Journal of econometrics. 1986;33(3):341-65. doi: [https://doi.org/10.1016/0304-4076\(86\)90002-3](https://doi.org/10.1016/0304-4076(86)90002-3).
27. Martin SS, Aday AW, Allen NB, Almarzooq ZI, Anderson CA, Arora P, et al. 2025 heart disease and stroke statistics: a report of US and global data from the American Heart Association. Circulation. 2025;151(8):e41-e660. <https://doi.org/10.1161/cir.0000000000001303>.
28. Ridout M, Hinde J, Demétrio CG. A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. Biometrics. 2001;57(1):219-23. <https://doi.org/10.1111/j.0006-341x.2001.00219.x>.
29. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. Nature Reviews Cardiology. 2018;15(9):505-22. <https://doi.org/10.1038/s41569-018-0064-2>.
30. Yang Z, Hardin JW, Addy CL. Score tests for zero-inflation in overdispersed count data. Communications in Statistics—Theory and Methods. 2010;39(11):2008-30. <https://doi.org/10.1080/03610920902948228>.
31. North BJ, Sinclair DA. The intersection between aging and cardiovascular disease. Circulation research. 2012;110(8):1097-108. <https://doi.org/10.1161/circresaha.111.246876>.
32. McCullagh P. Generalized linear models: Routledge; 2019.
33. Greene WH. Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. 1994.
34. Ungvari Z, Tarantini S, Donato AJ, Galvan V, Csiszar A. Mechanisms of vascular aging. Circulation research. 2018;123(7):849-67. <https://doi.org/10.1161/circresaha.118.311378>.
35. Cunha PG, Olsen MH. Vascular aging and cardiovascular disease. Early Vascular Aging (EVA): Elsevier; 2024. p. 19-32.
36. Min JK, Dunning A, Lin FY, Achenbach S, Al-Mallah M, Budoff MJ, et al. Age- and sex-related differences in all-cause mortality risk based on coronary computed tomography angiography findings: results from the International Multicenter CONFIRM (Coronary CT Angiography Evaluation for Clinical Outcomes: An International Multicenter Registry) of 23,854 patients without known coronary artery disease. Journal of the American College of Cardiology. 2011;58(8):849-60. <https://doi.org/10.1016/j.jacc.2011.02.074>.
37. Hoffmann U, Ferencik M, Udelson JE, Picard MH, Truong QA, Patel MR, et al. Prognostic value of noninvasive cardiovascular testing in patients with stable chest pain: insights from the PROMISE trial (Prospective Multicenter Imaging Study for Evaluation of Chest Pain). Circulation. 2017;135(24):2320-32. <https://doi.org/10.1161/circulationaha.116.024360>.
38. Song Y, Cao X, Zhang H, Tian X, Hua H, Ferdous M, et al. Construction of classification model and analysis of risk factors in patients with multi-vessel coronary artery disease. BMC Medical Informatics and Decision Making. 2026;26(1):136. <https://doi.org/10.1186/s12911-026-03437-w>.
39. Wang FM, Cainzos-Achirica M, Ballew S, Folsom AR, Wagenknecht LE, Howard-Claudio CM, et al. Abstract P213: Characterizing The Distribution Of Coronary Artery Calcification In The 75-and-older Population: The Atherosclerosis Risk In Communities Study. Circulation. 2021;143(Suppl. 1):AP213-AP. https://doi.org/10.1161/circ.143.suppl_1.P213.
40. Maas AH, Appelman YE. Gender differences in coronary heart disease. Netherlands Heart Journal. 2010;18(12):598-603. <https://doi.org/10.1007/s12471-010-0841-y>.
41. Vogel B, Acevedo M, Appelman Y, Merz CNB, Chieffo A, Figtree GA, et al. The Lancet women and cardiovascular disease Commission: reducing the global burden by 2030. The Lancet. 2021;397(10292):2385-438. [https://doi.org/10.1016/s0140-6736\(21\)00684-x](https://doi.org/10.1016/s0140-6736(21)00684-x).
42. Naoum C, Berman DS, Ahmadi A, Blanke P, Gransar H, Narula J, et al. Predictive value of age- and sex-specific nomograms of global plaque burden on coronary computed tomography angiography for major cardiac events. Circulation: Cardiovascular Imaging. 2017;10(3):e004896. <https://doi.org/10.1161/CIRCIMAGING.116.004896>.
43. Low Wang CC, Hess CN, Hiatt WR, Goldfine AB. Clinical update: cardiovascular disease in diabetes mellitus: atherosclerotic cardiovascular disease and heart failure in type 2 diabetes mellitus—mechanisms, management, and clinical considerations. Circulation. 2016;133(24):2459-502. <https://doi.org/10.1161/circulationaha.116.022194>.
44. Cosentino F, Grant PJ, Aboyans V, Bailey CJ, Ceriello A, Delgado V, et al. 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD: The Task Force for diabetes, pre-diabetes, and cardiovascular diseases of the European Society of Cardiology (ESC) and the European Association for the Study of Diabetes (EASD). European heart journal. 2020;41(2):255-323. <https://doi.org/10.1093/eurheartj/ehz486>.



45. Jia G, Hill MA, Sowers JR. Diabetic cardiomyopathy: an update of mechanisms contributing to this clinical entity. *Circulation research*. 2018;122(4):624-38. <https://doi.org/10.1161/circresaha.117.311586>.
46. Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease: an update. *Journal of the American college of cardiology*. 2004;43(10):1731-7. <https://doi.org/10.1016/j.jacc.2003.12.047>.
47. Banks E, Joshy G, Korda RJ, Stavreski B, Soga K, Egger S, et al. Tobacco smoking and risk of 36 cardiovascular disease subtypes: fatal and non-fatal outcomes in a large prospective Australian study. *BMC medicine*. 2019;17(1):128. <https://doi.org/10.1186/s12916-019-1351-4>.
48. Mortensen MB, Falk E. Primary prevention with statins in the elderly. *Journal of the American College of Cardiology*. 2018;71(1):85-94. <https://doi.org/10.1016/j.jacc.2017.10.080>.
49. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *European heart journal*. 2017;38(32):2459-72. <https://doi.org/10.1093/eurheartj/ehx144>.
50. Orooji A, Sahranavard T, Shakeri M-T, Tajfard M, Saffari SE. Application of the Truncated Zero-Inflated Double Poisson for Determining of the Effecting Factors on the Number of Coronary Artery Stenosis. *Computational and mathematical methods in medicine*. 2022;2022(1):5353539. <https://doi.org/10.1155/2022/5353539>.
51. Hastie T. The elements of statistical learning: data mining, inference, and prediction. springer; 2009.
52. Kuhn M, Johnson K. Applied predictive modeling: Springer; 2013.
53. James G. An introduction to statistical learning with applications in R. Springer; 2013.
54. Pitsha P, Chiruka RT, Marange CS. A Comparison of the Robust Zero-Inflated and Hurdle Models with an Application to Maternal Mortality. *Mathematical and Computational Applications*. 2025;30(5):95. <https://doi.org/10.3390/mca30050095>.
55. Young DS, Roemmele ES, Yeh P. Zero-inflated modeling part I: Traditional zero-inflated count regression models, their applications, and computational tools. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2022;14(1):e1541. <https://doi.org/10.1002/wics.1541>.
56. Bhaskar A, Thennarasu K, Philip M, Jaisoorya T. Regression models for count data with excess zeros: A comparison using survey data. *Quantitative Methods for Psychology*. 2023;19(1):1-13. <https://doi.org/10.20982/tqmp.19.1.p001>.

