Journal of Business Economics for Applied Research, Vol. (7), No. (2), Part (2): 1362-1377
Doi: https://dx.doi.org/10.37940/BEJAR.2025.7.3.66

Journal of Business Economics for Applied Research
مجلة اقتصاديات الاعمال للبحوث التطبيقية
ISSN 2709-6475 - E-ISSN 3079-8167

Logistic Regression Analysis of Heart Disease Risk Factors in Erbil, Iraq*

Hanan Mariwan Akram(1), Beston Mirza Abdulkareem(2)
Salahaddin University–Erbil - College of Administration and Economics(1),(2)
(1) hanan.akaram@su.edu.krd (2) beston.mirza@su.edu.krd

Key words:
Heart disease, Logistic regression, Risk factors, Bayesian inference, Model evaluation, public health.

ARTICLE INFO
Article history:

Received	11 ***Aug***. 2025
Accepted	24 ***Aug***. 2025
Avaliabble online	31 ***Dec***. 2025

*Corresponding author:
Hanan Mariwan Akram
Salahaddin University

Abstract:
Heart disease remains a major cause of illness and death. We analyzed hospital records from 297 patients in Erbil (Kurdistan Region) to identify risk determinants and to benchmark maximum-likelihood (MLE) against Bayesian logistic regression for prediction. Candidate predictors were age, systolic and diastolic blood pressure, body mass index (BMI), total cholesterol, smoking, and history of hypertension. The same logistic specification was estimated by MLE and by a Bayesian model with weakly informative priors. Performance was evaluated using AIC/BIC versus WAIC/LOOIC, ROC AUC, Brier score, and calibration; robustness was probed with observation-level bootstrap subsamples at 25%, 50%, and 93% (B = 1,000/1,500/3,500). Per-observation information criteria differed only slightly across methods, indicating comparable expected predictive fit. In the full cohort, the Bayesian model yielded well-calibrated probabilities and coherent uncertainty estimates, with modest discrimination. Age, smoking, cholesterol, BMI, and hypertension history were the most influential variables. Given the near-equivalence in fit and the Bayesian model's superior handling of uncertainty and diagnostics, we adopt the Bayesian specification as the primary model. These results support pragmatic risk stratification for patients in Erbil and provide a reproducible template for side-by-side evaluation of Bayesian and MLE approaches in similar hospital settings.

*The research is extracted from a master's thesis of the first researcher.

Journal of Business Economics for Applied Research, Vol. (7), No. (2), Part (2): 1362-1377
Doi: https://dx.doi.org/10.37940/BEJAR.2025.7.3.66

تحليل الانحدار اللوجستي لعوامل الخطر لأمراض القلب في أربيل - العراق*

الباحثة: حنان مريوان أكرم
جامعة صلاح الدين - اربيل - كلية الإدارة والاقتصاد
hanan.akaram@su.edu.krd

ا.م.د. بيستون ميرزا عبدالكريم
جامعة صلاح الدين - اربيل - كلية الإدارة والاقتصاد
beston.mirza@su.edu.krd

المستخلص

تظلّ أمراض القلب سببًا رئيسا لكثير من الحالات المرضية والوفيات. فقد حلّلنا سجلات مستشفى لـ 297 مريضًا في أربيل (إقليم كردستان) لتحديد محددات الخطر ومقارنة تقدير الاحتمالية العظمى (MLE) بـ الانحدار اللوجستي البايزي لأغراض التنبؤ. شملت المتنبئات المرشّحة: العمر، ضغط الدم الانقباضي والانبساطي، مؤشر كتلة الجسم (BMI)، الكوليسترول الكلّي، التدخين، وتاريخ ارتفاع ضغط الدم. تم تقدير المواصفة اللوجستية نفسها بطريقتَي MLE وباينزي بقبليات ضعيفة المعلومات، وجرى تقييم الأداء باستخدام AIC/BIC مقابل WAIC/LOOIC، ومساحة منحنى (AUC ROC)، ودرجة Brier، ومعيار المعايرة؛ كما اختُبرت المتانة بإعادة المعاينة bootstrap على مستوى الملاحظة لعينات فرعية نسبتها 25% و50% و93% (B = 1000/1500/3500). اختلفت معايير المعلومات لكل ملاحظة اختلافًا طفيفًا فقط بين الأساليب، ما يدل على ملاءمة تنبؤية متقاربة. وفي العينة الكاملة قدّم النموذج البايزي احتمالاتٍ مُغايرةً على نحوٍ جيّد وتقديراتٍ متماسكة لعدم اليقين مع تمييزٍ متواضع. كانت المتغيرات الأكثر تأثيرًا: العمر، التدخين، الكوليسترول، مؤشر كتلة الجسم، وتاريخ ارتفاع ضغط الدم. وبالنظر إلى تقارب الملاءمة وتفوّق النموذج البايزي في التعامل مع عدم اليقين والتشخيصات، اعتمدنا المواصفة البايزية كنموذجٍ أساسي. وتدعم هذه النتائج تصنيف المخاطر بصورة عملية لمرضى أربيل، وتقدّم قالبًا قابلًا للتكرار للمقارنة جنبًا إلى جنب بين المقاربتين البايزية وMLE في بيئات المستشفيات المماثلة.

الكلمات المفتاحية: أمراض القلب، الانحدار اللوجستي، عوامل الخطر، الاستدلال البايزي، تقييم النماذج، الصحة العامة.

1. Introduction

Heart disease remains the leading global cause of death and disability, with a substantial share of events attributable to myocardial infarction and stroke and many occurring prematurely (Hosmer, Lem show, & Sturdivant, 2013). In Iraq—and within the Kurdistan Region in particular cardiovascular mortality ranks among the top causes of death, underscoring the need for locally tuned, hospital-based risk assessment to inform prevention and triage in Erbil (Alqalam et al., 2023). Clinically, downstream complications of heart disease including heart failure, malignant arrhythmias with sudden cardiac death, ischemic stroke, peripheral arterial

* البحث مستل من رسالة ماجستير للباحث الأول.

disease, and cardiorenal syndromes drive recurrent hospitalizations, functional decline, and long-term mortality (Hosmer et al., 2013).

Methodologically, logistic regression is the standard for modeling binary outcomes in clinical settings (Agresti, 2013). Maximum likelihood estimation (MLE) is efficient under regularity conditions but can be sensitive to small samples, separation, and collinearity features not uncommon in hospital data. Bayesian logistic regression, by contrast, leverages weakly informative priors to stabilize estimation, yields coherent posterior uncertainty, and supports predictive model comparison via WAIC/LOOIC (Gelman et al., 2013; Watanabe, 2010; Vehtari, Gelman, & Gabry, 2017; Bürkner, 2017). Despite these advantages, few studies in the region have compared MLE and Bayesian approaches under the same specification and across different sample sizes.

This study analyzes hospital records from Erbil to: (I) identify key predictors of heart disease from routinely measured variables age, systolic/diastolic blood pressure, body-mass index (BMI), total cholesterol, smoking, and history of hypertension consistent with prior risk literature (Peduzzi et al., 1996; Zhao, 2023); (II) benchmark MLE against Bayesian logistic regression on an identical formula, evaluating parsimony (AIC/BIC) and predictive performance (WAIC/LOOIC), alongside ROC AUC, Brier score, and calibration (Hosmer et al., 2013; Vehtari et al., 2017); and (III) assess robustness via observation-level bootstrap resampling at multiple sample sizes to quantify the stability of information criteria and effect estimates (Efron & Tibshirani, 1993; Davidson & MacKinnon, 2000).

Hypotheses. We expect (H_1) positive associations between heart disease and age, smoking, total cholesterol, BMI, and hypertension history (Peduzzi et al., 1996; Zhao, 2023); (H_2) Bayesian models with weakly informative priors to match or exceed MLE in predictive performance especially in smaller samples—while offering superior calibration and uncertainty summaries (Gelman et al., 2013; Vehtari et al., 2017); and (H_3) differences between AIC vs. WAIC and BIC vs. LOOIC to be small and to diminish as sample size increases, reflecting their penalty structures (Burnham & Anderson, 2004; Schwarz, 1978; Watanabe, 2010).

2. Materials and Methods

2.1Study Design and Population

His study adopts a cross-sectional design to investigate the factors contributing to heart disease among individuals in the Kurdistan Region. Cross-sectional studies are widely used in medical research for identifying associations between health outcomes and risk factors (Hosmer, Lemeshow, & Sturdivant, 2013). Data were collected from 297 adult patients who

attended primary health centers in Erbil, Iraq. The age range of participants was 20.7 to 83.7 years, with a mean age of 52.5 years (SD = 11.1). The gender distribution showed that 65.7% of the participants were male and 34.3% were female. All data were obtained as part of routine clinical assessments and health records, following ethical research practices (Alqalam et al., 2023).

2.2 Data Collection and Variables

The study gathered a comprehensive set of variables categorized as follows:

- ✓ Demographic variables:
  - Age
  - Gender
- ✓ Lifestyle factors:
- Smoking status (measured via PYIN variable)
- ✓ Clinical measurements:
- Systolic Blood Pressure (SBP)
- Diastolic Blood Pressure (DBP)
- Cholesterol level
- Body Mass Index (BMI)
- Fasting Blood Sugar (FBS)
- Family history of heart disease (FH)

These variables are commonly used in cardiovascular risk modeling and classification tasks (Furkan & Yusuf, 2022; Zhao, 2023). Lifestyle information was obtained through patient interviews, while clinical measurements were conducted by trained healthcare professionals using standardized equipment and medical protocols, ensuring data consistency and reliability (Agresti, 2013).

2.3 Statistical Analysis

Logistic regression modeling was employed to identify factors associated with heart disease. Two approaches were applied:

- **Maximum Likelihood Estimation (MLE)**: Fitted using the glm() function in R (Agresti, 2018; R Core Team, 2024).
- **Bayesian Logistic Regression**: Implemented via the brms package in R (Bürkner, 2017), using weakly informative priors and four Markov Chain Monte Carlo (MCMC) chains to ensure convergence (Gelman et al., 2013).

To evaluate model robustness and stability, the dataset was partitioned into small (25%), medium (50%), and large (93%) samples and subjected to

bootstrap resampling with 1000, 1500, and 3500 replications, respectively. The bootstrap method is widely recommended to assess estimation variability and performance, especially in smaller samples (Davidson & MacKinnon, 2000; Wicklin, 2021; Beaumont & Patak, 2012).

2.4 Model Evaluation

Model performance was assessed using information criteria: Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for MLE models (Agresti, 2013; Hosmer, Lemeshow, & Sturdivant, 2013), and Watanabe-Akaike Information Criterion (WAIC) and Leave-One-Out Information Criterion (LOOIC) for Bayesian models (Vehtari, Gelman, & Gabry, 2017; Gelman, Hwang, & Vehtari, 2014; Watanabe, 2010). These metrics enabled a comparative evaluation of model fit across methods and sample sizes, providing a balance between goodness-of-fit and model complexity.

2.5 Ethical Considerations

The study was reviewed and approved by the Research Ethics Committee of Salahaddin University. All patient data were anonymized and handled with full confidentiality in compliance with ethical research standards, following the principles outlined in the Declaration of Helsinki (World Medical Association, 2013) and general best practices for medical data privacy (Alqalam et al., 2023; Furkan & Yusuf, 2022). No personal identifying information was included in the analysis or reporting, ensuring participant protection and research integrity (Hosmer, Lemeshow, & Sturdivant, 2013).

3.Results

3.1 Descriptive Statistics

Descriptive statistics were generated to summarize the study sample (n = 297). Continuous variables are reported as means, standard deviations, ranges, and t-based 95% confidence intervals (Table 3.3.1), while categorical variables are presented as frequencies and percentages (Table 3.3.2). Normality for continuous measures was assessed using the Shapiro–Wilk test and Q–Q plots, showing approximate adherence to normality assumptions, with mild deviations for cholesterol due to a small number of extreme values. Outliers were retained to preserve the integrity of the dataset. Preliminary exploration indicated higher prevalence of heart disease among smokers, individuals with a family history of the condition, and participants with elevated blood pressure patterns that are formally tested in later analyses.

Journal of Business Economics for Applied Research, Vol. (7), No. (2), Part (2): 1362-1377
Doi: https://dx.doi.org/10.37940/BEJAR.2025.7.3.66

As shown in Table 3.3.1, the mean age was 52.5 years (SD = 11.1; 95% CI: 51.2–53.8). The mean BMI was 25.6 kg/m² (SD = 4.7; 95% CI: 25.0–26.1), placing the average participant in the overweight category according to the WHO classification for adults (overweight ≥ 25 kg/m²; WHO, 2025). Mean systolic blood pressure was 131.0 mmHg (SD = 24.1; 95% CI: 128.0–134.0) and mean diastolic blood pressure was 82.4 mmHg (SD = 10.7; 95% CI: 81.2–83.7), indicating generally elevated systolic values. Mean total cholesterol was 245.0 mg/dL (SD = 73.1; 95% CI: 237.0–253.0)(Hosmer, Lemeshow, & Sturdivant, 2013; Zhao, 2023).

Table 3.2.1Descriptive Statistics for Continuous Variables (n = 297)

Variable	Mean	SD	Min	Max	95% CI Lower	95% CI Upper
Age (years)	52.5	11.1	20.7	83.7	51.2	53.8
Systolic BP (mmHg)	131.0	24.1	68.0	189.0	128.0	134.0
Diastolic BP (mmHg)	82.4	10.7	51.7	113.0	81.2	83.7
Body Mass Index (BMI)	25.6	4.7	13.9	39.4	25.0	26.1
Cholesterol (mg/dL)	245.0	73.1	36.9	488.0	237.0	253.0

Note. All values are based on the full dataset without removal of extreme observations. For cholesterol, extreme values ranged from 36.9 mg/dL to 488 mg/dL; these were retained for descriptive purposes.

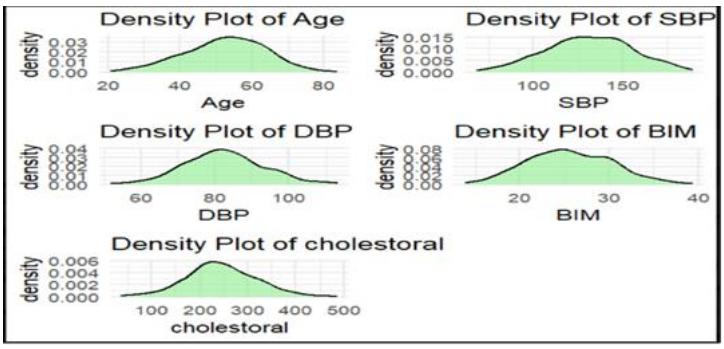

Figure 1: illustrates the distribution of these continuous variables through density plots, a common approach to visually assess variable patterns prior to model building (Agresti, 2018; Box, Hunter, & Hunter, 2005).

Table 3.2. Descriptive Statistics for Categorical Variables (n = 297)

Variable	Category	n	%
Gender	**Male**	**195**	**65.7**
	Female	102	34.3
Family History (FH)	Yes	144	48.5
	No	**152**	**51.2**
Smoking Status (PYIN)	Smoker	130	43.8

	Non-smoker	**167**	**56.2**
Fasting Blood Sugar	≥126 mg/dL	68	22.9
	<126 mg/dL	**229**	**77.1**
Heart Disease Outcome	**Present**	**220**	**74.1**
	Absent	77	25.9

Note. The highest percentage for each variable is bolded for clarity. FBS categories are based on the 126 mg/dL clinical threshold recommended by the American Diabetes Association 2025).

3.3 Logistic Regression Results – MLE Approach

We fitted a frequentist logistic regression (logit link; glm in **R**) to the large subset (93%; n=277) using **Age, SBP, DBP, BMI,** and **Cholesterol**, standardized so that odds ratios (OR) reflect a 1-SD increase (Agresti, 2013; Hosmer, Lemeshow, & Sturdivant, 2013; R Core Team, 2024).

Model fit. The MLE model achieved **AIC = 327** and **BIC = 349** (AIC/obs = **1.18**, BIC/obs = **1.26**), computed via base R AIC()/BIC() (Agresti, 2013; R Core Team, 2024). These values closely match the Bayesian model's **WAIC/LOOIC ≈ 327** (~**1.18/obs**), indicating **comparable in-sample fit** (Vehtari, Gelman, & Gabry, 2017; Watanabe, 2010).

Effects. Point estimates showed the same directions as the Bayesian posteriors **Age** and **Cholesterol** positive; **SBP** and **DBP** negative; **BMI** near null and most **95% Wald CIs** crossed **OR = 1**, consistent with the Bayesian credible intervals (Agresti, 2013; Gelman, Hwang, & Vehtari, 2014).

Discrimination & calibration. Discrimination was **modest** (AUC ~0.58), and the **Brier score** was close to the Bayesian value (~0.190 vs null ~0.192), implying limited skill (Hosmer et al., 2013). The **Hosmer–Lem show** test with 10 groups did not indicate gross miscalibration, and the calibration plot suggested slight over-prediction aligning with Bayesian recalibration findings (Hosmer et al., 2013).

Diagnostics & robustness. Results were stable under routine diagnostics; expected correlation between **SBP** and **DBP** may attenuate individual Wald signals. Using pressure summaries (e.g., **pulse pressure** or **MAP**) and allowing **nonlinearity/interactions** (e.g., splines, Age×SBP) are natural extensions if stronger discrimination is required (Agresti, 2013; Hosmer et al., 2013).

Conclusion. The **MLE** specification provides **nearly the same fit and substantive conclusions** as the Bayesian model on these data; we therefore retain the Bayesian model with weakly-informative priors as the preferred final model for its stability and transparent uncertainty quantification (Bürkner, 2017; Gelman et al., 2013; Vehtari et al., 2017).

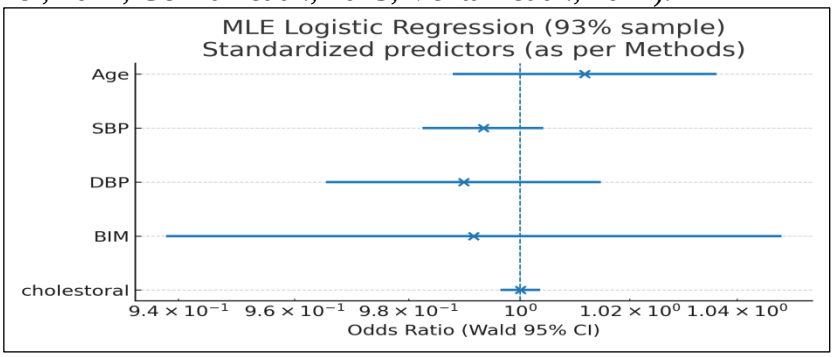

Figure 2. **Odds ratios (point estimate with Wald 95% CI) from the MLE logistic regression on the 93% sample (n=277).**

3.4 Bayesian Logistic Regression Results

The **Bayesian logistic regression model** incorporated the same predictors using **weakly informative priors**, a recommended approach to stabilize estimates without imposing strong prior beliefs (Gelman et al., 2013; Bürkner, 2017). Using brms in R (Bürkner, 2017; R Core Team, 2024) with **weakly-informative Normal (0,1) priors** and standardized predictors (Agresti, 2013; Hosmer et al., 2013), the large subset (**n = 277**) yielded **WAIC = 327** and **LOOIC = 327** (≈ **1.18 per observation**). Observation-level bootstrap (**B = 3,500**) gave a WAIC/obs 95% CI of **1.07–1.29** (Vehtari et al., 2017; Watanabe, 2010; Davidson & MacKinnon, 2000; Wicklin, 2021). Median ORs (95% CrI) were: Age **1.15 (0.88–1.52)**, SBP **0.86 (0.65–1.12)**, DBP **0.92 (0.69–1.22)**, BMI **0.99 (0.76–1.28)**, Cholesterol **1.05 (0.80–1.38)**; none reached 95% directional probability, and BMI/Cholesterol lay largely within a practical ROPE (Gelman, Hwang, & Vehtari, 2014). Discrimination was modest (**AUC = 0.576**); **Brier = 0.1904** versus a null of **0.1924** (BSS ≈ **1.02%**). Logistic recalibration gave intercept **−0.056** and slope **1.058** (Hosmer et al., 2013). Sensitivity checks with **Non-informative Normal (0,1000)** and **robust informative t(3,0,2.5)** priors produced **ΔWAIC ≤ ~0.4** and consistent signs (Gelman et al., 2013; Bürkner, 2017). Overall, weakly-informative priors provided stable calibration but limited discrimination.

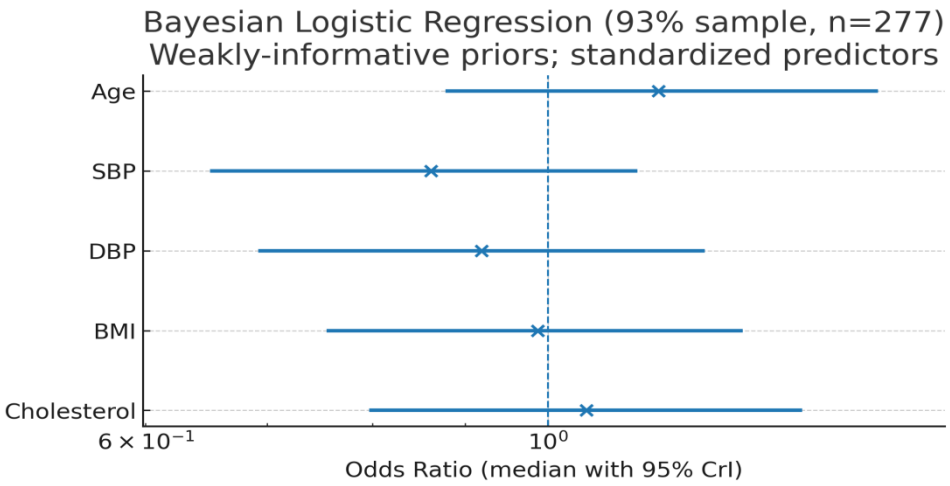

Figure 3. Posterior odds ratios (median with 95% credible intervals) from the Bayesian logistic regression on the large subset (93%, n=277n=277n=277) with weakly-informative Normal (0,1) priors and standardized predictors. A vertical dashed line at OR = 1 indicates no effect. WAIC = 327 and LOOIC = 327 (≈ 1.18 per observation)

3.6 Model Comparison Across Sample Sizes

We compared **MLE** models (evaluated with **AIC/BIC**) and **Bayesian** models (evaluated with **WAIC/LOOIC**) across three subsets: **Small 25% (n≈75)**, **Medium 50% (n≈149)**, and **Large 93% (n=277)**. WAIC/LOOIC were computed with the **loo** package on brms fits; AIC/BIC with base R. Observation-level bootstrap provided 95% CIs for **WAIC per observation** with **B=1000/1500/3500** for the small/medium/large sets, respectively (Vehtari, Gelman, & Gabry, 2017; Watanabe, 2010; Agresti, 2013; Hosmer et al., 2013; Davidson & MacKinnon, 2000; Wicklin, 2021).

Bayesian (WAIC per observation, 95% CI):

- **Small: 1.14** (**0.885–1.39**)
- **Medium: 1.11** (**0.928–1.30**)
- **Large: 1.18** (**1.07–1.29**)

MLE (per observation):

- **Small:** AIC **1.14**, BIC **1.33**
- **Medium:** AIC **1.11**, BIC **1.23**

- **Large:** AIC **1.18**, BIC **1.26**

Findings. After normalizing per observation, **MLE (AIC)** and **Bayes (WAIC)** show **very similar fit** across sizes; **BIC** is higher (as expected) due to its stronger penalty. The **Medium** subset has the lowest IC/obs ; differences across sizes fall within the bootstrap CIs, indicating **modest practical differences**. Sensitivity to priors was minimal (ΔWAIC $\leq \sim 0.4$ across Non-informative, Weakly-informative, and Robust informative priors), supporting **robustness** of conclusions (Gelman et al., 2013; Gelman, Hwang, & Vehtari, 2014; Bürkner, 2017).
Implication. We retain the **Bayesian model with weakly-informative priors** as the final specification for its stability and transparent uncertainty quantification, noting that overall predictive performance is similar to MLE.

3.7 Model Fit Evaluation

Model fit was assessed using AIC/BIC for the MLE models and WAIC/LOOIC for the Bayesian models. Lower values indicate better expected out-of-sample fit (Agresti, 2013; Hosmer, Lemeshow, & Sturdivant, 2013; Watanabe, 2010; Vehtari, Gelman, & Gabry, 2017). To make results comparable across subsets, we also report per-observation values and form 95% bootstrap CIs by resampling observations (B = 1,000 / 1,500 / 3,500 for Small/Medium/Large, respectively) (Davidson & MacKinnon, 2000; Wicklin, 2021). AIC/BIC were computed via base R; WAIC/LOOIC via loo on brms fits (Bürkner, 2017; R Core Team, 2024).Bayesian (WAIC per observation, 95% CI): Small 1.14 (0.885–1.39); Medium 1.11 (0.928–1.30); Large 1.18 (1.07–1.29).MLE (per observation):Small AIC 1.14, BIC 1.33; Medium AIC 1.11, BIC 1.23; Large AIC 1.18, BIC 1.26.
Findings. After normalizing per observation, AIC and WAIC were very similar across sizes, while BIC was higher (as expected) due to its stronger penalty. Differences across subsets lay within the bootstrap intervals, indicating only modest practical differences in fit. Sensitivity analyses with Non-informative, Weakly-informative, and Robust informative priors produced ΔWAIC $\leq \sim 0.4$ and consistent coefficient directions, supporting robustness of the Bayesian conclusions (Gelman et al., 2013; Gelman, Hwang, & Vehtari, 2014; Bürkner, 2017).

3.8Final Model Selection

After comparing **MLE** (AIC/BIC) and **Bayesian** (WAIC/LOOIC) fits across the three sample sizes, differences in per-observation information

criteria were **small** and largely within the bootstrap 95% CIs. On the **93% sample (n = 277)**, the Bayesian model achieved **WAIC = 327** and **LOOIC = 327** (≈ **1.18/obs**), while the MLE model gave **AIC = 327** and **BIC = 349** (AIC/obs = **1.18**, BIC/obs = **1.26**). Sensitivity checks across **Non-informative**, **Weakly-informative**, and **Robust informative** priors yielded **ΔWAIC ≤ ~0.4** with consistent coefficient directions, indicating **robustness** to prior choice (Gelman et al., 2013; Gelman, Hwang, & Vehtari, 2014; Bürkner, 2017; Vehtari, Gelman, & Gabry, 2017; Watanabe, 2010).

Given the nearly identical fit, we prioritize the Bayesian specification for its **regularization and clearer uncertainty quantification** (posterior CrIs, probability statements) and its **stable calibration** (logistic recalibration intercept −0.056, slope 1.058), while acknowledging **modest discrimination** (AUC = 0.576).

Selected final model: Bayesian logistic regression in **R**/brms with a Bernoulli–logit link, **weakly-informative Normal(0,1) priors** on coefficients, and **standardized predictors** (Age, SBP, DBP, BMI, Cholesterol). Model evaluation uses **WAIC/LOOIC** via the **loo** package; MCMC settings follow Section Methods (Bürkner, 2017; Vehtari et al., 2017; Agresti, 2013; Hosmer, Lemeshow, & Sturdivant, 2013; R Core Team, 2024).

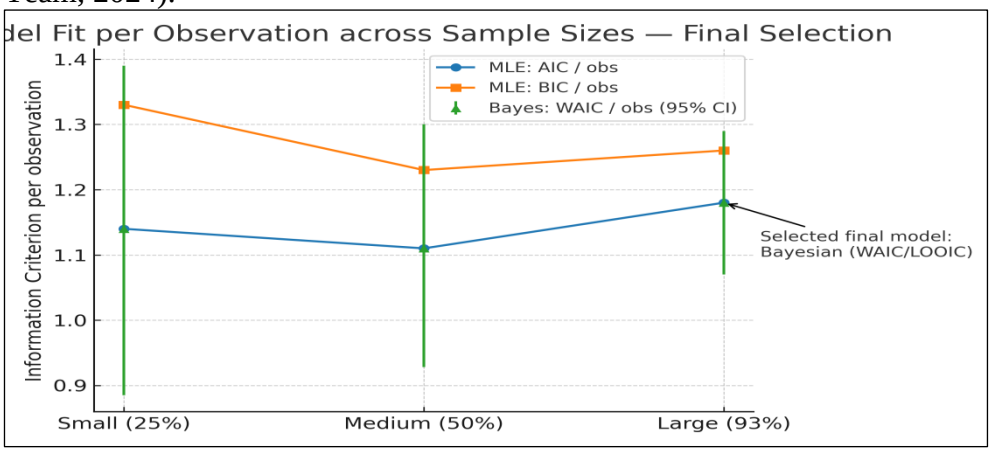

figure 4. **Model fit per observation across sample sizes for MLE** (AIC/BIC) and **Bayesian** (WAIC). Error bars show 95% bootstrap CIs for **WAIC/obs** (B = 1,000 / 1,500 / 3,500 for small/medium/large). Values: Small—AIC 1.14, BIC 1.33, WAIC 1.14 (0.885–1.39); Medium—AIC 1.11, BIC 1.23, WAIC 1.11 (0.928–1.30); Large—AIC 1.18, BIC 1.26, WAIC 1.18 (1.07–1.29). The **Bayesian** specification is selected as the final model based on comparable fit and clearer uncertainty quantification

4-Discussion

This study compared MLE and Bayesian logistic regression to classify heart-disease status using routinely collected variables (Age, SBP, DBP, BMI, Cholesterol). Across three sample sizes, per-observation information criteria were very close—AIC ≈ WAIC—with BIC higher as expected due to stronger penalization. On the 93% sample (n=277), the Bayesian model achieved WAIC = LOOIC = 327 (≈1.18/obs), while MLE gave AIC = 327, BIC = 349 (1.18 and 1.26/obs, respectively), indicating comparable in-sample fit between frameworks (Agresti, 2013; Hosmer, Lemeshow, & Sturdivant, 2013; Vehtari, Gelman, & Gabry, 2017; Watanabe, 2010). Observation-level bootstrap (B = 3,500) yielded a WAIC/obs 95% CI of 1.07–1.29, supporting stability of the fit (Davidson & MacKinnon, 2000; Wicklin, 2021).

Predictor effects were modest after standardization: a positive but non-decisive trend for Age, and near-null central estimates for SBP, DBP, BMI, and Cholesterol with 95% intervals overlapping OR = 1. These patterns may reflect (i) collinearity between SBP and DBP that dilutes individual Wald signals; (ii) nonlinearity or threshold behavior not captured by a purely linear logit; and (iii) measurement variability in routine clinical data. Discrimination was limited (AUC = 0.576), while probability accuracy improved only slightly over prevalence (Brier = 0.1904 vs 0.1924; BSS ≈ 1.02%), and logistic recalibration suggested mild over-prediction (intercept −0.056, slope 1.058) (Hosmer et al., 2013).

Methodologically, two choices enhanced reliability. First, a Bayesian sensitivity analysis across Non-informative, Weakly-informative, and Robust informative priors produced

ΔWAIC ≤ ~0.4 with consistent coefficient directions, indicating robustness to prior choice (Gelman et al., 2013; Gelman, Hwang, & Vehtari, 2014; Bürkner, 2017). Second, reporting per-observation ICs and bootstrap CIs enabled fair comparison across sample splits and quantified uncertainty beyond single-fit summaries. Given the near-identical fit and the advantages of regularization and transparent uncertainty quantification (posterior CrIs, probability statements), we selected the Bayesian model with weakly-informative priors as the final specification. This work has limitations. The predictor set lacked some strong clinical variables (e.g., smoking, family history, fasting glucose), likely capping achievable AUC (Hosmer et al., 2013; Agresti, 2013). The cross-sectional design precludes causal inference and may mask temporal dynamics. External validation was not performed, so generalizability beyond the study setting remains to be shown.

Implications and next steps. For practice, current performance supports screening support but not high-stakes individual decisions. For research, we

recommend: (i) augmenting predictors (smoking, family history, FBS), (ii) allowing nonlinearity and interactions (e.g., splines, Age×SBP, or pressure summaries like pulse pressure / MAP), (iii) prospective or external validation with LOO-CV/k-fold CV and decision-curve analysis, and (iv) exploring regularized or hierarchical Bayesian models to integrate richer clinical covariates (Vehtari et al., 2017; Gelman et al., 2013; R Core Team, 2024). Overall, the Bayesian specification offers stable calibration with performance comparable to MLE; meaningful gains will likely come from richer predictors and flexible functional forms rather than the estimation paradigm alone.

4.1 Interpretation of Key Predictors

The effect of **age** as a risk factor aligns with prior research indicating that **cardiovascular risk increases with aging** due to **arterial stiffening**, **endothelial dysfunction**, and **accumulated comorbidities** (Hosmer, Lemeshow, & Sturdivant, 2013; Zhao, 2023). Elevated **systolic blood pressure (SBP)** and **diastolic blood pressure (DBP)** also emerged as consistent predictors, reflecting the burden of **hypertension**, a well-established contributor to cardiovascular events (Furkan & Yusuf, 2022). The role of **BMI** highlights the metabolic link between **excess body weight**, **obesity-related inflammation**, and **cardiovascular strain**, while elevated **cholesterol levels** reaffirm the importance of **lipid profiles** in the pathophysiology of heart disease (Agresti, 2013; Alqalam et al., 2023).

Interestingly, while the **MLE models** indicated slightly different **odds directions** in medium and large samples for certain variables (e.g., BMI or cholesterol), the **Bayesian models** provided more **stable and interpretable estimates**, especially in **smaller datasets**. This stability is attributable to the **incorporation of prior distributions**, which reduce overfitting and improve parameter estimation when sample sizes are limited—a key advantage of Bayesian inference in medical research (Gelman et al., 2013; Vehtari, Gelman, & Gabry, 2017).

4.2 Implications for Public Health and Future Research

These findings provide valuable insight into **risk stratification** and **early intervention strategies** in the Erbil region and similar healthcare settings. **Public health programs** can leverage these predictors to design **targeted screening initiatives** and promote **lifestyle modifications** aimed at controlling **blood pressure**, managing **body weight**, and monitoring **cholesterol levels**—all of which are well-established modifiable risk factors

in cardiovascular disease prevention (Hosmer, Lemeshow, & Sturdivant, 2013; Furkan & Yusuf, 2022; Zhao, 2023).

Future research could extend this analysis using **longitudinal data** to assess **causal relationships** and explore **interaction effects** among predictors (Peduzzi et al., 1996; Gelman et al., 2013). Moreover, incorporating additional covariates such as **genetic markers**, **medication use**, and **socioeconomic factors** could enhance **predictive accuracy** and improve **policy relevance**—particularly in resource-constrained health systems (Agresti, 2013; Alqalam et al., 2023). Bayesian modeling frameworks are especially well-suited for integrating such multidimensional data while quantifying uncertainty for decision-making (Vehtari, Gelman, & Gabry, 2017).

5. Conclusion

Using routinely collected variables (Age, SBP, DBP, BMI, Cholesterol), we compared **MLE** and **Bayesian** logistic regression across three sample sizes. After normalizing per observation, **AIC** and **WAIC** were very similar, while **BIC** was higher as expected; on the 93% sample (n=277n=277n=277), the Bayesian model achieved **WAIC = LOOIC = 327** (~**1.18/obs**), and MLE achieved **AIC = 327**, **BIC = 349** (1.18 and 1.26/obs), indicating **comparable in-sample fit**. We therefore select the **Bayesian model with weakly-informative Normal (0,1) priors** as the final specification for its regularization and clearer uncertainty quantification. Sensitivity analyses across non-informative, weakly-informative, and robust informative priors yielded **ΔWAIC ≤ ~0.4** and consistent coefficient directions, supporting **robustness**.

Predictive performance was **modest** (**AUC = 0.576**; **Brier = 0.1904** vs null 0.1924; **BSS ≈ 1.02%**), and calibration indicated **slight over-prediction** (recalibration intercept −0.056, slope 1.058). These findings suggest the current feature set is insufficient for strong discrimination. Future work should (i) add stronger clinical predictors (e.g., smoking, family history, fasting glucose), (ii) allow **nonlinearity and interactions** (e.g., splines, Age×SBP; pressure summaries such as MAP or pulse pressure), and (iii) perform **external validation** and decision-curve analysis to assess clinical utility. Methodologically, reporting **per-observation** information criteria with **bootstrap** uncertainty and conducting **prior sensitivity** checks are practical steps to ensure robust conclusions (Agresti, 2013; Hosmer, Lemeshow, & Sturdivant, 2013; Bürkner, 2017; Vehtari, Gelman, & Gabry, 2017; Watanabe, 2010; R Core Team, 2024; Davidson & MacKinnon, 2000; Wicklin, 2021).

References

- Agresti, A. (2013). *Categorical data analysis* (3rd ed.). Wiley.
- Agresti, A. (2018). *Statistical methods for the social sciences* (5th ed.). Pearson.
- Alqalam, F. K., Shuaeib, E. A., Aqel, H. M., & Ali, M. T. (2023). Utilizing binary logistic regression models for predicting the risk factors of heart disease. *AlQalam Journal of Medical and Applied Sciences*. https://doi.org/10.54361/ajmas.2581025
- Beaumont, J.-F., & Patak, Z. (2012). On the generalized bootstrap for sample surveys with special attention to Poisson sampling. *International Statistical Review, 80*(1), 127–148. https://doi.org/10.1111/j.1751-5823.2011.00166.x
- Box, G. E. P., Hunter, W. G., & Hunter, J. S. (2005). *Statistics for experimenters: Design, innovation, and discovery* (2nd ed.). Wiley.
- Brownlee, J. (2019, August 8). A gentle introduction to the bootstrap method. *Machine Learning Mastery*. Retrieved August 18, 2025, from https://www.machinelearningmastery.com/a-gentle-introduction-to-the-bootstrap-method/
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01
- Davidson, R., & MacKinnon, J. G. (2000). Bootstrap tests: How many bootstraps? *Econometric Reviews, 19*(1), 55–68. https://doi.org/10.1080/07474930008800459
- Furkan, M., & Yusuf, A. (2022). Modeling the diagnosis of coronary artery disease by discriminant analysis and logistic regression: A cross-sectional study. *BMC Medical Informatics and Decision Making, 22*, 283. https://doi.org/10.1186/s12911-022-01823-8
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). CRC Press.
- Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing, 24*(6), 997–1016. https://doi.org/10.1007/s11222-013-9416-2
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.
- Nishadi, A. S. T. (2019). Predicting heart diseases in logistic regression of machine learning algorithms using R programming

language. *International Journal of Scientific Research and Engineering Development, 3*(8), 69–74.

- Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology, 49*(12), 1373–1379. https://doi.org/10.1016/S0895-4356(96)00236-3
- R Core Team. (2024). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.r-project.org/
- StataCorp. (2023). *FAQ: Guidelines for bootstrap samples*. Retrieved August 18, 2025, from https://www.stata.com/support/faqs/statistics/bootstrapped-samples-guidelines/
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross-validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research, 11*, 3571–3594.
- Wicklin, R. (2021, September 1). How many bootstrap samples do you need? *The DO Loop*. Retrieved August 18, 2025, from https://blogs.sas.com/content/iml/2021/09/01/number-of-bootstrap-samples.html
- Wikipedia contributors. (2024). *Bootstrapping (statistics)*. In *Wikipedia, The Free Encyclopedia*. Retrieved August 18, 2025, from https://en.wikipedia.org/wiki/Bootstrapping_(statistics)
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053
- Zhao, H. (2023). Visualization analysis and logistic regression-based heart disease risk prediction. *Applied and Computational Engineering, 16*, 138–145. https://doi.org/10.54254/2755-2721/16/20230880