



Available online at <http://scik.org>

Commun. Math. Biol. Neurosci. 2026, 2026:78

<https://doi.org/10.28919/cmbn/9936>

ISSN: 2052-2541

A BAYESIAN BINARY QUANTILE REGRESSION APPROACH WITH ADAPTIVE LASSO FOR HYPERTENSION MODELING: EVIDENCE FROM BOOTSTRAP CONSISTENCY ANALYSIS

LILIS HARIANTI HASIBUAN^{1,2}, FERRA YANUAR^{1,*}, DODI DEVIANTO¹, MAIYASTRI¹

¹ Department of Mathematics and Data Sciences, Andalas University, Padang, Indonesia

² Department of Mathematics, Universitas Islam Negeri Imam Bonjol, Padang, Indonesia

Copyright © 2026 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: Hypertension remains one of the most prevalent and serious non-communicable disease worldwide due to its strong association with cardiovascular, and renal complications. This study aimed to develop a robust statistical model to identify significant predictors of hypertension status using the Bayesian Binary Quantile Regression with Adaptive Lasso (BBALQR) method. Given the complexity of health data, BBALQR offers a flexible and robust alternative to traditional models, especially in the presence of outliers and binary responses. The dataset comprises 653 patient records from Arosuka Hospital in Solok, West Sumatra, Indonesia. The proposed model was evaluated across five quantile levels $\tau = 0.05, 0.25, 0.55, 0.75$ and, 0.95 , and parameter stability was assessed through Bootstrap resampling with 10, 25, and 50 replications. Among the tested quantiles, $\tau = 0.05$ yielded the lowest mean squared error (MSE), and greater Accuracy value and Press's Q, indicating the best model performance. Furthermore, Bootstrap analysis confirmed the consistency and reliability of parameter estimates within the 95% confidence interval. Significant predictors of hypertension include age, body weight, cholesterol, and blood sugar levels. The findings

*Corresponding author

E-mail address: ferrayanuar@sci.unand.ac.id

Received April 21, 2026

suggest that BBALQR, combined with Bootstrap validation, offers a reliable modeling framework for binary health outcomes with non-normal and heteroscedastic data characteristics.

Keywords: Bayesian; quantile regression; adaptive lasso; bootstrap

2020 AMS Subject Classification: 62F15.

1. INTRODUCTION

Hypertension remains one of the most prevalent and serious non-communicable disease worldwide due to its strong association with cardiovascular, and renal complications [1]. It is often termed the “silent killer” because of its typically asymptomatic progression, leading to high rates of undetected and uncontrolled cases. In Indonesia, the prevalence of hypertension among adults has risen steadily, reaching 30.8% in 2023, with certain regions such as West Sumatra reporting rates as high as 35.6% [2]. This growing trend underscores the urgent need for robust statistical models to assess and predict hypertension risk accurately across diverse populations. Conventional regression methods, such as Ordinary Least Squares (OLS), assume normality, homoscedasticity, and the absence of multicollinearity [3],[4], [5]. However, these assumptions are often violated in health data, especially when the response variable is binary or affected by outliers. Quantile regression, introduced [6],[7] provides a more flexible framework by estimating conditional quantiles of the response variable, enabling deeper insights into data heterogeneity[8]. Yet, classical quantile regression is not designed for binary outcomes and does not include built in variable selection mechanisms.

In recent years, Bayesian has been widely used to model response variables with various distributions [9], such as the inverse of an asymmetric exponential [9], Asymmetric Laplace Distribution [10], binary quantile regression has been extended using a Bayesian framework, allowing for the modelling of binary outcomes via latent continuous variables and the incorporation of prior knowledge [11]. To address the challenge of high dimensional predictors and the need for robust variable selection, the Least Absolute Shrinkage and Selection Operator (LASSO) has been adapted to Bayesian quantile models by [10]. This study employs the Bayesian

Binary Quantile Regression with Adaptive Lasso (BBALQR), which enhances both estimation precision and sparsity in high dimensional settings [11].

To evaluate the consistency and stability of parameter estimates, we apply the Bootstrap method [12] a resampling based technique that estimates the sampling distribution of a statistic without relying on parametric assumptions. By combining BBALQR with Bootstrap validation, this study aims to produce a statistically sound and practically relevant model for hypertension prediction. Specifically, our objectives are: (1) to model hypertension status using BBALQR, and (2) to assess the robustness of the parameter estimates through Bootstrap based consistency analysis. The dataset, consisting of 653 patient records from Arosuka Hospital in Solok, Indonesia, provides a contextual application of the proposed method in a real-world clinical setting. Analysis using this method is rarely performed on health data, and the combination of several methods to obtain a more robust estimate, as well as consistency using Bayesian Bootstrap, is still rarely done.

2. PRELIMINARIES

2.1. Data set

The data used in this study are secondary data taken from the polyclinic of Arosuka Solok Hospital, West Sumatra in 2024 as many as 653 patients. The variables used in this study consist of dependent variables (Y), namely hypertension status with Binary categories that are dichotomous with two categories (hypertension = 1) and (not hypertension = 0) Then the independent variables are factors that are assumed to affect the status of hypertension, including gender (X_{1D1}) with category male and female by [13], [14] Age (X_2) in years [15], smoking (X_{3D1}) with category smoking and not smoking [16], Body weight (X_4) in Kg [17]; Cholesterol (X_5) in decilitre (mg/dL) [18]; [19]; [20], Triglyceride (X_6) in decilitre (mg/dL) [12] and Blood Sugar levels (X_7) in mg% [17].

The independent variable of gender, smoking is an independent variable with a type of category so that the independent variable is first converted into a dummy variable. The process of forming research dummy variables is presented in Table 1 below.

Table 1. Variable Dummy

Gender (X_1)	(X_{1D1})
Male	1
Female	0
Smoking Status (X_3)	(X_{3D1})
Smoking	1
Not Smoking	0

Table 1 informs that 43% were male and 36% were smoking. Table 2 below presents a description of the categorical data used in this study as independent variables.

Table 2. Statistics Descriptive of Categorical Variables

Variable	Category	Frequency	Percentage (%)
Gender (X_1)	Male	281	43%
	Female	372	57%
Smoking (X_3)	Smoking	235	36%
	Not Smoking	318	64%

Furthermore, information on independent variables with numerical data is presented in Table 3 below.

Table 3. Statistics Descriptive of Numerical Variables

Variable	Mean	Min	Max
Age (X_2)	51.8	12	94
Body Weight (X_4)	60.5	26	96
Cholesterol X_5	213	173	385
Triglyceride X_6	115.5	34	433
Blood Sugar Levels X_7	129.8	70	370

2.2. Quantile Regression

Given the general equation for quantile regression model [3]:

$$y_i = x_i' \beta + e_i, \quad (1)$$

With y_i is the dependent variable at the i^{th} sampel, x_i is a $k \times 1$ vector of independent variables (possibly, the first term of x_i is 1 in case an intercept is needed), β is a $k \times 1$ is a vector of coefficient variables, and e_i is the error term. The above equation (1) is made equal to zero, that

is $\int_{-\infty}^0 f_p(e_i) de_i = \tau, i = 1, \dots, n$. The study conducted [6] shows that the estimation of coefficient variables in the quantile regression equation $\widehat{\beta}_\tau$ is obtained by minimizing the equation presented [21], [22]:

$$\sum_{i=1}^n \rho_\tau(y_i - x_i' \beta), \quad (2)$$

with $\rho_\tau(\cdot)$ is the test function defined by:

$$\rho_\tau(t) = \frac{|t| - (2\tau - 1)t}{2}. \quad (3)$$

Since the test function (2) is not differentiable at the origin, there is no solution for the coefficient vector β . However, minimizing (2) can be computed by the algorithm found in [23]. The test function (2) is directly related to the Asymmetric Laplace Distribution (ALD). Random variable e with ALD distribution with density functions [24]:

$$f_\tau(e) = \tau(1 - \tau) \exp(-\rho_\tau(e)). \quad (4)$$

With σ could be a scale parameter, p decides the quantile level and ρ is the test function as in condition (2). Minimizing the the loss function (2) can be accomplished by maximizing the likelihood function (4). The ALD distribution is one of the continuous probability distributions. Suppose Z is a random variable with exponential distribution ($Z \sim \exp(1)$) and U is a random variable with standard normal distribution $U \sim N(0,1)$. If e is an ALD distributed random variable then e can be expressed in the following equation:

$$e = \theta z + pu\sqrt{z} \quad (5)$$

where $\theta = \frac{1-2\tau}{(1-\tau)\tau}$ and $p^2 = \frac{2}{(1-\tau)\tau}$. Based on equation (5) the likelihood function used in parameter β estimation for the τ^{th} quantile in the Bayesian quantile regression analysis

2.3. Bayesian Binary Quantile Regression Adaptive Lasso (BBALQR)

Next, we will discuss the combination of Bayesian quantile regression by adding a regulator parameter, namely adaptive lasso. The idea of this study is a continuation of the results of the study developed by [10]; [3] who used the lasso regulator parameter. The model that will be built by

Bayesian Adaptive Lasso binary quantile regression goes through two stages, the first is modelling with dichotomous response data or often called binary quantile regression. Then the second estimates the value of the parameters that build the model. Adaptive Lasso has an advantage over lasso because Adaptive Lasso adjusts each parameter. From this, we believe that the model obtained from the combination of Bayesian binary quantile regression and Adaptive Lasso accurate. Model of Binary regression is written as follows [25], [26]:

$$y_i^* = x_i' \beta + e_i,$$

$$y_i = \begin{cases} 1, & \text{if } y_i^* \geq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

With y_i is the observed data respon of i^{th} sampel determined by the latent unobserved response y_i^* .

Discussions related to binary quantile regression have been discussed in several studies, namely [27]; [10] which can be used as a guide. The most important point from the above discussion is that since quantile regression can model with non-normal errors, in this case binary quantile regression handles response data that is clearly non-normal because the data is binary. In addition, the Bayesian approach is strongly supported because the distribution of the initial data can be used as information to obtain more robust predictor variable coefficients [9]. Moreover, the Bayesian approach can deal with a high dimensional predictor space and this rarely the case for optimization algorithm that are normally used for frequentist quantile regression. A recent exception here is [28]. Other frequentist binary quantile regression can be found [29],[30].

However, all these frequentists are difficult to optimize because the absence of first order conditions that can be exploited. [29] tried to solve the difficult optimization problem [27] by smoothing the objective function, but then ran into difficulties concerning tuning band width parameter. In the context of variable selection, however the latter is crucial. In many studies the goal is to find a small set of relevant variables. The variable selections approach proposed adaptive least absolute shrinkage and selection operator Adaptive Lasso together for binary quantile regression model are ideally suited to tackle the problem of predictor dimension reduction in binary data sets.

From a mathematical standpoint, the Adaptive Lasso estimates of the binary quantile regression coefficient can be computed by [31], [32]:

$$\beta_{ALASSO} = \min_{\beta \in \mathbb{R}} \sum_{i=1}^n \rho_{\tau}(y_i^* - x_i' \beta) + \sum_{j=1}^k \lambda_j |\beta_j|, \quad (7)$$

where λ_i is a non-negative variable penalty coefficient, $i = 1, 2, \dots, k$. Based on equation (5) the likelihood function used in parameter β estimation for the τ^{th} quantile in the Bayesian quantile regression analysis is formulated in equation (8) as follows:

$$L(y_i^* | \beta, \sigma, v) = (\prod_{i=1}^n (\sigma v_i)^{-\frac{1}{2}}) \left(\exp \left(-\frac{(y_i^* - (x_i' \beta_{\tau} + \theta v_i))^2}{2p^2 \sigma v_i} \right) \right). \quad (8)$$

The fully conditional distribution of y_i^* is a mixture of two truncated normal distributions [10]:

$$y_i^* | \beta, \sigma, v = \begin{cases} N(x_i' \beta_{\tau} + \theta v_i, p^2 \sigma v_i) I(y_i^* > 0), & \text{if } y_i = 1 \\ N(x_i' \beta_{\tau} + \theta v_i, p^2 \sigma v_i) I(y_i^* \leq 0), & \text{otherwise} \end{cases}. \quad (9)$$

In Bayesian analysis, we need a prior distribution to construct parameter estimates. Prior distribution for $\beta_{\tau}, \lambda_j^2, \zeta, \sigma, s, v, \delta$ used in Bayesian Adaptive Lasso binary quantile regression is as follows [11]:

$$\begin{aligned} f(\beta | \lambda_j^2, s_j) &= \prod_{j=1}^k \int_0^{\infty} \frac{1}{\sqrt{2\pi s_j}} \exp \left(-\frac{\beta_j^2}{2s_j} \right) \frac{\lambda_j^2}{2} \exp \left(-\frac{\lambda_j^2}{2} s_j \right) ds_j, \\ f(\zeta | \delta) &= 1, \\ f(\sigma) &= \sigma^{a_1-1} \exp(-a_2 \sigma), \\ f(s_j | \lambda_j^2) &= \frac{\lambda_j^2}{2} \exp \left(-\frac{\lambda_j^2}{2} s_j \right), \\ f(v_i | \sigma) &= \sigma \exp(-v_i \sigma), \\ f(\delta | \zeta, \lambda_j^2) &= \frac{(\zeta \lambda_j^2)^{\delta}}{\Gamma(\delta)} \end{aligned} \quad (10)$$

with $s = (s_1, \dots, s_k)$, $i = 1, 2, \dots, k$, $v = (v_1, \dots, v_n)$, $\sigma > 0$, $a_1 > 0$, $a_2 > 0$, $\lambda_j^2 \geq 0$, $\zeta > 0$, $\delta > 0$. According to Bayesian theorem, the posterior distribution is the product of the likelihood function and the prior distribution. Therefore referring to [10]:

$$f(\beta | y_i^*, s_j, v_i, \sigma) \sim N(\bar{\beta}_j, \bar{\sigma}_j),$$

$$\begin{aligned}
& \text{with, } \bar{\sigma}_j^2 = \left(\sigma \sum_1^n \frac{x_{ij}^2}{2v_i} + \frac{1}{s_j} \right)^{-1}, \quad \bar{\beta}_j = \bar{\sigma}_j^2 \sigma \sum_{i=1}^n \frac{(y_i^* - \sum_{l \neq j} x_{il} \beta_l - \zeta v_i)}{2v_i} \\
& f((\sigma | y_i^*, s_j, v_i, \sigma) \sim \text{Gamma} \left(\frac{3}{2}n + a_1, \sum_{i=1}^n \left(\frac{(y_i^* - \zeta v_i)^2}{4v_i} + p(1-p)v_i + b_1 \right) \right), \\
& f(\lambda_j^2 | s_j, \sigma) \sim \text{Gamma}(1 + a_0, \frac{s_j \cdot \sigma}{2} + b_0), \\
& f(s_j | \lambda_j^2, \beta, \sigma) \sim \text{GIG} \left(\frac{1}{2} + \beta_j^2, \sigma \lambda_j^2 \right), \\
& f(v^{-1} | y_i^*, \beta, \sigma) \sim \text{GIG} \left(\frac{1}{2}, \frac{\sigma}{2}, \frac{\sigma(y_i^* - x_i' \beta)^2}{2} \right), \\
& f(s_j | y_i^*, \beta, \sigma) \sim \text{GIG} \left(\frac{1}{2}, \frac{\sigma}{2}, \frac{\sigma(y_i^* - x_i' \beta)^2}{2} \right).
\end{aligned} \tag{11}$$

The multiple roots method suggested by [33]; [22] will be utilized to generate sample, and the restriction condition $\|\beta\|_1$ will be imposed value. The steps are follows: initiate the process with a preliminary value $\beta, v, \sigma, \lambda_j^2 (j = 1, 2, \dots, k)$; From the following source $f(y_i^* | y, \beta, v_i, \sigma)$ generate $y_i^*, i = 1, 2, \dots, n$; from the following source $f(v^{-1} | y_i^*, \beta, \sigma)$ generate $v^{-1}, i = 1, 2, \dots, n$; from the following source $f((\sigma | y_i^*, s_j, v_i, \sigma)$ generate σ ; from $f(\beta_j | y_i^*, s_j, v_i, \sigma)$ generate β_j ; and the constraint condition $\beta = 1, j = 1, 2, \dots, k$; from $f(s_j | \lambda_j^2, \beta, \sigma)$ generate $s_j, j = 1, 2, \dots, k$; from $f(\lambda_j^2 | s_j, \sigma)$ generate $\lambda_j^2, j = 1, 2, \dots, k$; return to the second to seventh stage.

2.4. Metode Bootstrap

The model that has been constructed using the BBALQR method will be analysed for the consistency of the coefficients of the predictor variables using the Bootstrap method. Bootstrapping is resampling approach that allows estimating the probabilistic properties of the target statistic by randomly sampling with replacement from data [34];[35].

Estimate each parameter value of Bayesian binary quantile regression, Bayesian Lasso binary quantile regression and Bayesian Adaptive Lasso binary quantile regression at quantiles 0.05, 0.25, 0.55, 0.75, and 0.95 using the Bootstrap method with the number of replications used being 10 times, 25 times and 50 times.

Bootstrap sample formation can be done with the following steps. The first, constructs samples from paired data (x_n, y_n) randomly drawn with probability $\frac{1}{n}$. Bootstrap samples are considered as random samples of size n . The data obtained is the original data coming from the population with repetition. Second, Modelling Bayesian Adaptive Lasso binary quantile regression. Based on the hypothesized models, the parameters $\widehat{\beta}_j(\tau)$ will be estimated for each quantile τ or denoted by $\widehat{\beta}_j(\tau)$ with $j = 1, 2, \dots, p$. Third, Repeat steps first and second B boots so as to obtain B estimated parameter values or $\widehat{\beta}_{b,j}(\tau)$ with $b = 1, 2, \dots, B$.

After that proceed (first, second, and third) continued to calculate the statistical value for each replication of the Bootstrap method. Calculate the mean and variance for the Bootstrap method statistics $\widehat{\beta}_b(\tau)$. The mean of the Bootstrap parameter is:

$$\overline{\widehat{\beta}_b}(\tau) = \frac{1}{B} \sum_{b=1}^B \widehat{\beta}_{b,j}(\tau). \quad (12)$$

As for the variance of the Bootstrap method statistics, namely:

$$Var \widehat{\beta}_j(\tau) = \frac{1}{B-1} \sum_{b=1}^B (\widehat{\beta}_{b,j}(\tau) - \overline{\widehat{\beta}_b}(\tau))^2, \quad (13)$$

where $j = 1, 2, \dots, p$. The final step is to determine the Bootstrap confidence interval for each replication.

3. MAIN RESULTS

In this section we will estimate the model parameters using BBALQR method. The resulting model is as much as the selected quantile. The selected quantile is 0.05; 0.25; 0.55; 0.75; 0.95. In this study, the model hypothesis was presented in hypertension status as follows:

$$\begin{aligned} \text{Logit}(\text{Hypertension Status}_i) & \\ &= \beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Age}_i + \beta_3 \text{Smoking}_i + \beta_4 \text{Body Weight}_i \\ &\quad + \beta_5 \text{Cholesterol}_i + \beta_6 \text{Triglyceride}_i + \beta_7 \text{Blood sugar levels}_i. \end{aligned} \quad (14)$$

Estimation results of model parameter estimation using Bayesian Binary Adaptive Lasso Quantile Regression (BBALQR) present in Table 4 below.

Table 4. Coefficient estimated for Hypertension Status model using BBALQR

Independent Variable	Estimate Parameter of BBALQR in quantile				
	$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.55$	$\tau = 0.75$	$\tau = 0.95$
β_1 (Gender)	0.369(8.163)	-2.441(16.920)	-1.378(6.795)	2.516(7.036)	2.637(6.042)
β_2 (Age)	0.859 (0.986)*	0.864 (0.991)*	1.088 (0.611)*	0.608 (0.984)*	0.608 (0.984)*
β_3 (Smoking)	8.610(3.015)	1.124(8.817)	0.711(6.795)	13.118(21.223)	1.379(3.266)
β_4 (Body weight)	0.716 (0.769)*	0.847 (0.927)*	0.575 (0.638)*	0.716 (0.547)*	0.804 (0.354)*
β_5 (Cholesterol)	0.254 (0.907)*	0.319 (0.930)	0.468(0.825)	0.567 (0.967)	0.616(0.915)*
β_6 (Triglyceride)	0.140 (0.780)	0.403(0.941)	0.699 (0.748)*	0.721 (0.753)*	0.709 (0.758)*
β_7 (Blood sugar levels)	0.491 (1.019)*	0.755 (0.944)*	0.989 (0.206)*	1.073(0.519)*	1.094(0.236)*
*Significant at 5% level					

Based on Table 4, the independent variables that affect hypertension status are age, body weight, cholesterol, triglyceride, and blood sugar levels. Age variable is significant in all quantiles. Body weight variable is significant in all quantiles. Cholesterol variable is significant in 0.05 and 0.95 quantiles. Triglyceride variable is significant in quantile 0.55, 0.75, and 0.95. Blood Sugar Levels variable is significant in all quantiles. Furthermore, the significant independent variables are used as a reference in estimating parameters with the Bayesian Adaptive Lasso binary quantile regression (BBALQR) method. The results of parameter estimation with the BBALQR Stage 2 method. The results of parameter estimation with BBALQR Stage 2 are presented in the following Table 5 below:

Table 5. Coefficient estimated for Hypertension Status model using BBALQR Stage 2

Independent Variable	Estimate Parameter of BBALQR in quantile				
	$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.55$	$\tau = 0.75$	$\tau = 0.95$
β_2 (Age)	0.838 (0.876)*	0.853 (0.991)*	1.078 (0.621)*	0.618 (0.884)*	0.607 (0.974)*
β_4 (Body weight)	0.715 (0.759)*	0.826 (0.918)*	0.548 (0.629)*	0.706 (0.557)*	0.810 (0.352)*
β_5 (Cholesterol)	0.253 (0.904)*	0.316 (0.929)	0.480(0.818)*	0.557 (0.962)	0.615(0.913)*
β_6 (Triglyceride)	0.141 (0.781)	0.413(0.951)	0.697 (0.746)*	0.720 (0.751)*	0.718 (0.749)*
β_7 (Blood sugar levels)	0.481 (1.009)*	0.765 (0.954)*	0.981 (0.202)*	1.083(0.529)*	1.084(0.136)*
*Significant at 5% level					

Based on Table 5, it is known that at quantile 0.05 the variables of age, body weight, cholesterol and blood sugar levels have a statistically significant effect on hypertension status. Furthermore, at quantile 0.25 the variables of age, body weight and blood sugar level have a

statistically significant effect on hypertension status. Furthermore, at quantiles 0.55 and 0.75 the variables age, body weight and blood sugar levels have a significant effect on hypertension status. Finally at quantile 0.95 all independent variables are statistically significant in influencing hypertension status. The significance level used is 5%. After modelling using BBALQR, we will continue to test the goodness of the resulting model by calculating the Mean Square Error (MSE) value. The results of the MSE calculation for each selected quantile are shown in Table 6 below.

Table 6. Mean Square Error (MSE) for BBALQR selected quantile

Quantile	MSE
$\tau = 0.05$	0.282
$\tau = 0.25$	0.530
$\tau = 0.55$	0.402
$\tau = 0.75$	0.430
$\tau = 0.95$	0.435

Table 6 shows that the 0.05 quantile produces the smallest MSE value compared to other quantiles. It can be concluded that the model at quantile 0.05 is the best model in estimating parameters with the BBALQR method. Thus, the conjecture model for parameter estimation of the BBALQR model at quantile 0.05 is as follows:

$$\text{Logit}(y_i) = 0.838X_2 + 0.715X_4 + 0.253X_5 + 0.481X_7. \quad (15)$$

Based on the equation (14) above, it can be interpreted by looking at the odds ratio value for each independent variable. The odds ratio values are presented in Table 7 below.

Table 7. Odd Ratio Value of Independent Variable

Independent Variables	Mean $(\hat{\beta})$	Odds Ratio
Age (X_2)	0.838	2.312
Body weight (X_4)	0.715	2.044
Cholesterol (X_5)	0.253	1.287
Blood sugar levels (X_7)	0.481	1.617

Based on Table 7, it can be interpreted that for every one year's increase in age, the probability of developing hypertension becomes 2.312 greater than the previous one, assuming other variables are constant. Every one kg increase in body weight, the probability of developing hypertension becomes 2,044 greater than the previous one, assuming other variables are constant. Every 1 mg/dL increase in cholesterol, the probability of having hypertension becomes 1,287 greater than the previous one, assuming other variables are constant. Then every 1mg% increase

in blood sugar levels, the probability of having hypertension becomes 1,617 greater than the previous one, assuming other variables are constant.

Furthermore, the accuracy of the parameter values that have been generated from the Bayesian Adaptive Lasso Binary quantile regression method at each selected quantile is tested whether it has produced the true and acceptable value using Bootstrap. In estimating parameters using the Bootstrap method, the number of replications used is 10, 25 and 50 times. Then from each replication will be compared which replication is able to produce the best guess for modelling hypertension status. The criterion for this stage is the replication that produces the smallest Bootstrap confidence interval width. Then from each replication, it will be compared which replication is able to produce the best estimation model in modelling hypertension status. The criterion for this stage is the replication that produces the smallest Bootstrap confidence interval (CI) width. The BBALQR Bootstrap parameter estimation results for each quantile are presented in Tables 8, 9 10, 11 and 12.

Table 8. The Parameters Estimated and CI Width of 95% Using BBALQR Bootstrap Quantile 0.05

Independent Variables	Estimated Mean	Estimation <i>Bootstrap</i>	Bootstrap CI 95%		Bootstrap CI Width 95%
			Lower	Upper	
B =10 Replication					
Age (X_2)	0.859*	0.14141	-0.13328	0.63742	0.77068
Body weight (X_4)	0.716*	0.60375	0.26958	0.88701	0.61742
Cholesterol (X_5)	0.254*	0.38415	0.08696	0.92508	0.83812
Triglyceride (X_6)	0.140	0.34671	0.05935	0.94230	0.88295
Blood sugar levels (X_7)	0.491*	0.48183	0.09672	1.03563	0.93888
B =25 Replication					
Age (X_2)	0.859*	0.13265	-0.11109	0.62964	0.51885
Body weight (X_4)	0.716*	0.55059	0.30632	0.90830	0.60198
Cholesterol (X_5)	0.254*	0.38769	0.08448	0.91320	0.82872
Triglyceride (X_6)	0.140	0.33704	0.05752	0.89900	0.84148
Blood sugar levels (X_7)	0.491*	0.49502	0.09332	1.02208	0.92876
B =50 Replication					
Age (X_2)	0.859*	0.12130	-0.10185	0.62524	0.52338
Body weight (X_4)	0.716*	0.58428	0.296434	0.94359	0.64715
Cholesterol (X_5)	0.254*	0.38536	0.083492	0.92455	0.84106
Triglyceride (X_6)	0.140	0.32236	0.054874	0.91064	0.85576
Blood sugar levels (X_7)	0.491*	0.49058	0.101192	1.03348	0.93229

A BAYESIAN BINARY QUANTILE REGRESSION

Table 9. The Parameters Estimated and CI Width of 95% CI Using BBALQR Bootstrap Quantile 0.25

Independent Variables	Estimated	Estimation	Bootstrap CI 95%		Bootstrap CI
	Mean	<i>Bootstrap</i>	Lower	Upper	Width 95%
B =10 Replication					
Age (X_2)	0.864*	0.2756	-0.00686	0.96688	0.96001
Body weight (X_4)	0.847*	0.5859	0.32869	1.06090	0.73221
Cholesterol (X_5)	0.319	0.4769	0.08696	1.00448	0.91752
Triglyceride (X_6)	0.403	0.5131	0.08952	1.01230	0.92278
Blood sugar levels (X_7)	0.755*	0.6636	0.14368	1.05550	0.91182
B =25 Replication					
Age (X_2)	0.864*	0.25443	-0.01547	0.97190	0.95643
Body weight (X_4)	0.847*	0.61483	0.34264	1.03920	0.34264
Cholesterol (X_5)	0.319	0.47996	0.09884	1.00280	0.90396
Triglyceride (X_6)	0.403	0.48936	0.07839	0.98836	0.07839
Blood sugar levels (X_7)	0.755*	0.65678	0.12644	1.06116	0.12644
B =50 Replication					
Age (X_2)	0.864*	0.24956	-0.01462	0.97965	0.96503
Body weight (X_4)	0.847*	0.64643	0.33411	1.05363	0.71951
Cholesterol (X_5)	0.319	0.47162	0.08306	1.00186	0.91874
Triglyceride (X_6)	0.403	0.48645	0.07652	0.99825	0.92173
Blood sugar levels (X_7)	0.755*	0.63863	0.12041	1.04213	0.92178

Table 10. The Parameters Estimated and CI Width of 95% Using BBALQR Bootstrap Quantile 0.55

Independent Variables	Estimate d Mean	Estimation <i>Bootstrap</i>	Bootstrap CI 95%		Bootstrap CI Width 95%
			Lower	Upper	
B =10 Replication					
Age (X_2)	1.088*	0.38537	0.06053	1.0254	0.96487
Body weight (X_4)	0.575*	0.64610	0.40807	1.0950	0.68693
Cholesterol (X_5)	0.468	0.58031	0.13928	0.9940	0.85472
Triglyceride (X_6)	0.699*	0.66092	0.17568	1.0134	0.83772
Blood sugar levels (X_7)	0.989*	0.88736	0.39571	1.0738	0.67809
B =25 Replication					
Age (X_2)	1.088*	0.37203	0.06874	1.01919	0.95045
Body weight (X_4)	0.575*	0.68964	0.48195	1.08316	0.60121
Cholesterol (X_5)	0.468	0.58796	0.15062	1.00480	0.85419
Triglyceride (X_6)	0.699*	0.65018	0.19173	1.00952	0.81779
Blood sugar levels (X_7)	0.989*	0.88566	0.43319	1.07729	0.64410

B =50 Replication					
Age (X_2)	1.088*	0.36535	0.06531	1.02762	0.96231
Body weight (X_4)	0.575*	0.72270	0.46251	1.07284	0.61033
Cholesterol (X_5)	0.468	0.57716	0.14196	1.00288	0.86092
Triglyceride (X_6)	0.699*	0.64970	0.18013	1.01484	0.83471
Blood sugar levels (X_7)	0.989*	0.85713	0.40616	1.06031	0.65415

Table 11. The Parameters Estimated and CI Width of 95% Using BBALQR Bootstrap Quantile 0.75

Independent Variables	Estimated Mean	Estimation <i>Bootstrap</i>	Bootstrap CI 95%		Bootstrap CI Width 95%
			Lower	Upper	
B =10 Replication					
Age (X_2)	0.951*	0.41928	0.06892	1.02946	0.96053
Body weight (X_4)	0.716*	0.70055	0.41787	1.09069	0.67282
Cholesterol (X_5)	0.567	0.6168	0.16943	1.01051	0.84108
Triglyceride (X_6)	0.721*	0.70479	0.23365	1.01554	0.78189
Blood sugar levels (X_7)	1.073*	0.98047	0.67359	1.11372	0.44013
B =25 Replication					
Age (X_2)	0.951*	0.40666	0.10664	1.01335	0.90671
Body weight (X_4)	0.716*	0.74321	0.49807	1.03687	0.29179
Cholesterol (X_5)	0.567	0.63192	0.20627	0.98768	0.78140
Triglyceride (X_6)	0.721*	0.69266	0.26933	1.00565	0.73632
Blood sugar levels (X_7)	1.073*	0.98296	0.68840	1.11392	0.42552
B =50 Replication					
Age (X_2)	0.951*	0.40443	0.06491	1.03526	0.97035
Body weight (X_4)	0.716*	0.75832	0.50934	1.05932	0.54998
Cholesterol (X_5)	0.567	0.62719	0.18194	0.99630	0.81436
Triglyceride (X_6)	0.721*	0.70300	0.25209	1.01272	0.76063
Blood sugar levels (X_7)	1.073*	0.96424	0.67270	1.09825	0.42555

Table 12. The Parameters Estimated and CI Width of 95% Using BBALQR Bootstrap Quantile 0.95

Independent Variables	Estimated Mean	Estimation <i>Bootstrap</i>	Bootstrap CI 95%		Bootstrap CI Width 95%
B =10 Replication					
Age (X_2)	0.608*	0.38155	0.02815	1.0396	1.01145
Body weight (X_4)	0.804*	0.66605	0.32443	1.0609	0.73647
Cholesterol (X_5)	0.616*	0.62971	0.18504	1.1507	0.96568
Triglyceride (X_6)	0.709*	0.72582	0.27820	1.0277	0.74950
Blood sugar levels (X_7)	1.094*	1.10811	0.92834	1.2523	0.32396

A BAYESIAN BINARY QUANTILE REGRESSION

		B =25 Replication			
Age (X_2)	0.608*	0.35870	0.01516	1.00704	0.99188
Body weight (X_4)	0.804*	0.72048	-0.94328	1.17925	0.23660
Cholesterol (X_5)	0.616*	0.63873	0.14040	1.02420	0.88381
Triglyceride (X_6)	0.709*	0.71488	0.22857	1.02930	0.70072
Blood sugar levels (X_7)	1.094*	1.08366	0.89326	1.19409	0.30083
		B =50 Replication			
Age (X_2)	0.608*	0.36964	-0.00670	1.02762	1.02092
Body weight (X_4)	0.804*	0.74409	-0.26670	1.15342	0.88668
Cholesterol (X_5)	0.616*	0.63631	0.17024	1.07838	0.90812
Triglyceride (X_6)	0.709*	0.72487	0.26946	1.02769	0.75826
Blood sugar levels (X_7)	1.094*	1.05997	0.90918	1.22386	0.31437

Based on Tables 8, 9, 10, 11 and 12 it can be seen that the estimated values obtained using the BBALQR method have produced values that are almost the same as the values obtained using Bootstrap for all model parameters in each replication. Furthermore, the parameter estimation results using BBALQR for all model parameters are already in the Bootstrap confidence interval for each replication. Thus, it is concluded that the model parameter estimation values obtained using the BBALQR method have produced correct and acceptable values.

Furthermore, from the table above, it is obtained that the smallest Bootstrap BBALQR confidence interval width for all parameters in each selected quantile is at replication 25. Thus, it can be concluded that the best number of replications in each selected quantile for hypertension status data at Arosuka Solok Hospital, West Sumatra with the Bootstrap BBALQR method is at replication 25 times.

3. CONCLUSION

This study contributes to the growing body of statistical research addressing the complexities of modelling binary health outcomes by proposing the Bayesian Binary Adaptive Lasso Quantile Regression (BBALQR) method as an effective and flexible solution. The analysis, based on real-world data from Arosuka Hospital in Solok, Indonesia, demonstrates that BBALQR not only captures distributional heterogeneity across different quantiles of hypertension risk but also performs simultaneous variable selection through its adaptive shrinkage mechanism.

Among the five evaluated quantile levels, the model at the lower quantile $\tau = 0.05$ exhibited superior predictive performance, as evidenced by the lowest mean squared error (MSE). This finding suggests that early indicators of hypertension risk, particularly in patients at the tail end of the risk spectrum, can be effectively captured by focusing on extreme quantiles. Key predictors consistently identified across quantiles include age, body weight, cholesterol, and blood sugar levels, reinforcing their clinical relevance.

The incorporation of Bootstrap resampling further validated the robustness and stability of the parameter estimates. The Bootstrap confidence intervals confirmed that BBALQR yields statistically reliable results, with 25 replications offering the most efficient balance between precision and computational feasibility. These results enhance the methodological credibility of BBALQR as a data-driven decision-making tool in clinical epidemiology and health risk stratification.

Overall, the proposed BBALQR framework proves to be a statistically rigorous and practically applicable approach, particularly in settings characterized by binary outcomes, non-normality, and high-dimensional predictor spaces. Future research may expand this work by exploring hierarchical or multilevel BBALQR models to accommodate data with nested structures, or by integrating clinical priors in a more hierarchical Bayesian context. Additionally, comparative evaluations with alternative machine learning-based classifiers could further benchmark the practical utility of BBALQR in medical diagnostics and preventive health analytics.

ACKNOWLEDGEMENTS

This research was funded by *Penelitian Disertasi Doktor* (PDD) Batch I Universitas Andalas Grant No. 87/UN16.19/PT.01.03/PDD/2025.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

- [1] World Health Organization, Global Report on Hypertension 2025: High Stakes: Turning Evidence Into Action,

- World Health Organization, 2025. <https://www.who.int/publications/i/item/9789240115569>.
- [2] L.R. Putri, M. Azam, A.A. Nisa, A.I. Fibriana, P. Kanthawee, et al., Prevalence and Risk Factors of Hypertension among Young Adults: An Indonesian Basic Health Survey, *Open Public Health J.* 18 (2025), e18749445361291. <https://doi.org/10.2174/0118749445361291241129094132>.
- [3] L.H. Hasibuan, F. Yanuar, D. Devianto, M. Maiyastri, R. Rudiyanto, Comparison between Bayesian Quantile Regression and Bayesian Lasso Quantile Regression for Modeling Poverty Line with Presence of Heteroscedasticity in West Sumatra, *Barekeng J. Ilmu Mat. Terap.* 19 (2025), 1587-1596. <https://doi.org/10.30598/barekengvol19iss3pp1587-1596>.
- [4] L.H. Hasibuan, S. Musthofa, Penerapan Metode Regresi Linear Sederhana untuk Prediksi Harga Beras di Kota Padang, *JOSTECH* 2 (2022), 85-95. <https://doi.org/10.15548/jostech.v2i1.3802>.
- [5] A. Sholih, L. Hasibuan, I. Rianjaya, Analisis Regresi Logistik Ordinal terhadap Faktor-Faktor yang Mempengaruhi Predikat Kelulusan Mahasiswa Sarjana UIN Imam Bonjol Padang, *Math. Appl. J.* 6 (2024), 99-110. <https://doi.org/10.15548/map.v6i2.10075>.
- [6] R. Koenker, G. Bassett, Regression Quantiles, *Econometrica* 46 (1978), 33-50. <https://doi.org/10.2307/1913643>.
- [7] L. Hasibuan, F. Yanuar, D. Devianto, M. Maiyastri, Modeling Classification of Stunting Toddler Height Using Bayesian Binary Quantile Regression with Penalized LASSO, *Mathline* 10 (2025), 373-385. <https://doi.org/10.31943/mathline.v10i2.928>.
- [8] L.H. Hasibuan, F. Yanuar, V.P. Harahap, L. Qalbi, Performance Quantile Regression and Bayesian Quantile Regression in Dealing with Non-normal Errors (Case Study on Simulated Data), *Numer. J. Mat. Pendidik. Mat.* 8 (2024), 46-54. <https://doi.org/10.25217/numerical.v8i2.4922>.
- [9] L.H. Hasibuan, F. Yanuar, D. Devianto, M. Maiyastri, A. Apriona, Bayesian Method for Linear Regression Modeling: Simulation Study with Normality Assumption, *J. Sci. Technol.* 5 (2025), 81-92. <https://doi.org/10.15548/jostech.v5i1.10994>.
- [10] D.F. Benoit, R. Alhamzawi, K. Yu, Bayesian Lasso Binary Quantile Regression, *Comput. Stat.* 28 (2013), 2861-2873. <https://doi.org/10.1007/s00180-013-0439-0>.
- [11] R. Alhamzawi, K. Yu, D.F. Benoit, Bayesian Adaptive Lasso Quantile Regression, *Stat. Model.* 12 (2012), 279-297. <https://doi.org/10.1177/1471082X1101200304>.

- [12] M. Bozorg, A. Bracale, P. Caramia, G. Carpinelli, M. Carpita, et al., Bayesian Bootstrap Quantile Regression for Probabilistic Photovoltaic Power Forecasting, *Prot. Control Mod. Power Syst.* 5 (2020), 21.
<https://doi.org/10.1186/s41601-020-00167-7>.
- [13] F. Meinert, C. Thomopoulos, R. Kreutz, Sex and Gender in Hypertension Guidelines, *J. Hum. Hypertens.* 37 (2023), 654-661. <https://doi.org/10.1038/s41371-022-00793-8>.
- [14] L. Cohen, G.C. Curhan, J.P. Forman, Influence of Age on the Association between Lifestyle Factors and Risk of Hypertension, *J. Am. Soc. Hypertens.* 6 (2012), 284-290. <https://doi.org/10.1016/j.jash.2012.06.002>.
- [15] Y.T.G. Arum, Hipertensi pada Penduduk Usia Produktif (15-64 Tahun), *Higeia J. Public Health Res. Dev.* 3 (2019), 345–356.
- [16] K. Gao, X. Shi, W. Wang, The Life-Course Impact of Smoking on Hypertension, Myocardial Infarction and Respiratory Diseases, *Sci. Rep.* 7 (2017), 4330. <https://doi.org/10.1038/s41598-017-04552-5>.
- [17] J. Poorolajal, F. Farbakhsh, H. Mahjub, A. Bidarafsh, E. Babaee, How Much Excess Body Weight, Blood Sugar, or Age Can Double the Risk of Hypertension?, *Public Health* 133 (2016), 14-18.
- [18] S. L, M.L.F. Kumalasari, E. Kusumawati, E.N. Andyarini, Hubungan Kadar Kolesterol dengan Hipertensi pada Pegawai di Fakultas Psikologi dan Kesehatan UIN Sunan Ampel, Indones. *J. Health Sci.* 4 (2020), 10-15.
<https://doi.org/10.24269/ijhs.v4i1.2321>.
- [19] L. Yekeen, R. Sanusi, A. Ketiku, Prevalence of Obesity and High Level of Cholesterol in Hypertension: Analysis of Data from the University College Hospital, Ibadan, *Afr. J. Biomed. Res.* 6 (2010), 129-132.
<https://doi.org/10.4314/ajbr.v6i3.54040>.
- [20] B.M. Egan, J. Li, S. Qanungo, et al. Blood Pressure and Cholesterol Control in Hypertensive Hypercholesterolemic Patients, *Circulation* 128 (2013), 29-41.
<https://doi.org/10.1161/CIRCULATIONAHA.112.000500>.
- [21] C. Davino, M. Furno, D. Vistocco, *Quantile Regression: Theory and Applications*, Wiley, 2013.
- [22] L.H. Hasibuan, F. Yanuar, D. Devianto, M. Maiyastri, Hypertension Status Model with Bayesian Adaptive LASSO Binary Quantile Regression Using Asymmetric Laplace Distribution, *IAENG Int. J. Comput. Sci.* 52 (2025), 2865-2872.
- [23] R. Koenker, K.F. Hallock, Quantile Regression, *J. Econ. Perspect.* 15 (2001), 143-156.
<https://doi.org/10.1257/jep.15.4.143>.

- [24] K. Yu, R.A. Moyeed, Bayesian Quantile Regression, *Stat. Probab. Lett.* 54 (2001), 437-447.
[https://doi.org/10.1016/S0167-7152\(01\)00124-9](https://doi.org/10.1016/S0167-7152(01)00124-9).
- [25] Z. Ma, Z. Han, M. Wang, Bayesian Variable Selection and Estimation in Binary Quantile Regression Using Global-Local Shrinkage Priors, *Commun. Stat. Simul. Comput.* 54 (2023), 1707-1721.
<https://doi.org/10.1080/03610918.2023.2293638>.
- [26] L. Hasibuan, F. Yanuar, D. Devianto, Maiyastri, Bayesian Binary Quantile Regression Adaptive Lasso for Modelling Hypertension: a Bootstrap Based Consistency Analysis, *Commun. Stat. Case Stud. Data Anal. Appl.* (2026), 1-24. <https://doi.org/10.1080/23737484.2026.2671866>.
- [27] C.F. Manski, Semiparametric Analysis of Binary Response from Response-based Samples, *J. Econom.* 31 (1986), 31-40. [https://doi.org/10.1016/0304-4076\(86\)90054-0](https://doi.org/10.1016/0304-4076(86)90054-0).
- [28] S. Zheng, QBoost: Predicting Quantiles with Boosting for Regression and Binary Classification, *Expert Syst. Appl.* 39 (2012), 1687-1697. <https://doi.org/10.1016/j.eswa.2011.06.060>.
- [29] G. Kordas, Smoothed Binary Regression Quantiles, *J. Appl. Econom.* 21 (2006), 387-407.
<https://doi.org/10.1002/jae.843>.
- [30] F. Yanuar, H. Yozza, A. Zetra, Modified Quantile Regression for Modeling the Low Birth Weight, *Front. Appl. Math. Stat.* 8 (2022), 890028. <https://doi.org/10.3389/fams.2022.890028>.
- [31] Y. Wu, Y. Liu, Variable Selection in Quantile Regression, *Stat. Sin.* 19 (2009), 801-17.
<http://www.jstor.org/stable/24308857>.
- [32] F. Yanuar, A.S. Deva, A. Zetra, Maiyastri, Length of Hospital Stay Model of COVID-19 Patients with Quantile Bayesian with Penalty LASSO, *Commun. Math. Biol. Neurosci.* 2023 (2023), 23.
<https://doi.org/10.28919/cmbn/7881>.
- [33] B.I. Yun, Transformation Methods for Finding Multiple Roots of Nonlinear Equations, *Appl. Math. Comput.* 217 (2010), 599-606. <https://doi.org/10.1016/j.amc.2010.05.094>.
- [34] T. Hastie, R. Tibshirani, J. Friedman, Boosting and Additive Trees, in: *Springer Series in Statistics*, Springer, New York, pp. 337-387, 2008. https://doi.org/10.1007/978-0-387-84858-7_10.
- [35] F. Yanuar, A. Zetra, Length-of-Stay of Hospitalized COVID-19 Patients Using Bootstrap Quantile Regression, *IAENG Int. J. Appl. Math.* 51 (2021), IJAM_51_3_40.