Sains Malaysiana 55(4)(2026): 756-773
http://doi.org/10.17576/jsm-2026-5504-14

Bayesian Logistic Regression to Explore the Role of Complete Blood Count in Kidney Disease Mortality

(Regresi Logistik Bayesian untuk Meneroka Peranan Kiraan Darah Lengkap dalam Kematian Penyakit Ginjal)

DG SITI NURISYA SAHIRAH AG ISHA[1], NURLIYANA JUHAN[2,*], YONG ZULINA ZUBAIRI[3], NORNAZIRAH AZIZAN[4], CHONG MUN HO[1], ABU SAYED MD. AL MAMUN[5] & QIN ZHI LEE[6]

[1]*Faculty of Science and Technology, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia*
[2]*Preparatory Centre for Science and Technology, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia*
[3]*Institute for Advanced Studies, University of Malaya, 50603 Kuala Lumpur, Malaysia*
[4]*Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia*
[5]*Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh*
[6]*Queen Elizabeth Hospital, Locked Bag No. 2029, 88586 Kota Kinabalu, Sabah, Malaysia*

Received: 4 July 2025/Accepted: 20 April 2026

ABSTRACT

Kidney disease is a major global health challenge, ranking fifth in Malaysia and ninth worldwide as a leading cause of death in 2021. This growing burden highlights the need for cost-effective tools to support early identification of patients at risk of mortality. The complete blood count (CBC) is an affordable, widely used diagnostic test, while Bayesian methods offer advantages for incorporating prior knowledge and quantifying uncertainty. However, the use of CBC parameters with Bayesian approaches for mortality prediction among kidney disease patients in Malaysia remains limited. This study aimed to develop a risk stratification model for kidney disease mortality using CBC data and Bayesian logistic regression (BLR). A retrospective study was conducted using data from 5,158 patients with kidney disease treated at Queen Elizabeth I Hospital. The final multivariate BLR model identified 13 significant predictors of mortality. The strongest predictors were low haemoglobin, high mean platelet volume (MPV), and high neutrophil-to-lymphocyte ratio (NLR), followed by high white blood cells (WBC), and hospitalisation history. The model demonstrated good calibration and discrimination, with an area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve (AUPRC) greater than 0.8, supporting its reliability for mortality risk stratification. These findings suggest that combining CBC parameters with demographic information may improve early detection and clinical decision-making, particularly in resource-limited settings.

Keywords: Bayesian logistic regression; complete blood count; kidney disease; mortality; risk stratification

ABSTRAK

Penyakit buah pinggang merupakan cabaran kesihatan global yang utama, menduduki tempat kelima di Malaysia dan kesembilan di peringkat dunia sebagai penyebab utama kematian pada tahun 2021. Beban yang semakin meningkat ini menekankan keperluan terhadap alat diagnostik yang kos efektif untuk memudahkan pengesanan awal. Kiraan darah lengkap (KDL) merupakan ujian diagnostik yang berpatutan dan digunakan secara rutin, manakala kaedah Bayesian menawarkan kelebihan dengan menggabungkan pengetahuan terdahulu dan memberikan pengkuantitian ketakpastian yang lebih tepat. Walau bagaimanapun, penggunaan parameter KDL dan Bayesian bagi meramal kematian dalam kalangan pesakit buah pinggang di Malaysia masih terhad. Kajian ini bertujuan untuk membangunkan model stratifikasi risiko bagi kematian akibat penyakit buah pinggang menggunakan KDL dan regresi logistik Bayesian (BLR). Kajian retrospektif ini menganalisis data daripada 5,158 pesakit buah pinggang dari Hospital Queen Elizabeth I. Model BLR multivariat akhir mengenal pasti 13 peramal kematian yang signifikan. Peramal yang paling kuat ialah hemoglobin rendah, min isi padu platelet (MPV) yang tinggi dan nisbah neutrofil kepada limfosit (NLR) yang tinggi, diikuti oleh sel darah putih (WBC) yang tinggi dan sejarah kemasukan ke hospital. Model menunjukkan prestasi penentukuran dan diskriminasi yang baik, dengan kawasan bawah lengkungan dalam menerima ciri operasi (AUROC) dan kawasan di bawah lengkung penarikan balik ketepatan (AUPRC) melebihi 0.8, sekali gus menyokong kebolehpercayaannya untuk stratifikasi risiko kematian. Penemuan ini mencadangkan bahawa penggabungan parameter KDL dengan maklumat demografi berpotensi meningkatkan pengesanan awal pesakit berisiko dan keputusan klinikal, khususnya dalam persekitaran dengan sumber terhad.

Kata kunci: Bayesian regresi logistik; kiraan darah lengkap; kematian; penyakit buah pinggang; stratifikasi risiko

INTRODUCTION

Kidney disease is a significant worldwide health challenge, resulting in elevated morbidity and mortality rates (Francis et al. 2024). It has significantly increased, ascending from the 19th to the ninth leading cause of mortality globally, with fatalities increasing by 95% from 2000 to 2021 (World Health Organisation 2024a). In Malaysia, it ranked as the fifth leading cause of death in 2021 (World Health Organisation 2024b). The incidence of chronic kidney disease (CKD) is projected to become the fifth leading cause of death worldwide by 2040 (Francis et al. 2024). This increasing burden emphasises the necessity for cost-effective diagnostic tools to enable the early identification of mortality risk to halt disease development, mitigate complications, and eventually reduce death rates among afflicted individuals.

Among various biomarkers, complete blood count (CBC) parameters have attracted attention for their potential prognostic value in predicting adverse outcomes in patients with kidney disease (Balasa 2023; Luo et al. 2024; Umeres-Francia et al. 2022). CBC may also predict sepsis in burn patients (Kim et al. 2024a) and coronavirus disease (COVID-19) cases (Omar et al. 2023; Yadav et al. 2023). It is frequently executed, cost-effective, and easily accessible in healthcare settings (Agnello et al. 2021). CBC results are usually classified into three components: white blood cells (WBC), red blood cells (RBC), and platelets (Agnello et al. 2021; Seo & Lee 2022). These results provide essential information on inflammatory markers (Fan et al. 2017; Uludag & Arikan 2021; Wei et al. 2024) and anaemia-related indicators (Dratch et al. 2024; Kim et al. 2024b; Li et al. 2024; Lindholm et al. 2022). These parameters help determine the onset of kidney disease, shape its progression, and assess related health outcomes (Fan et al. 2017; Uludag & Arikan 2021; Wei et al. 2024). Despite its advantages, there has been limited research in Malaysia, especially in Sabah, on the use of CBC to predict mortality in kidney disease patients. This gap underscores the need for studies assessing the prognostic value of CBC parameters in predicting mortality risk in this group.

In developing risk prediction models for mortality among kidney disease patients, including early detection, frequentist methods such as logistic regression (LR) are widely employed globally (Ganapathy et al. 2022; Lukman, Abdullah & Rachman 2021). Despite its widespread use, LR relies on several assumptions, such as linearity, multicollinearity, and the absence of influential outliers (Aw et al. 2021; Ramalu et al. 2022), which are often underreported and inadequately addressed in medical research (Ganapathy et al. 2022). Moreover, LR presents notable limitations in handling uncertainty, managing correlated or nonlinear data, accommodating small sample sizes, and maintaining stable parameter estimates, particularly when applied to complex clinical datasets (Aw et al. 2021; Gosho et al. 2025; Juhan et al. 2020a; Sabanayagam et al. 2023).

Bayesian logistic regression (BLR) presents a robust alternative to traditional methods by integrating prior knowledge and probability distributions into the model estimation process, thereby improving both interpretability and predictive performance (Ganapathy et al. 2022; Juhan et al. 2020a; Lukman, Abdullah & Rachman 2021). This approach enables more flexible parameter estimation, accommodates uncertainty in model coefficients, and produces credible intervals instead of conventional confidence intervals (Ganapathy et al. 2022; Juhan et al. 2020a; Lukman, Abdullah & Rachman 2021). Previous studies have shown that this approach can accurately perform diagnosis and prediction of diseases or health conditions in many countries for a wide variety of problems, such as mortality of patients (Juhan et al. 2020b; Wong & Ismail 2016), infectious diseases (Ma et al. 2022; Tsheten et al. 2024), and non-communicable diseases (Juhan et al. 2020a). Nevertheless, these previous studies have been based on worldwide samples from Vietnam, the Philippines, Singapore, and other states in Malaysia, suggesting limitations to implementing the Bayesian approach in the context of kidney disease mortality detection, including in Sabah.

Given the growing burden of kidney disease, two important gaps remain in the existing literature. First, limited evidence is available on the prognostic value of CBC parameters for mortality prediction among kidney disease patients in Malaysia, particularly in Sabah. Second, although BLR has shown potential for disease prediction and prognosis, its application for mortality risk stratification in kidney disease patients in this setting remains underexplored. Addressing these gaps is important because cost-effective, accessible, and reliable risk assessment tools are needed to support early identification of high-risk patients and improve clinical decision-making. Therefore, this study aims to address these gaps by developing a BLR-based risk stratification model for mortality among kidney disease patients using CBC parameters and demographic profiles, and by evaluating its predictive performance. Specifically, the study has two primary objectives: 1) To develop a BLR model for predicting mortality risk among kidney disease patients using CBC parameters and demographic profiles and 2) To evaluate the predictive performance of BLR in assessing mortality risk.

METHODOLOGY

The study commenced with retrieving secondary data on patients with kidney disease from Queen Elizabeth I Hospital (QEH1). The following stage was data preparation, which focused on cleansing, division, and encoding. The third stage involved identifying potential predictors and testing for collinearity. This was followed by implementing BLR to identify the associated mortality factors. Lastly, calibration, discrimination, and posterior predictive accuracy were used to evaluate the performance of the model. For clarity, a flowchart summarising the overall study workflow has been provided in Figure 1. All statistical analyses were conducted using R (version 4.4.1) and IBM SPSS (version 29.0).

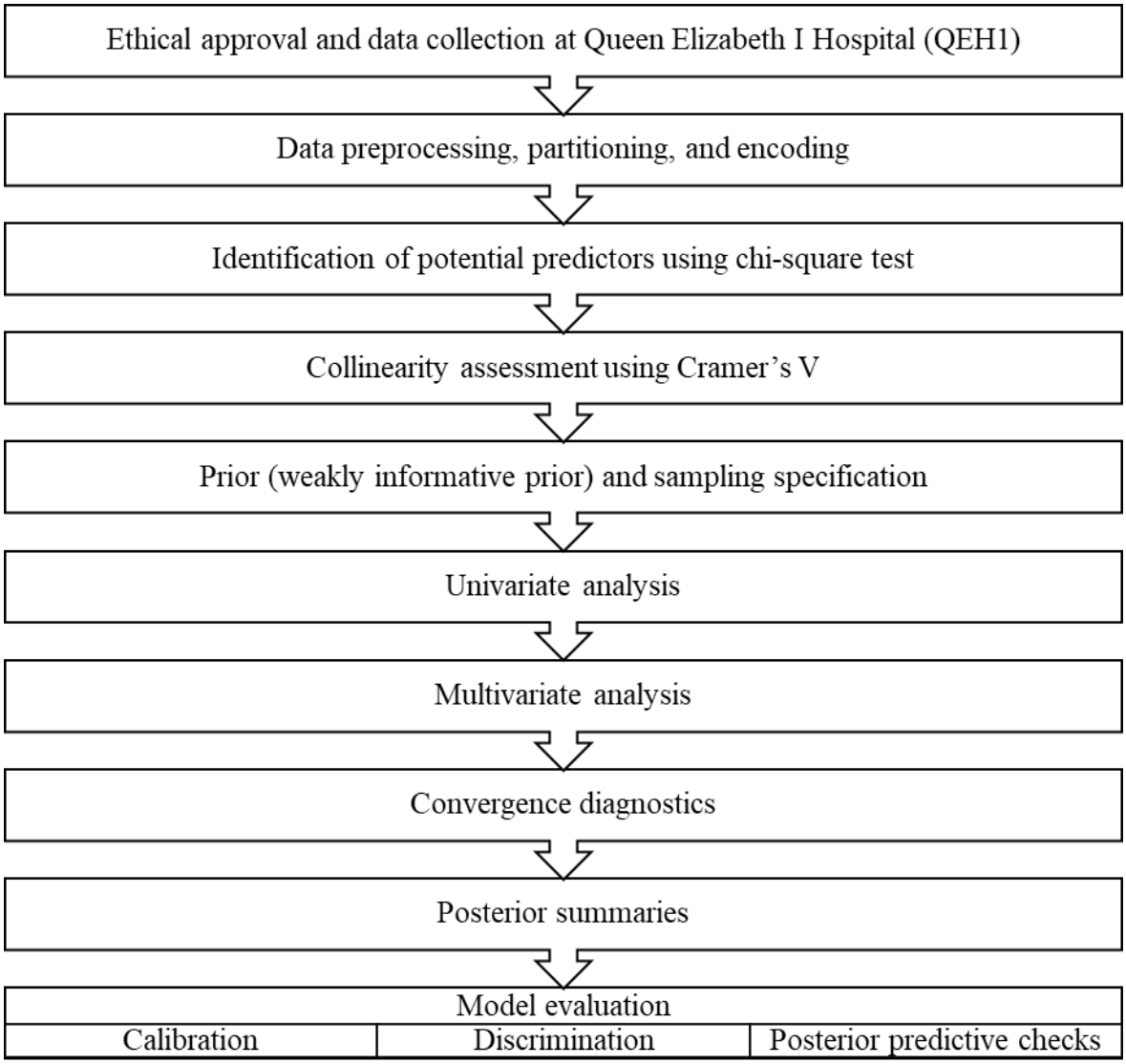

FIGURE 1. Flowchart of the overall study workflow

RESEARCH DESIGN AND DATA SOURCES

This retrospective study examined 8,514 patients with kidney disease from QEH1. Secondary data, including patients' demographic information, CBC results, and patient outcomes, were gathered between January 1, 2018, and December 31, 2022, with no exclusion criteria. Ethical approval was secured from the Universiti Malaysia Sabah Medical Research Ethics Committee and the Medical Research and Ethics Committee (MREC) of the Ministry of Health, Malaysia.

DATA PREPROCESSING, PARTITIONING, AND ENCODING

The data cleaning procedure was carried out, retaining a final dataset of 5,158 patients used in the analysis. The dataset was then divided into two subsets: training (70%) and testing (30%), using stratified random selection. The training dataset was used for model building and internal validation using ten-fold cross-validation. The overall model performance (discrimination and calibration) was assessed by averaging the performance measures across all iterations (Nguyen et al. 2023). The testing dataset was used to externally validate the model's performance and generalisability (Nguyen et al. 2023).

The demographic characteristics included in this research were age, gender, and ethnicity. Age was divided into two categories: <60 and ≥60 years. Ethnicity was divided into four categories: Bajau, Bruneian-Malay, Chinese, Kadazan-Dusun, and Others, which reflect the four main ethnic groups in Sabah (Sabah State Government 2024). Furthermore, hospitalisation history was classified as either yes or no.

Each CBC parameter was divided into three categories: low, normal, and high. These categories followed established normal reference ranges from multiple reputable guidelines (Dean 2005; Drăgoescu et al. 2021;

Khan & Ullah 2020; Pogorzelska et al. 2020; Thombare et al. 2023; Wulandari & Hapsari 2024). This approach helped to streamline the model (Mahachai et al. 2023). The MPV/Platelet Ratio (MPVPL) is a computed variable derived from CBC parameters and lacks a standardised reference range. This study used a median cutoff of 0.0378 to classify MPVPL into low and high categories. These categorical variables were coded as dummy variables. The reference groups were age <60, Bruneian-Malay, male, no hospitalisation history, and normal levels of CBC parameters (except MPVPL, which was set to a low level).

IDENTIFICATION OF POTENTIAL PREDICTORS AND COLLINEARITY

This study applied Pearson's chi-square test of independence to identify factors significantly associated with the outcome. Variables displaying a statistically significant relationship ($p < 0.05$) were selected as potential candidates for further analysis. To address collinearity concerns, a correlation matrix was constructed using Cramer's V to identify highly correlated categorical predictors, and those exhibiting strong correlations were excluded from subsequent analyses.

BAYESIAN LOGISTIC REGRESSION (BLR)

Bayesian logistic regression (BLR) was implemented using the *brms* package in R to model binary outcomes based on Bayesian principles (Bürkner 2017; Bürkner et al. 2024; Marine & Zewde 2024). It combines prior information with a likelihood function and the resulting posterior distribution to estimate parameters and quantify uncertainty (Marine & Zewde 2024). Univariate BLR was initially used to assess the effect of each predictor variable on the outcome. Predictors showing credible evidence of association, based on posterior estimates with 95% credible intervals (CrIs) excluding 0, equivalent to odds ratios with 95% CrIs excluding 1, were entered into the multivariate BLR model. The model was then refined iteratively to retain only variables with credible evidence of association. The general form of the multivariate logistic model is shown in Equation (1):

$$\mathrm{y_i} = \mathrm{logit}\ (\pi_\mathrm{i}) = \log\left(\frac{\pi_\mathrm{i}}{1-\pi_\mathrm{i}}\right) = \beta_0 + \beta_1 \mathrm{X_{i1}} + \ldots + \beta_\mathrm{j}\mathrm{X_{ij}} \quad \text{for i} = 1,2,3,\ldots,\mathrm{n},\ \mathrm{j} = 1,\ 2,\ 3,\ \ldots,\ \mathrm{k} \tag{1}$$

where y_i is observed binary outcome (survivor vs non-survivor) for the i-th observation, X_{ij} represents the j-th predictor (demographic, hospitalisation history, CBC parameters) for the i-th observation. $\beta_0 \ldots . \beta_j$ are regression coefficients, n is the total number of observations, and k number of predictors. In BLR, prior distributions are assigned to each regression coefficient to represent prior beliefs or knowledge about their values. This study used normal distribution as expressed in Equation (2):

$$\mathrm{P}(\beta_\mathrm{j}\,|\mu_\mathrm{j},\ \sigma_\mathrm{j}^2) = \frac{1}{\sqrt{2\pi\sigma_\mathrm{j}^2}}\exp\left\{-\frac{1}{2\sigma_\mathrm{j}^2}(\beta_\mathrm{j}-\mu_\mathrm{j})^2\right\} \tag{2}$$

The likelihood function (Equation 3), based on the Bernoulli distribution, describes the probability of the observed outcomes given the predictors and coefficients:

$$\mathrm{P}\,(\mathrm{y}|\ \beta,\ \mathrm{X}) = \prod_{\mathrm{i}=1}^{\mathrm{n}} \pi_\mathrm{i}^{\mathrm{y_i}}(1-\pi_\mathrm{i})^{1-\mathrm{y_i}} \tag{3}$$

Using Bayes' theorem, the posterior distribution is obtained by combining the likelihood with the joint prior distribution of the regression coefficients, as shown in Equation (4):

$$P(\beta \mid y, \mathrm{X}) \propto P(y \mid \beta, \mathrm{X}).\ P(\beta) \tag{4}$$

where $P(\beta)$ denotes the joint prior distribution of the regression coefficients; $P(y|\beta, \mathrm{X})$ is the likelihood of the data given the parameters; and $P(\beta|y, \mathrm{X})$ is the posterior distribution, which reflects updated beliefs about β after observing the data.

Priors Used: Weakly Informative Priors

Weakly informative normal priors were used for the univariate and multivariate analyses. Specifically, a normal prior with mean zero and standard deviation 10, N(0,10) was applied to the intercept, whereas N(0,2) priors were assigned to the remaining regression coefficients, following the recommendations of Bui et al. (2024), Egidi, Pauli and Torelli (2022), Gelman et al. (2008), Gosho et al. (2025), and Lemoine (2019). These choices were grounded in both statistical theory and practical modelling considerations. As Lemoine (2019) showed, such priors yield more interpretable and stable results on the probability scale than overly flat alternatives, such as Normal (0,100), which can inflate posterior estimates and reduce inferential reliability. These selections were also consistent with Gelman et al. (2008), who propose Normal (0,2.5) priors for regression coefficients and Normal (0,10) for intercepts to promote robust inference while avoiding strong prior influence (Egidi, Pauli & Torelli 2022).

The normal (0,10) prior assigned to the intercept allows for a broad range of plausible baseline log odds without being overly diffuse. In contrast, the Normal (0,2) priors assigned to the remaining regression coefficients introduce moderate regularisation that helps mitigate overfitting, prevent extreme or unstable estimates, and reduce convergence issues while still allowing the data to meaningfully inform posterior inference. This prior places about 95% of its mass between −3.92 and 3.92 on the log-odds scale (odds ratio from ~0.02 to 50.4), offering enough flexibility to capture meaningful associations while avoiding implausible estimates. These priors are well-suited to medical research, where effect sizes are typically modest, and models must remain stable and data driven.

Moreover, normal distribution was used as a prior for the regression coefficients, as these coefficients are unbounded and can take any value from the real number line (Gunn-Sandell et al. 2024; Johnson, Ott & Dogucu 2022). Normal distribution, defined over the entire range from negative to positive infinity, is a natural and mathematically convenient choice (Gunn-Sandell et al. 2024; Johnson, Ott & Dogucu 2022). Centring the prior at zero allocates equal probability mass above and below zero, supporting more conservative and stable estimates than flat priors (Farhin & Ali Khan 2022). Besides, weakly informative priors accounted for limited detailed prior evidence when existing research suggests potential associations; such priors offer a balanced way to reflect this uncertainty while still incorporating meaningful prior assumptions into the analysis (Larson et al. 2023).

Sampling method: No-U-Turn Sampler (NUTS)

The posterior distributions of the regression coefficients were estimated using the No-U-Turn Sampler (NUTS), a Markov Chain Monte Carlo (MCMC) method that extends Hamiltonian Monte Carlo (HMC) (Farhin & Ali Khan 2022; Gelman et al. 2025; Hoffman & Gelman 2014). NUTS is the default sampler in *Stan* and, consequently, in *brms*, which serves as its R interface (Bürkner 2017). Compared to traditional algorithms such as Metropolis-Hastings and Gibbs sampling, NUTS demonstrates substantially faster convergence, particularly in high-dimensional settings, and performs effectively regardless of the priors used (Bürkner 2017; Fanconi et al. 2023; Farhin & Ali Khan 2022; Hoffman & Gelman 2014; Mattos & Ramos 2022).

Univariate and Multivariate Sampling Specifications

In Bayesian analysis, the number of iterations, warm-up length, and number of chains are chosen empirically. This depends on model complexity, sampling chain behaviour, convergence, and the specific MCMC algorithm (Depaoli & van de Schoot 2017; Wang et al. 2024). Running multiple chains is widely recommended to improve diagnostic accuracy and ensure reliable convergence (Gunn-Sandell et al. 2024; Mattos & Ramos 2022).

While there are no strict rules, Depaoli and van de Schoot (2017) suggested using at least 10,000 iterations each for the warm-up and post-warm-up phases, noting that substantially longer chains may be required to achieve convergence. Larson et al. (2023) further noted that warm-up is often set at approximately 10% of the total iterations, although no fixed standard exists. Other studies suggested that at least 1000 and up to a million iterations were commonly used for estimation (Juhan et al. 2020a, 2020b; Lukman, Abdullah & Rachman 2021; Tsheten et al. 2024; Wong & Ismail 2016).

Following these general guidelines, the initial sampling specifications in this study were informed by previous methodological recommendations and empirical practices reported in the literature (Depaoli & van de Schoot, 2017; Juhan et al. 2020a, 2020b; Larson et al. 2023; Lukman, Abdullah & Rachman 2021; Tsheten et al. 2024; Wong & Ismail 2016). The final number of iterations used for the univariate and multivariate models was determined empirically based on model complexity and the need to achieve satisfactory convergence and stable posterior estimates. Accordingly, three chains of 10,000 iterations each, including 1,000 warm-up iterations, were used for the univariate models. For the multivariate model, three chains of 90,000 iterations were run, with 18,000 warm-up iterations per chain, producing 216,000 post-warm-up samples for inference.

MODEL CONVERGENCE

Model convergence is essential in Bayesian analysis to ensure that the computational algorithm has produced stable posterior distributions suitable for summary and inference (Larson et al. 2023). The model convergence was assessed using both visual and numerical methods. The visual assessments included trace plots, Kernel density estimate (KDE) plots, autocorrelation plots, and histograms of the posterior distributions, common approaches in prior research (Bui et al. 2024; Depaoli & van de Schoot 2017; Gunn-Sandell et al. 2024; Larson et al. 2023; Marine & Zewde 2024). Numerical assessments included the Gelman-Rubin diagnostic (R-hat), effective sample size (ESS), Raftery-Lewis diagnostic, Heidelberger-Welch diagnostic (stationarity and half-width tests), and Geweke diagnostic, as recommended by several studies (Bui et al. 2024; Depaoli & van de Schoot 2017; Larson et al. 2023; Mattos & Ramos 2022).

POSTERIOR SUMMARISES

Posterior summaries were reported using the posterior means as point estimates, standard error, and Monte Carlo (MC) error, alongside 95% credible intervals (CrIs), derived from the 2.5th and 97.5th percentiles of the posterior distribution, representing the 95% most plausible effect sizes according to the prior data and model (Bui et al. 2024; Gunn-Sandell et al. 2024; Marine & Zewde 2024; Wolde et al. 2024).

MODEL EVALUATION

Calibration and Discrimination

As Bayesian models generate a full predictive distribution for each observation, the posterior predictive mean was used as the risk prediction for performance evaluation (Fanconi et al. 2023; Johnson, Ott & Dogucu 2022). These mean probabilities were used to compute calibration and discrimination metrics. Calibration was assessed using the Average Absolute Error (AAE), Brier score, and Spiegelhalter's Z-test. Both AAE and the Brier score quantify the difference between predicted probabilities and observed outcomes, with AAE measuring the average absolute difference and the Brier score measuring the mean squared difference (Han et al. 2022; Huang et al. 2020; Lindhiem et al. 2018). Spiegelhalter's Z-test evaluates

whether the predicted probabilities differ significantly from the observed outcomes (Huang et al. 2020; Lindhiem et al. 2018). Lower values (closer to zero) for AAE and the Brier score indicate better calibration (Han et al. 2022; Huang et al. 2020; Lindhiem et al. 2018). In contrast, for Spiegelhalter's Z-test, a non-significant p-value indicates no evidence of lack of fit and thus good calibration (Huang et al. 2020; Lindhiem et al. 2018).

For model discrimination evaluation, this study initially assessed it through the area under the receiver operating characteristic curve (AUROC) and the area under the Precision-Recall Curve (AUPRC). AUC, obtained from ROC and PRC curves, provides a single metric ranging from 0 to 1 that reflects overall classifier performance across different thresholds. Generally, an AUC of 0.7 to <0.8 denotes acceptable, 0.8 to <0.9 is excellent, and 0.9 or higher indicates outstanding predictive performance (Wong et al. 2022).

Given the imbalanced nature of the dataset (3,795 'survivors' and 1,363 'non-survivors'), this study also used unbiased evaluation metrics, including F1 score and balanced accuracy (BAcc). No undersampling or oversampling techniques were applied to preserve the integrity of the original datasets (Ting, Zakariah & Mohd Yusri 2022). The F1 Score and BAcc provide robust performance evaluation for imbalanced datasets and are commonly used for this purpose (Huynh et al. 2022; Mohd Faizal et al. 2023; Thölke et al. 2023; Ting, Zakariah & Mohd Yusri 2022). The values of these metrics range from 0 to 1, where 0 denotes worst performance, and 1 signifies best performance (Chicco, Tötsch & Jurman 2021).

Posterior Predictive Accuracy

Both graphical and quantitative assessments were conducted to determine how effectively the Bayesian model predicts future observations and reflects the underlying data-generating process. The posterior predictive check (PPC), a graphical method, is a post-estimation diagnostic used to evaluate the adequacy of Bayesian models after fitting (Farhin & Ali Khan 2022). This study assessed PPC using the pp_check() function from the *brms* package (Bürkner et al. 2024) to evaluate how well the model reproduces the observed data by comparing the empirical distribution of the observed outcomes to replicated data generated from the posterior predictive distribution (Farhin & Ali Khan 2022; Kruschke 2021). A good fit is indicated by substantial overlap between the observed and replicated distributions, suggesting that the simulated data behave similarly to the observed outcomes and that the model performs well in predicting mortality (Farhin & Ali Khan 2022).

For quantitative evaluation, the Watanabe-Akaike Information Criterion (WAIC) and Leave-One-Out Cross-Validation Information Criterion (LOOIC) were employed, both computed using the waic() and loo() functions applied to the fitted model from the *brms* package. WAIC and LOOIC are fully Bayesian metric that assesses how well a model can predict new data (Gelman et al. 2025; Larson et al. 2023). Both have several advantages over simpler estimates of predictive error, such as the Deviance Information Criterion (DIC) and Akaike Information Criterion (AIC) (Gabry 2024). Lower WAIC and LOOIC values suggest the model has better predictive performance on unseen or future data (Bürkner 2017; Gelman et al. 2025; Larson et al. 2023; Mattos & Ramos 2022; Wolde et al. 2024).

RESULTS AND DISCUSSION

DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF PATIENTS

A total of 5,158 patients were included in the analysis, comprising 3,795 survivors and 1,363 non-survivors. Gender distribution was similar between survivors and non-survivors, and a higher proportion of non-survivors were hospitalised (72.3%) compared to survivors (27.4%) (Table 1). Patients with age ≥60 years comprised 49.7% of non-survivors versus 36.2% of survivors. Ethnic distribution showed slight variation across groups, with Bruneian-Malay as the most represented. The proportions were similar across the overall sample, training, and testing datasets, although these are not shown in the table.

ASSOCIATION BETWEEN PATIENT OUTCOMES AND PREDICTORS AND COLLINEARITY CHECK

Although not shown here, the study used the chi-square test to examine the associations between patient outcomes and various factors within the training dataset. Significant associations with patient outcomes were found for age, hospitalisation history, ethnicity, white blood cell (WBC), neutrophil, lymphocyte, monocyte, mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), red cell distribution width (RDW), platelet count, platelet-large cell ratio (PLCR), nucleated red blood cell (NRBC), mean platelet volume (MPV), haemoglobin, red blood cell (RBC), haematocrit (HCT), neutrophil to lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), plateletcrit (PCT), and MPV-to-platelet ratio (MPVPL). Eosinophil, basophil, platelet distribution width (PDW), and gender were not significant and were excluded, leaving 22 variables for further analysis.

Subsequently, collinearity among the 22 variables in the training dataset (though not shown here) was assessed using Cramer's V, applying thresholds of 0.350 for degrees of freedom (df) = 2 and 0.500 for df = 1 to identify strong associations. Several variable pairs exceeded these thresholds, indicating high collinearity. To minimise redundancy and avoid multicollinearity in further analyses, the following highly correlated variables were excluded: neutrophil, lymphocyte, MCH, PLCR, RBC, HCT, and PCT. This process resulted in a refined set of 15 variables for univariate analysis.

TABLE 1. Demographic and clinical characteristics of patients

Patient characteristics, n (%)	Survivors (n= 3,795)	Non-survivors (n= 1,363)	Patient (n=5,158)
1. Gender			
Male	2010 (53)	731 (53.6)	2741 (53.1)
Female	1785 (47)	632 (46.4)	2417 (46.9)
2. Hospitalisation history			
No	2755 (72.6)	377 (27.7)	3132 (60.7)
Yes	1040 (27.4)	986 (72.3)	2026 (39.3)
3. Age			
< 60	2419 (63.8)	686 (50.3)	3105 (60.2)
≥ 60	1376 (36.2)	677 (49.7)	2053 (39.8)
4. Ethnicity			
Bajau	575 (15.1)	216 (15.8)	791 (15.3)
Bruneian-Malay	1027 (27.1)	357 (26.2)	1384 (26.8)
Chinese	635 (16.7)	179 (13.1)	814 (15.8)
Kadazan-Dusun	783 (20.6)	258 (18.9)	1041 (20.2)
Others	775 (20.4)	353 (25.9)	1128 (21.9)

CONVERGENCE DIAGNOSTICS

To ensure the reliability and interpretability of the results, model convergence was assessed using both visual (Supplement A) and numerical (Supplement B) diagnostics to confirm the adequacy of the sampling settings and computational algorithm for yielding stable posterior distributions suitable for inference (Larson et al. 2023). For all parameters, the trace plots (Figure S.1) demonstrated good mixing across the three chains, the Kernel density estimate (KDE) plots (Figure S.2) showed smooth and symmetric distributions with substantial overlap among chains, and the autocorrelation plots (Figure S.3) showed a rapid drop-off in autocorrelation with increasing lags, together suggesting adequate sampling with minimal dependence between successive iterations (Gunn-Sandell et al. 2024; Larson et al. 2023; Marine & Zewde 2024). Moreover, the histograms (Figure S.4) showed no gaps or unusual patterns, suggesting that the MCMC sampling was stable and that the posterior estimates were reliable (van de Schoot et al. 2020).

For numerical assessments (Table S.1), all parameters exhibited Gelman-Rubin R-hat values close to 1, and ESS were sufficiently large, implying well-mixed chains and stable posterior estimates (Gelman et al. 2025; Kruschke 2021; Mattos & Ramos 2022). The Raftery-Lewis diagnostic showed that the suggested number of iterations (N), including warm-up (burn-in) (M), was consistently lower than the actual number used, with dependence factors generally near 2 (< 5), further supporting sampling adequacy (Plummer et al. 2024). The Heidelberger-Welch diagnostic showed that all parameters passed both the stationarity and half-width tests, and Geweke's diagnostic yielded non-significant Z-scores ($p > 0.05$), suggesting no evidence of non-convergence or autocorrelation issues (Gerber & Scharler 2024; Plummer et al. 2024). Both assessments demonstrated that the selected sampling strategy produced reliable and well-converged posterior estimates, validating the modelling approach and supporting the robustness of the findings.

FINALISED MULTIVARIATE LOGISTIC REGRESSION

Once convergence was confirmed, posterior summaries were examined to identify CBC variables significantly associated with mortality in kidney disease. Although not shown, the univariate analysis indicated that all 15 independent variables were significant and were therefore included in the multivariate analysis. The initial multivariate model showed that ethnicity and NRBC did not demonstrate credible evidence of association with mortality, as their estimated effects had 95% credible intervals that included the null value. Therefore, these variables were excluded during model refinement to obtain a more parsimonious final model.

Table 2 presents the findings of the finalised multivariate analysis. Several risk factors were associated with increased mortality. The strongest predictors were low haemoglobin (OR = 8.6568, 95% CrI: 5.7388-13.2219), high MPV (OR = 7.6742, 95% CrI: 5.2407-11.3161), and high NLR (OR = 7.2951, 95% CrI: 5.2251-10.2550). These were followed by high WBC (OR = 6.1067, 95% CrI: 4.4971-8.3386), hospitalisation history (OR = 5.4937, 95% CrI: 4.2681-7.0959), and high RDW (OR = 4.6279, 95% CrI: 3.5498-6.0459). Moderate associations were also

observed for age (≥60 years), high MCV, high PLR, and high monocyte count, while elevated MPVPL showed a relatively modest effect.

In contrast, variables such as low monocyte count, low NLR, low MCV, low RDW, and low WBC, as well as high haemoglobin and high MCHC, did not show credible evidence of association with the outcome. This is because the 95% credible intervals for their estimated coefficients and odds ratios included 0 and 1, respectively. Additionally, the Monte Carlo (MC) error for each parameter of interest was less than 5% of its respective posterior standard error. This finding confirms the convergence and reliability of the posterior estimates (Marine & Zewde 2024).

These findings align with and extend the existing literature on the prognostic relevance of haematological markers in kidney disease. For instance, low haemoglobin, a hallmark of anaemia in kidney disease, results primarily from a shortened lifespan of red blood cells (RBC) and iron deficiency. This leads to fewer RBCs in circulation, decreased oxygen-carrying capacity, and increased tissue hypoxia, which contributes to higher mortality in patients with acute kidney injury (AKI), chronic kidney disease (CKD), and those on dialysis (Gameiro & Lopes 2019; Kim et al. 2024b; Sheng et al. 2020; Yang et al. 2023). Additionally, reduced MCHC, a marker of hypochromia and impaired oxygen-carrying capacity, has been inversely associated with all-cause mortality in patients with CKD and AKI (Kim et al. 2024b; Li et al. 2024; Yang et al. 2023). Additionally, a high red cell RDW indicates variability in RBC size due to disrupted erythropoiesis and increased RBC destruction. This marker has been consistently linked to mortality and CKD progression through mechanisms involving inflammation, oxidative stress, endothelial dysfunction, and malnutrition (Agnello et al. 2021; Kim et al. 2024b; Kimura et al. 2023; Kor et al. 2018; Yang et al. 2023). Similarly, elevated MCV, reflecting macrocytosis and RBC abnormalities, has been associated with increased mortality in CKD and haemodialysis patients. This condition was likely due to oxidative stress, impaired oxygen delivery, nutritional deficiencies, and endothelial dysfunction (Dratch et al. 2024; Kor et al. 2018; Lindholm et al. 2022).

TABLE 2. Results of the finalised multivariate analysis (training sample, n = 3,612)

Variable	Estimate	Standard error	MC error	Estimate		Odd Ratio	Odd ratio	
				Q2.5	Q97.5		Q2.5	Q97.5
Intercept	-8.3747	0.3401	0.0009	-9.0574	-7.7204	0.0002	0.0001	0.0004
Hospitalised	1.7036	0.1297	0.0003	1.4512	1.9595	5.4937	4.2681	7.0959
Age (≥ 60)	1.0874	0.1307	0.0003	0.8331	1.3447	2.9667	2.3004	3.8369
WBC (Low)	0.1827	0.2749	0.0005	-0.36	0.7163	1.2005	0.6977	2.0468
WBC (High)	1.8094	0.1575	0.0003	1.5034	2.1209	6.1067	4.4971	8.3386
Monocyte (Low)	-0.2936	0.3469	0.0006	-0.9785	0.3796	0.7456	0.3759	1.4617
Monocyte (High)	1.2277	0.1763	0.0003	0.8828	1.5757	3.4133	2.4176	4.8342
MCV (Low)	0.044	0.1539	0.0003	-0.2575	0.3449	1.045	0.773	1.4119
MCV (High)	1.4134	0.357	0.0006	0.7123	2.1128	4.1098	2.0387	8.2711
MCHC (Low)	0.8667	0.1773	0.0003	0.5202	1.2151	2.3789	1.6823	3.3707
MCHC (High)	0.3626	0.3782	0.0007	-0.3836	1.0957	1.4371	0.6814	2.9912
RDW (Low)	0.4335	0.9442	0.0017	-1.5115	2.2119	1.5427	0.2206	9.1328
RDW (High)	1.5321	0.1354	0.0003	1.2669	1.7994	4.6279	3.5498	6.0459
Platelet (Low)	0.9828	0.2137	0.0004	0.5654	1.4022	2.6719	1.7601	4.0642
Platelet (High)	0.432	0.2021	0.0004	0.0373	0.8289	1.5404	1.038	2.2908
MPV (High)	2.0379	0.1963	0.0004	1.6565	2.4262	7.6742	5.2407	11.3161
HGB (Low)	2.1583	0.2133	0.0005	1.7473	2.5819	8.6568	5.7388	13.2219
HGB (High)	0.4667	0.8244	0.0015	-1.1671	2.073	1.5947	0.3113	7.9487
NLR (Low)	-0.7342	0.5697	0.001	-1.8909	0.3379	0.4799	0.1509	1.4019
NLR (High)	1.9872	0.1725	0.0004	1.6535	2.3278	7.2951	5.2251	10.255
PLR (Low)	0.5306	0.2188	0.0005	0.1029	0.9597	1.7	1.1084	2.6108
PLR (High)	1.367	0.1538	0.0003	1.0674	1.6694	3.9236	2.9078	5.309
MPVPL (High)	0.3851	0.1541	0.0003	0.0831	0.6877	1.4698	1.0867	1.9892

In terms of platelet-related parameters, MPV was a strong predictor of mortality, as larger, more reactive platelets were associated with systemic inflammation, enhanced thrombotic potential, and poor vascular outcomes (Balasa 2023; Gameiro & Lopes 2019; Ma et al. 2023; Wang et al. 2023). Elevated MPVPL, which integrates MPV and platelet count and is a marker of a heightened inflammatory state, has been significantly associated with higher mortality. This condition was likely due to its link with impaired clotting balance and heightened thrombotic risk (Çeleğen & Çeleğen 2024; Ma et al. 2023). This study further found a unique relationship between PLR and mortality, in which both high and low PLR were associated with increased mortality in kidney disease patients (Jeon et al. 2023; Umeres-Francia et al. 2022). Similarly, both low and high platelet counts were significantly associated with increased mortality in kidney disease patients, as thrombocytopenia impairs haemostasis and immune defence, while thrombocytosis promotes thrombotic events and vascular inflammation. These together contribute to heightened cardiovascular and infectious risks (Liao et al. 2022; Seo & Lee 2022; Zhao et al. 2024).

Markers of systemic inflammation also showed strong associations with mortality. Elevated NLR, suggesting an imbalance between pro-inflammatory neutrophils and anti-inflammatory lymphocytes. It is a marker of inflammation and has been consistently associated with increased mortality in patients with kidney disease due to its role in promoting oxidative stress, vascular damage and cardiovascular complications (Liao et al. 2022; Luo et al. 2024; Umeres-Francia et al. 2022; Zhao et al. 2024). Additionally, high WBC count, another inflammation marker, has been linked with increased mortality and kidney function decline in patients with AKI, peritoneal dialysis, and CKD, likely reflecting chronic inflammatory and uremic conditions (Fan et al. 2017; Gameiro & Lopes 2019; Guo et al. 2024; Uludag & Arikan 2021). Elevated monocyte counts have also been associated with higher mortality and CKD progression, driven by monocyte activation induced by uremic toxins, which promote systemic inflammation, oxidative stress, vascular inflammation, and cardiovascular calcification (Oh et al. 2022; Wei et al. 2024).

In addition to haematological parameters, patients' characteristics, such as hospitalisation history, were predictors of mortality in kidney disease patients. This might be because the prior hospital admissions often reflect poor health conditions and cardiovascular or infection-related complications, all of which contribute to poor survival outcomes (Arif et al. 2017; Schulman et al. 2023; Yip et al. 2024). Besides, advanced age (≥60 years) was significantly associated with increased mortality in kidney disease patients. This was consistent with prior studies showing that older individuals face a higher risk due to greater comorbidity burden, progressive kidney decline, and stronger links between anaemia and adverse outcomes in this population (Hashmi, Benjamin & Lappin 2023; Kim et al. 2024b; Yang et al. 2023).

EVALUATION OF MODEL PERFORMANCE

In addition to identifying significant predictors, the model's performance in calibration, discrimination, and posterior predictive accuracy was evaluated and demonstrated robust results. Model calibration and discrimination performance were evaluated via 10-fold cross-validation and summarised in Tables 3 and 4. The calibration assessment confirmed that the model's predicted probabilities were well aligned with observed outcomes. Across both datasets, low AAE and Brier score values were observed, indicating good calibration and minimal deviation between predicted and observed probabilities (Han et al. 2022; Huang et al. 2020; Lindhiem et al. 2018).

Additionally, the Spiegelhalter Z test suggested no significant evidence of miscalibration (Huang et al. 2020; Lindhiem et al. 2018). These findings denoted that the model's predicted probabilities were well-calibrated and reliable across both datasets. The Hosmer-Lemeshow test was not used to assess model calibration in this study, as it offers only a broad indication of calibration and is highly sensitive to sample size (Binuya et al. 2022; Huang et al. 2020; Lindhiem et al. 2018).

Discrimination results (Table 4) demonstrated excellent discriminatory ability with high AUROC and AUPRC values greater than 0.8. In addition, to assess other discrimination metrics, most performance metrics reached or approached 0.8, demonstrating excellent discriminative ability to distinguish survivors from non-survivors across both training and testing datasets (Chicco, Tötsch & Jurman 2021; Nguyen et al. 2023; Wong et al. 2022).

The model's ability to reproduce observed data and generalise new data was evaluated through graphical and quantitative methods. As shown in Figure 2, the PPC illustrates a strong overlap between the predicted probabilities of the observed data (y) and the replicated data (y^{rep}) generated from the posterior predictive distribution, suggesting that the simulated data behave similarly to the observed outcomes and that the model effectively captured the underlying data structure (Farhin & Ali Khan 2022; Kruschke 2021).

Quantitatively, the WAIC and LOOIC values obtained during internal validation (1827.951 and 1828.07, respectively) supported the model's fit. In the external validation dataset, WAIC and LOOIC decreased to 817.8629 and 817.8965, respectively, suggesting strong generalisability and consistent predictive accuracy on unseen data (Bürkner 2017; Gelman et al. 2025; Larson et al. 2023; Mattos & Ramos 2022; Wolde et al. 2024). All these results confirm the model's suitability for reliable prediction and reinforce its applicability to external data.

TABLE 3. Assessment of model calibration

No.	Measure	Internal validation	External validation
1.	AAE	0.1520	0.1626
2.	Brier Score	0.0758	0.0805
3.	Spiegelhalter Z (*p*-value)	-0.0229 (0.5066)	-0.4011 (0.6898)

TABLE 4. Assessment of model discrimination

No.	Measure	Internal validation	External validation
1.	AUROC	0.9554	0.9528
2.	AUPRC	0.8294	0.8313
3.	F1 Score	0.7947	0.7754
4.	Balanced accuracy	0.8556	0.8435

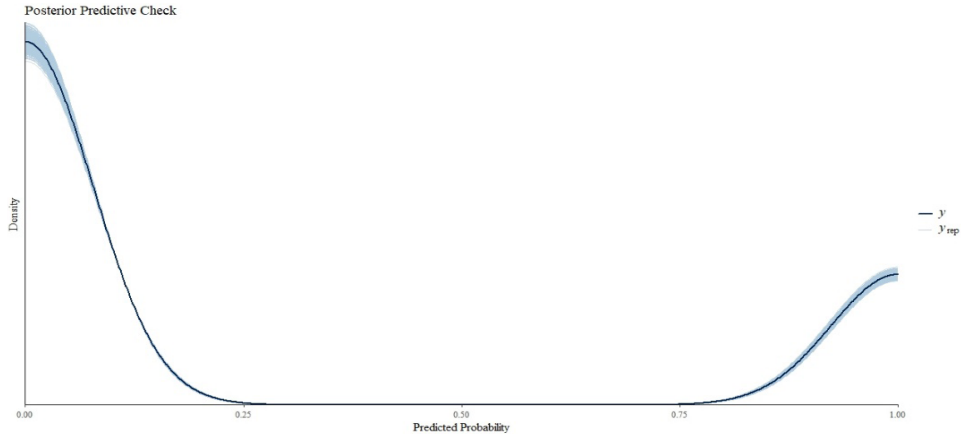

FIGURE 2. Posterior predictive check

LIMITATION AND FUTURE DIRECTION

Despite the advantages of the Bayesian framework, such as transparent uncertainty quantification and prior knowledge incorporation, this study has limitations. Specifically, it relied on secondary data, restricting variable selection to those already available, and focused solely on CBC parameters and limited demographic characteristics. Furthermore, other potential predictors, such as family history, lifestyle, and comorbidities, were not included due to data unavailability. Additionally, the observational design precludes any causal inference.

Regardless of these limitations, the clinical implications are notable. CBC parameters are inexpensive, easily accessible, and routinely collected, making them well-suited for integration into prognostic models to support early risk stratification and decision-making. Future research should explore dynamic modelling strategies, such as hierarchical and survival models, to capture time-to-event outcomes. Bayesian networks may also help show complex interdependencies among variables and improve predictive accuracy in clinical risk modelling. In addition, future research should consider primary data collection or access to comprehensive and multimodal datasets that include laboratory, clinical, behavioural, comorbidities, and family medical history to provide a more holistic and accurate understanding of patient outcomes. Besides, it is recommended that the hospital's data management centre integrate these variables into its database infrastructure to enhance the quality and scope of future research.

CONCLUSIONS

This study evaluated the prognostic utility of CBC parameters in predicting mortality among kidney disease patients using BLR. Thirteen variables, including age, hospitalisation status, and markers of red cell morphology, platelet activity, and systemic inflammation, emerged as significant predictors and valuable indicators for mortality risk stratification. The model demonstrated excellent

performance, with strong calibration, discrimination, and predictive accuracy across datasets. These results affirm that the model was well-calibrated and effectively differentiated between survivors and non-survivor patients, indicating its reliability and generalisability.

These findings highlight the practical value of combining CBC parameters with demographic information for mortality risk assessment among kidney disease patients, particularly in Sabah. The model's predicted probabilities may support the development of an automated risk-detection system that classifies patients into lower- and higher-risk groups using routinely available CBC, demographic, and patient profile data. Such a system may help clinicians identify patients who require closer monitoring and earlier management. As CBC is affordable and widely accessible, this approach may be especially useful in resource-limited settings, where low-cost tools are needed to support timely risk detection and clinical decision-making.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Ministry of Higher Education Malaysia for funding this research through the Fundamental Research Grant Scheme (FRGS) (Project Code: FRGS/1/2023/STG06/UMS/02/2). The authors also thank Queen Elizabeth Hospital for its collaboration and for providing the data used in this study. Artificial intelligence was used only for limited language editing, specifically grammar checking and improving sentence flow. Supplementary data are available from the corresponding author upon request.

REFERENCES

Agnello, L., Giglio, R.V., Bivona, G., Scazzone, C., Gambino, C.M., Iacona, A., Ciaccio, A.M., Lo Sasso, B. & Ciaccio, M. 2021. The value of a complete blood count (CBC) for sepsis diagnosis and prognosis. *Diagnostics* 11(10): 1881. https://doi.org/10.3390/diagnostics11101881

Arif, F.M., Sumida, K., Molnar, M.Z., Potukuchi, P.K., Lu, J.L., Hassan, F., Thomas, F., Siddiqui, O.A., Gyamlani, G.G., Kalantar-Zadeh, K. & Kovesdy, C.P. 2017. Early mortality associated with inpatient versus outpatient hemodialysis initiation in a large cohort of US veterans with incident end-stage renal disease. *Nephron* 137(1): 15-22. https://doi.org/10.1159/000473704

Aw, J.Y.H., Clarke, N.E., Mayfield, H., Lau, C.L., Richardson, A. & Nery, S.V. 2021. Novel statistical approaches to identify risk factors for soil-transmitted helminth infection in Timor-Leste. *International Journal for Parasitology* 51(9): 729-739. https://doi.org/10.1016/j.ijpara.2021.01.005

Balasa, M.J. 2023. Accuracy of platelet indices as predictor of in-hospital mortality in patients with acute kidney injury requiring renal replacement therapy. *Nephrology Dialysis Transplantation* 38(Suppl 1). https://doi.org/10.1093/ndt/gfad063d_5304

Binuya, M.A.E., Engelhardt, E.G., Schats, W., Schmidt, M.K. & Steyerberg, E.W. 2022. Methodological guidance for the evaluation and updating of clinical prediction models: A systematic review. *BMC Medical Research Methodology* 22(1): 316. https://doi.org/10.1186/s12874-022-01801-8

Bui, M., Groothuis-Oudshoorn, C.G.M., Witteveen, A. & Hegeman, J.H. 2024. Palliative non-operative management in geriatric hip fracture patients: When would surgeons abstain from surgery? *Journal of Clinical Medicine* 13(6): 1594. https://doi.org/10.3390/jcm13061594

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

Bürkner, P.C., Gabry, J., Weber, S., Johnson, A., Modrak, M., Badr, H.S., Weber, F., Vehtari, A., Ben-Shachar, M.S., Rabel, H., Mills, S.C., Wild, S., Popov, V., Kosmidis, I., Schneider, B., Kallioinen, N., Silva, W.J., Carvalho, L. & Pekin, S. 2024. brms: Bayesian regression models using 'Stan' (R package version 2.22.0). *CRAN*. https://cran.r-project.org/web/packages/brms/brms.pdf

Çeleğen, K. & Çeleğen, M. 2024. Prognostic significance of mean platelet volume to platelet count ratio in pediatric patients with acute kidney injury. *The Turkish Journal of Pediatrics* 6(3): 354-363. https://doi.org/10.24953/turkjpediatr.2024.4514

Chicco, D., Tötsch, N. & Jurman, G. 2021. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Mining* 14. https://doi.org/10.1186/s13040-021-00244-z

Dean, L. 2005. *Blood Groups and Red Cell Antigens* [Internet]. Bethesda (MD): National Center for Biotechnology Information (US). https://www.ncbi.nlm.nih.gov/books/NBK2261/

Depaoli, S. & van de Schoot, R. 2017. Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological Methods* 22(2): 240-261. https://doi.org/10.1037/met0000065

Dratch, A., Kleine, C.E., Streja, E., Soohoo, M., Park, C., Hsiung, J.T., Rhee, C.M., Obi, Y., Molnar, M.Z., Kovesdy, C.P. & Kalantar-Zadeh, K. 2019. Mean corpuscular volume and mortality in incident hemodialysis patients. *Nephron* 141(3): 188-200. https://doi.org/10.1159/000495726

Drăgoescu, A.N., Pădureanu, V., Stănculescu, A.D., Chiuțu, L.C., Tomescu, P., Geormăneanu, C., Pădureanu, R., Iovănescu, V.F., Ungureanu, B.S., Pănuș, A. & Drăgoescu, O.P. 2021. Neutrophil to lymphocyte ratio (NLR) - A useful tool for the prognosis of sepsis in the ICU. *Biomedicines* 10(1): 75. https://doi.org/10.3390/biomedicines10010075

Egidi, L., Pauli, F. & Torelli, N. 2022. Avoiding prior–data conflict in regression models via mixture priors. *Canadian Journal of Statistics* 50(2): 491-510. https://doi.org/10.1002/cjs.11637

Fan, F., Jia, J., Li, J., Huo, Y. & Zhang, Y. 2017. White blood cell count predicts the odds of kidney function decline in a Chinese community-based population. *BMC Nephrology* 18(1): 190. https://doi.org/10.1186/s12882-017-0608-4

Fanconi, C., de Hond, A., Peterson, D., Capodici, A. & Hernandez-Boussard, T. 2023. A Bayesian approach to predictive uncertainty in chemotherapy patients at risk of acute care utilization. *EBioMedicine* 92: 104632. https://doi.org/10.1016/j.ebiom.2023.104632

Farhin, S. & Ali Khan, A. 2022. Bayesian survival analysis of acute encephalitis syndrome with censoring mechanism using brms package. *Journal of Statistics Applications & Probability* 11(3): 963-980. https://doi.org/10.18576/jsap/110319

Francis, A., Harhay, M.N., Ong, A.C.M., Tummalapalli, S.L., Ortiz, A., Fogo, A.B., Fliser, D., Roy-Chaudhury, P., Fontana, M., Nangaku, M., Wanner, C., Malik, C., Hradsky, A., Adu, D., Bavanandan, S., Cusumano, A., Sola, L., Ulasi, I., Jha, V., American Society of Nephrology, European Renal Association & International Society of Nephrology. 2024. Chronic kidney disease and the global public health agenda: An international consensus. *Nature Reviews Nephrology* 20(7): 473-485.

Gabry, J. 2024. loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models (Version 2.8.0) [R package]. https://cloud.r- project.org/web/packages/loo/loo.pdf

Gameiro, J. & Lopes, J.A. 2019. Complete blood count in acute kidney injury prediction: A narrative review. *Annals of Intensive Care* 9: 87. https://doi.org/10.1186/s13613-019-0561-4

Ganapathy, S., Harichandrakumar, K.T., Penumadu, P., Tamilarasu, K. & Nair, N.S. 2022. Comparison of Bayesian, frequentist and machine learning models for predicting the two-year mortality of patients diagnosed with squamous cell carcinoma of the oral cavity. *Clinical Epidemiology and Global Health* 17: 101145. https://doi.org/10.1016/j.cegh.2022.101145

Gelman, A., Jakulin, A., Pittau, M.G. & Su, Y.S. 2008. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* 2(4): 1360-1383. https://doi.org/10.1214/08-AOAS191

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. 2025. *Bayesian Data Analysis*. 3rd ed. New York: Chapman and Hall. https://sites.stat.columbia.edu/gelman/book/BDA3.pdf

Gerber, G. & Scharler, U.M. 2024. Jump around: Selecting Markov Chain Monte Carlo parameters and diagnostics for improved food web model quality and ecosystem representation. *Ecological Informatics* 84: 102865. https://doi.org/10.1016/j.ecoinf.2024.102865

Gosho, M., Ishii, R., Nagashima, K., Noma, H. & Maruo, K. 2025. Determining the prior mean in Bayesian logistic regression with sparse data: A nonarbitrary approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 74(1): 126-141. https://doi.org/10.1093/jrsssc/qlae048

Gunn-Sandell, L.B., Bedrick, E.J., Hutchins, J.L., Berg, A.A., Kaizer, A.M. & Carlson, N.E. 2024. A practical guide to adopting Bayesian analyses in clinical research. *Journal of Clinical and Translational Science* 8(1): e3. https://doi.org/10.1017/cts.2023.689

Guo, S., Yang, L., Zhu, X., Zhang, X., Sun, Z., Meng, L., Wang, Y., Li, J., Cheng, S., Zhuang, X. & Cui, W. 2024. Risk factors of different mortality periods in older patients with end-stage renal disease undergoing urgent-start peritoneal dialysis: A retrospective observational study. *BMC Geriatrics* 24(1): 343. https://doi.org/10.1186/s12877-024-04931-4

Han, W., Allameen, N.A., Ibrahim, I., Dhanasekaran, P., Mengling, F. & Lahiri, M. 2022. Hospital admission risk stratification of patients with gout presenting to the emergency department. *Clinical Rheumatology* 41(6): 1801-1807. https://doi.org/10.1007/s10067-021-05902-5

Hashmi, M.F., Benjamin, O. & Lappin, S.L. 2023. *End-Stage Renal Disease* [updated 2023 Aug 28]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK499861/

Hoffman, M.D. & Gelman, A. 2014. The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *arXiv* [Preprint]. https://arxiv.org/pdf/1111.4246.pdf

Huang, Y., Li, W., Macheret, F., Gabriel, R.A. & Ohno-Machado, L. 2020. A tutorial on calibration measurements and calibration models for clinical prediction models. *Journal of the American Medical Informatics Association* 27(4): 621-633. https://doi.org/10.1093/jamia/ocz228

Huynh, Q.T.V., Le, N.Q.K., Huang, S.Y., Ho, B.T., Vu, T.H., Pham, H.T.M., Pham, A.L., Hou, J.W., Nguyen, N.T.K. & Chen, Y.C. 2022. Development and validation of clinical diagnostic model for girls with central precocious puberty: Machine-learning approaches. *PLOS ONE* 17(1): e0261965. https://doi.org/10.1371/journal.pone.0261965

Jeon, Y.H., Jeon, Y., Jung, H.Y., Choi, J.Y., Park, S.H., Kim, C.D., Kim, Y.L., Cho, J.H. & Lim, J.H. 2023. Platelet-to- lymphocyte ratio and in-hospital mortality in patients with AKI receiving continuous k i d n e y replacement therapy: A retrospective observational cohort study. *Kidney Medicine* 5(6): 100642. https://doi.org/10.1016/j.xkme.2023.100642

Johnson, A.A., Ott, M.Q. & Dogucu, M. 2022. *Bayes Rules!: An Introduction to Applied Bayesian Modeling*. 1st ed. Boca Raton: Chapman and Hall/CRC. https://doi.org/10.1201/9780429288340

Juhan, N., Zubairi, Y.Z., Khalid, Z.M. & Zuhdi, A.S.M. 2020a. A comparison between Bayesian and frequentist approach in the analysis of risk factors for female cardiovascular disease patients in Malaysia. *ASM Science Journal* https://doi.org/10.32802/asmscj.2020.sm26(1.1)

Juhan, N., Zubairi, Y.Z., Khalid, Z.M. & Mahmood Zuhdi, A.S. 2020b. Risk factors of mortality among male patients with cardiovascular disease in Malaysia using Bayesian analysis. *Iranian Journal of Public Health* 49(9): 1642-1649. https://doi.org/10.18502/ijph.v49i9.4080

Khan, M.I. & Ullah, I. 2020. Diagnostic importance of mean platelet volume, platelet distribution width and platelet large cell ratio as screening tool in immune thrombocytopenia. *Porto Biomedical Journal* 5(6): e094. https://doi.org/10.1097/j.pbj.0000000000000094

Kim, M., Kym, D., Park, J., Yoon, J., Cho, Y.S., Hur, J., Chun, W. & Yoon, D. 2024a. Big data insights into the diagnostic values of CBC parameters for sepsis and septic shock in burn patients: A retrospective study. *Scientific Reports* 14: 800. https://doi.org/10.1038/s41598-023-50695-z

Kim, Y.H., Lee, W., Kim, K.Y., Kim, Y., Ko, A., Weon, B., Lee, J., Jin, W., Kim, D.K., Kim, Y.S., Lim, C.S., Lee, J.P.; Korean Association for the study of Renal Anemia, artificial Intelligence (KARAI) 2024b. The estimated mediating roles of anemia-related variables in the association between kidney function and mortality: A National Health and Nutrition Examination Survey (NHANES) study. *Scientific Reports* 14(1): 6621. https://doi.org/10.1038/s41598-024-56877-7

Kimura, H., Tanaka, K., Saito, H., Iwasaki, T., Kazama, S., Shimabukuro, M., Asahi, K., Watanabe, T. & Kazama, J.J. 2023. Impact of red blood cell distribution width–albumin ratio on prognosis of patients with CKD. *Scientific Reports* 13(1): 15774. https://doi.org/10.1038/s41598-023-42986-2

Kor, C.T., Hsieh, Y.P., Chang, C.C. & Chiu, P.F. 2018. The prognostic value of interaction between mean corpuscular volume and red cell distribution width in mortality in chronic kidney disease. *Scientific Reports* 8(1): 11870. https://doi.org/10.1038/s41598-018-19881-2

Kruschke, J.K. 2021. Bayesian analysis reporting guidelines. *Nature Human Behaviour* 5(10): 1282-1291. https://doi.org/10.1038/s41562-021-01177-7

Larson, C., Kaplan, D., Girolamo, T., Kover, S.T. & Eigsti, I.M. 2023. A Bayesian statistics tutorial for clinical research: Prior distributions and meaningful results for small clinical samples. *Journal of Clinical Psychology* 79(11): 2602-2624. https://doi.org/10.1002/jclp.23570

Lemoine, N.P. 2019. Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos* 128(7): 912-928. http://dx.doi.org/10.1111/oik.05985

Li, D., Wang, A., Li, Y., Ruan, Z., Zhao, H., Li, J., Zhang, Q. & Wu, B. 2024. Nonlinear relationship of red blood cell indices (MCH, MCHC, and MCV) with all-cause and cardiovascular mortality: A cohort study in U.S. adults. *PLoS ONE* 19(8): 1-13. https://doi.org/10.1371/journal.pone.0307609

Liao, J., Wei, D., Sun, C., Yang, Y., Wei, Y. & Liu, X. 2022. Prognostic value of the combination of neutrophil-to-lymphocyte ratio, monocyte-to-lymphocyte ratio and platelet-to-lymphocyte ratio on mortality in patients on maintenance hemodialysis. *BMC Nephrology* 23(1): 393. https://doi.org/10.1186/s12882-022-03020-1

Lindhiem, O., Petersen, I.T., Mentch, L.K. & Youngstrom, E.A. 2018. The importance of calibration in clinical psychology. *Assessment* 27(4): 840-854. https://doi.org/10.1177/1073191117752055

Lindholm, B., Zhang, X., Chen, Z., Qureshi, A.R.T., Danielson, K., Mukai, H., Suliman, M., Heimbürger, O., Chen, J., Stenvinkel, P., Barany, F.P. 2022. MO673: Association of mean corpuscular volume with mortality in kidney failure patients: Opposite findings in Chinese and Swedish kidney failure patients. *Nephrology Dialysis Transplantation* 37(Suppl 3): gfac078.010. https://doi.org/10.1093/ndt/gfac078.010

Lukman, P.A., Abdullah, S. & Rachman, A. 2021. Bayesian logistic regression and its application for hypothyroid prediction in post-radiation nasopharyngeal cancer patients. *Journal of Physics: Conference Series* 1725(1): 012010. https://doi.org/10.1088/1742-6596/1725/1/012010

Luo, J., Zhou, Y., Song, Y., Wang, D., Li, M., Du, X., Kang, J., Ye, P. & Xia, J. 2024. Association between the neutrophil-to-lymphocyte ratio and in-hospital mortality in patients with chronic kidney disease and coronary artery disease in the intensive care unit. *European Journal of Medical Research* 29(1): 260. https://doi.org/10.1186/s40001-024-01850-3

Ma, J., Vongpradith, A., Ledesma, J.R., Novotney, A., Yi, S., Lim, K., Hay, S.I., Murray, C.J.L. & Kyu, H.H. 2022. Progress towards the 2020 milestones of the End TB Strategy in Cambodia: Estimates of age- and sex-specific TB incidence and mortality from the Global Burden of Disease Study 2019. *BMC Infectious Diseases* 22(1): 904. https://doi.org/10.1186/s12879-022-07891-5

Ma, L., Han, Q., Sun, F., Zhu, K. & Sun, Q. 2023. Mean platelet volume/platelet count ratio as a predictor of both incidence and prognosis of acute ischemic stroke in hemodialysis patients. *International Journal of General Medicine* 16: 3985-3994. https://doi.org/10.2147/IJGM.S416817

Mahachai, N., Washirasaksiri, C., Ariyakunaphan, P., Kositamongkol, C., Sitasuwan, T., Tinmanee, R., Auesomwang, C., Sayabovorn, N., Chaisathaphol, T., Phisalprapa, P., Charatcharoenwitthaya, P. & Srivanichakorn, W. 2023. Clinical predictive score for identifying metabolic dysfunction-associated steatotic liver disease in individuals with prediabetes using transient elastography. *Journal of Clinical Medicine* 12(24): 7617. https://doi.org/10.3390/jcm12247617

Marine, B.T. & Zewde, M.G. 2024. Modern contraceptive utilization among street beggar women in the Sidama region of Ethiopia using Bayesian model approach. *Reproductive Health* 21: 163. https://doi.org/10.1186/s12978-024-01905-y

Mattos, D.I. & Ramos, É.M.S. 2022. Bayesian paired comparison with the *bpcs* package. *Behavior Research Methods* 54(4): 2025-2045. https://doi.org/10.3758/s13428-021-01714-2

Mohd Faizal, A.S., Hon, W.Y., Thevarajah, T.M., Khor, S.M. & Chang, S-W. 2023. A biomarker discovery of acute myocardial infarction using feature selection and machine learning. *Medical & Biological Engineering & Computing* 61(10): 2527-2541. https://doi.org/10.1007/s11517-023-02841-y

Nguyen, L.P., Tung, D.D., Nguyen, D.T., Le, H.N., Tran, T.Q., Binh, T.V. & Pham, D.T.N. 2023. The utilization of machine learning algorithms for assisting physicians in the diagnosis of diabetes. *Diagnostics* 13(12): 2087. https://doi.org/10.3390/diagnostics13122087

Oh, E.S., You, Z., Nowak, K.L. & Jovanovich, A.J. 2022. Association of monocyte count and monocyte/lymphocyte ratio with the risk of cardiovascular outcomes in patients with CKD. *Kidney360* 3(4): 657-665. https://doi.org/10.34067/KID.0007922021

Omar, M.S.F., Md Noor, S., Seman, Z. & Mohd Tajuddin, S.N. 2023. Full blood count parameters in COVID-19 patients with disease severity, patient outcome and vaccination status. *Malaysian Journal of Medicine and Health Sciences* 19(Suppl 16): 16-23. https://doi.org/10.47836/mjmhs.19.s16.4

Plummer, M., Best, N., Cowles, K., Vines, K., Sarkar, D., Bates, D., Almond, R. & Magnusson, A. 2024. *coda*: Output analysis and diagnostics for MCMC (Version 0.19-4.1) [R package]. *CRAN*. https://cran.r-project.org/web/packages/coda/coda.pdf

Pogorzelska, K., Krętowska, A., Krawczuk-Rybak, M. & Sawicka-Żukowska, M. 2020. Characteristics of platelet indices and their prognostic significance in selected medical conditions: A systematic review. *Advances in Medical Sciences* 65(2): 310-315. https://doi.org/10.1016/j.advms.2020.05.002

Ramalu, R.R., Kandasamy, R., Raffiq, A., Idris, Z. & Adnan, J.S. 2022. Risk factors associated with outcome of spontaneous intracerebral haemorrhage: Hospital Kuala Lumpur experience. *Malaysian Journal of Medical Sciences* 29(1): 76-90. https://doi.org/10.21315/mjms2022.29.1.8

Sabah State Government. 2024. *People & History* [online]. https://sabah.gov.my/en/content/people-history (Accessed on 23 February 2025).

Sabanayagam, C., He, F., Nusinovici, S., Li, J., Lim, C., Tan, G. & Cheng, C.Y. 2023. Prediction of diabetic kidney disease risk using machine learning models: A population-based cohort study of Asian adults. *eLife* 12: 81878. https://doi.org/10.7554/eLife.81878

Schulman, I.H., Chan, K., Der, J.S., Wilkins, K.J., Corns, H.L., Sayer, B., Ngo, D.A., Eggers, P., Norton, J., Shah, N., Mendley, S., Parsa, A., Star, R.A. & Kimmel, P.L. 2023. Readmission and mortality after hospitalization with acute kidney injury. *American Journal of Kidney Diseases* 82(1): 63-74.e1. https://doi.org/10.1053/j.ajkd.2022.12.008

Seo, I.H. & Lee, Y.J. 2022. Usefulness of complete blood count (CBC) to assess cardiovascular and metabolic diseases in clinical settings: A comprehensive literature review. *Biomedicines* 10(11): 2697. https://doi.org/10.3390/biomedicines10112697

Sheng, K., Zhang, P., Yao, X., Li, J., He, Y. & Chen, J. 2020. Prognostic machine learning models for first-year mortality in incident hemodialysis patients: Development and validation study. *JMIR Medical Informatics* 8(10): e20578. https://doi.org/10.2196/20578

Ting, C.Y., Zakariah, H. & Mohd Yusri, Y.Z. 2022. Probabilistic risk assessment of COVID-19 patients at COVID-19 assessment centre. *International Journal of Technology* 13(6): 1193. https://doi.org/10.14716/ijtech.v13i6.5882

Thombare, D., Bhalerao, A., Dixit, P., Joshi, S. & Dapkekar, P. 2023. Neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio in antenatal women with pre-eclampsia: A case-control study. *Cureus* 15(6): e40338. https://doi.org/10.7759/cureus.40338

Thölke, P., Mantilla-Ramos, Y.J., Abdelhedi, H., Maschke, C., Dehgan, A., Harel, Y., Kemtur, A., Berrada, L.M., Sahraoui, M., Young, T., Pépin, A.B., El Khantour, C., Landry, M., Pascarella, A., Hadid, V., Combrisson, E., O'Byrne, J. & Jerbi, K. 2023. Class imbalance should not throw you off balance: Choosing the right classifiers and performance metrics for brain decoding with imbalanced data. *NeuroImage* 277: 120253. https://doi.org/10.1016/j.neuroimage.2023.120253

Tsheten, T., Alene, K.A., Restrepo, A.C., Kelly, M., Lau, C., Clements, A.C.A., Gray, D.J., Daga, C., Mapalo, V.J., Espino, F.E. & Wangdi, K. 2024. Risk mapping and socio-ecological drivers of soil-transmitted helminth infections in the Philippines: A spatial modelling study. *The Lancet Regional Health – Western Pacific* 43: 100974. https://doi.org/10.1016/j.lanwpc.2023.100974

Uludag, K. & Arikan, T. 2021. Is white blood cell count associated with mortality in peritoneal dialysis patients? A retrospective single-center analysis. *Cureus* 13(11): e19728. https://doi.org/10.7759/cureus.19728

Umeres-Francia, G.E., Rojas-Fernández, M.V., Herrera-Añazco, P. & Benites-Zapata, V.A. 2022. Neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio as risk factors for mortality in Peruvian adults with chronic kidney disease. *Renal Replacement Therapy* 8: 30. https://doi.org/10.1186/s41100-022-00420-9

van de Schoot, R., Veen, D., Smeets, L., Winter, S.D. & Depaoli, S. 2020. A tutorial on using the WAMBS checklist to avoid the misuse of Bayesian statistics. In *Small Sample Size Solutions: A Guide for Applied Researchers and Practitioners*, edited by van de Schoot, R. & Miočević, M. Abingdon: Routledge. pp. 30-49. https://doi.org/10.4324/9780429273872

Wang, Z., Nie, J., Song, X., Denson, L.A. & Hyams, J.S. 2024. A comparison of Bayesian and frequentist approaches to incorporating clinical and biological information for the prediction of response to standardized pediatric colitis therapy. *PLoS ONE* 19(3): e0295814. https://doi.org/10.1371/journal.pone.0295814

Wang, J., Ma, X., Si, X., Wu, M. & Han, W. 2023. Mean platelet volume and the association with all-cause mortality and cardiovascular mortality among incident peritoneal dialysis patients. *BMC Cardiovascular Disorders* 23(1): 543. https://doi.org/10.1186/s12872-023-03551-x

Wei, L., Mao, S., Liu, X. & Zhu, C. 2024. Association of systemic inflammation response index with all-cause mortality as well as cardiovascular mortality in patients with chronic kidney disease. *Frontiers in Cardiovascular Medicine* 11: 1363949. https://doi.org/10.3389/fcvm.2024.1363949

Wolde, H.F., Clements, A.C., Gilmour, B. & Alene, K.A. 2024. Spatial co-distribution of tuberculosis prevalence and low BCG vaccination coverage in Ethiopia. *Scientific Reports* 14(1): 31561. https://doi.org/10.1038/s41598-024-68549-7

Wong, R.S. & Ismail, N.A. 2016. An application of Bayesian approach in modeling risk of death in an intensive care unit. *PLoS ONE* 11(3): e0151949. https://doi.org/10.1371/journal.pone.0151949

Wong, J.E., Yamaguchi, M., Nishi, N., Araki, M. & Wee, L.H. 2022. Predicting overweight and obesity status among Malaysian working adults with machine learning or logistic regression: Retrospective comparison study. *JMIR Formative Research* 6(12). https://doi.org/10.2196/40404

World Health Organization. 2024a. *The Top 10 Causes of Death* [online].https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (Accessed on 21 February 2025).

World Health Organization. 2024b. Data – Malaysia [online]. https://data.who.int/countries/458 (Accessed on 21 February 2025).

Wulandari, E. & Hapsari, R.A.F. 2024. Innovative hematology analysis using menstrual blood. *Medical Devices: Evidence and Research* 17: 285-293. https://doi.org/10.2147/MDER.S468500

Yadav, R., Yadav, V., Pokhriyal, S., Zahid, U. & Gandhi, A. 2023. Retrospective observational study of complete blood count (CBC) parameters and ICU mortality of COVID-19 disease in Delta variant and Omicron variant in a community-based hospital in New York City. *Cureus* 15(2): e34894. https://doi.org/10.7759/cureus.34894

Yang, K., Yang, N., Sun, W., Dai, L., Jin, J., Wu, J. & He, Q. 2023. The association between albumin and mortality in patients with acute kidney injury: A retrospective observational study. *BMC Nephrology* 24(1): 332. https://doi.org/10.1186/s12882-023-03323-x

Yip, W., Ng, S.H.X., Kaur, P., George, P.P., Guan, J.H.C., Lee, G., Koh, T.J.K., Tan, W.S. & Hum, A.Y.M. 2024. Risk factors for short-term all-cause mortality in patients with end-stage renal disease: A scoping review. *BMC Nephrology* 25(1): 71. https://doi.org/10.1186/s12882-024-03503-3

Zhao, X., Karaboyas, A., Gan, L., Hou, F.F., Liang, X., Chen, X., Chen, Y., Ni, Z., Pecoits-Filho, R. & Zuo, L. 2024. Platelet count has a U-shaped association with mortality in hemodialysis patients. *Scientific Reports* 14(1): 26572. https://doi.org/10.1038/s41598-024-77718-7

*Corresponding author; email: liyana87@ums.edu.my

Supplementary Materials: Bayesian Logistic Regression to Explore the Role of Complete Blood Count in Kidney Disease Mortality

Supplement A (Visual Convergence Diagnostics)

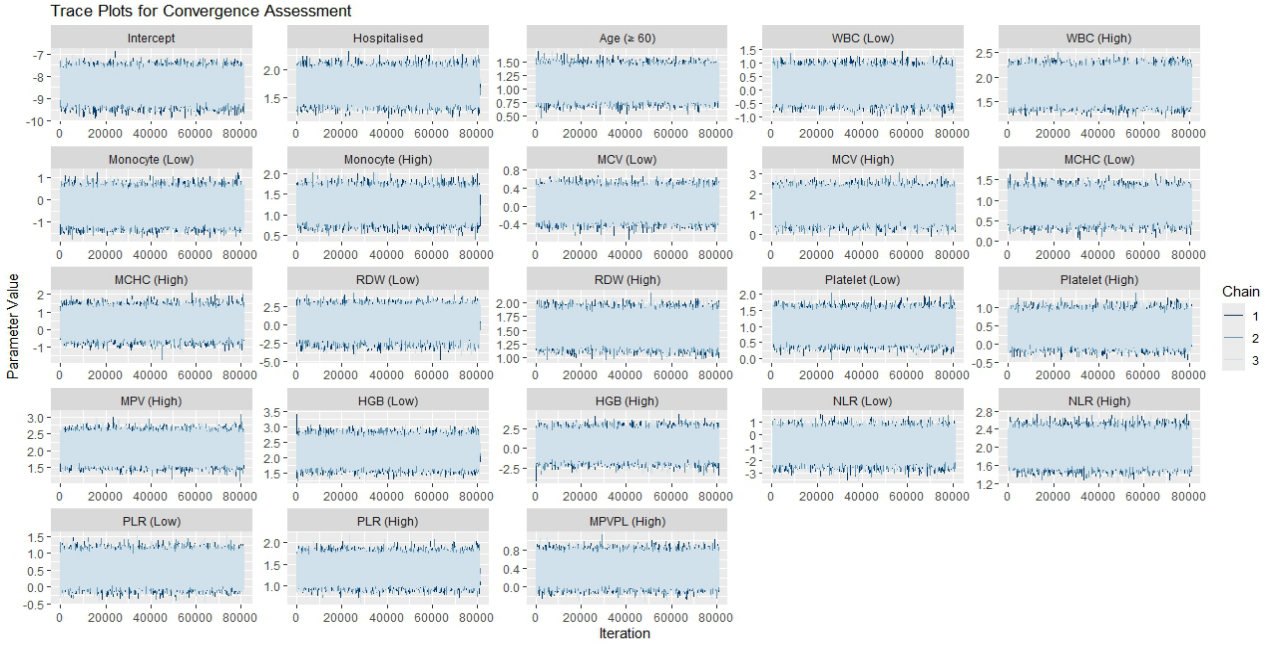

FIGURE S.1. Trace plots

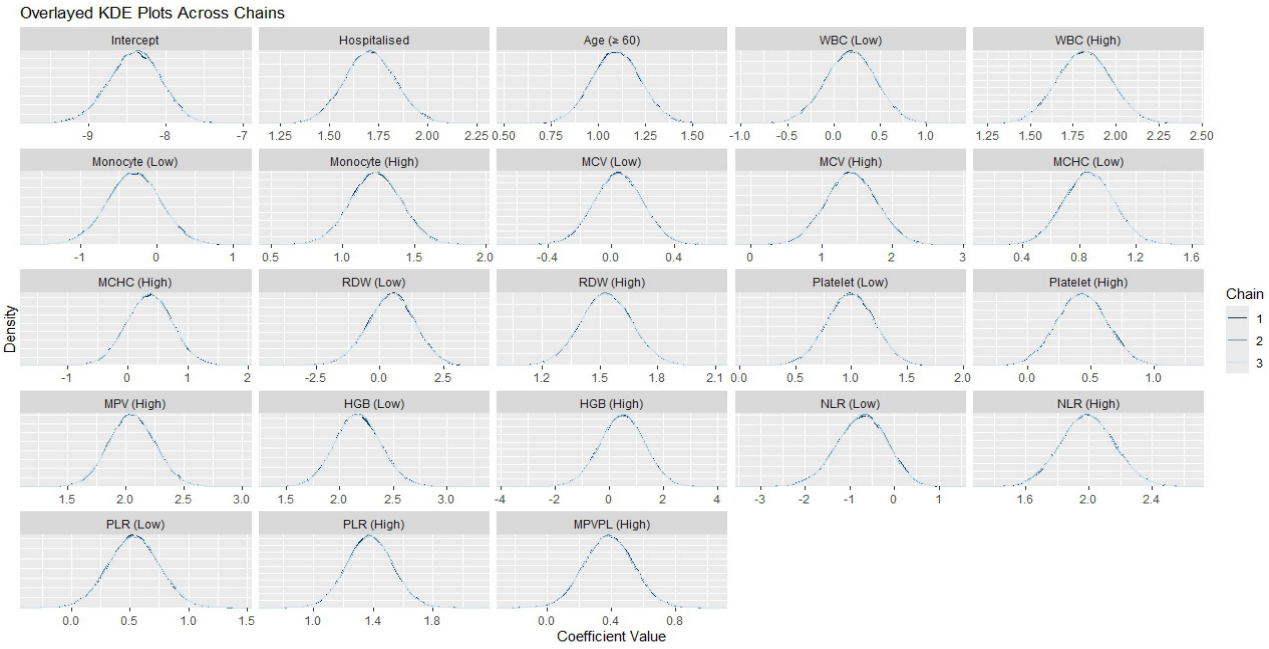

FIGURE S.2. Kernel density estimate (KDE) plots

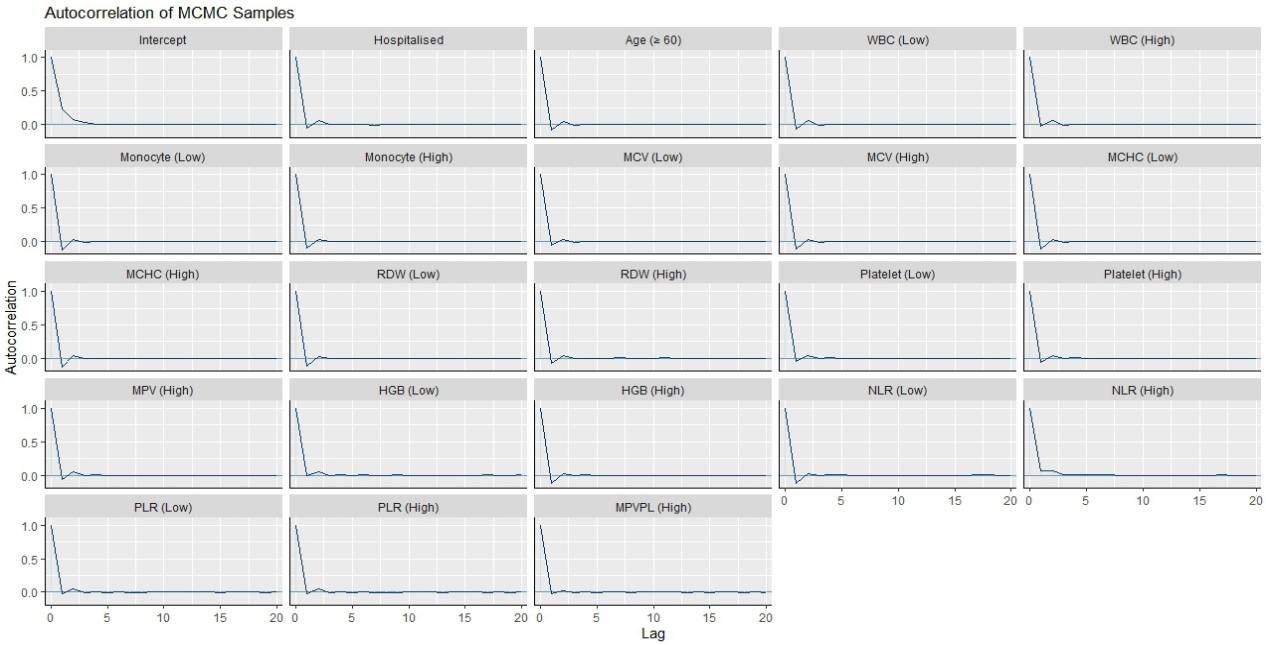

FIGURE S.3. Autocorrelation plots

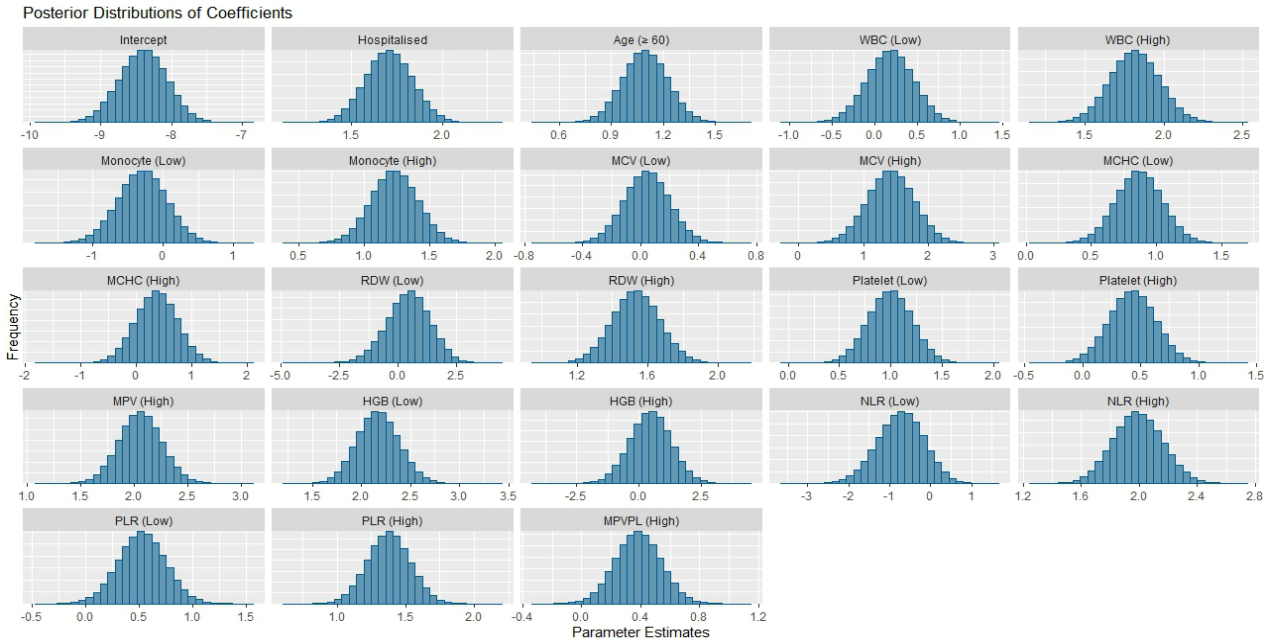

FIGURE S.4. Histograms of the posterior distributions

Supplement B (Numerical Convergence Diagnostics)

TABLE S.1. Numerical convergence diagnostics

No.	Parameter	Gelman-Rubin R-hat	Effective Sample Size (ESS)		Raftery-Lewis Diagnostic				Heidelberger-Welch Diagnostic					Geweke's Diagnostic	
			ESS bulk	ESS tail	Warm-up (Burn-in) (M)	Total (N)	Lower Bound (Nmin)	Depen Factor (I)	Stationarity Test	*P*-Value	Halfwidth Test	Mean	Halfwidth	Z-score	*P*-value
1.	Intercept	1.0000144	146420.7293	145872.2705	6	8196	3746	2.19	passed	0.3137	passed	-8.3747	0.0017	0.8824	0.3776
2.	Hospitalised	1.0000089	249207.4084	167593.3629	6	8096	3746	2.16	passed	0.3816	passed	1.7036	0.0005	0.1933	0.8467
3.	Age (≥ 60)	1.0000234	263160.0304	168222.811	6	8274	3746	2.21	passed	0.4729	passed	1.0874	0.0005	-0.7064	0.48
4.	WBC (Low)	1.0000077	254224.8356	172283.0834	6	8138	3746	2.17	passed	0.6303	passed	0.1827	0.0011	-0.3656	0.7147
5.	WBC (High)	1.0000087	220954.3276	170060.0717	4	7780	3746	2.08	passed	0.36	passed	1.8094	0.0007	0.5944	0.5522
6.	Monocyte (Low)	0.9999967	292664.9319	166639.4019	6	8286	3746	2.21	passed	0.6474	passed	-0.2936	0.0012	1.2631	0.2065
7.	Monocyte (High)	1.0000272	283931.5499	171076.5335	6	8266	3746	2.21	passed	0.4762	passed	1.2277	0.0006	-0.0558	0.9555
8.	MCV (Low)	1.0000231	265285.367	171229.7744	6	8094	3746	2.16	passed	0.1732	passed	0.044	0.0006	0.7414	0.4585
9.	MCV (High)	1.0000022	304587.6843	161653.5898	6	8452	3746	2.26	passed	0.2197	passed	1.4134	0.0013	-0.6194	0.5357
10.	MCHC (Low)	0.9999963	283212.1852	167230.0158	6	8274	3746	2.21	passed	0.1944	passed	0.8667	0.0007	1.6818	0.0926
11.	MCHC (High)	1.0000176	306328.9451	167066.7147	6	8478	3746	2.26	passed	0.6586	passed	0.3626	0.0013	-1.3194	0.187
12.	RDW (Low)	1.0000124	303115.0406	151245.4491	6	8868	3746	2.37	passed	0.6864	passed	0.4335	0.0034	-1.1185	0.2633
13.	RDW (High)	1.0000120	255857.724	170634.8609	6	8140	3746	2.17	passed	0.623	passed	1.5321	0.0005	-1.8696	0.0615
14.	Platelet (Low)	1.0000004	234432.4534	169718.8805	6	8026	3746	2.14	passed	0.9646	passed	0.9828	0.0009	-0.6222	0.5338
15.	Platelet (High)	1.0000079	253538.3598	171931.7526	4	8086	3746	2.16	passed	0.3793	passed	0.432	0.0008	-1.6107	0.1072
16.	MPV (High)	0.9999990	247936.4777	169535.8682	6	8250	3746	2.2	passed	0.799	passed	2.0379	0.0008	0.0789	0.9371
17.	HGB (Low)	0.9999927	223490.9616	169068.7282	4	7672	3746	2.05	passed	0.6364	passed	2.1583	0.0009	-1.288	0.1977
18.	HGB (High)	1.0000179	296832.3917	165010.133	6	8158	3746	2.18	passed	0.8642	passed	0.4667	0.0029	-0.0981	0.9219
19.	NLR (Low)	1.0000280	303388.0311	159132.2176	6	8354	3746	2.23	passed	0.8678	passed	-0.7342	0.002	0.162	0.8713
20.	NLR (High)	1.0000115	186702.2998	168173.608	6	7914	3746	2.11	passed	0.8826	passed	1.9872	0.0008	-0.1376	0.8906
21.	PLR (Low)	1.0000096	224110.6282	171578.0201	6	8312	3746	2.22	passed	0.7041	passed	0.5306	0.0009	0.1667	0.8676
22.	PLR (High)	0.9999994	224927.3447	174708.9346	6	7848	3746	2.1	passed	0.2075	passed	1.367	0.0006	0.6084	0.5429
23.	MPVPL (High)	1.0000438	243646.8763	166978.4161	6	8104	3746	2.16	passed	0.0739	passed	0.3847	0.0008	-0.4157	0.6776