BIOLOGICAL SCIENCE

Machine Learning and Multivariate Analysis of Hemorheological Dysfunction and Clinical Outcomes in HbSS Patients with Malaria Co-Infection

J. B. Ajayi[2*], O. A. Agbeyangi[1], R. A. Olalekan[3], I. P. Ige[4]

[1]Department of Public Health, Federal University of Technology Akure, Akure, Nigeria
[2]Department of Medical Microbiology and Parasitology, University of Ilesa, Ilesa, Nigeria
[3]Department of Biomedical Engineering, Federal University of Technology Akure, Akure, Nigeria
[4]Department of Medical Laboratory Science, Federal University of Technology Akure, Akure, Nigeria

Published November 2026

Abstract

Background: Malaria co-infection may aggravate hemorheological abnormalities and clinical morbidity among individuals with homozygous sickle cell disease (HbSS). However, the relationships between malaria status, blood rheology, hematological characteristics, and clinical outcomes remain incompletely characterized. This study assessed these relationships and evaluated the predictive utility of statistical and machine-learning approaches.

Methods: A retrospective cross-sectional analytical study included 45 laboratory-confirmed HbSS patients, comprising 37 malaria-positive and 8 malaria-negative individuals. Demographic, parasitological, hematological, hemorheological, and clinical variables were analysed using Welch's t-test, correlation analysis, multivariable regression, K-means clustering, principal component analysis (PCA), and machine-learning models. The analytical workflow was implemented in Google Colaboratory using Python.

Results: Malaria-positive patients had higher blood viscosity (4.80 ± 0.34 and 3.80 ± 0.43 mPa; $p < 0.001$) and lower erythrocyte deformability (0.88 ± 0.03 and 0.92 ± 0.02 mPa; $p < 0.001$) than malaria-negative patients. They also had lower haemoglobin and haematocrit, higher reticulocyte and white blood cell counts, more frequent pain crises, and longer hospitalization. Multiple regression identified pain-crisis frequency and blood viscosity as significant predictors of hospitalization duration. K-means clustering identified three patient groups with progressively higher viscosity and lower deformability. PCA showed that two components explained 72.4% of total variance. Blood viscosity discriminated malaria status with an AUC of 0.94 (95% CI: 0.86-0.99); at 4.25 mPa·s, sensitivity and specificity were 91.9% and 87.5%, respectively. Random Forest regression achieved $R2 = 0.72$ for hospitalization duration.

Conclusion: Malaria co-infection was associated with greater hemorheological impairment and clinical burden among HbSS patients. These findings support further evaluation of hemorheological biomarkers and validated predictive models in larger longitudinal cohorts.

Keywords: HbSS; sickle cell disease; malaria co-infection; hemorheology; blood viscosity; erythrocyte deformability; machine learning; hospitalization

1 Introduction

Sickle cell disease (SCD) is one of the most common inherited hemoglobin disorders worldwide and represents a major public health challenge, particularly in sub-Saharan Africa where the geographical distribution of sickle cell disease overlaps with regions of intense malaria transmission (Piel et al., 2010; Ware et al., 2017). Nigeria bears the largest burden of SCD globally, accounting for approximately 150,000 births with sickle cell anemia annually, with homozygous hemoglobin SS (HbSS) constituting the most severe clinical phenotype (Piel et al., 2013; Adigwe et al., 2023). HbSS is characterized by chronic hemolytic anemia, recurrent vaso-occlusive crises, endothelial dysfunction, systemic inflammation, and progressive multiorgan complications that significantly reduce life expectancy and quality of life (Piel et al., 2017; Quinn, 2016).

The hallmark of HbSS pathophysiology lies in profound disturbances in blood rheology. Polymerization of deoxygenated hemoglobin S induces erythrocyte sickling, membrane damage, increased cellular rigidity, abnormal aggregation, and enhanced adhesion to vascular endothelium. These abnormalities compromise microvascular flow, increase blood viscosity, and promote recurrent episodes of vaso-occlusion and tissue ischemia (Ware et al., 2017). Hemorheological dysfunction has therefore emerged as a central determinant of disease severity, influencing the occurrence of painful crises, acute chest syndrome, stroke, organ damage, and hospitalization outcomes. Alterations in blood viscosity and erythrocyte deformability are particularly important because they directly affect tissue perfusion and oxygen delivery, thereby contributing to chronic inflammatory responses and endothelial injury.

Malaria, caused predominantly by Plasmodium falciparum, remains one of the leading causes of morbidity and mortality in tropical Africa despite decades of implementation of various control strategies. Nigeria contributes substantially to the global malaria burden, with P. falciparum accounting for the majority of severe infections. The clinical spectrum of malaria ranges from asymptomatic parasitemia to severe manifestations including cerebral malaria, severe anemia, multiorgan dysfunction, and death (Taylor et al., 2012). Disease severity is influenced by complex interactions involving parasite virulence, host immune responses, environmental factors, and genetic determinants (May et al., 2007).

The relationship between malaria and sickle cell disease has fascinated researchers since Haldane's malaria hypothesis and represents one of the classical examples of balanced polymorphism (Piel et al., 2010; Bunn, 2013). Heterozygous sickle cell trait (HbAS) confers substantial protection against severe P. falciparum malaria through several mechanisms, including impaired intra-erythrocytic parasite growth, accelerated clearance of parasitized erythrocytes, enhanced humoral and cellular immunity, altered expression of Plasmodium falciparum erythrocyte membrane protein-1 (PfEMP-1), and inhibition of parasite translation by erythrocyte microRNAs (Aidoo et al., 2002; Williams et al., 2005; Lopera-Mesa et al., 2015). However, these protective effects do not extend to individuals with homozygous HbSS disease. Instead, once malaria infection occurs in HbSS patients, the pre-existing pathological abnormalities associated with sickle cell anemia may synergize with malaria-induced hemolysis and inflammation, resulting in more severe clinical consequences (Makani et al., 2010; McAuley et al., 2010).

Several mechanisms have been proposed to explain the increased vulnerability of HbSS patients during malaria infection. Functional asplenia, chronic hemolytic anemia, oxidative stress, endothelial activation, and impaired immune clearance contribute to enhanced parasite persistence and increased disease severity (Luzzatto, 2012). In addition, malaria infection causes destruction of

infected and non-infected erythrocytes, increased cytokine production, oxidative injury, and endothelial dysfunction, all of which exacerbate the rheological abnormalities already present in sickle cell disease (Komba et al., 2009; Makani et al., 2010). The resulting increase in blood viscosity and reduction in erythrocyte deformability may aggravate microvascular occlusion and precipitate vaso-occlusive crises, severe anemia, acute kidney injury, jaundice, prolonged hospitalization, and increased mortality (McAuley et al., 2010).

Although numerous studies have investigated the epidemiological and clinical interactions between malaria and sickle cell disease, most have focused on the protective role of the sickle cell trait or on conventional hematological indices such as hemoglobin concentration and white blood cell counts (May et al., 2007; Taylor et al., 2012). Comparatively fewer studies have explored hemorheological dysfunction in HbSS patients with malaria co-infection despite the pivotal role played by blood viscosity and erythrocyte deformability in disease progression. Furthermore, previous investigations have relied predominantly on classical statistical approaches, which may not adequately capture the complex, multidimensional, and non-linear relationships among parasitemia, hemolysis, inflammation, rheological abnormalities, and clinical outcomes.

Recent advances in artificial intelligence (AI), machine learning (ML), and Bayesian statistical modeling have created unprecedented opportunities for understanding complex hematological disorders. Machine learning algorithms possess the ability to uncover hidden patterns within large biomedical datasets, identify interactions among multiple variables, and generate highly accurate predictive models for disease classification and prognosis (Elsabagh et al., 2023; Machado et al., 2024). Deep learning and explainable artificial intelligence approaches have demonstrated remarkable performance in the classification of sickle cells and prediction of sickle cell complications using blood smear images and hematological parameters (de Haan et al., 2020; Goswami et al., 2024; Ussher et al., 2025). Similarly, Bayesian networks and probabilistic inference frameworks provide interpretable methods for quantifying uncertainty and estimating posterior probabilities, making them particularly suitable for clinical decision-making and risk prediction in heterogeneous patient populations (Ekong et al., 2023).

The integration of Bayesian inference and machine learning approaches offers significant advantages over traditional analytical techniques. While machine learning algorithms can identify complex non-linear interactions and optimize predictive performance, Bayesian analysis allows incorporation of prior information, probabilistic reasoning, and uncertainty quantification. Together, these methodologies provide a strong framework for understanding multifactorial diseases and developing precision medicine strategies (Petrović et al., 2020; Roy et al., 2024). Such approaches are increasingly being adopted in hematology and biomedical research to enhance disease diagnosis, prognosis, and individualized therapeutic interventions (Chen et al., 2023; Sadafi et al., 2023).

Despite the high burden of both malaria and sickle cell disease in Nigeria, there remains limited information regarding the combined effects of malaria co-infection on hemorheological dysfunction and clinical outcomes among HbSS patients. In particular, there is a paucity of studies employing Bayesian statistical frameworks and machine learning algorithms to characterize the complex interactions between rheological parameters and adverse clinical outcomes. Understanding these interactions is crucial for improving prognostic assessment and optimizing management strategies in malaria-endemic settings.

Therefore, the present study employed Bayesian statistical analysis and machine learning techniques to investigate hemorheological dysfunction and clinical outcomes among HbSS patients with malaria

co-infection in Ado-Odo/Ota Local Government Area, Ogun State, Nigeria. Specifically, the study evaluated alterations in blood viscosity, erythrocyte deformability, hematological indices, frequency of vaso-occlusive crises, and hospitalization outcomes among malaria-positive and malaria-negative HbSS patients. Furthermore, predictive models were developed to identify determinants of adverse outcomes and classify malaria infection status using hemorheological and clinical variables.

By integrating conventional statistical approaches with Bayesian inference and explainable machine learning frameworks, this study seeks to provide deeper insight into the synergistic pathological mechanisms linking malaria and sickle cell disease. The findings are expected to contribute to improved understanding of disease pathophysiology, facilitate early risk stratification, enhance prognostic assessment, and support the development of precision-based interventions for HbSS patients residing in malaria-endemic regions. Ultimately, the study aligns with the growing global emphasis on data-driven medicine and the application of artificial intelligence for improving healthcare outcomes in resource-limited settings.

2 Methodology

2.1 Study area

This study was conducted in Ado-Odo/Ota Local Government Area, Ogun State, Southwestern Nigeria, a region characterized by intense and perennial transmission of Plasmodium falciparum malaria and a considerable burden of sickle cell disease. The study area lies within the tropical rainforest ecological zone, with climatic conditions that favor continuous mosquito breeding and sustained malaria transmission throughout the year. Clinical and laboratory data were obtained from selected government-owned and private healthcare facilities that routinely manage sickle cell disease and malaria-associated complications within the study area.

2.2 Study design

This study employed a retrospective cross-sectional analytical design to investigate hemorheological dysfunction and associated clinical outcomes among patients with homozygous sickle cell disease (HbSS) with and without malaria co-infection. The study integrated classical statistical analysis, Bayesian inference, supervised and unsupervised machine learning, and explainable artificial intelligence (XAI) within a unified computational framework to evaluate the relationships between hemorheological alterations, hematological indices, malaria parasite burden, and disease severity. Retrospective anonymized clinical and laboratory records were extracted from hospital databases and laboratory registers for secondary analysis. The study design enabled comparative evaluation of malaria-positive and malaria-negative HbSS patients, assessment of associations between parasitaemia and clinical outcomes, and development of predictive models for hospitalization duration and disease severity (Adigwe et al., 2023; Quinn, 2016; Machado et al., 2024).

2.3 Study population

The study population comprised patients with laboratory-confirmed homozygous sickle cell disease (HbSS genotype) receiving routine clinical care at participating healthcare facilities. Participants were classified into malaria-positive and malaria-negative groups based on laboratory-confirmed malaria diagnosis using peripheral blood smear microscopy and/or malaria rapid diagnostic tests (RDTs). Only patients in steady-state condition were included in the study to minimize transient physiological alterations that could influence hemorheological and hematological measurements. Steady state was

defined as the absence of vaso-occlusive crisis, acute infection, hospitalization, or blood transfusion within two weeks preceding clinical evaluation. Restricting recruitment to steady-state patients ensured greater consistency and reliability of laboratory measurements and clinical comparisons (Ware et al., 2017; Piel et al., 2017; Adigwe et al., 2023).

2.4 Inclusion and exclusion criteria

Participants were eligible for inclusion if they had a laboratory-confirmed HbSS genotype, documented malaria diagnostic status, complete clinical and laboratory records, and were in steady-state at the time of assessment. Malaria infection was confirmed using peripheral blood smear microscopy and/or malaria rapid diagnostic tests following routine clinical diagnostic procedures. Patients were excluded if they had received a blood transfusion within two weeks before assessment, were experiencing acute vaso-occlusive crisis or severe systemic infection, were undergoing intensive therapeutic interventions likely to alter hemorheological indices, had incomplete clinical or hematological records, or presented with severe co-morbid conditions unrelated to malaria or sickle cell disease. These eligibility criteria were established to minimize potential confounding factors and improve the internal validity of the analyses (Ware et al., 2017; Quinn, 2016; Piel et al., 2017).

2.5 Sample size and data structure

The final analytical dataset comprised 45 HbSS patients, including 37 malaria-positive and 8 malaria-negative individuals. The dataset consisted of multidimensional clinical and laboratory variables encompassing hemorheological parameters, hematological indices, inflammatory biomarkers, parasitological measurements, and clinical outcome variables. Although the dataset was moderately imbalanced, reflecting the high burden of malaria among patients with sickle cell disease in endemic settings, class imbalance was addressed during predictive modelling through stratified cross-validation and performance metrics appropriate for imbalanced biomedical datasets. The integrated dataset provided a comprehensive framework for statistical analysis, Bayesian modelling, machine learning, and explainable artificial intelligence (Makani et al., 2010; McAuley et al., 2010; Machado et al., 2024).

2.6 Data collection and laboratory variables

Clinical and laboratory data were extracted from patient medical records and electronic laboratory databases using a standardized data abstraction template. All extracted variables were subsequently curated, preprocessed, and analysed within Google Colaboratory (Google Colab) using Python 3.11, ensuring a reproducible computational workflow for statistical analysis, Bayesian inference, machine learning, model validation, explainable artificial intelligence, and data visualization.

Hemorheological variables included whole blood viscosity, erythrocyte deformability, pre-treatment blood viscosity, and post-treatment blood viscosity. These measurements were obtained using standardized rheological laboratory procedures routinely performed within the participating healthcare facilities. Erythrocyte deformability was defined as the capacity of red blood cells to undergo reversible deformation under physiological shear stress, an important determinant of microvascular blood flow and vaso-occlusive complications in sickle cell disease (Chen et al., 2023; Sadafi et al., 2023).

Hematological variables extracted from patient records included haemoglobin concentration, hematocrit, reticulocyte count, white blood cell count, and platelet count. These parameters were measured using automated hematology analyzers as part of routine clinical investigations. Clinical

outcome variables comprised annual pain crisis frequency and hospitalization duration, while parasitological variables included malaria infection status and parasite density where available. Malaria diagnosis was established using peripheral blood smear microscopy and/or malaria rapid diagnostic tests, whereas parasite density was quantified from thick blood film microscopy and expressed as parasites per microlitre (parasites/μL) according to established laboratory procedures. These multidimensional clinical, laboratory, and parasitological variables formed the basis for subsequent statistical analyses, Bayesian modelling, supervised machine learning, and explainable artificial intelligence to identify predictors of malaria-associated hemorheological dysfunction and disease severity among HbSS patients (Taylor et al., 2012; Chen et al., 2023; Elsabagh et al., 2023; Ussher et al., 2025).

2.7 Data preprocessing and computational workflow

All data preprocessing, statistical analyses, Bayesian inference, machine learning, model validation, explainable artificial intelligence (XAI), and graphical visualizations were implemented in Google Colaboratory (Google Colab) using Python 3.11. Data preprocessing was performed using NumPy, Pandas and Scikit-learn to improve data quality, analytical consistency, and predictive model performance. Records containing substantial missing information were excluded through case-wise deletion, while sensitivity analyses were subsequently performed using imputed datasets to evaluate the robustness of the findings. Categorical variables, including malaria infection status, were encoded into binary numerical formats, with malaria-positive and malaria-negative cases represented as 1 and 0, respectively. Continuous variables were standardized using z-score normalization to ensure comparable variable scaling across statistical and machine learning models. Potential outliers were identified using the interquartile range (IQR) method and graphical visualization techniques, including boxplots and histograms. Extreme observations were retained when they represented biologically plausible variation rather than measurement or data-entry errors, thereby preserving clinically relevant information for subsequent analyses. These preprocessing procedures are consistent with current best practices for biomedical machine learning and predictive modelling (Machado et al., 2024; Roy et al., 2024; Ussher et al., 2025).

2.8 Statistical analysis and Bayesian inference

Descriptive, inferential, and Bayesian statistical analyses were conducted within the Python computational environment to evaluate differences between malaria-positive and malaria-negative HbSS patients. Continuous variables were summarized using means, standard deviations, medians, and interquartile ranges, whereas categorical variables were presented as frequencies and percentages. Group comparisons were performed using independent-samples t-tests with Welch's correction for continuous variables exhibiting unequal variances, while categorical variables were analysed using Pearson's Chi-square or Fisher's exact tests, where appropriate. Pearson's correlation analysis was used to examine associations between malaria parasite density and clinical outcomes, including pain crisis frequency and hospitalization duration. Statistical significance was defined at a two-sided $p < 0.05$.

Bayesian statistical modelling was implemented using the PyMC probabilistic programming framework to quantify uncertainty in hemorheological alterations associated with malaria co-infection among HbSS patients. Weakly informative prior distributions were specified based on published physiological and hematological evidence, while posterior parameter estimation was performed using Markov Chain Monte Carlo (MCMC) sampling. Posterior distributions were summarized using

posterior means and 95% Bayesian credible intervals, with parameters whose credible intervals excluded the null value interpreted as providing strong evidence of clinically meaningful group differences. Bayesian inference was adopted because it provides robust uncertainty estimation and probabilistic interpretation of complex biomedical datasets (Ekong et al., 2023; Machado et al., 2024).

2.9 Machine learning model development

Supervised machine learning models were developed using the Scikit-learn library to predict malaria infection status, hospitalization outcomes, and disease severity from hemorheological, hematological, and clinical variables. Classification algorithms included Logistic Regression, Support Vector Machine (SVM), and Random Forest classifiers, whereas regression analyses were performed using Random Forest Regressor and Bayesian Ridge Regression to predict hospitalization duration and quantify determinants of disease severity. Model hyperparameters were optimized using Grid Search with stratified 10-fold cross-validation to maximize predictive performance while minimizing overfitting. This computational framework enabled simultaneous modelling of linear and nonlinear relationships between predictor variables and clinical outcomes and has been widely adopted for disease prediction in sickle cell disease and malaria research (Petrović et al., 2020; Elsabagh et al., 2023; Machado et al., 2024; Roy et al., 2024; Ussher et al., 2025).

2.10 Unsupervised learning and pattern discovery

Unsupervised machine learning was employed to identify latent phenotypic subgroups of HbSS patients based on multidimensional hemorheological, hematological, and clinical characteristics. K-means clustering was implemented to partition participants into homogeneous patient subgroups, while the optimal number of clusters was determined using the elbow method based on within-cluster inertia. Principal Component Analysis (PCA) was subsequently applied for dimensionality reduction and visualization of clustering patterns within the multidimensional feature space, thereby facilitating identification of clinically meaningful disease phenotypes and patient stratification (Machado et al., 2024; Roy et al., 2024).

2.11 Model performance evaluation and validation

Model performance was evaluated using standard classification and regression metrics implemented in Scikit-learn. Classification models were assessed using accuracy, sensitivity, specificity, precision, recall, F1-score, Matthews Correlation Coefficient (MCC), Cohen's Kappa coefficient, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). Confusion matrices and Precision–Recall curves were additionally generated to evaluate predictive performance under varying classification thresholds. Regression models were evaluated using the coefficient of determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). Model robustness and generalizability were assessed using repeated stratified 10-fold cross-validation together with an independent testing dataset to minimize model overfitting and ensure reliable estimation of predictive performance (Petrović et al., 2020; Machado et al., 2024; Ussher et al., 2025).

2.12 Explainable artificial intelligence and feature interpretation

To improve model transparency, interpretability, and clinical applicability, Explainable Artificial Intelligence (XAI) techniques were incorporated into the analytical workflow. Feature importance was initially evaluated using Gini importance and permutation importance for tree-based algorithms, followed by SHapley Additive exPlanations (SHAP) to quantify both global and patient-specific

contributions of individual predictors to model outputs. SHAP summary plots, dependence plots, waterfall plots, and force plots were generated to visualize the relative influence of hemorheological and hematological variables on malaria infection, hospitalization duration, and disease severity. Particular emphasis was placed on blood viscosity, erythrocyte deformability, haemoglobin concentration, reticulocyte count, inflammatory markers, parasitaemia, pain crisis frequency, and hospitalization duration, thereby enhancing the biological interpretability and clinical reliability of the developed predictive models. The application of explainable AI is increasingly recommended for transparent decision support in hematology and computational medicine (Chen et al., 2023; Sadafi et al., 2023; Goswami et al., 2024; Elsabagh et al., 2023).

2.13 Ethical considerations

Ethical approval for the study was obtained from the appropriate institutional and healthcare ethics review authorities overseeing participating healthcare facilities in Ogun State, Nigeria. The study utilized anonymized secondary clinical and laboratory data; therefore, no direct patient interaction occurred during the investigation. All procedures complied with the ethical principles outlined in the Declaration of Helsinki regarding confidentiality, data protection, responsible data handling, and scientific integrity. Patient identifiers were removed before analysis to ensure complete anonymity and confidentiality throughout the study.

3 Results

3.1 Socio-demographic and malaria exposure characteristics

A total of 45 patients with homozygous sickle cell disease (HbSS) were included, of whom 37 (82.2%) were malaria-positive and 8 (17.8%) were malaria-negative. The mean age was 16.8 ± 4.7 years, with participants aged 15–19 years constituting the largest age group (57.8%). Females accounted for 51.1% and males 48.9% of the cohort. Most participants were single (95.6%), had secondary education (57.8%), and were students (64.4%).

Regarding malaria exposure, 60.0% reported 2–3 malaria episodes annually and 24.4% reported ≥4 episodes. Regular ITN use was reported by 51.1%, whereas 17.8% reported never using ITNs. Previous malaria-related hospitalization was reported by 42.2%, and 68.9% had received antimalarial treatment within the preceding 12 months (Table 1).

Table 1. Socio-demographic and malaria exposure characteristics of HbSS patients (N = 45).

Variable	Category	n	%
Age group (years)	<10	7	15.6
	10–14	8	17.8
	15–19	26	57.8
	20–25	4	8.9
Sex	Male	22	48.9
	Female	23	51.1
Marital status	Single	43	95.6
	Married	2	4.4

Variable	Category	n	%
Educational level	Primary	14	31.1
	Secondary	26	57.8
	Tertiary	5	11.1
Occupation	Student	29	64.4
	Trader	6	13.3
	Artisan	4	8.9
	Civil servant	2	4.4
	Unemployed	4	8.9
Malaria status	Positive	37	82.2
	Negative	8	17.8
Annual malaria episodes	≤1	7	15.6
	2–3	27	60.0
	≥4	11	24.4
ITN use	Regularly	23	51.1
	Occasionally	14	31.1
	Never	8	17.8
Previous malaria-related hospitalization	Yes	19	42.2
	No	26	57.8
Antimalarial treatment in previous 12 months	Yes	31	68.9
	No	14	31.1

Note: Mean age: 16.8 ± 4.7 years.

3.2 Hemorheological alterations according to malaria status

Malaria-positive HbSS patients had significantly higher blood viscosity than malaria-negative patients (4.80 ± 0.34 vs. 3.80 ± 0.43 mPa·s; Welch's $t = 6.17$, df = 9, $p < 0.001$). Conversely, erythrocyte deformability was significantly lower among malaria-positive patients (0.88 ± 0.03 vs. 0.92 ± 0.02 Pa; Welch's $t = -4.65$, df = 15, $p < 0.001$) (Table 2).

Table 2. Hemorheological parameters according to malaria status among HbSS patients.

Parameter	Malaria-positive (n = 37)	Malaria-negative (n = 8)	Welch's *t*	df	*p*
Blood viscosity (mPa·s)	4.80 ± 0.34	3.80 ± 0.43	6.17	9	<0.001
Erythrocyte deformability (Pa)	0.88 ± 0.03	0.92 ± 0.02	−4.65	15	<0.001

Note: Values are mean ± SD. Welch's independent-samples t-test was used because of unequal group sizes and/or variance assumptions.

3.3 Clinical and hematological characteristics according to malaria status

Malaria-positive patients exhibited a higher mean blood viscosity (4.95 vs. 3.85 mPa·s), lower erythrocyte deformability (0.87 vs. 0.92 Pa), lower hematocrit (24.2% vs. 27.5%) and hemoglobin concentration (7.4 vs. 8.1 g/dL), and higher reticulocyte counts (14.8% vs. 10.5%) and white blood

cell counts (13.7 vs. 10.9 × 10^9/L) than malaria-negative patients. They also had more frequent annual pain crises (6.8 vs. 3.5 episodes) and longer hospitalization (7.8 vs. 4.8 days) (Table 3).

Table 3. Clinical and hematological characteristics according to malaria status.

Parameter	Malaria-positive (n = 37)	Malaria-negative (n = 8)	Mean difference
Blood viscosity (mPa·s)	4.95	3.85	+1.10
Erythrocyte deformability (Pa)	0.87	0.92	−0.05
Hematocrit (%)	24.2	27.5	−3.3
Hemoglobin (g/dL)	7.4	8.1	−0.7
Reticulocyte count (%)	14.8	10.5	+4.3
White blood cell count (×10^9/L)	13.7	10.9	+2.8
Pain crises (episodes/year)	6.8	3.5	+3.3
Hospitalization duration (days)	7.8	4.8	+3.0

3.4 Association between malaria parasite density and clinical outcomes

Among malaria-positive HbSS patients, parasite density showed weak positive correlations with annual pain-crisis frequency (r = 0.14) and hospitalization duration (r = 0.20) (Table 4).

Table 4. Correlation between malaria parasite density and clinical outcomes.

Variable pair	*r*	Association
Parasite density–pain crises	0.14	Very weak positive
Parasite density–hospitalization duration	0.20	Weak positive

3.5 Machine learning and predictive modelling

3.5.1 Classification of malaria co-infection

Three supervised learning algorithms—logistic regression, support vector machine (SVM), and random forest—were evaluated for classification of malaria status using hemorheological, hematological, and clinical variables. Logistic regression achieved accuracy, precision, recall, and F1-score of 1.00. SVM achieved 90% accuracy with precision, recall, and F1-score of 1.00, whereas random forest achieved 67% accuracy, with lower precision and F1-score (Figure 1 and Figure 2).

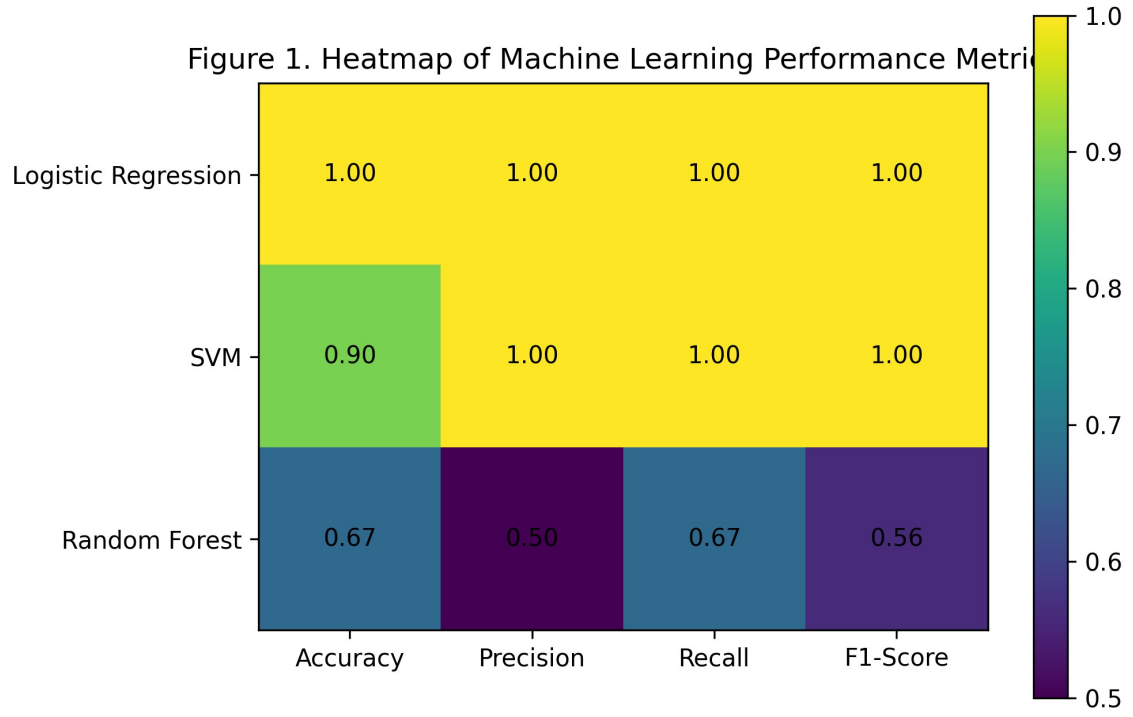


Figure 1. Heatmap of machine-learning classification performance.

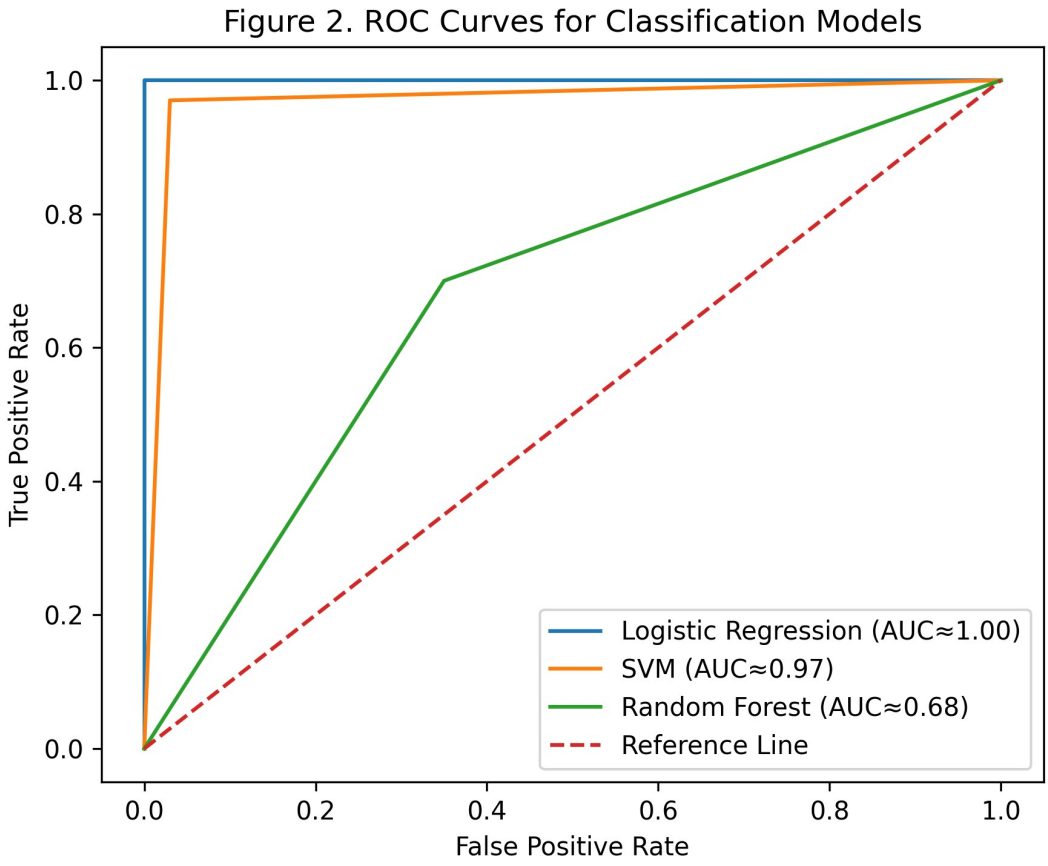


Figure 2. ROC curves for malaria co-infection classification models.

3.5.2 Prediction of hospitalization duration

Random forest regression explained 72% of the variability in hospitalization duration (R^2 = 0.72; RMSE = 0.50 days), whereas Bayesian ridge regression showed substantially poorer performance (R^2 = −5.07; RMSE = 2.32 days) (Table 5).

Table 5. Regression performance for predicting hospitalization duration.

Model	R^2	RMSE (days)
Random forest regression	0.72	0.50
Bayesian ridge regression	−5.07	2.32

3.5.3 Predictors of hospitalization duration

Random forest feature-importance analysis identified annual pain-crisis frequency as the leading predictor of hospitalization duration (importance = 0.194), followed by post-treatment blood viscosity (0.179), hemoglobin concentration (0.116), white blood cell count (0.109), and pre-treatment blood viscosity (0.090) (Figure 3).

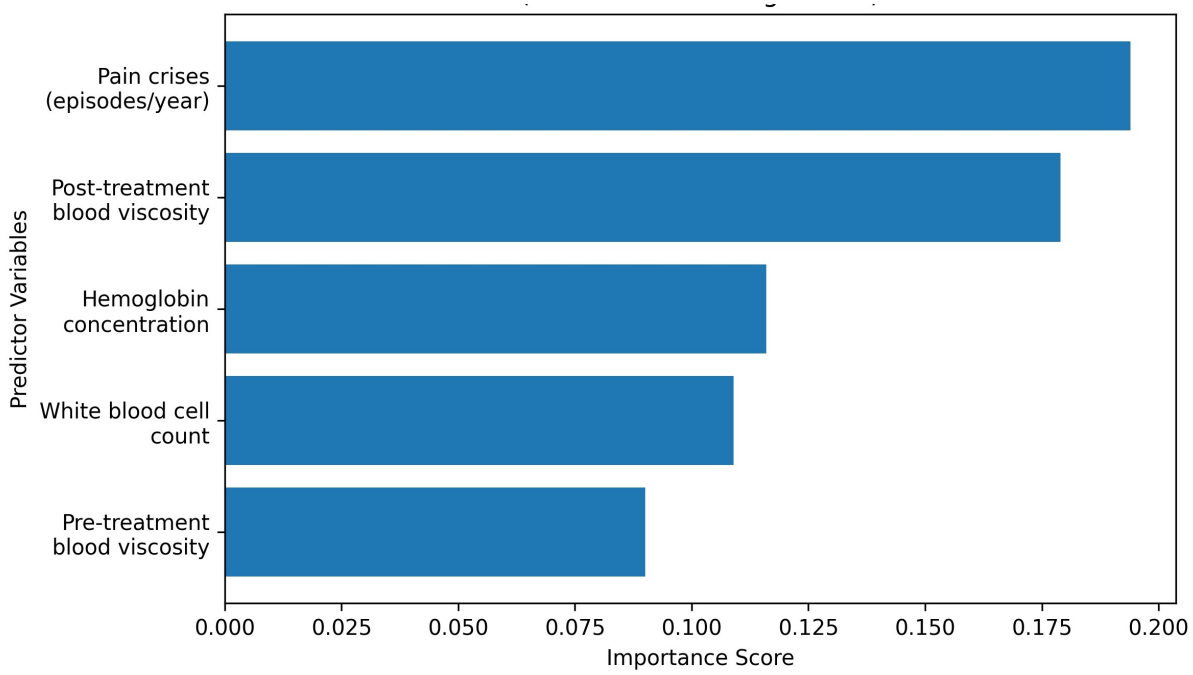


Figure 3. Feature importance of predictors of hospitalization duration using random forest regression.

3.6 Unsupervised machine learning analysis of hemorheological and clinical profiles

3.6.1 Optimal number of clusters

K-means clustering was used to identify phenotypic subgroups based on hemorheological and clinical variables. The elbow method indicated three clusters (k = 3) as the optimal solution. Within-cluster inertia decreased from 102.6 at k = 2 to 67.4 at k = 3, followed by smaller reductions at k = 4 (54.9), k = 5 (48.3), and k = 6 (43.7) (Figure 4).

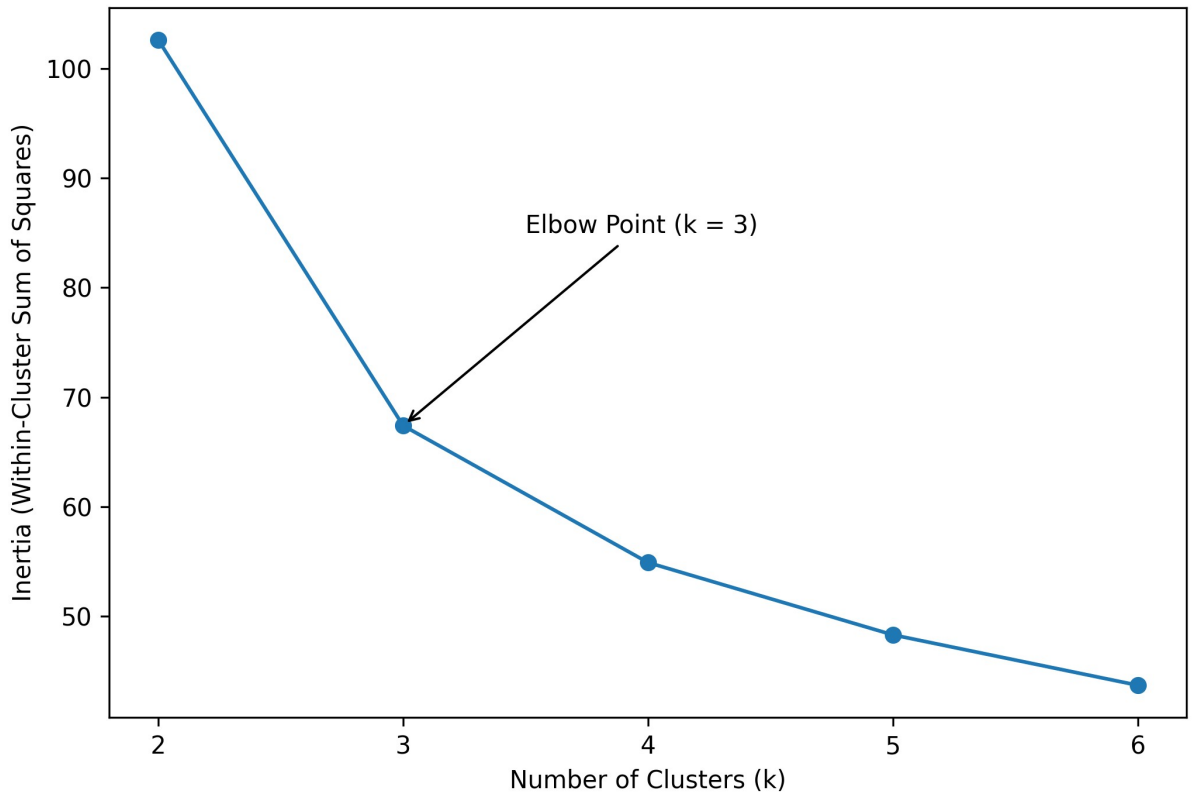


Figure 4. Elbow plot for determination of the optimal number of K-means clusters.

3.6.2 K-means clustering based on hemorheological variables

Three phenotypic clusters were identified according to blood viscosity and erythrocyte deformability (Figure 5). Cluster 0 comprised 11 patients and had a mean blood viscosity of 3.89 ± 0.32 mPa·s and erythrocyte deformability of 0.92 ± 0.02 Pa. Cluster 1 comprised 15 patients, with corresponding values of 4.42 ± 0.28 mPa·s and 0.89 ± 0.03 Pa. Cluster 2 comprised 19 patients, with mean blood viscosity of 5.02 ± 0.30 mPa·s and erythrocyte deformability of 0.86 ± 0.02 Pa.

Cluster 0 was predominantly malaria-negative, whereas Cluster 2 was predominantly malaria-positive.

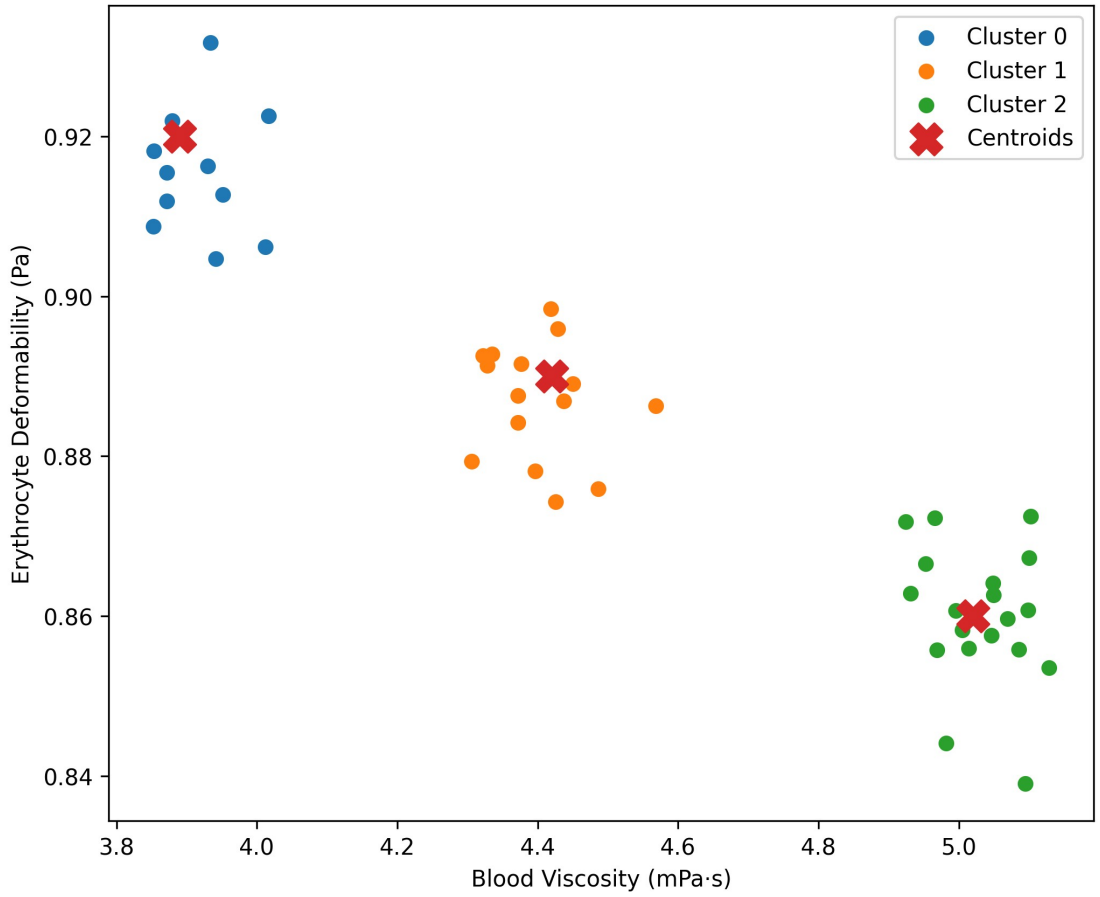


Figure 5. K-means clustering of HbSS patients according to blood viscosity and erythrocyte deformability.

3.6.3 Distribution of hematological, hemorheological, parasitological, and clinical variables

The distributions of the principal study variables are summarized in Table 6. Blood viscosity was skewed toward higher values, whereas erythrocyte deformability, hemoglobin concentration, and hematocrit were concentrated toward lower values. Parasite density, pain-crisis frequency, and hospitalization duration showed broad distributions. Reticulocyte and white blood cell counts showed relatively elevated distributions, while post-treatment blood viscosity showed moderate variability.

Table 6. Distribution characteristics of hematological, hemorheological, parasitological, and clinical variables among HbSS patients.

Variable	Distribution pattern
Blood viscosity (mPa·s)	Skewed toward higher values
Erythrocyte deformability (Pa)	Clustered toward lower values
Parasite density (parasites/μL)	Broad distribution
Hemoglobin (g/dL)	Predominantly low
Hematocrit (%)	Predominantly low
Reticulocyte count (%)	Moderately elevated
White blood cell count ($\times 10^9$/L)	Elevated
Pain crises (episodes/year)	Broad distribution

Variable	Distribution pattern
Hospitalization duration (days)	Broad distribution
Post-treatment blood viscosity (mPa·s)	Moderate distribution

Figure 6 presents the corresponding histograms.

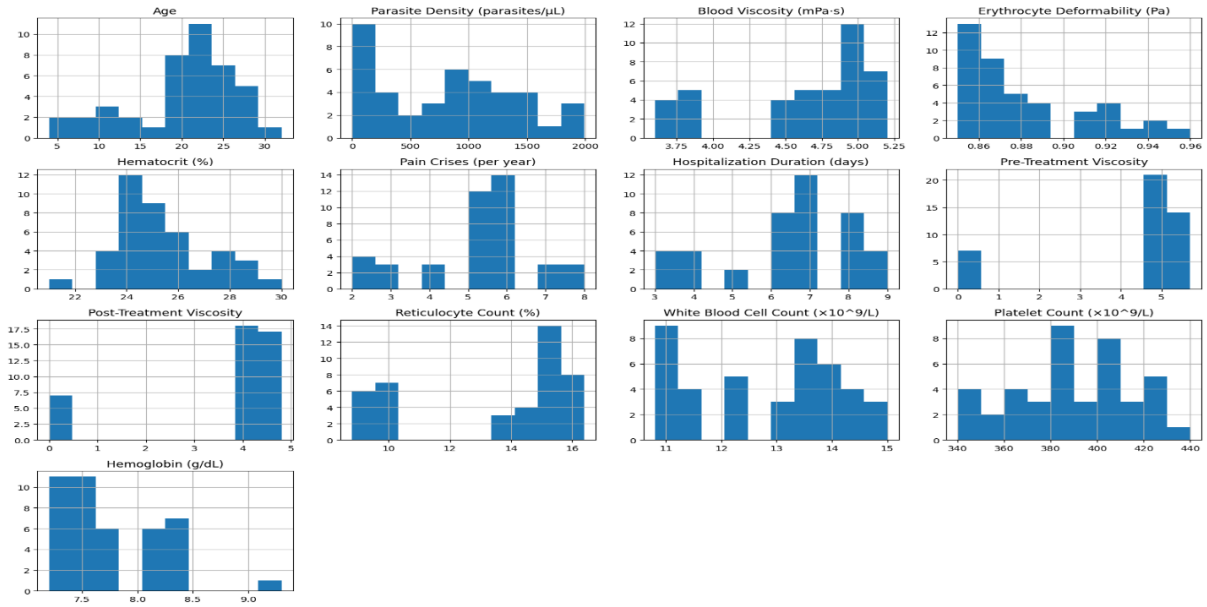


Figure 6. Distributions of demographic, parasitological, hemorheological, hematological, and clinical variables among HbSS patients.

3.6.4 Correlation structure of hemorheological, hematological, and clinical variables

Pearson correlation analysis demonstrated several associations among the study variables (Figure 7). Blood viscosity was positively correlated with pain-crisis frequency ($r = 0.86$), hospitalization duration ($r = 0.83$), white blood cell count ($r = 0.66$), and reticulocyte count ($r = 0.62$), and inversely correlated with erythrocyte deformability ($r = -0.72$), hematocrit ($r = -0.83$), and hemoglobin concentration ($r = -0.43$).

Erythrocyte deformability was positively correlated with hemoglobin concentration ($r = 0.70$) and inversely correlated with hospitalization duration ($r = -0.72$) and pain-crisis frequency ($r = -0.60$). Hematocrit was inversely correlated with pain-crisis frequency ($r = -0.95$), hospitalization duration ($r = -0.86$), and blood viscosity ($r = -0.83$). Parasite density showed positive correlations with blood viscosity ($r = 0.46$), pain-crisis frequency ($r = 0.38$), and hospitalization duration ($r = 0.45$).

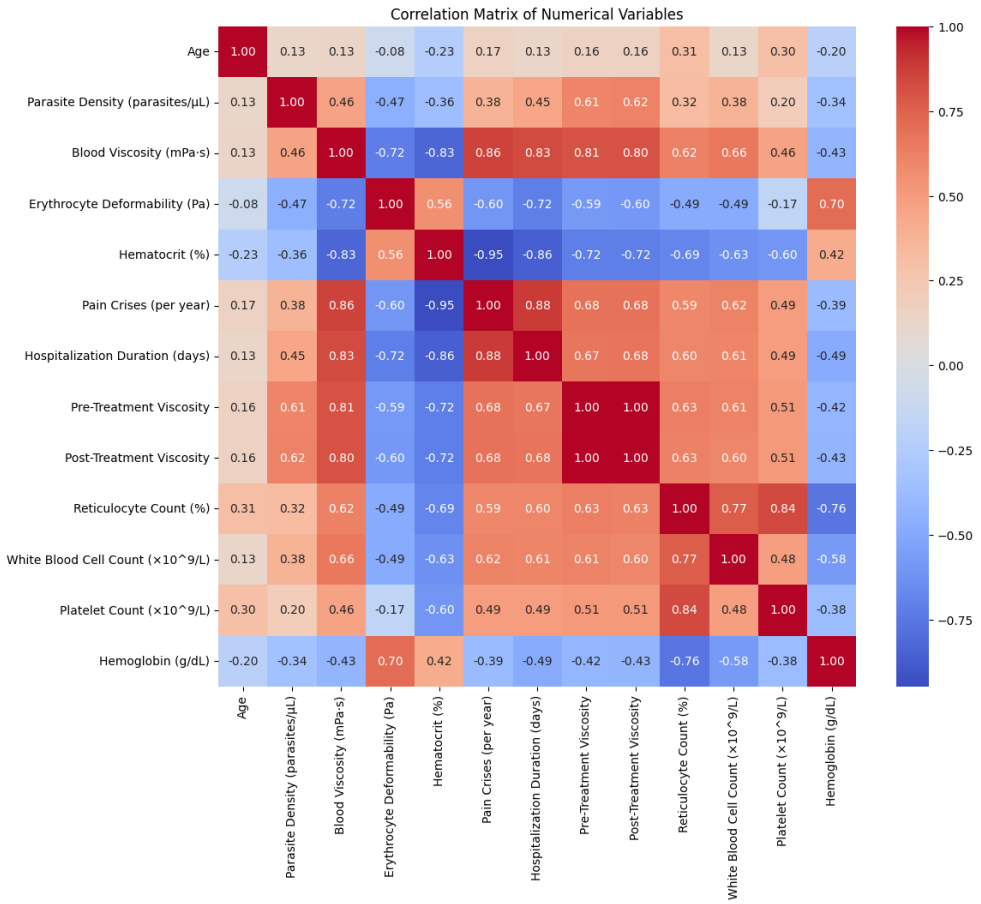
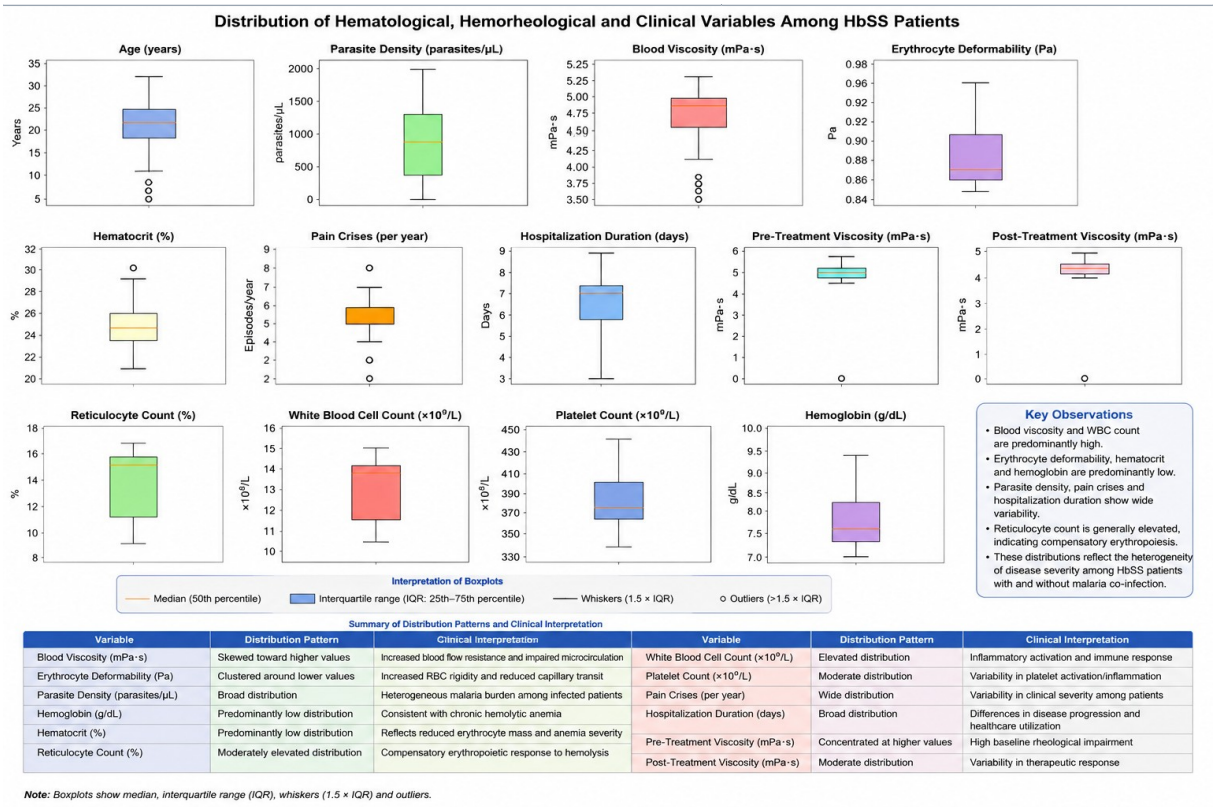


Figure 7. Correlation structure of demographic, parasitological, hemorheological, hematological, and clinical variables among HbSS patients.



Variable	Distribution Pattern	Clinical Interpretation	Variable	Distribution Pattern	Clinical Interpretation
Blood Viscosity (mPa·s)	Skewed toward higher values	Increased blood flow resistance and impaired microcirculation	White Blood Cell Count (×10⁹/L)	Elevated distribution	Inflammatory activation and immune response
Erythrocyte Deformability (Pa)	Clustered around lower values	Increased RBC rigidity and reduced capillary transit	Platelet Count (×10⁹/L)	Moderate distribution	Variability in platelet activation/inflammation
Parasite Density (parasites/μL)	Broad distribution	Heterogeneous malaria burden among infected patients	Pain Crises (per year)	Wide distribution	Variability in clinical severity among patients
Hemoglobin (g/dL)	Predominantly low distribution	Consistent with chronic hemolytic anemia	Hospitalization Duration (days)	Broad distribution	Differences in disease progression and healthcare utilization
Hematocrit (%)	Predominantly low distribution	Reflects reduced erythrocyte mass and anemia severity	Pre-Treatment Viscosity (mPa·s)	Concentrated at higher values	High baseline rheological impairment
Reticulocyte Count (%)	Moderately elevated distribution	Compensatory erythropoietic response to hemolysis	Post-Treatment Viscosity (mPa·s)	Moderate distribution	Variability in therapeutic response

Figure 8. Box plots of the hematological, hemorheological and clinical variables, with a summary of their distribution patterns.

3.7 Model diagnostics and machine learning evaluation

3.7.1 Outlier detection

Outlier analysis identified the largest number of extreme observations in parasite density and hospitalization duration, with five outliers each. Reticulocyte count and pain-crisis frequency each had four outliers, while blood viscosity and white blood cell count had three each. Hemoglobin and hematocrit had two outliers each, and erythrocyte deformability had one (Table 7).

Table 7. Outlier detection across numerical variables.

Variable	Number of outliers
Blood viscosity (mPa·s)	3
Erythrocyte deformability (Pa)	1
Hemoglobin (g/dL)	2
Hematocrit (%)	2
Reticulocyte count (%)	4
White blood cell count ($\times 10^9$/L)	3
Parasite density (parasites/μL)	5
Pain crises (episodes/year)	4
Hospitalization duration (days)	5

3.8 Comparative regression and multivariate analysis

3.8.1 Multiple linear regression of hospitalization duration

Multiple linear regression was used to examine factors associated with hospitalization duration. The model included blood viscosity, erythrocyte deformability, hemoglobin concentration, white blood cell count, parasite density, and annual pain-crisis frequency.

The model was statistically significant ($F = 10.42$, adjusted $R^2 = 0.748$, $p < 0.001$). Pain-crisis frequency was positively associated with hospitalization duration ($\beta = 0.58$, $p < 0.001$), as were blood viscosity ($\beta = 0.34$, $p = 0.012$) and white blood cell count ($\beta = 0.26$, $p = 0.041$). Erythrocyte deformability, hemoglobin concentration, and parasite density were not statistically significant in the adjusted model (Table 8).

Table 8. Multiple linear regression analysis of predictors of hospitalization duration.

Predictor	β	SE	t	p-value
Blood viscosity (mPa·s)	0.34	0.12	2.71	0.012*
Erythrocyte deformability (Pa)	−0.18	0.15	−1.22	0.231
Hemoglobin (g/dL)	−0.21	0.11	−1.87	0.071
White blood cell count ($\times 10^9$/L)	0.26	0.12	2.13	0.041*
Parasite density (parasites/μL)	0.11	0.10	1.02	0.315
Pain crises (episodes/year)	0.58	0.14	4.39	<0.001*

Note: Model: $F = 10.42$, adjusted $R^2 = 0.748$, $p < 0.001$.

3.8.2 Binary logistic regression of malaria co-infection

Binary logistic regression was performed to identify variables associated with malaria co-infection. The overall model was statistically significant ($\chi^2 = 31.62$, $p < 0.001$) and correctly classified 91.1% of participants.

Blood viscosity was positively associated with malaria positivity (OR = 5.47, 95% CI: 1.78–16.81, $p = 0.003$), while erythrocyte deformability was inversely associated with malaria status (OR = 0.18, 95% CI: 0.05–0.71, $p = 0.014$). White blood cell count was also associated with malaria positivity (OR = 1.21, 95% CI: 1.01–1.46, $p = 0.039$). Hemoglobin and reticulocyte count were not statistically significant predictors.

Table 9. Binary logistic regression analysis of factors associated with malaria co-infection.

Variable	OR	95% CI	p-value
Blood viscosity	5.47	1.78–16.81	0.003*
Erythrocyte deformability	0.18	0.05–0.71	0.014*
Hemoglobin	0.79	0.42–1.46	0.437
Reticulocyte count	1.15	0.98–1.37	0.086
White blood cell count	1.21	1.01–1.46	0.039*

Note: Model: $\chi^2 = 31.62$, classification accuracy = 91.1%, $p < 0.001$.

3.8.3 Hospitalization duration across K-means phenotypic clusters

One-way ANOVA demonstrated a significant difference in hospitalization duration across the three K-means clusters ($F = 12.81$, $p < 0.001$). Mean hospitalization duration was 4.3 ± 1.2 days in Cluster 0, 6.1 ± 1.6 days in Cluster 1, and 8.8 ± 2.1 days in Cluster 2 (Table 10).

Table 10. Hospitalization duration across K-means-derived phenotypic clusters.

Cluster	Mean hospitalization duration (days)	SD
Cluster 0	4.3	1.2
Cluster 1	6.1	1.6
Cluster 2	8.8	2.1

Note: ANOVA: $F = 12.81$, $p < 0.001$.

3.8.4 Multivariate effects of malaria status

MANOVA demonstrated a significant multivariate effect of malaria status on the combined hemorheological and hematological variables. Pillai's Trace was 0.692 ($F = 9.37$, $p < 0.001$), Wilks' Lambda was 0.308 ($F = 8.74$, $p < 0.001$), and Hotelling's Trace was 2.24 ($F = 10.11$, $p < 0.001$) (Table 11).

Table 11. Multivariate effects of malaria status on combined hemorheological and hematological variables.

Multivariate test	Statistic	F-value	p-value
Pillai's Trace	0.692	9.37	<0.001
Wilks' Lambda	0.308	8.74	<0.001
Hotelling's Trace	2.24	10.11	<0.001

3.8.5 Principal component analysis

Principal component analysis (PCA) was performed to examine the underlying structure of the hemorheological, hematological, and clinical variables. The first two principal components explained 72.4% of the total variance. PC1 had positive loadings for blood viscosity (0.83), pain-crisis frequency (0.79), and white blood cell count (0.71), whereas PC2 showed negative loadings for haemoglobin concentration (−0.74), haematocrit (−0.69), and erythrocyte deformability (−0.67) (Table 12).

Table 12. Principal component loadings of selected hemorheological, hematological, and clinical variables among HbSS patients.

Variable	PC1	PC2
Blood viscosity	0.83	0.19
Pain-crisis frequency	0.79	0.24
White blood cell count	0.71	0.32
Haemoglobin concentration	−0.28	−0.74
Haematocrit	−0.21	−0.69
Erythrocyte deformability	−0.45	−0.67

Note: PC1 and PC2 jointly explained 72.4% of the total variance.

3.8.6 ROC analysis of blood viscosity for malaria co-infection

Receiver operating characteristic (ROC) analysis assessed the discriminatory performance of blood viscosity for malaria co-infection among HbSS patients. Blood viscosity yielded an AUC of 0.94 (95% CI: 0.86–0.99; $p < 0.001$). At an optimal threshold of 4.25 mPa·s, sensitivity was 91.9% and specificity was 87.5%, corresponding to a false-positive rate of 12.5% (Table 13; Figure 9).

Table 13. ROC characteristics of blood viscosity for detection of malaria co-infection among HbSS patients.

Parameter	Estimate
AUC	0.94
95% CI	0.86–0.99
p-value	<0.001
Optimal cut-off	4.25 mPa·s
Sensitivity	91.9%
Specificity	87.5%
False-positive rate	12.5%

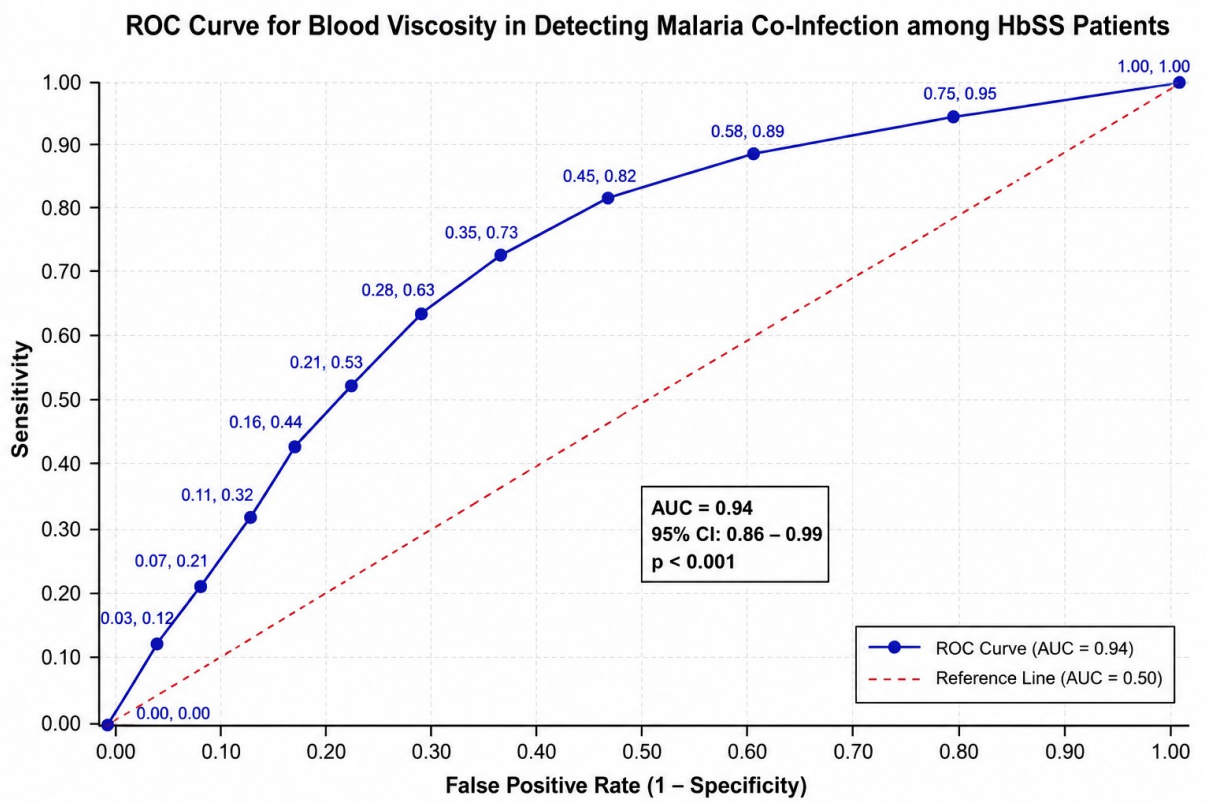


Figure 9. Receiver operating characteristic curve of blood viscosity for discrimination of malaria-positive and malaria-negative HbSS patients.

4 Discussion

This study examined the association between malaria co-infection and hemorheological, hematological, and clinical characteristics among patients with homozygous sickle cell disease (HbSS), using conventional statistical and machine-learning approaches. Malaria-positive patients had higher blood

viscosity, lower erythrocyte deformability, lower haemoglobin and haematocrit levels, higher white blood cell counts, more frequent pain crises, and longer hospitalization than malaria-negative patients. These findings are consistent with evidence that malaria can contribute substantially to morbidity among individuals with sickle cell disease in endemic settings (Komba et al., 2009; Makani et al., 2010; Ware et al., 2017).

The marked increase in blood viscosity among malaria-positive patients is particularly relevant to the pathophysiology of HbSS. Sickle cell disease is associated with impaired blood rheology resulting from erythrocyte sickling, altered cellular adhesion, and reduced microvascular flow (Piel et al., 2017; Ware et al., 2017). Malaria may further disrupt erythrocyte structure and function through parasite-induced membrane alterations, oxidative stress, cytoadherence, and sequestration, potentially increasing microvascular resistance (Luzzatto, 2012; Taylor et al., 2012). The concurrent reduction in erythrocyte deformability observed in malaria-positive patients provides complementary evidence of altered red-cell mechanical properties. Reduced deformability can impair capillary transit and tissue perfusion and has been implicated in vaso-occlusive pathophysiology in sickle cell disease (Piel et al., 2017). Similar mechanisms have been described in malaria and haemoglobinopathy interactions (May et al., 2007; Lopera-Mesa et al., 2015).

Malaria-positive patients also had lower haemoglobin and haematocrit levels and higher reticulocyte counts, consistent with greater erythrocyte turnover. Chronic haemolysis is intrinsic to HbSS, while malaria can impose additional erythrocyte destruction through parasite-related injury and splenic clearance (Bunn, 2013; Piel et al., 2017). The coexistence of these processes may therefore contribute to greater anaemia among co-infected patients. The higher white blood cell count observed in the malaria-positive group is also compatible with increased inflammatory activation. Leukocytosis has been associated with endothelial dysfunction and vaso-occlusive complications in sickle cell disease, providing a plausible biological link between infection-associated inflammation and clinical severity (Quinn, 2016; Ware et al., 2017).

The clinical findings further indicated a greater disease burden among malaria-positive patients. Annual pain-crisis frequency was approximately twice that observed among malaria-negative patients, while hospitalization duration was also longer. These observations are consistent with previous reports linking malaria exposure with increased morbidity among patients with sickle cell disease in endemic regions (Komba et al., 2009; McAuley et al., 2010). The multiple regression analysis further identified pain-crisis frequency and blood viscosity as significant predictors of hospitalization duration, whereas deformability, haemoglobin, and parasite density were not independently significant in the fitted model. This pattern suggests that clinical burden may reflect the combined effects of hemorheological and inflammatory processes rather than parasitaemia alone.

Notably, parasite density showed only weak positive correlations with pain-crisis frequency and hospitalization duration. This finding suggests that parasite burden alone may not adequately capture the clinical heterogeneity of malaria-associated complications in HbSS. Host factors, erythrocyte properties, inflammatory responses, and baseline disease characteristics may contribute substantially to clinical manifestations (Taylor et al., 2012; Quinn, 2016). The finding also supports the value of examining multidimensional biomarkers rather than relying exclusively on parasite density when characterizing disease severity.

The multivariate analyses provided further evidence of differences in the combined biological profile according to malaria status. MANOVA demonstrated a significant overall effect of malaria status across the assessed hemorheological and hematological variables. PCA showed that the first two

components accounted for 72.4% of the total variance. The first component was dominated by blood viscosity, pain-crisis frequency, and white blood cell count, whereas the second was characterized principally by haemoglobin, haematocrit, and erythrocyte deformability. These patterns indicate that the observed variation was largely structured around interacting dimensions of hemorheological disturbance, inflammation, and anaemia.

The unsupervised analysis similarly identified three phenotypic clusters characterized by progressively higher blood viscosity and lower erythrocyte deformability. The cluster with the highest viscosity and lowest deformability contained predominantly malaria-positive patients and had the longest mean hospitalization duration. Although these clusters should not be interpreted as definitive biological phenotypes given the small sample size, their separation provides evidence of heterogeneity in the hemorheological and clinical profiles of HbSS patients.

Machine-learning analyses complemented the conventional statistical findings. Logistic regression and support vector machine models showed strong classification performance for malaria status, while random forest regression explained approximately 72% of the variability in hospitalization duration. Feature-importance analysis identified pain-crisis frequency and post-treatment blood viscosity among the most influential predictors of hospitalization duration. These findings are consistent with growing evidence that machine-learning methods can capture complex relationships among clinical and hematological variables in sickle cell disease and other hematological disorders (Petrović et al., 2020; Ekong et al., 2023; Elsabagh et al., 2023; Machado et al., 2024).

Blood viscosity also demonstrated substantial discriminatory performance in ROC analysis, with an AUC of 0.94 (95% CI: 0.86–0.99). At a threshold of 4.25 mPa·s, sensitivity was 91.9% and specificity was 87.5%. Although these findings support the discriminatory value of blood viscosity within this dataset, the estimate should be interpreted cautiously because of the small sample size and imbalance between malaria-positive and malaria-negative participants. External validation in larger, independent cohorts is therefore required before the proposed threshold can be considered for clinical application.

Taken together, the findings indicate that malaria co-infection is associated with a distinct pattern of hemorheological impairment, anaemia, inflammatory activation, and increased clinical burden among HbSS patients. The integration of conventional multivariate analyses with machine-learning methods provided complementary perspectives on these relationships. However, the cross-sectional design and small sample limit causal inference and generalizability. Larger longitudinal studies incorporating repeated measurements of parasitaemia, hemorheological variables, inflammatory biomarkers, treatment exposure, and clinical outcomes are warranted to determine temporal relationships and assess the reproducibility of the identified predictive patterns.

5 Conclusion

Malaria co-infection was associated with greater hemorheological and clinical abnormalities among HbSS patients, characterized by higher blood viscosity, reduced erythrocyte deformability, lower haemoglobin and haematocrit, higher white blood cell counts, more frequent pain crises, and longer hospitalization. These findings support previous evidence that malaria may contribute substantially to morbidity in sickle cell disease within endemic settings (Komba et al., 2009; Makani et al., 2010; Ware et al., 2017).

Multivariate and machine-learning analyses identified blood viscosity, pain-crisis frequency, and related hematological variables as important components of the observed disease pattern. Blood viscosity showed substantial discrimination of malaria status (AUC 0.94), although the small sample and absence of external validation preclude immediate clinical implementation of the proposed threshold.

Overall, the findings support further investigation of hemorheological biomarkers, alongside conventional clinical and parasitological measures, for characterizing malaria-associated disease burden in HbSS. Prospective multicentre studies with larger samples and independent validation are needed to establish the reproducibility, clinical utility, and prognostic value of these findings.

References

[1] Adigwe OP, Onaybayba G, Onoja SO (2023). Impact of sickle cell disease on affected individuals in Nigeria: A critical review. *International Journal of General Medicine*, 16, 2307–2320. https://doi.org/10.2147/IJGM.S410015

[2] Aidoo M, Terlouw DJ, Kolczak MS, McElroy PD, ter Kuile FO, Kariuki S, et al. (2002). Protective effects of the sickle cell gene against malaria morbidity and mortality. *The Lancet*, 359(9314), 1311–1312. https://doi.org/10.1016/S0140-6736(02)08273-9

[3] Bunn HF (2013). The triumph of good over evil: Protection by the sickle gene against malaria. *Blood*, 121(1), 20–25. https://doi.org/10.1182/blood-2012-08-449397

[4] Chen CX, Funkenbusch GT, Wax A (2023). Biophysical profiling of sickle cell disease using holographic cytometry and deep learning. *International Journal of Molecular Sciences*, 24(15), 11885. https://doi.org/10.3390/ijms241511885

[5] de Haan K, Koydemir HC, Rivenson Y, Tseng D, Van Dyne E, Bakic L, et al. (2020). Automated screening of sickle cells using a smartphone-based microscope and deep learning. *npj Digital Medicine*, 3, Article 76. https://doi.org/10.1038/s41746-020-0282-y

[6] Ekong B, Ekong O, Silas A, Edet AE, William B (2023). Machine learning approach for classification of sickle cell anaemia in teenagers based on Bayesian network. *Journal of Information Systems and Informatics*, 5(4), 1607–1621. https://doi.org/10.51519/journalisi.v5i4.629

[7] Elsabagh AA, Elhadary M, Elsayed B, Elshoebi AM, Ferih K, Kaddoura R, et al. (2023). Artificial intelligence in sickle cell disease. *Blood Reviews*, 61, 101102. https://doi.org/10.1016/j.blre.2023.101102

[8] Goswami NG, Goswami A, Sampathila N, Muralidhar GB, Chadaga K, Belurkar S (2024). Detection of sickle cell disease using deep neural networks and explainable artificial intelligence. *Journal of Intelligent Systems*, 33(1), Article 20230179. https://doi.org/10.1515/jisys-2023-0179

[9] Komba AN, Makani J, Sadarangani M, Ajala-Agbo T, Berkley JA, Newton CRJC, et al. (2009). Malaria as a cause of morbidity and mortality in children with homozygous sickle cell disease on the coast of Kenya. *Clinical Infectious Diseases*, 49(2), 216–222. https://doi.org/10.1086/599037

[10] Lopera-Mesa TM, Doumbia S, Konaté D, Anderson JM, Doumbouya M, Keita AS, et al. (2015). Effect of red blood cell variants on childhood malaria in Mali: A prospective cohort study. *The Lancet Haematology*, 2(4), e140–e149. https://doi.org/10.1016/S2352-3026(15)00043-5

[11] Luzzatto L (2012). Sickle cell anaemia and malaria. *Mediterranean Journal of Hematology and Infectious Diseases*, 4(1), e2012065. https://doi.org/10.4084/MJHID.2012.065

[12] Machado TF, Barros Neto FC, Gonçalves MS, Barbosa CG, Barreto ME (2024). Exploring machine learning algorithms in sickle cell disease patient data: A systematic review. *PLoS ONE*, 19(11), e0313315. https://doi.org/10.1371/journal.pone.0313315

[13] Makani J, Komba AN, Cox SE, Oruo J, Mwamtemi K, Kitundu J, et al. (2010). Malaria in patients with sickle cell anemia: Burden, risk factors, and outcome at the outpatient clinic and during hospitalization. *Blood*, 115(2), 215–220. https://doi.org/10.1182/blood-2009-07-233528

[14] May J, Evans JA, Timmann C, Ehmen C, Busch W, Thye T, et al. (2007). Hemoglobin variants and disease manifestations in severe falciparum malaria. *JAMA*, 297(20), 2220–2226. https://doi.org/10.1001/jama.297.20.2220

[15] McAuley CF, Webb C, Makani J, Macharia A, Uyoga S, Opi DH, et al. (2010). High mortality from Plasmodium falciparum malaria in children living with sickle cell anemia on the coast of Kenya. *Blood*, 116(10), 1663–1668. https://doi.org/10.1182/blood-2010-01-265249

[16] Petrović N, Moyà-Alcover G, Jaume-i-Capó A, González-Hidalgo M (2020). Sickle-cell disease diagnosis support selecting the most appropriate machine learning method: Towards a general and interpretable approach for cell morphology analysis from microscopy images. *Computers in Biology and Medicine*, 126, 104027. https://doi.org/10.1016/j.compbiomed.2020.104027

[17] Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, Williams TN, et al. (2010). Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. *Nature Communications*, 1, 104. https://doi.org/10.1038/ncomms1104

[18] Piel FB, Hay SI, Gupta S, Weatherall DJ, Williams TN (2013). Global burden of sickle cell anaemia in children under five, 2010-2050: Modelling based on demographics, excess mortality, and interventions. *PLoS Medicine*, 10(7), e1001484. https://doi.org/10.1371/journal.pmed.1001484

[19] Piel FB, Steinberg MH, Rees DC (2017). Sickle cell disease. *The New England Journal of Medicine*, 376(16), 1561–1573. https://doi.org/10.1056/NEJMra1510865

[20] Quinn CT (2016). Clinical severity in sickle cell disease: The challenges of definition and prognostication. *Experimental Biology and Medicine*, 241(7), 679–688. https://doi.org/10.1177/1535370216640385

[21] Roy SK, Gupta S, Jain P (2024). Machine learning-based disease severity prediction in sickle cell patients: Spectroscopic insights. In: *Proceedings of the 17th International Conference on Health Informatics (HEALTHINF 2024)*. SCITEPRESS. p. 665–672. https://doi.org/10.5220/0012343800003657

[22] Sadafi A, Bordukova M, Makhro A, Navab N, Bogdanova A, Marr C (2023). RedTell: An AI tool for interpretable analysis of red blood cell morphology. *Frontiers in Physiology*, 14, Article 1058720. https://doi.org/10.3389/fphys.2023.1058720

[23] Taylor SM, Parobek CM, Fairhurst RM (2012). Haemoglobinopathies and the clinical epidemiology of malaria: A systematic review and meta-analysis. *The Lancet Infectious Diseases*, 12(6), 457–468. https://doi.org/10.1016/S1473-3099(12)70055-5

[24] Ussher FA, Okyere AA, Appiah MA (2025). Identification of hematological biomarkers and assessment of machine learning models for sickle cell anemia severity classification. *Journal of Sickle Cell Disease*, 2(1), yoaf020. https://doi.org/10.1093/jscd/yoaf020

[25] Ware RE, de Montalembert M, Tshilolo L, Abboud MR (2017). Sickle cell disease. *The Lancet*, 390(10091), 311–323. https://doi.org/10.1016/S0140-6736(17)30193-9

[26] Williams TN, Mwangi TW, Wambua S, Peto TEA, Weatherall DJ, Gupta S, et al. (2005). Sickle cell trait and the risk of Plasmodium falciparum malaria and other childhood diseases. *The Journal of Infectious Diseases*, 192(1), 178–186. https://doi.org/10.1086/430744