STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **15**, April 2026, pp 3014–3028.
Published online in International Academic Press (www.IAPress.org)

Adult Dengue Fever in Bangladesh: A One-Sample Rank-Based Test of Hematologic Location Against Healthy References

Nabila A. Alsharif*, Inaam Aboud Hussain, Ethar Hussain Jawad

Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq

Abstract Dengue fever is a mosquito-borne viral infection that produces characteristic abnormalities in routine blood tests, yet these hematologic changes are typically analysed one parameter at a time rather than as an integrated multivariate profile. This study investigated whether the joint hematologic profile of adult dengue patients in Bangladesh is systematically displaced from healthy-adult reference values. We analysed a cohort of laboratory-confirmed adult dengue cases from a Bangladeshi hospital, focusing on four core hematologic indices: haemoglobin, white blood cell count, platelet count, and platelet distribution width (PDW). External adult reference means were used to define a healthy location vector, and robust multivariate inference was carried out using the rank-based one-sample location test of Utts and Hettmansperger. Sex-specific (male, female) and pooled (all adults) analyses were performed following careful data cleaning, outlier diagnostics, and verification of strong deviations from multivariate normality.
Across all sex-specific and pooled analyses, the same stable multivariate profile emerged: haemoglobin, white blood cell count, and platelet count were consistently lower than their healthy reference means, whereas PDW was higher, indicating increased platelet-size variability. The Utts–Hettmansperger test strongly rejected the null hypothesis of equality with the healthy reference vector in every analysis, documenting a large and directionally coherent displacement of the dengue cohort in four-dimensional hematologic space. Taken together, these results provide, robust rank-based evidence that adult dengue fever in Bangladesh is associated with a biologically interpretable and stable multivariate shift in core blood indices, integrating leukopenia, thrombocytopenia, and altered platelet morphology into a single statistical summary.

Keywords Dengue fever, Bangladesh, Robust multivariate location test, Utts–Hettmansperger test, Center-outward ranks and signs test, Hematologic profile, Platelet distribution width, White blood cell count.

AMS 2010 subject classifications 62M10, 93A30

DOI: 10.19139/soic-2310-5070-3446

1. Introduction

Dengue fever is an arboviral disease caused by four closely related dengue viruses (DENV-1–4; genus Flavivirus) and transmitted primarily by Aedes aegypti and Aedes albopictus. Clinical expression ranges from mild febrile illness to severe dengue with plasma leakage, hemorrhage, and shock. The global burden continues to rise; recent WHO and U.S. Centers for Disease Control and Prevention (CDC) updates document millions of infections annually and sustained high levels of transmission in many regions. About half of the world's population is now at risk, with an estimated 100–400 million infections each year. Dengue occurs in tropical and subtropical climates worldwide, mostly in urban and semi-urban areas. While many DENV infections are asymptomatic or produce only mild illness, some cases progress to severe disease and can be fatal. There is no specific antiviral treatment for dengue or severe dengue, and early detection with access to appropriate clinical management substantially reduces fatality rates [1, 2].

*Correspondence to: Nabila A. Alsharif (Email: nabila_alsharif@coadec.uobaghdad.edu.iq). Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

The term “dengue” is commonly linked to a Spanish rendering of the Kiswahili phrase ki-denga pepo, while the colloquial label “break-bone fever” was popularized by Benjamin Rush in the late 18th century, reflecting long-standing recognition of the syndrome’s myalgia–arthralgia pattern [3].

Bangladesh experienced its worst recorded season in 2023, with 321,179 reported infections and 1,705 deaths, and surveillance has remained heightened since. These figures underscore the need for reliable hematologic characterization to support bedside assessment and population-level monitoring [4].

The present analysis uses a curated clinical dataset from Upazila Health Complex, Kalai, Joypurhat (Bangladesh), publicly released on Mendeley Data (1,003 rows; 9 variables), recording Age, Sex, Haemoglobin, WBC Count, Differential Count, RBC panel, Platelet count, and PDW in standard laboratory units suitable for comparative hematology [5, 6].

From a methodological standpoint, the multivariate location analysis is restricted to continuous complete blood count (CBC) measurements. Multivariate rank-based location procedures, including spatial sign and rank methods, are primarily formulated for continuous measurements and rely on ordering information that is well defined in such settings. Incorporating binary variables can introduce extensive ties, which may affect the efficiency and interpretability of rank-based multivariate procedures. These considerations motivate focusing the location vector on continuous hematologic indices [7].

This study aims to determine whether the adult hematologic location vector characterizing dengue fever in Bangladesh differs, in a statistically meaningful way, from healthy-adult reference values. The four-dimensional mean vector comprises Haemoglobin (g/dL), White Blood Cell count (WBC, cells/μL), Platelet count (cells/μL), and Platelet Distribution Width (PDW, %). The null hypothesis is equality to external healthy-adult means in matching units; the two-sided alternative allows deviation in any component. Inference is conducted using the Utts–Hettmansperger robust multivariate rank-score test, whose statistic has an asymptotic χ^2 distribution with four degrees of freedom, together with component-wise signed rank-score means describing the direction of deviation hematologic displacement. This design is chosen for robustness to heavy tails, ties, and modest violations of multinormal assumptions while targeting clinically interpretable complete blood count (CBC) indices in adult dengue, which are known to shift through thrombocytopenia, leukopenia, and platelet-index changes during acute illness.

2. Literature Review

The one sample multivariate location problem asks whether a population’s location vector equals a specified reference. When normality is doubtful or outliers are present, robust rank based procedures provide a practical alternative to Hotelling’s T squared. Utts and Hettmansperger introduced a robust class of one sample multivariate rank score tests with a chi squared null reference and strong resistance to extreme observations; this framework underpins the Utts–Hettmansperger multivariate one sample rank score test used in the present study [8].

Peters and Randles proposed an affine invariant signed rank test for the one sample location problem. Their statistic can be viewed as a modification of the multivariate sign test using interdirections, and they compared its performance with competing procedures through asymptotic efficiency and Monte Carlo studies. The method is robust and performs well when the underlying distribution is light tailed, while remaining competitive with Hotelling’s T squared under multivariate normality [9].

Hettmansperger, Möttönen, and Oja then developed affine invariant extensions of the one sample signed rank test and related Hodges–Lehmann type estimates, together with the necessary distribution theory and efficiency comparisons [10].

Möttönen, Oja, and Tienari examined the efficiency of multivariate spatial sign and spatial rank procedures for the one sample location setting. Their results clarify power trade offs relative to parametric benchmarks and help guide method choice when deviations from normality are expected [11].

Oja and Randles provided a comprehensive review of multivariate nonparametric tests for the one-sample location problem, situating spatial signs and spatial ranks within a unified inferential framework. Their review clarifies

the construction, interpretation, and comparative advantages of rank-based multivariate procedures, particularly in settings where classical parametric assumptions such as multivariate normality are doubtful or violated [7].
Bhattacharya and Ghosal presented Bayesian nonparametric tests for multivariate location in both one sample and two sample settings. Their approach builds a testing procedure from posterior credible regions for the spatial median and evaluates local asymptotic power as well as finite sample behavior through simulations [12].
More recently, Hlubinka and Hudecová investigated one-sample multivariate location testing based on center-outward signs and ranks under central symmetry, building on the measure-transport framework of multivariate ranks and signs. Their work establishes the theoretical properties and empirical power of center-outward procedures across a range of distributional settings and provides a modern, affine-invariant alternative for robust multivariate inference [13].
Within this statistical context, contemporary dengue hematology datasets provide a compact panel of core blood indices in infected cohorts, suitable for multivariate location analysis. The Bangladesh benchmark dataset, which reports haemoglobin, white blood cell count, platelet count, and platelet distribution width, is adopted here as the basis for the present analysis [5]. In parallel, the machine learning literature has explored predictive modeling from similar laboratory features. For example, Sarker, Tiang, and Nahid combined swarm intelligence-based feature selection with an extreme gradient boosting classifier and reported model interpretability through SHAP and DiCE analyses [14].

3. Data Description

The publicly available dengue dataset from Upazila Health Complex, Kalai, Joypurhat (Bangladesh) contains 1,003 records and nine variables recorded in standard laboratory units [5]. Table 1 below lists the variables with units, types, and observed coding/levels.

Table 1. Variable Inventory (Original Dataset)

	Variable (dataset label)	**Full name/definition**	**Type**	**Coding/Levels (as observed)**	**Unit**
1	Age	Patient's age	Numeric	-	years
2	Sex	Gender of the patient	Categorical	Child, Male, Female	none
3	Haemoglobin	Haemoglobin levels in the blood	Numeric	-	g/dL
4	WBC Count	Total white blood cell count	Numeric	-	cells/µL
5	Differential Count	Distribution of different types of white blood cells	Categorical/Binary	0, 1	none
6	RBC PANEL	Detailed red blood cell measurements, including counts and morphology	Categorical/Binary	0, 1	none
7	Platelet Count	Total number of platelets in the blood	Numeric	-	cells/µL
8	PDW	Platelet distribution width, which indicates variability in platelet size	Numeric	-	%
9	Final Output	Case status indicator (as provided)	Categorical/Binary	0, 1	none

WBC and platelet counts are recorded as absolute counts per μL (e.g., 4,300/μL), which is numerically equivalent to $4.3 \times 10^3/\mu$L [6].
In Table 2, descriptive statistics for numeric variables, with counts, means, standard deviations, medians, minimum and maximum values.

Table 2. Descriptive statistics for numerical variables (original dataset)

	Variable	n	Missing	Mean	SD	Median	Min	Max
1	Age	1003	0	42.25	20.97	41.0	3	120
2	Haemoglobin	1003	0	13.70	1.48	13.7	11	25
3	WBC Count	979	24	4319.82	2333.30	3200.0	2000	10900
4	Platelet Count	986	17	114202.72	89000.98	92000.0	10000	500000
5	PDW	984	19	22.74	14.62	17.7	1	215

Table 3 shows descriptive statistics for the categorical variables with levels, count and percent, computed directly from the original dataset without any filtering or preprocessing.

Table 3. Categorical variables

	Variable	n	Level	Count	Percent
1	Sex	1003	Child	35	3.49
			Male	518	51.64
			Female	450	44.87
2	Differential Count	1003	1	941	93.82
			0	62	6.18
3	RBC PANEL	1003	1	940	93.72
			0	63	6.28
4	Final Output	1003	1	669	66.70
			0	320	31.90
			(Missing)	14	1.40

4. Data preparation

All analyses were based on the original dataset of 1,003 hospital encounters and nine variables extracted from the hospital information system. The variable *Final Output* is the diagnostic field supplied in the medical records; in this study, the value 1 was interpreted as "dengue infection" and the value 0 as "non-dengue". Missing diagnostic information in *Final Output* was coded as NA in the original file.
The analytic cohort of infected adults was derived in several sequential steps. First, all observations with *Final Output* = 0 (non-infected) were removed (320 records in total), including 3 individuals whose recorded age was $\geq$100 years. Second, 14 records with missing diagnostic status (*Final Output* = NA) were excluded. Third, among the remaining infected records we removed 26 observations whose sex was coded as "Child", because the present analysis is restricted to adults. Fourth, among patients coded as Male or Female we deleted 44 records with age < 18 years. Finally, in the subset of infected adults we required complete information for the core analytic variables; 30 records with at least one missing value in haemoglobin, WBC count, platelet count, or PDW were removed. Altogether, these steps eliminated 434 records, leaving an analytic cohort of 569 infected adults.
All subsequent diagnostic checks and location tests were performed on this cohort of 569 infected adults. The primary multivariate location vector comprised the four continuous hematologic markers: haemoglobin (g/dL), total white blood cell (WBC) count (cells/μL), platelet count (cells/μL), and platelet distribution width (PDW, %). Age and the categorical variables (sex, differential count, RBC panel, and final diagnostic output) were not included

in the four-dimensional location vector because they are not continuous hematologic biomarkers and no external multivariate reference means are available for them. Instead, these variables were used to define the infected adult cohort, to stratify results (Male/Female), and to provide descriptive summaries, while the Uts-Hettmaperger test was restricted to the four continuous blood markers for which clinically interpretable reference values exist.
The multivariate location analysis was deliberately restricted to four continuous hematologic biomarkers (haemoglobin, WBC count, platelet count, and PDW) for which external reference means are available. The remaining variables from the original dataset, age, sex, differential count, and RBC panel, were used for cohort definition, stratification, or descriptive summaries only, and were excluded from the primary p-vector for methodological reasons. Table 4 summarizes these variables, their definitions and units, and the specific rationale for exclusion from the multivariate location test.

5. Data diagnostics and verification of assumptions

Before applying the Utts–Hettmansperger multivariate location test, we examined the empirical distribution of the four hematologic variables in the analytic cohort. Figure 1 displays univariate boxplots for haemoglobin, WBC count, platelet count, and PDW. The boxes show fairly concentrated distributions for haemoglobin and WBC count, with medians near the centre of the interquartile ranges and only modest asymmetry. Platelet counts exhibit a wider spread and a visibly right-skewed upper tail, consistent with substantial variability in platelet responses among infected adults. The PDW boxplot shows the strongest deviation from symmetry, with a compact central box but several isolated high values, including one extreme observation far above the main cloud. These features suggest that the joint distribution is unlikely to be multivariate normal and that robust methods are appropriate.

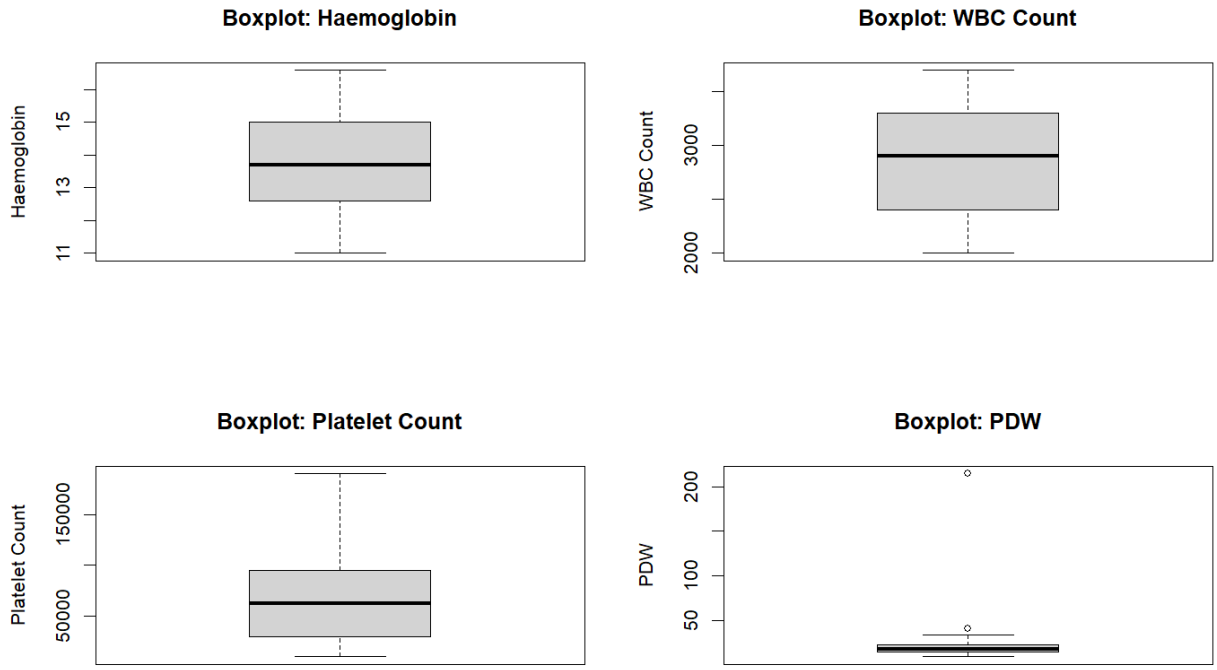

Figure 1. Univariate boxplots for Haemoglobin, WBC Count, platelet Count and PDW

Table 4. Excluded variables from the test

	variable	Full name/definition	Unit	Reason
1	Age	Patient's age	years	a demographic covariate not a hematology biomarker for the prima aim
2	Sex	Gender of the patient	none	categorical; handled instead via sex-specific Haemoglobin reference means in μ_0 and confirmed by optional sex-stratified sensitivity analyses
3	ferential Count	Distribution of different types of white blood cells	none	Differential Count is binary in this dataset; including binary components introduces heavy ties, adds little multivariate information, and complicates the large-sample χ^2 approximation and componentwise interpretation in rank-based multivariate location tests.
4	RBC PANEL	Detailed red blood cell measurements, including counts and morphology	none	RBC panel is binary in this dataset; including binary components introduces heavy ties, adds little multivariate information, and complicates the large-sample χ^2 approximation and componentwise interpretation in rank-based multivariate location tests.

Pairwise relationships and linear dependencies among the four markers are illustrated in Figure 2, which presents a scatterplot matrix with grey histograms on the diagonal and pairwise scatterplots in the lower panels. Individual observations are plotted as orange points against the grey histograms, and the upper panels display the corresponding Pearson correlation coefficients, and black points denote extreme or potential outlying values detected by the robust distance diagnostics. Across all pairs, the absolute correlations remain small (all values are well below 0.20), indicating only weak linear dependence between the markers. In particular, the association between PDW and platelet count is positive but modest, while haemoglobin shows almost no linear correlation

with WBC or platelet count. This weak dependence structure supports the use of a four-dimensional location test but also implies that no single pair of markers dominates the multivariate signal.

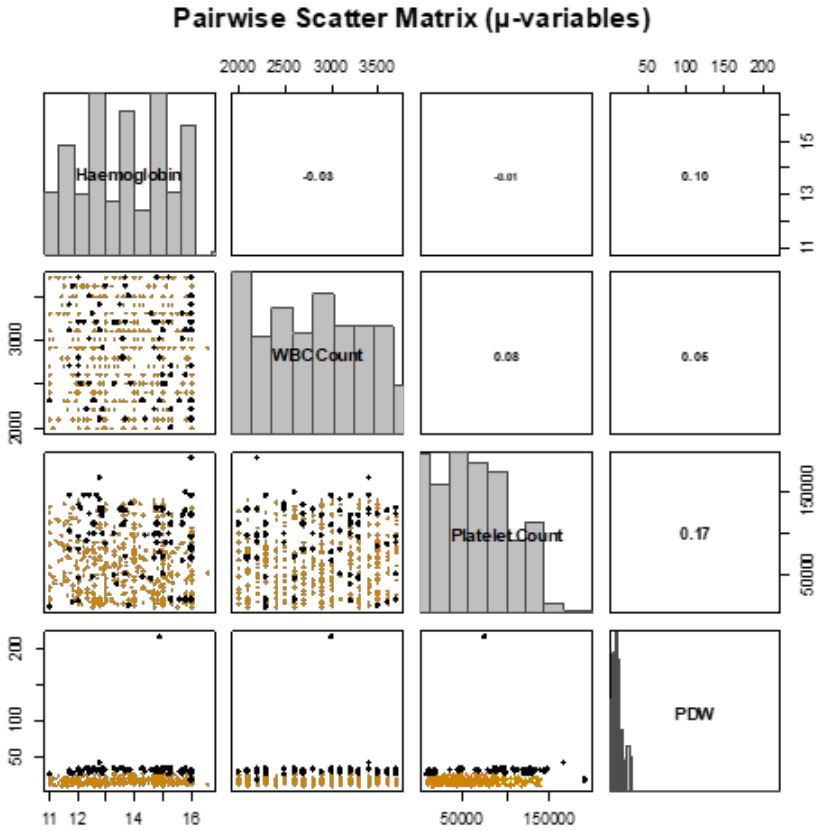

Figure 2. Pairwise relationships among the four hematologic variables (Haemoglobin, WBC Count, Platelet Count, and PDW)

To quantify departures from multivariate normality, we computed Mardia's multivariate skewness and kurtosis [15] statistics for the four-dimensional marker vector in the analytic cohort; the results are summarized in Table 5. Both the skewness χ^2 statistic (with 20 degrees of freedom) and the kurtosis Z statistic lie far in the upper tails of their respective reference distributions, yielding p-values effectively equal to zero. Thus, the joint distribution deviates strongly from multivariate normality in both shape and tail behaviour, again motivating the use of a rank-based robust location test rather than classical multivariate normal theory.

Table 5. Mardia's multivariate skewness and kurtosis (test statistics and p-values)

Metric	Value	Test statistics	df	p-value
Mardia skewness (b1,p)	154.6274	14663.83 (χ^2)	20	$< 2.2 \times 10^{-16}$
Mardia kurtosis (b2,p)	250.402	389.7497 (Z)	–	$< 2.2 \times 10^{-16}$

Potential multivariate outliers were investigated by comparing classical Mahalanobis distances (MD^2) with robust Mahalanobis distances based on the minimum covariance determinant (MCD) estimator [16, 17]. Figure 3 shows

the resulting distance–distance (DD) plot for the square-root distances, with points classified as non-outliers (black), robust-only outliers (blue), and outliers according to both criteria (red). The majority of observations cluster near the origin and fall below the $\chi^2_{0.975}$ cutoff in both metrics, indicating a large core of homogeneous cases. However, the robust distances identify a substantial number of additional outlying observations that are not flagged by the classical distances, reflecting sensitivity to leverage points in the heavy-tailed PDW and platelet distributions. Two observations are extreme under both metrics, corresponding to very unusual combinations of PDW and platelet values. These diagnostics confirm the presence of influential points and heavy tails, which are later addressed through sensitivity analyses (deleting robust outliers and Winsorizing extreme values) without altering the primary Utts–Hettmansperger conclusions.

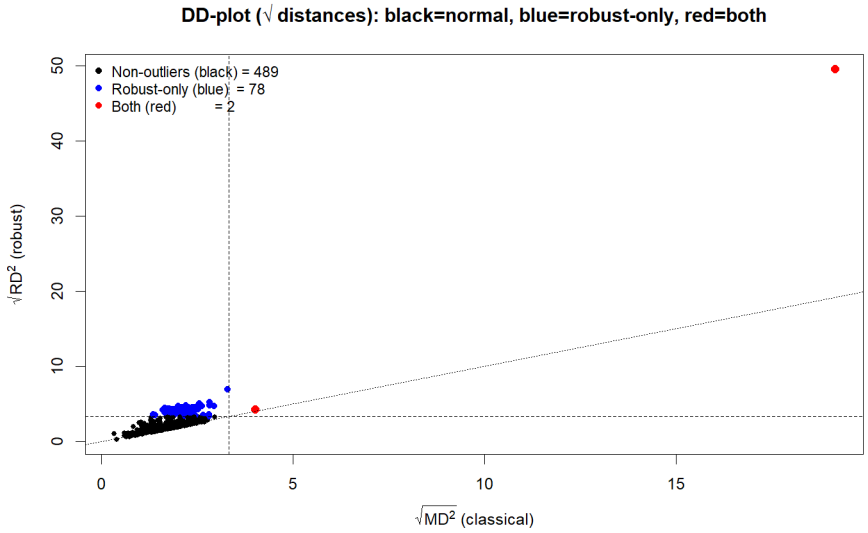

Figure 3. (DD) plot for the square-root distances

6. Utts–Hettmansperger multivariate location test

This study employs the one-sample multivariate rank-score test of Utts and Hettmansperger (UH) [8] to assess whether the hematologic location vector of adult dengue patients differs from a fixed healthy-adult reference. The data matrix consists of n observations on $p = 4$ variables, ordered as haemoglobin (Hb), white blood cell (WBC) count, platelet count, and platelet distribution width (PDW), all expressed in units matched to their corresponding reference values.

Let X denote the $n \times p$ observation matrix and let μ_0 be the fixed reference location. Centered observations are defined as

$$Z = X - 1\mu_0^\top,$$

where 1 is the n-vector of ones. For each observation i and variable j, the scaled rank of the absolute centered value is computed as

$$U_{ij} = \frac{\text{rank}\{|Z_{ij}|\}}{n+1}.$$

Scores are generated using the UH winsorization parameter $\nu \in [0, 1)$, with score-generating function

$$g(u;\nu) = \min\left(\frac{u}{1-\nu}, 1\right).$$

Signed marginal scores are then defined by attaching the sign of Z_{ij} to the transformed rank,

$$T_{ij} = \text{sign}(Z_{ij})\ g(U_{ij};\nu), \quad i = 1,\dots,n,\ j = 1,\dots,p.$$

Let T_i denote the p-dimensional score vector for observation i, and define the mean score vector as

$$\hat{h} = \frac{1}{n}\sum_{i=1}^{n} T_i.$$

The estimated score–covariance matrix $\hat{\Gamma}$ places the constant

$$c_\nu = \frac{1+2\nu}{3}$$

on the diagonal and uses sample cross-moments off the diagonal,

$$\hat{\Gamma}_{rs} = \frac{1}{n}\sum_{i=1}^{n} T_{ir}T_{is},\ r \neq s.$$

The UH test statistic is the Hotelling-type quadratic form

$$H^2 = n\hat{h}^\top \hat{\Gamma}^{-1}\hat{h},$$

which has an asymptotic χ^2_p null distribution. Reporting follows the original formulation: we present H^2, the degrees of freedom ($p = 4$), and the associated p-value.
To bracket tail-robustness in practice, two values of ν are reported. The choice $\nu = 0$ yields pure signed-rank scores without winsorization, while $\nu = 0.30$ down-weights the largest absolute ranks by capping $g(u;\nu)$ earlier, providing increased protection against heavy tails at a small efficiency cost under ideal symmetry.

7. Center-outward sign-and-rank test

As a confirmatory analysis, we apply the symmetrized-sample (SYM) Wilcoxon-type center-outward sign-and-rank test of Hubinka and Hudecová (2024) [13]. This procedure builds on the measure-transport formulation of multivariate ranks and signs [18] and yields a quadratic chi-square statistic for one-sample location testing.
Step 1. Centering at the null location
Let

$$X_1, ..., X_n \in \mathbb{R}^p$$

denote the observed sample and let $\mu_0 \in \mathbb{R}^p$ be the reference location. Define the centered observations

$$Y_i = X_i - \mu_0, \quad i = 1, ..., n.$$

The null hypothesis

$$H_0 : \mu = \mu_0$$

is equivalent to central symmetry of the distribution of Y_i about the origin.
Step 2. Construction of the augmented sample
Form the augmented sample

$$X_a = \{Y_1, ..., Y_n, -Y_1, ..., -Y_n\},$$

which contains $2n$ points. Under H_0, this augmented sample is exactly centrally symmetric, enabling the construction of distribution-free center-outward ranks and signs.
Step 3. Symmetric grid on the unit ball
Let $G \subset \mathbb{R}^p$ be a regular grid on the unit ball and define the symmetric grid

$$G_{2n} = G \cup (-G),$$

satisfying

$$\sum_{g \in G_{2n}} g = 0.$$

This symmetry ensures that the resulting ranks and signs are centered at the origin under the null hypothesis.
Step 4. Empirical center-outward map
Compute the empirical center-outward distribution function $F^{(n)}_{\pm,a}$ by solving an optimal assignment (via the Hungarian algorithm) between the augmented sample X_a and the grid G_{2n}. This yields mapped points

$$F^{(n)}_{\pm,a}(Y_i), F^{(n)}_{\pm,a}(-Y_i), \quad i = 1, ..., n.$$

Step 5. Retaining the original sample
Only the images of the original centered observations are retained:

$$F^{(n)}_{\pm,a}(Y_1), ..., F^{(n)}_{\pm,a}(Y_n).$$

These mapped points jointly encode center-outward ranks (radial components) and center-outward signs (directional components).
Step 6. Center-outward signs and ranks
For each Y_i, define the center-outward sign

$$S_{\pm,a}(Y_i) = \frac{F^{(n)}_{\pm,a}(Y_i)}{\|F^{(n)}_{\pm,a}(Y_i)\|}$$

and the corresponding rank

$$R_i = \|F^{(n)}_{\pm,a}(Y_i)\|.$$

In this study, only the Wilcoxon-type (rank-weighted) quadratic statistic is reported as a single confirmatory center-outward test.
Step 7. Test statistic
The Wilcoxon-type quadratic statistic is defined as

$$Q_F = 3p\|T_{(F,a)}\|^2, \quad T(F,a) = \frac{1}{n}\sum_{i=1}^{n} R_i \; S_{\pm,a}(Y_i).$$

Here, "Wilcoxon-type" refers to the rank-weighted center-outward sign statistic $T_{(F,a)}$. This multivariate statistic reduces to the classical one-sample Wilcoxon signed-rank test when $p = 1$.
Step 8. Asymptotic reference distribution
Under H_0,

$$Q_F \xrightarrow{d} \chi^2_p.$$

Rejection of H_0 indicates a multivariate location shift away from μ_0.
Step 9. Choice of grid resolution (shell construction)

Let n_R denote the number of radial shells. If n is divisible by n_R, allocate

$$n_S = \frac{n}{n_R}$$

grid points per shell. Otherwise, allocate shell sizes n_i such that

$$\sum_i n_i = n,$$

with the n_i differing by at most one. This construction ensures a nearly uniform empirical distribution over the unit ball.

The center-outward one-sample location problem admits both randomized split-sample and symmetrized-sample constructions. We report the symmetrized-sample statistic because it yields a direct quadratic chi-square form based on center-outward signs and rank-weighted scores computed from the augmented sample $\{Y_i, -Y_i\}$, avoiding random splitting while retaining distribution-free inference under the symmetry-based null.

8. Hypotheses

The inferential objective is to determine whether the adult dengue cohort differs in multivariate hematologic location from healthy-adult reference centers defined in matched units and a fixed variable order (Hb in g/dL; WBC and platelet count in cells per microliter; PDW in percent). Three two-sided one-sample tests are conducted: one for all adults, one for males, and one for females. For each analysis group $G \in \{\text{All,Male,Female}\}$, with μ denoting the unknown population location vector, the hypotheses are

$$H_0 : \mu = \mu_{0,G} \text{ versus } H_1 : \mu \neq \mu_{0,G}.$$

The reference vectors $\mu_{0,G}$ are grounded in widely cited adult clinical intervals and regionally verified reference ranges. Adult haemoglobin reference intervals are typically 13.5–17.5 g/dL for males and 12.0–16.0 g/dL for females; accordingly, the sex-specific Hb centers are fixed at the midpoints 15.5 and 14.0 g/dL. The adult WBC interval of $4.5 - 11.0 \times 10^9/L$ (equivalently $4.5 - 11.0 \times 10^3/\mu L$) yields a midpoint of 7,750/μL. The adult platelet count interval of $150 - 400 \times 10^3/\mu L$ yields a midpoint of 275,000/μL. These conventional intervals and midpoints are documented in NCBI Bookshelf clinical resources [21] and are consistent with large reference-interval studies from the Gulf region [20, 21].

For PDW, a center of 13.3% is adopted based on an analyzer-specific study using the ABX Penta 120 automated hematology analyzer, which reported a median of 13.3% with a 10.0–17.9% reference interval [22]. This choice anchors the PDW target to a peer-reviewed, device-specific source and matches the scale used in the present dataset. In the all-adults analysis, haemoglobin uses a sex-averaged midpoint, while the remaining components use common adult reference values. The reference vectors used in the analyses are therefore

$$\mu_{0,\text{Male}} = (15.5, 7750, 275000, 13.3)^T,$$

$$\mu_{0,\text{Female}} = (14.0, 7750, 275000, 13.3)^T,$$

$$\mu_{0,\text{All}} = (14.75, 7750, 275000, 13.3)^T.$$

These reference centers respect the physiological sex difference in haemoglobin while pooling WBC and platelet targets across sexes, and they retain PDW on the percent scale used in the data. Their clinical plausibility is supported by NCBI complete blood count reference tables [21] and by population-based adult reference-interval studies from Oman and Saudi Arabia conducted under International Federation of Clinical Chemistry and Laboratory Medicine and Clinical and Laboratory Standards Institute protocols [20, 21].

9. Results

Table 6 summarizes the outcomes of the Utts–Hettmansperger (UH) one-sample multivariate location test applied to the adult dengue cohort, together with male- and female-specific subsets. For each group, the test statistic $H^2(\nu)$ was computed under two winsorization parameters, $\nu = 0$ and $\nu = 0.3$, corresponding respectively to the pure signed-rank and the moderately tail-robust versions of the UH procedure. The standardized multivariate effect size is reported as $\delta^2 = \frac{H^2}{n}$.

Table 6. Utts–Hettmansperger Multivariate Location Test Results (Three Hypotheses, $v = 0$ and 0.3)

Cohort	**n**	$\boldsymbol{\nu}$	$\boldsymbol{H^2(\nu)}$	$\boldsymbol{\delta^2 = H^2/n}$	**df**	**p-value**	**Direction of shift (Hb, WBC, Platelet, PDW)**
All adults	569	0.0	509.226	0.895	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
		0.3	516.403	0.908	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
Male adults	313	0.0	286.075	0.914	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
		0.3	289.305	0.924	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
Female adults	256	0.0	224.858	0.878	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
		0.3	229.383	0.896	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑

Across all analyses, the null hypothesis of equality between the observed hematologic location vector μ and the healthy-adult reference vector μ_0 is decisively rejected. The standardized effect sizes δ^2 remain close to unity, indicating large and systematic multivariate departures from the reference centers rather than sampling variability. The direction column in Table 6 reveals a coherent shift pattern: haemoglobin, white blood cell count, and platelet count are uniformly lower, whereas platelet distribution width (PDW) is higher. This configuration persists in both sexes, demonstrating that the observed displacement is not driven by sex-specific differences but reflects the hematologic impact of dengue infection. All UH tests are highly significant ($p < 2 \times 10^{-16}$, $df = 4$).
Consistently, the SYM Wilcoxon-type direction vector (after MAD scaling) indicates the same joint shift pattern, Hb↓, WBC↓, Platelet↓, PDW↑, for the pooled, male-only, and female-only cohorts.

10. Discussion

The UH results summarized in Table 6 demonstrate that adult dengue fever in Bangladesh is associated with a strong and coherent multivariate displacement of core hematologic indices relative to healthy-adult reference centers. Across pooled and sex-specific analyses, haemoglobin, white blood cell count, and platelet count are consistently shifted downward, while PDW is shifted upward. The magnitude of the standardized effect size ($\delta^2 \approx 1$) indicates a large, coordinated multivariate departure rather than an isolated marginal effect.
From a clinical standpoint, this joint pattern aligns with established dengue pathophysiology. Bone-marrow suppression and peripheral platelet destruction contribute to leukopenia and thrombocytopenia, while inflammatory and megakaryocytic responses manifest as increased platelet-size variability, captured here by elevated PDW.
The sensitivity analyses reported in Table 7 show that these conclusions are not driven by extreme observations or distributional irregularities. Whether robustly flagged observations are removed (scenario B), PDW values are winsorized (scenario C), or all observations are retained (scenario A), both the magnitude and the direction of the multivariate displacement remain stable. This invariance highlights the practical robustness of the UH framework in the presence of heavy tails and leverage points, particularly for platelet-related indices.

Table 7. Sensitivity analysis (All adults only): Utts–Hettmansperger Multivariate Location Test $H^2(\nu)$ across scenarios A–C

Scenario	Cohort	n	ν	$H^2(\nu)$	$\delta^2 = H^2/n$	df	p-value	Direction of shift (Hb, WBC, Platelet, PDW)
A – Baseline	All adults	569	0.0	509.226	0.895	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
			0.3	516.403	0.908	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
B – Drop Robust		489	0.0	431.839	0.883	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
			0.3	439.542	0.899	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
C – Winsorized		569	0.0	509.249	0.895	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
			0.3	516.420	0.908	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑

The center-outward signs-and-ranks results reported in Table 8 provide an independent, affine-invariant confirmation of the UH findings. Across the baseline, outlier-deletion, and winsorization scenarios, the symmetrized Wilcoxon-type quadratic statistic Q_F remains extremely large with $df = p = 4$ and $p < 2 \times 10^{-16}$, both in the full cohort and in sex-stratified analyses. The associated direction pattern derived from the MAD-scaled Wilcoxon-type vector $T_{(F,a)}$ is consistent across all, male-only, and female-only cohorts: Hb↓, WBC↓, Platelet↓, and PDW↑.

Across all scenarios, platelet count shows a negative displacement and PDW a positive displacement, consistent with increased platelet-size variability as a stable feature of the dengue hematologic profile.

Because the reference vector μ_0 is compiled from published healthy-adult intervals rather than a contemporaneous local control cohort, inference is specifically about displacement from those external healthy centers, not about differences relative to locally sampled controls.

In addition to the global SYM statistic, componentwise directions were examined using the Wilcoxon-type center-outward vector $T_{(F,a)}$. After robust MAD scaling to avoid dominance by platelet magnitude, the directional pattern remained consistent across pooled, male-only, and female-only cohorts: haemoglobin, WBC count, and platelet count shifted downward relative to μ_0, while PDW shifted upward. This directional agreement mirrors the UH rank-score signs. Table 8 reports the single SYM Wilcoxon-type quadratic statistic Q_F, used as the confirmatory center-outward test.

Taken together, the concordance between the UH analyses (Tables 6–7) and the center-outward procedure (Table 8) indicates that the observed hematologic shift is not an artifact of marginal scaling, rank construction, or distributional asymmetry. Instead, it reflects a genuine and robust multivariate signature of dengue infection.

11. Limitations

Several limitations should be noted. First, the healthy reference vector μ_0 was constructed from published external reference intervals rather than a locally sampled Bangladeshi control cohort; consequently, inference quantifies displacement from standard healthy centers rather than from contemporaneous matched controls. Second, the data originate from a single clinical site and represent hospital-based encounters, which may limit generalizability to other settings or to community-managed dengue cases. Third, PDW exhibited extreme values in a small number of records; although sensitivity analyses (outlier deletion and winsorization) did not alter the global conclusions, PDW reference targets remain analyzer-dependent.

Table 8. Sensitivity analysis for the SYM center-outward sign-and-rank one-sample location test across scenarios A–C (All/Male/Female)

Scenario	**Cohort**	**n**	Q_F **(Wilcoxon-type)**	**df**	**p-value**	**Direction of shift (Hb, WBC, Platelet, PDW)**
A – Baseline	All adults	569	323.157	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
	Male adults	313	162.700	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
	Female adults	256	121.731	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
B – Drop Robust	All adults	489	261.835	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
	Male adults	268	140.967	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
	Female adults	221	106.159	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
C – Winsorized	All adults	569	323.157	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
	Male adults	313	162.700	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑
	Female adults	256	121.731	4	$< 2 \times 10^{-16}$	Hb↓, WBC↓, Platelet↓, PDW↑

12. Conclusion

This study provides robust rank-based evidence, distribution-free under the central-symmetry working model for the SYM center-outward procedure, that adult dengue fever in Bangladesh is characterized by a pronounced and directionally coherent multivariate shift in core hematologic indices relative to healthy-adult reference values. Using the Utts–Hettmansperger one-sample multivariate rank-score test, we show that the joint hematologic location vector differs decisively from reference centers, with reduced haemoglobin, white blood cell count, and platelet count accompanied by increased platelet distribution width.

The magnitude and direction of this displacement are stable across pooled and sex-specific analyses and remain unchanged under multiple sensitivity scenarios, underscoring robustness to outliers and heavy-tailed distributions. An independent center-outward signs-and-ranks test confirms the same qualitative conclusions within an affine-invariant, transport-based framework, strengthening confidence that the observed hematologic signature is not method-dependent.

Overall, these findings demonstrate that dengue fever produces a coordinated multivariate hematologic response that can be effectively captured using modern rank-based location tests. The combined use of classical robust rank-score methods and contemporary center-outward procedures offers a powerful and interpretable strategy for analyzing clinical datasets that deviate markedly from multivariate normality, providing a complementary inferential layer to routine laboratory summaries when joint biomarker behavior is of primary scientific interest.

REFERENCES

1. World Health Organization (WHO), *Dengue and severe dengue*, Geneva: World Health Organization, 2024. [Online]. [Accessed: 2025]. Available: https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue.
2. Centers for Disease Control and Prevention (CDC), *About dengue*, Atlanta, GA: CDC, 2024. [Online]. [Accessed: 2025]. Available: https://www.cdc.gov/dengue/about/index.html.
3. Centers for Disease Control and Prevention (CDC). Etymology of 'dengue' and 'break-bone fever', *Etymologia: Dengue. Emerg Infect Dis.*, vol. 12, no. 6, p. 893, 2006, doi:10.3201/eid1206.et1206.
4. M. N. Hasan, M. Rahman, M. Uddin, S. A. A. Ashrafi, K. M. Rahman, K. K. Paul, et al., *The 2023 fatal dengue outbreak in Bangladesh highlights a paradigm shift of geographical distribution of cases*, *Epidemiol. Infect.*, vol. 153, p. e3, 2025, doi:10.1017/S0950268824001791.
5. M. A. S. Nirob, A. K. M. F. K. Siam, P. Bishshash, S. Mia, T. I. Tareq, M. Ashikuzzaman, *Dengue Fever Hematological Dataset: Clinical Insights for Improved Diagnosis and Patient Management*, *Mendeley Data*, V1, 2024, doi:10.17632/6fsrsk3mb8.1.
6. M. Asaduzzaman, O. Islam, M. A. S. Nirob, M. M. H. Mim, A. Mahmud, *A benchmark dataset for analyzing hematological responses to dengue fever in Bangladesh*, *Data Brief*, vol. 57, p. 111030, 2024, doi:10.1016/j.dib.2024.111030.
7. H. Oja and R. H. Randles, *Multivariate nonparametric tests*, *Stat. Sci.*, vol. 19, no. 4, pp. 598–605, 2004, doi:10.1214/088342304000000558.
8. J. M. Utts and T. P. Hettmansperger, *A robust class of tests and estimates for multivariate location*, *J. Am. Stat. Assoc.*, vol. 75, no. 372, pp. 939–946, 1980, doi:10.1080/01621459.1980.10477576.
9. D. Peters and R. H. Randles, *A multivariate signed-rank test for the one-sample location problem*, *J. Am. Stat. Assoc.*, vol. 85, no. 410, pp. 552–557, 1990, doi:10.1080/01621459.1990.10476234.
10. T. P. Hettmansperger, J. Möttönen, and H. Oja, *Affine-invariant multivariate one-sample signed-rank tests*, *J. Am. Stat. Assoc.*, vol. 92, no. 440, pp. 1591–1600, 1997, doi:10.1080/01621459.1997.10473681.
11. J. Möttönen, H. Oja, and J. Tienari, *On the efficiency of multivariate spatial sign and rank tests*, *Ann. Stat.*, vol. 25, no. 2, pp. 542–552, 1997, doi:10.1214/aos/1031833663.
12. I. Bhattacharya and S. Ghosal, *Bayesian nonparametric tests for multivariate locations*, *J. Stat. Plan. Inference*, vol. 219, pp. 1–12, 2022, doi:10.1016/j.jspi.2021.10.005.
13. D. Hlubinka and Š. Hudecová, *One-sample location tests based on center-outward signs and ranks*, in *Recent Advances in Econometrics and Statistics*, J. G. De Gooijer, Ed. Cham: Springer Nature Switzerland, 2024, pp. 29–48, doi:10.1007/978-3-031-61853-6_2.
14. P. Sarker, J. J. Tiang, and A. A. Nahid, *Dengue fever detection using swarm intelligence and XGBoost classifier: an interpretable approach with SHAP and DiCE*, *Information*, vol. 16, no. 9, p. 789, 2025, doi:10.3390/info16090789.
15. K. V. Mardia, *Measures of multivariate skewness and kurtosis with applications*, *Biometrika*, vol. 57, no. 3, pp. 519–530, 1970, doi:10.1093/biomet/57.3.519.
16. P. J. Rousseeuw and K. Van Driessen, *A fast algorithm for the minimum covariance determinant estimator*, *Technometrics*, 1999, doi:10.1080/00401706.1999.10485670.
17. P. J. Rousseeuw and B. C. Van Zomeren, *Unmasking multivariate outliers and leverage points*, *J. Am. Stat. Assoc.*, 1990, doi:10.1080/01621459.1990.10474944.
18. M. Hallin, E. del Barrio, J. A. Cuesta-Albertos, and C. Matrán, *Distribution and quantile functions, ranks and signs in dimension d: a measure transportation approach*, *Ann. Stat.*, vol. 49, no. 2, pp. 1139–1165, 2021, doi:10.1214/20-AOS1996.
19. National Library of Medicine (US), *NCBI Bookshelf*, Bethesda, MD: National Center for Biotechnology Information. [Online]. [Accessed: 2025]. Available: https://www.ncbi.nlm.nih.gov/books/.
20. A. Al-Mawali, A. D. Pinto, R. Al-Busaidi, R. H. Al-Lawati, and M. Morsi, *Comprehensive haematological indices reference intervals for a healthy Omani population: first comprehensive study in Gulf Cooperation Council (GCC) and Middle Eastern countries based on age, gender and ABO blood group comparison*, *PLoS One*, vol. 13, no. 4, p. e0194497, 2018,doi:10.1371/journal.pone.0194497.
21. A. Borai, I. Al Alwan, S. Alzahrani, T. Alwasaidi, A. Altowijri, A. Alghamdi, et al., *Establishment of reference intervals for hematological parameters of adult population in the western region of Saudi Arabia*, *PLoS One*, vol. 18, no. 2, p. e0281494, 2023, doi:10.1371/journal.pone.0281494.
22. M. G. Farias, E. G. Schunck, S. Dal Bó, and S. M. de Castro, *Definition of reference ranges for the platelet distribution width (PDW): a local need*, *Clin. Chem. Lab. Med.*, vol. 48, no. 2, pp. 255–257, 2010, doi:10.1515/CCLM.2010.035.