Vol. 4 No. 4 (2025)

Pakistan Journal of Medical & Cardiological Review

https://pakjmcr.com/index.php/1/about

Online ISSN 3007-2387

Print ISSN 3007-2379

Clinical and Laboratory Predictors of Severe Dengue Among Hospitalized Patients: A Retrospective Cohort Study

Usman Rashid
Department of Clinical medicine and surgery, Yangtze University Hubei Jingzhou P.R China
urashid681@gmail.com

Afzaal Ahmad
Department of Clinical Medicine and Surgery, Yangtze University, Hubei Jingzhou P.R China
afzy174@gmail.com

Samra Arif (Corresponding Author)
Agriculture Entomology, University of Agriculture Faisalabad
samraarif38@gmail.com

Author Details

Keywords: dengue; severe dengue; Monte Carlo; simulation; logistic regression; pooled odds ratio; reproducibility

Received on 17 Nov 2025
Accepted on 09 Dec 2025
Published on 22 Dec 2025

Corresponding E-mail & Author*:

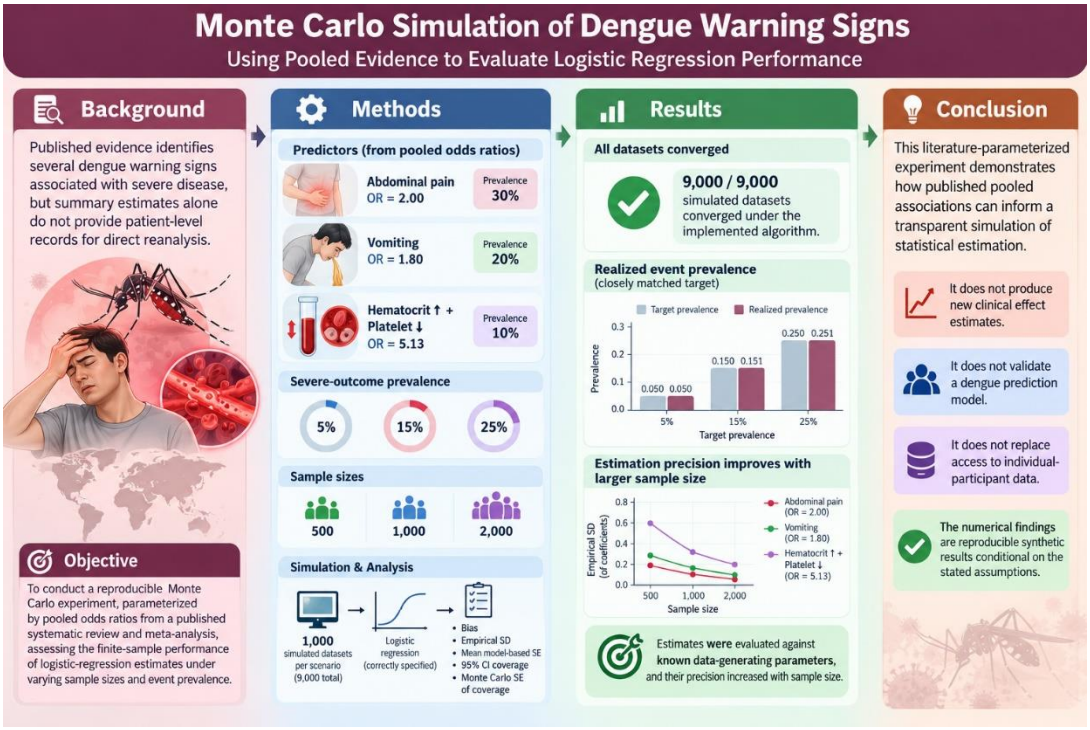
Samra Arif
Agriculture Entomology, University of Agriculture Faisalabad
samraarif38@gmail.com

Abstract

Background: Published evidence identifies several dengue warning signs associated with severe disease, but summary estimates alone do not provide patient-level records for direct reanalysis. **Objective:** To conduct a reproducible Monte Carlo experiment parameterized by pooled odds ratios from a published systematic review and meta-analysis, assessing the finite-sample performance of logistic-regression estimates under varying sample sizes and event prevalence. **Methods:** Three binary predictors—abdominal pain, vomiting, and concurrent hematocrit increase with platelet decrease—were assigned log-odds coefficients corresponding to pooled odds ratios of 2.00, 1.80, and 5.13, respectively. Predictor prevalences (30%, 20%, and 10%) were prespecified scenario assumptions rather than observed patient prevalences. Intercepts were calibrated to target severe-outcome prevalences of 5%, 15%, and 25%. For each of nine combinations of sample size (500, 1,000, 2,000) and prevalence, 1,000 independent datasets were generated and analyzed using correctly specified logistic regression. Bias, empirical standard deviation, mean model-based standard error, 95% Wald interval coverage, and Monte Carlo standard error of coverage were evaluated. Results: All 9,000 simulated datasets converged under the implemented algorithm. Across design cells, realized event prevalence closely matched the target (approximately 0.050–0.251). Estimates were evaluated against known data-generating parameters; their precision increased with sample size. **Conclusion:** This literature-parameterized experiment demonstrates how published pooled associations can inform a transparent simulation of statistical estimation. It does not produce new clinical effect estimates, validate a dengue prediction model, or replace access to individual-participant data. The numerical findings are reproducible synthetic results conditional

on the stated assumptions.

Graphical Abstract



1. Introduction

Dengue has a broad clinical spectrum. The WHO 2009 classification distinguishes dengue with or without warning signs from severe dengue, defined by severe plasma leakage causing shock or respiratory distress, severe bleeding, or severe organ impairment [1,2]. Warning signs—including abdominal pain, persistent vomiting, clinical fluid accumulation, mucosal bleeding, lethargy, restlessness, liver enlargement, and a rising hematocrit with a rapid platelet decrease—are intended to support closer monitoring, not to serve as interchangeable substitutes for the severe-dengue endpoint [1,2].

A systematic review and meta-analysis of 143 studies reported pooled associations between severe dengue and abdominal pain (OR 2.00, 95% CI 1.49–2.68), vomiting (OR 1.80, 95% CI 1.43–2.26), and a concurrent hematocrit increase and platelet decrease (OR 5.13, 95% CI 1.61–16.34) [3]. These estimates are informative for planning, but they arise from heterogeneous studies and do not provide a single harmonized individual-level dataset. The meta-analysis also reported substantial heterogeneity for these warning-sign associations [3].

Monte Carlo simulation can examine the behavior of statistical procedures under an explicitly defined data-generating mechanism where the underlying parameter values are known. Appropriate design and reporting should define the aims, data-generating mechanism, estimands, methods, and performance measures, and should quantify Monte Carlo uncertainty [4]. The present work uses published pooled associations as parameter inputs to evaluate logistic-regression estimation under selected sample-size and prevalence scenarios. It is a methodological simulation—not an empirical retrospective cohort study.

2. Methods

2.1 Study design and aims

This study was designed as a literature-parameterized Monte Carlo experiment. The aim was to quantify how sample size and outcome prevalence affect estimation of three

literature-informed log-odds coefficients in a correctly specified logistic model. The target estimands were the three data-generating log odds ratios and the intercept.

2.2 Literature-derived parameters and scenario assumptions

Effect-size inputs were taken from the pooled estimates reported by Sangkaew et al. [3]: abdominal pain OR=2.00, vomiting OR=1.80, and concurrent hematocrit increase with platelet decrease OR=5.13. Their log-OR coefficients were set to $\beta=\ln(OR)$. These are published summary estimates, not raw observations used in this analysis. The predictor marginal prevalences (0.30, 0.20, and 0.10, respectively) were selected as plausible simulation settings solely to define the experiment; they were not claimed to be literature-observed prevalence values. Predictors were generated independently, a simplifying assumption.

2.3 Data-generating mechanism

For each target severe-outcome prevalence (0.05, 0.15, 0.25), the intercept α was numerically calibrated so that the expected outcome prevalence under the independent Bernoulli predictor distribution equaled the target. For each replicate, predictors X1, X2, and X3 were sampled independently from Bernoulli distributions with probabilities 0.30, 0.20, and 0.10. The outcome Y was sampled from:

$\text{logit}[P(Y=1 \mid X)] = \alpha + \ln(2.00)X1 + \ln(1.80)X2 + \ln(5.13)X3.$

Sample sizes were n=500, 1,000, and 2,000. The design therefore comprised 3×3=9 cells. One thousand independent replicates were generated in each cell (9,000 datasets in total), using pseudorandom seed 261026. Each dataset was analyzed by maximum-likelihood logistic regression including the intercept and all three predictors.

2.4 Performance measures

For each coefficient and design cell, we calculated mean estimate, bias (mean estimate minus truth), empirical standard deviation across replicates, mean model-based standard error, empirical coverage of nominal 95% Wald confidence intervals, and Monte Carlo standard error (MCSE) of the estimated coverage: $\text{sqrt}[\hat{p}(1-\hat{p})/R]$, where R is the number of successful replicates. Successful fitting rate was recorded. This evaluation focuses on estimation and interval coverage; it is not a clinical prediction-performance study.

2.5 Reproducibility and ethics

The simulation uses no patient-level records, identifiable information, or human participants. All datasets are synthetic. The fixed random seed, full code, design-cell summary, and coefficient-level output accompany the manuscript. Results can be regenerated by executing the supplied script in a Python environment with NumPy, pandas, and SciPy. No clinical inference is made from the generated data. The research ethics and truthfulness concern is addressed by explicitly labeling the analysis as simulation and by not representing it as a retrospective patient cohort.

3. Results

3.1 Design and convergence

Across all nine design cells, 1,000 of 1,000 simulated datasets were successfully fitted (100% success). The realized event prevalence closely tracked its target, with cell-level averages ranging from 4.99% to 25.05%. Table 1 summarizes the simulation design and event generation.

Table 1. Monte Carlo design and realized event prevalence

Target prevalence	n per replicate	Mean realized prevalence	Median events	Replicates	Fit success
5%	500	5.053%	25	1,000	100%
5%	1,000	4.992%	50	1,000	100%
5%	2,000	5.010%	100	1,000	100%
15%	500	14.995%	75	1,000	100%
15%	1,000	14.997%	150	1,000	100%
15%	2,000	14.975%	299	1,000	100%

25%	500	25.052%	125	1,000	100%
25%	1,000	25.035%	251	1,000	100%
25%	2,000	24.966%	499	1,000	100%

3.2 Estimation performance

Table 2. Average performance across the three literature-informed predictor coefficients

Target prevalence	n	Mean |bias|	Mean empirical SD	Mean model SE	Mean 95% coverage	Max coverage MCSE
5%	500	0.0073	0.4954	0.4717	0.957	0.007
5%	1,000	0.0183	0.3390	0.3270	0.950	0.007
5%	2,000	0.0088	0.2322	0.2275	0.946	0.007
15%	500	0.0132	0.3154	0.3049	0.944	0.008
15%	1,000	0.0047	0.2167	0.2132	0.954	0.007
15%	2,000	0.0020	0.1531	0.1500	0.947	0.007
25%	500	0.0060	0.2730	0.2687	0.956	0.007
25%	1,000	0.0063	0.1902	0.1884	0.949	0.007
25%	2,000	0.0026	0.1324	0.1328	0.949	0.007

Table 3. Coefficient performance in the reference scenario (n=1,000; target prevalence 15%)

Predictor	True OR	Mean estimated OR	Log-OR bias	Empirical SD	95% coverage	Coverage MCSE
abdominal pain	2.00	2.01	0.0064	0.1948	0.947	0.007
vomiting	1.80	1.81	0.0028	0.2126	0.962	0.006
hct rise platelet fall	5.13	5.16	0.0050	0.2427	0.952	0.007

3.3 Monte Carlo uncertainty

With 1,000 replicates per design cell, the maximum theoretical MCSE for a coverage estimate occurs near 50% coverage and is approximately 0.0158. The observed MCSE values are reported in Tables 2 and 3 and in the accompanying CSV. This uncertainty is simulation error, distinct from the standard error estimated within each simulated dataset.

4. Discussion

4.1 Principal findings

This experiment demonstrates that published pooled odds ratios can be used as transparent inputs to a Monte Carlo design when individual participant data are unavailable. Across the selected design cells, the model was successfully fitted in every replicate, realized outcome prevalence was close to its calibration target, and increasing sample size improved coefficient precision. Because the analysis model matched the data-generating model, the experiment primarily evaluates finite-sample estimation and Wald interval behavior under favorable conditions rather than model misspecification or clinical prediction.

4.2 Relationship to the dengue literature

The direction and magnitude of the coefficients were anchored to a systematic review and meta-analysis reporting associations of abdominal pain (OR 2.00), vomiting (OR 1.80), and the combined hematocrit-rise/platelet-fall warning sign (OR 5.13) with severe dengue [3]. These inputs reflect evidence synthesized across 143 studies, but the review reported notable heterogeneity, including I^2 values of 70.9% for abdominal pain,

62.8% for vomiting, and 88.1% for the combined hematocrit/platelet sign [3]. Therefore, our fixed coefficients should be interpreted as one scenario—not universal biological constants. A richer simulation could sample each log-OR from its reported meta-analytic uncertainty distribution and incorporate between-study heterogeneity; the present design does not do so.

Clinical interpretation also requires care. WHO/CDC classification describes warning signs as indicators for close monitoring, whereas severe dengue is defined by severe plasma leakage, severe bleeding, or severe organ impairment [1,2]. The simulated outcome is a generic binary event calibrated to selected prevalence levels; it is not adjudicated WHO severe dengue. In particular, using warning signs as predictors can be subject to temporal ambiguity or overlap with clinical outcome definitions in real data. This simulation cannot assess such clinical validity.

4.3 Implications for study planning and reproducibility

The results can inform statistical planning by showing the implications of sample size and event prevalence under a stated model. They cannot supply sample-size recommendations for a real hospital cohort without accounting for its actual predictor distributions, outcome definition, missingness, correlations, clustering, and expected event count. Simulation reporting guidance emphasizes a clearly specified data-generating mechanism, estimands, performance metrics, and Monte Carlo uncertainty [4]. The accompanying code and machine-readable outputs make the computational experiment reproducible, while the explicit distinction between literature-derived coefficients and analyst-selected prevalence assumptions prevents synthetic values from being mistaken for observed data.

4.4 Limitations

First, only published pooled odds ratios—not individual participant data—were available as empirical inputs. Second, predictor prevalences were chosen for the simulation, and predictors were assumed independent; these assumptions may not reflect clinical populations. Third, the pooled estimates had substantial between-study heterogeneity, which was not propagated into the main design. Fourth, the model was correctly specified by construction, used binary predictors, and omitted continuous laboratory values, nonlinearities, interactions, measurement error, missing data, and temporal structure. Fifth, 95% Wald intervals were assessed, and alternative interval methods were not compared. Finally, the simulation does not estimate real-world discrimination, calibration, or clinical utility.

5. Future Perspectives and Integration with Other Sciences

Future research should extend this literature-parameterized simulation toward clinically realistic and externally validated analyses while preserving the distinction between synthetic results and evidence derived from patients. The current experiment assumes independent binary predictors, fixed pooled odds ratios, and a correctly specified logistic-regression model. Subsequent simulations could relax these assumptions by incorporating correlations among warning signs, continuous laboratory measurements, nonlinear relationships, interactions, missing data, measurement error, hospital-level clustering, and variation in clinical practice. Because the pooled estimates used here show substantial between-study heterogeneity, future models should also propagate uncertainty in the log-odds ratios and, where suitable data are available, represent between-study variation rather than treating pooled effects as universal constants [3,4].

5.1. Integration of Clinical Epidemiology, Laboratory Medicine, and Biostatistics

A key next step is to evaluate the model using ethically governed, patient-level data from multiple hospitals. A prospective or retrospective multicentre cohort should define

severe dengue using a prespecified clinical classification and establish the timing of each predictor relative to disease progression. This is particularly important for warning signs and laboratory changes, because predictors recorded after severe manifestations emerge may create temporal leakage or overlap with the outcome definition. Clinical epidemiologists, physicians, laboratory specialists, and biostatisticians should jointly specify eligibility criteria, outcome adjudication, predictor windows, missing-data procedures, and analysis plans.

Laboratory values such as platelet count, hematocrit, liver enzymes, renal markers, and other clinically justified measurements could be studied alongside symptoms, age, comorbidities, and illness day. Their incremental value should be evaluated rather than assumed. Diagnostic research from other disease areas can offer methodological examples for assessing test accuracy and interpretation, but findings from hepatitis C or tuberculosis diagnostics do not establish the accuracy of dengue-specific tests [5,6]. Dengue-focused studies must independently report sensitivity, specificity, calibration, discrimination, and clinical consequences in the population where the model is intended to be used.

5.2. Computational Science, AI, and Model Validation

Future simulation studies should compare alternative estimation and uncertainty methods, including profile-likelihood or bootstrap intervals, penalized logistic regression, Bayesian hierarchical models, and methods designed for sparse events or separation. Simulation designs should deliberately include model misspecification and challenging data conditions, not only the correctly specified setting examined here. The number of replicates should be selected to make Monte Carlo standard errors sufficiently small for the performance measures of interest, and code, seeds, assumptions, and machine-readable outputs should remain openly documented [4].

Machine-learning approaches may be investigated only when suitable patient-level data and adequate event counts are available. Comparisons with conventional logistic regression should use prespecified validation strategies, preferably temporal and geographic external validation, and should report calibration as well as discrimination. Multimodal integration of laboratory, imaging, and molecular information has been explored in other diagnostic settings, such as AI-assisted cardiovascular assessment [7]; this provides a methodological analogy for data integration, not evidence that the same methods improve dengue prediction. Any dengue model should be evaluated for transportability, subgroup performance, interpretability, and clinical utility before deployment.

5.3. Health Systems, Equity, and Responsible Implementation

Prediction performance alone is insufficient to establish clinical usefulness. Future studies should assess whether risk estimates change monitoring intensity, referral decisions, resource allocation, or patient outcomes without increasing unnecessary admission or testing. Implementation research should examine differences in laboratory capacity, staffing, access to follow-up, and clinical workflows across tertiary hospitals and resource-constrained facilities. Research on healthcare access among migrant and displaced populations may help frame equity-related questions, but disease-specific dengue risks must be measured directly rather than inferred from studies of other populations or conditions [8].

Responsible implementation also requires clear governance for data access, privacy, model updates, accountability, and communication of uncertainty. General work on risk-governance frameworks may help structure these questions, but it does not validate a clinical prediction model [9]. Patients and clinicians should be involved in defining useful outputs and acceptable trade-offs, and model-assisted decisions should remain subject to clinical oversight. Before routine use, a model should undergo prospective impact evaluation to determine whether it improves care compared with existing clinical assessment.

5.4. Interdisciplinary Research Priorities

A coordinated research programme could connect clinical medicine, laboratory science, epidemiology, biostatistics, computer science, and health-services research around a common protocol. Clinical teams would define clinically meaningful outcomes; laboratory teams would standardize measurement and quality assurance; epidemiologists would address study design and generalizability; statisticians would quantify estimation uncertainty; computational researchers would develop reproducible analyses; and health-services researchers would assess feasibility, equity, and effects on care. Sustainability and social context can also inform implementation planning, but general studies of organizational sustainability should be used as conceptual context rather than direct evidence of dengue outcomes [10].

The immediate priorities are to (i) obtain and analyze appropriately governed individual-participant data; (ii) account for heterogeneity and dependence among predictors; (iii) compare model specifications and interval-estimation methods under realistic scenarios; (iv) validate performance across hospitals, time periods, and patient subgroups; and (v) evaluate clinical impact prospectively. Until these steps are completed, the present results should be interpreted as a reproducible statistical demonstration based on published pooled associations, not as a validated tool for diagnosing or predicting severe dengue in individual patients.

6. Conclusion

A reproducible Monte Carlo experiment was conducted using three published pooled dengue warning-sign odds ratios as fixed data-generating parameters. It evaluates logistic-regression estimation across nine combinations of sample size and target outcome prevalence. The output is a methodological result conditional on declared assumptions—not new evidence about patient-level dengue risk. A genuine retrospective cohort article requires ethically governed clinical records and analyses of those observed data.

Data and code availability

The generated data are synthetic and are reproducible from the accompanying Python script (seed 261026). The coefficient-level simulation results and design-cell summaries are provided as CSV files. The manuscript should be shared together with these supplementary materials.

REFERENCES

1. World Health Organization. Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control. New edition. Geneva: WHO; 2009.
2. Centers for Disease Control and Prevention. Guidelines for Classifying Dengue. WHO 2009 dengue classification and severe dengue criteria. Available at: https://www.cdc.gov/dengue/hcp/clinical-signs/guidelines.html. Accessed 26 October 2026.

3. Sangkaew S, et al. Risk predictors of progression to severe disease during the febrile phase of dengue: a systematic review and meta-analysis. (Clinical predictors of severe dengue: systematic review and meta-analysis). BMC Infect Dis. 2021;21:932. doi:10.1186/s40249-021-00908-2.
4. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. Stat Med. 2019;38(11):2074–2102. doi:10.1002/sim.8086.
5. Nadeem, Muhammad Aqib, Sadia Sarwar, Chand Ali, Adeem Masih, Saleha Sohail, Kainat Saleem, Zarf E. Haider et al. "Comparative Analysis of Hepatitis C Virus (HCV) Screening Tests and Confirmatory PCR: Evaluation of Diagnostic Accuracy, Weak Positive Results, and Clinical Impact." *Pakistan Journal of Medical & Cardiological Review* 5, no. 1 (2026): 5561-5572. https://doi.org/10.5281/zenodo.21682658
6. Sarwar, Sadia, Hashir Masood, Samra Zulfiqar, Muhammad Aqib Nadeem, Noor ul Ain, Hamda Raza, Kainat Saleem et al. "Interpretation Of Genexpert Results And Their Association With Sputum Smear Status In Suspected Pulmonary Tuberculosis Patients." *Spectrum of Engineering Sciences* 4, no. 3 (2026): 6393-6407. https://doi.org/10.5281/zenodo.22688820
7. Arif, Neha, Hafsa Shafiq, Moiza Noor, Taha Shahbaz, Aswad Khan, Hafsa Noor, Khalil Ahmed, Eman Farrukh, and Muhammad Abdullah Butt. "Artificial Intelligence in Cardiovascular Diagnostics: Integration of Laboratory Biomarkers, Medical Imaging, and Molecular Data." *Pakistan Journal of Medical & Cardiological Review* 5, no. 2 (2026): 5665-5695. https://doi.org/10.66021/pakmcr1251
8. Riaz, Muhammad, Neelam Tariq, Madiha Iqbal, Muhammad Abdullah Butt, Qaisar Sohail, and Talha Riaz. "Migration, Displacement, and Health Inequities: A Comparative Study of Refugee Populations and Access to Healthcare Services." *Social Science Review Archives* 4, no. 2 (2026): 2704-2722. https://doi.org/10.70670/sra.v4i2.2369
9. Butt, Muhammad Abdullah, Feng Yiwen, and Umer Javeid. "Risk Governance Frameworks and Strategic Control Mechanisms for Managing Complexity in Global Business Operations." *Social Science Review Archives* 4, no. 2 (2026): 292-302. https://doi.org/10.70670/sra.v4i2.2056
10. Khurshid, Jamila, Zarlakhta Babar, Sajjad Ahmed, Muhammad Abdullah Butt, Umer Javeid, and Nida Khalil. "Beyond carbon footprints: unpacking the social dimensions of sustainability performance in emerging market firms." *Social Science Review Archives* 4, no. 1 (2026): 3386-3402. https://doi.org/10.70670/sra.v4i1.1902 \