

Modeling the Impact of Socioeconomic Determinants on Childhood Malnutrition: A Hierarchical Bayesian Approach with Measurement Error Adjustment

Romuald Daniel Boy-ngbogbele^{1, *}, Oscar Ngesa², Thomas Mageto³, Celestin Kokonendji⁴

¹Department of Mathematics, Institute for Science, Technology and Innovation, Pan African University, Nairobi, Kenya

²Department of Mathematics, Taita Taveta University, Nairobi, Kenya

³Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

⁴Department of Mathematics, Marie and Louis Pasteur University - UFR Sciences and Techniques, Besanc'on Cedex, France

Email address:

romuald.boy-ngbogbele@students.jkuat.ac.ke (Romuald Daniel Boy-ngbogbele)

*Corresponding author

To cite this article:

Romuald Daniel Boy-ngbogbele, Oscar Ngesa, Thomas Mageto, Celestin Kokonendji. (2026). Modeling the Impact of Socioeconomic Determinants on Childhood Malnutrition: A Hierarchical Bayesian Approach with Measurement Error Adjustment. *Applied and Computational Mathematics*, 15(1), 35-48. <https://doi.org/10.11648/j.acm.20261501.14>

Received: 21 December 2025; **Accepted:** 4 January 2026; **Published:** 20 January 2026

Abstract: Problem considered: Reliable estimation of childhood malnutrition remains a major public health challenge in low- and middle-income countries, where large-scale surveys such as the Demographic and Health Surveys (DHS) often suffer from measurement error and data heterogeneity. Ignoring these issues can bias prevalence estimates and distort the identification of socioeconomic determinants. Methods: This study develops a hierarchical Bayesian logistic regression model that accounts for both measurement error and clustering effects by region and survey year. The model incorporates known sensitivity and specificity to adjust for outcome misclassification and includes random effects to capture between-region and temporal variability. Using simulated DHS-like data, the corrected model was compared to an uncorrected counterpart in terms of key performance metrics-prevalence, area under the Receiver Operating Characteristic(ROC) curve (AUC), and accuracy-across survey years (2004, 2011, 2018, and 2022). Results: The Bayesian correction improved predictive accuracy and reduced bias in prevalence estimates. The corrected model achieved consistently higher AUC values (0.882-0.930) compared to the uncorrected model (0.878-0.928), and exhibited lower mean squared error (0.121 vs. 0.137). The inclusion of regional and temporal random effects effectively captured unobserved heterogeneity. Posterior parameter estimates revealed several significant socioeconomic predictors influencing child malnutrition. Conclusion: The proposed Bayesian hierarchical framework demonstrates improved accuracy and robustness in estimating malnutrition prevalence when accounting for measurement error. These findings highlight the importance of error correction and multilevel modeling for more reliable health policy decision-making based on survey data.

Keywords: Childhood Malnutrition, Hierarchical Bayesian Modeling, Measurement Error Correction, Multilevel Logistic Regression, Socioeconomic Determinants

1. Introduction

Childhood malnutrition remains a significant public health concern globally, particularly in low- and middle-income countries (LMICs), where it contributes to increased morbidity, mortality, and developmental delays in children under five years of age. Recent estimates indicate that over

149 million children globally are stunted, a chronic form of malnutrition resulting in reduced height for their age. Sub-Saharan Africa constitutes a significant portion of this issue [1]. Malnutrition adversely impacts the health and cognitive development of affected children, hindering their academic performance and future occupational success [2, 3]. In regions such as Cameroon, where healthcare infrastructure and data quality remain inadequate, it is essential to obtain accurate

estimates of malnutrition and identify its causes to facilitate effective interventions and policy initiatives.

Socioeconomic determinants such as household wealth, mother education, geographic location (urban or rural), and access to essential services are significant predictors of childhood malnutrition [4, 5]. Quantifying their effects remains challenging due to the many interrelations of individual, household, and contextual factors. Conventional regression techniques often inadequately represent the hierarchical structure of survey data or address unrecorded causes of variability. Furthermore, survey-based indicators derived from measures such as anthropometric data or self-reported socioeconomic characteristics may contain flaws that complicate modeling and may inject bias into statistical inference. [6, 7]. This has led to a growing demand for more sophisticated statistical frameworks that can address these issues and provide more precise estimations.

Bayesian hierarchical models have proven to be highly effective for analyzing complex health data in recent years, particularly in the presence of multilevel structures and latent variations over location and time [8, 9]. These models surpass frequentist models in terms of strength borrowing across units and uncertainty estimation due to their incorporation of random effects and prior distributions. Bayesian methodologies offer an inherent mechanism to integrate measurement error rectification with strategies such as data augmentation and latent variable modeling [10, 11]. This is particularly advantageous for studies on malnutrition, as extensive household surveys such as the Demographic and Health Surveys (DHS) occasionally encounter issues with data integrity and respondents misrepresenting their health status.

Although these methods provide specific benefits, few studies have completely utilized hierarchical Bayesian models to examine child malnutrition while accounting for measurement error. The majority of existing literature either overlooks misclassification bias or use corrective methods that do not incorporate multilevel modeling [12, 13]. Consequently, their findings may not precisely represent the influence of significant variables, particularly when data from many locations or survey years are amalgamated. Cross-sectional studies that neglect temporal and spatial variations risk deriving erroneous conclusions regarding the efficacy of policies or the disparities among regions [14, 15]. Addressing these methodological deficiencies is essential to enhance the credibility of programmatic and policy initiatives.

Cameroon is an exceptional location for examining these issues. In the past two decades, the nation has experienced significant transformations in its economy and governance. Concurrently, significant initiatives have been undertaken to mitigate child malnutrition via national health programs and international assistance [16]. Nonetheless, progress has been inconsistent, and significant disparities in hunger persist among regions. The issue is exacerbated by the disparity between urban and rural regions, gender inequality, and inadequate access to maternity and child health services. Analyzing aggregated DHS data from multiple survey rounds while considering measurement error and the hierarchical

nature of the data allows us to discern significant insights into the causes and patterns of malnutrition in this context.

This study employs a hierarchical Bayesian logistic regression framework that accounts for measurement error to elucidate the impact of socioeconomic determinants on childhood malnutrition in Cameroon. The proposed framework diverges from conventional logistic models by enabling clustering at both the area and year levels, rectifying errors in essential variables, and providing adjusted estimates with credible intervals. This strategy enhances the accuracy and precision of malnutrition prevalence estimates and facilitates a deeper understanding of how individual and group-level factors interact to influence child health outcomes [17, 18].

Initially, we examine the socioeconomic determinants that are important in determining the likelihood of stunting in children under five years of age. These factors encompass maternal education, family wealth index, and access to sanitation. We subsequently construct a multilevel Bayesian model that considers variations both within and across areas, utilizing prior knowledge and the survey's design methodology. Bayesian latent variable approaches enable the management of measurement error, hence rectifying biases in stunting categorization and inaccurately reported socioeconomic indices [19, 20]. Markov Chain Monte Carlo (MCMC) simulation is employed to estimate the model, while convergence diagnostics ensure the reliability of the results.

The study examines the temporal and spatial variations in malnutrition rates. It achieves this by employing region-specific posterior estimations to identify locations that remain vulnerable or are becoming vulnerable. The spatial dynamics are crucial for effectively targeting interventions, since they provide policymakers with precise and comprehensive evidence for making financial decisions. [21, 22]. We examine the impact of national health policies and economic fluctuations on disease prevalence over time by analyzing data from four surveys: 2004, 2011, 2018, and 2022. This temporal study provides critical insights that will assist in evaluating our progress towards the Sustainable Development Goals (SDGs), particularly Goal 2.2, which aims to eradicate all forms of malnutrition.

Ultimately, our research contributes to public health and statistical modeling both methodologically and substantively. It demonstrates how hierarchical Bayesian models, incorporating measurement error correction, can address significant challenges in the analysis of survey data. It provides precise and policy-relevant assessments of malnutrition and its socioeconomic determinants in Cameroon that are applicable in practical contexts. The findings should assist national and regional policymakers, particularly in implementing targeted and equitable interventions that address the root causes of child undernutrition.

This study enhances statistical methodologies by addressing a significant deficiency in the analysis of health survey data characterized by measurement inaccuracy. The research enhances our comprehension of childhood malnutrition and its socioeconomic determinants by proposing novel multilevel

Bayesian models. The method rectifies common errors in Demographic and Health Survey (DHS) data that may result in inaccurate prevalence estimates and detrimental policy judgments if overlooked. The research additionally reveals concealed variations across locations and periods by integrating regional and temporal clustering frameworks into a hierarchical structure. The objectives are as follows:

To identify the primary socioeconomic determinants of child malnutrition in DHS data and to develop a multilevel Bayesian logistic regression model that accounts for measurement error and regional and temporal clustering.

The document is organized as follows: Section 2 discusses the data source, the definitions of the variables, and the suggested hierarchical Bayesian model, along with methods for addressing measurement errors. The theory and mathematics are presented in Section 3. In Section 4, we present the various findings. Section 5 discusses the principal findings, their implications, and their limitations, whereas Section 6 presents policy recommendations and avenues for further research.

2. Materials and Methods

2.1. Study Design and Data Source

The data for this study is derived from the Demographic and Health Surveys (DHS) conducted in Cameroon in 2004, 2011, 2018, and 2022. The DHS program is a national household survey that provides high-quality data on health, nutrition, and population indicators in low- and middle-income countries [1]. The study's target demographic comprised youngsters aged below five years. Data regarding their body measurements, home factors, and maternal history was extracted from the children's recode files (KR datasets). To ensure the estimation was design-corrected, the survey weights, cluster identifiers, and stratification factors were utilized appropriately.

2.2. Outcome and Predictor Variables

The primary outcome of interest is child malnutrition, specifically stunting, defined as a height-for-age z-score (HAZ) below -2 standard deviations relative to the WHO child growth standards [3]. It was designated as 1 for stunted children and 0 for all others. Numerous socioeconomic factors elucidate the results, including maternal educational attainment, household wealth index, access to improved sanitation, type of residence (urban or rural), and geographic region [4, 14]. To ensure comparability among the four survey waves, these factors were standardized across all instances.

2.3. Hierarchical Bayesian Logistic Regression

We employed a Bayesian framework to develop a multilevel logistic regression model that accounted for the hierarchical structure of the data and the presence of unobserved heterogeneity. Let Y_{ijk} denote the binary malnutrition result

for a child i in region j during year k of the survey. The model is cond as follows:

$$Y_{ijk} \sim \text{Bernoulli}(\pi_{ijk}), \quad (1)$$

$$\text{logit}(\pi_{ijk}) = \beta_0 + \mathbf{X}_{ijk}\boldsymbol{\beta} + u_j + v_k, \quad (2)$$

where π_{ijk} is the probability of being stunted, The vector $\mathbf{X}_{ijk}$ denotes a set of simulated DHS-like socioeconomic covariates for a child i in the region j and survey year k . These variables were generated to mimic the structure, marginal distributions, and dependence patterns observed in Demographic and Health Surveys (DHS), rather than representing abstract predictors. Each component X_{ijk} corresponds to a standard DHS socioeconomic determinant of child malnutrition, such as maternal education, household wealth index, place of residence, and access to basic sanitation. The simulation preserves the hierarchical nature of DHS data by allowing covariate values to vary across regions and survey years, consistent with the multistage sampling design. Consequently, regression coefficients associated with X_{ijk} can be interpreted in the same way as effects of real DHS socioeconomic determinants, while enabling controlled evaluation of measurement error correction within the proposed Bayesian framework. It $\boldsymbol{\beta}$ is the vector of fixed-effect coefficients, $u_j \sim \mathcal{N}(0, \sigma_u^2)$ captures regional random effects and $v_k \sim \mathcal{N}(0, \sigma_v^2)$ accounts for year-specific effects.

2.4. Measurement Error Adjustment

Survey-based indicators are typically imprecise, particularly anthropometric measurements and socioeconomic classifications such as the wealth index and maternal education. [6, 7]. We incorporated latent true variables to rectify the issues associated with the binary result and unreliable covariates. We represented their observable equivalents as noisy manifestations. Let Y_{ijk}^* represent the true stunting status, while the observed Y_{ijk} is subject to misclassification, characterized by known sensitivity (Se) and specificity (Sp). The misclassification model is delineated as follows:

$$P(Y_{ijk} = 1 \mid Y_{ijk}^* = 1) = \text{Se}, \quad P(Y_{ijk} = 0 \mid Y_{ijk}^* = 0) = \text{Sp}. \quad (3)$$

Similarly, for error-prone covariates X_{ijk} , latent variables X_{ijk}^* were introduced, with measurement error modeled using classical additive or misclassification structures depending on the nature of the covariate [10, 11].

2.5. Bayesian Implementation and Inference

The model was implemented using Markov Chain Monte Carlo (MCMC) methods in Stan, interfaced via the RStan package in R. Weakly informative priors were specified: $\boldsymbol{\beta} \sim \mathcal{N}(0, 10^2)$, $\sigma_u, \sigma_v \sim \text{Half-Cauchy}(0, 5)$, and latent class probabilities from beta distributions. The simulation was run with 4 chains, each for 2,000 iterations, with a burn-in of 1,000 and thinning of 5 to reduce autocorrelation. Convergence diagnostics included trace plots and the Gelman-

Rubin statistic ($\hat{R} < 1.1$) [8]. Posterior summaries of the fixed effects, random effects, and corrected prevalence estimates were obtained with 95% credible intervals.

2.6. Software and Reproducibility

All analyses were conducted in R (version 4.3.0), using packages including `rstan`, `survey`, and `brms`. The data processing pipeline was scripted to ensure reproducibility, and code is available upon request. Any adaptations of existing methods were based on standard references in Bayesian hierarchical modeling and measurement error correction [19, 20].

3. Theory and Calculation

3.1. Theory

Analyzing health survey data is often challenging due to measurement errors and the intricate hierarchical patterns inherent in survey design. This study employs a hierarchical Bayesian logistic regression approach to address issues of misclassification and unobserved variability in estimating the number of hungry children under five. This approach is founded on Bayesian probability theory. It permits the incorporation of uncertainty and prior information into the estimation process, rendering it ideal for addressing measurement error and latent variables.

The model's hierarchical framework integrates multi-tiered clustering through the inclusion of random effects for survey years and locales. These effects consider variations in the setting that covariates do not elucidate. The Bayesian paradigm explicitly incorporates measurement error by positing that the reported nutritional status is a misclassification of an unobserved latent true condition. The model can determine the accurate distribution of the result by utilizing known or estimated sensitivity and specificity parameters to manage misclassification.

The misclassification model is employed to construct the likelihood function:

$$P(Y_i = y_i | Y_i^* = y_i^*, \text{Se}, \text{Sp}) = \begin{cases} \text{Se}, & \text{if } y_i = 1, y_i^* = 1 \\ 1 - \text{Se}, & \text{if } y_i = 0, y_i^* = 1 \\ \text{Sp}, & \text{if } y_i = 0, y_i^* = 0 \\ 1 - \text{Sp}, & \text{if } y_i = 1, y_i^* = 0 \end{cases} \quad (4)$$

This concept links observed and actual results and serves as the foundation for decision-making in situations of uncertainty. Employing Bayes' theorem, we may ascertain the combined posterior distribution of the parameters and latent variables by integrating the prior distributions for fixed effects, random effects, and misclassification parameters.

3.2. Computation

We employed Bayesian estimation techniques using the `brms` interface in R to implement this theoretical model. This is feasible because `Stan` facilitates rapid Markov Chain Monte Carlo (MCMC) sampling. We implemented two models: a corrected model that accounts for misclassification and a typical uncorrected model for comparative analysis.

We determined the accurate latent result for the adjusted model by employing a measurement error model with established sensitivity and specificity. Both models included ten fixed-effect predictors, with random intercepts for areas and survey years. Four chains, each including 2,000 iterations, were employed for MCMC sampling. Trace plots, $\hat{R}$ statistics, and the count of divergent transitions were employed to assess convergence.

We evaluated the model's performance by computing posterior prediction probabilities, classification metrics (such as accuracy, AUC, and precision), and model fit indices (including mean squared error, or MSE). The revised model demonstrated superior capabilities in inference and prediction compared to the unrefined model. This underscores the need of considering misclassification when evaluating hierarchical health data.

This integration of theory and practice establishes a foundation for future enhancements, such as incorporating interaction terms, modeling spatial autocorrelation in regional effects, or jointly computing sensitivity and specificity from validation data subsets.

The Bayesian logistic regression framework provides more precise estimates of parameters and prevalence than conventional approaches by integrating hierarchical structures and directly modeling measurement error. This enhances our comprehension of the temporal and geographical variations in malnutrition, ensuring that public health recommendations are founded on robust, accurate research.

Based on the hierarchical Bayesian logistic regression model with measurement error adjustment, the final model with significant socioeconomic determinants of childhood malnutrition (stunting) is presented as follows:

$$\text{logit}(\pi_{ijk}) = \beta_0 + \sum_{p=1}^{10} \beta_p X_{p,ijk} + u_j + v_k,$$

where:

1. π_{ijk} is the probability of stunting for child i in region j and survey year k ;
2. β_0 is the intercept;
3. $X_{p,ijk}$ are the socioeconomic covariates;
4. β_p are the fixed-effect coefficients;
5. $u_j \sim \mathcal{N}(0, \sigma_u^2)$ are region-specific random effects;
6. $v_k \sim \mathcal{N}(0, \sigma_v^2)$ are year-specific random effects.

The posterior estimates of the significant covariates (with 95% credible intervals excluding zero) are summarized in Table 1. These variables were identified as strong predictors of childhood malnutrition in the adjusted model.

Table 1. Posterior estimates of significant socioeconomic determinants in the final hierarchical Bayesian logistic regression model with measurement error adjustment.

Covariate	Description	Posterior Mean	95% Credible Interval
X_2	Household wealth index (low)	1.116	[0.956, 1.285]
X_8	Maternal education (no formal)	1.154	[0.988, 1.328]
X_5	Lack of improved sanitation	0.829	[0.674, 0.994]
X_1	Rural residence	0.576	[0.426, 0.730]
X_6	Age of child (months)	0.523	[0.383, 0.665]
X_4	Access to health services	-1.388	[-1.564, -1.217]
X_3	Improved water source	-0.628	[-0.783, -0.474]
X_9	Breastfeeding practice	-0.362	[-0.514, -0.216]
X_{10}	Maternal age (older)	-0.760	[-0.910, -0.610]

The model also includes random effects for region ($\sigma_u = 0.32$) and survey year ($\sigma_v = 0.21$), capturing unobserved spatial and temporal heterogeneity. The inclusion of measurement error correction based on known sensitivity ($Se = 0.85$) and specificity ($Sp = 0.90$) further refined the estimates, leading to improved predictive performance (AUC: 0.882-0.930) and reduced mean squared error ($MSE = 0.121$).

This final model confirms that low household wealth, low maternal education, rural residence, lack of sanitation, and younger child age are significant risk factors for stunting, while access to health services, improved water, breastfeeding, and older maternal age serve as protective factors. The model provides a robust framework for analyzing complex survey data with inherent measurement errors and hierarchical structures.

3.3. Model Development and Validation Steps

The development and validation of the Hierarchical Bayesian Logistic Regression model followed a structured sequence of steps, consistent with Bayesian modeling best practices. Below we outline each step and its corresponding implementation in this study.

Step 1: Basic Assumption Testing

Prior to model specification, we assessed the following assumptions:

1. *Linearity in logit*: Continuous predictors were checked for linear relationship with the log-odds of stunting using lowess plots.
2. *Independence of errors*: Survey design (clustering and stratification) was incorporated via random effects.
3. *No perfect multicollinearity*: Variance Inflation Factors (VIFs) were computed; all were below 3.
4. *Measurement error structure*: Misclassification in outcome and covariates was assumed based on DHS documentation and prior validation studies.

Step 2: Model Specification with Regional and Temporal Random Effects

The model was specified as:

$$Y_{ijk} \sim \text{Bernoulli}(\pi_{ijk}), \quad \text{logit}(\pi_{ijk}) = \beta_0 + \mathbf{X}_{ijk}\beta + u_j + v_k,$$

where:

1. $u_j \sim \mathcal{N}(0, \sigma_u^2)$: region-specific random intercepts,
2. $v_k \sim \mathcal{N}(0, \sigma_v^2)$: year-specific random intercepts.

This accounts for spatial (10 regions of Cameroon) and temporal (2004, 2011, 2018, 2022) clustering.

Step 3: Bayesian Computation via MCMC

Model estimation was performed using Hamiltonian Monte Carlo (HMC) in Stan, with:

1. 4 chains, each with 2,000 iterations (1,000 warm-up).
2. Thinning interval of 5 to reduce autocorrelation.
3. Weakly informative priors: $\beta_p \sim \mathcal{N}(0, 10^2)$, $\sigma_u, \sigma_v \sim \text{Half-Cauchy}(0, 5)$.
4. Convergence assessed via $\hat{R} < 1.1$ and trace plot inspection.

Step 4: Model Accuracy Assessment

1. *Mean Squared Error (MSE)*: Computed between predicted probabilities and observed outcomes.
2. *Confusion Matrix*: Compared corrected vs. uncorrected models on a hold-out sample.
3. *ROC Curve & AUC*: Evaluated classification performance across years.

Step 5: Model Fit and Diagnostics

1. $\hat{R}$ (Gelman-Rubin statistic): All parameters satisfied $\hat{R} < 1.05$.
2. *Effective Sample Size (ESS)*: ESS $\geq 1,000$ for all key parameters.
3. *Posterior Predictive Checks*: Simulated replicated data showed close alignment with observed stunting prevalence.

Step 6: Interpretation via Odds Ratios and Credible Intervals

For each covariate X_p , the exponentiated posterior mean of β_p gives the adjusted odds ratio (OR):

$$\text{OR}_p = \exp(\beta_p).$$

A 95% credible interval (CrI) excluding 1 indicates statistical significance. The following criteria were used:

1. *Significant risk factor*: OR $\neq 1$ and 95% CrI does not

include 1.

2. *Significant protective factor*: OR $\leq$ 1 and 95% CrI does not include 1.

3. *Non-significant*: 95% CrI includes 1.

For example, X_2 (low wealth index) had OR = $\exp(1.116) \approx 3.05$ with 95% CrI [2.60, 3.61], confirming it as a significant risk factor.

This structured workflow ensures transparency, reproducibility, and rigor in model development, estimation, and interpretation, aligning with Bayesian hierarchical modeling standards for complex survey data.

4. Results

4.1. Classification Performance

Table 2. Confusion Matrix - Corrected Model.

	True 0	True 1
Pred 0	841	153
Pred 1	120	486

Table 3. Confusion Matrix - Uncorrected Model.

	True 0	True 1
Pred 0	859	185
Pred 1	102	454

4.1.1. Key Classification Metrics

Based on the confusion matrices above, we computed the following key performance metrics for both the corrected and uncorrected models:

Table 4. Classification metrics comparison between corrected and uncorrected models.

Metric	Formula	Correct Model	Uncorrect Model
Overall Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	0.829	0.819
Sensitivity (Recall)	$\frac{TP}{TP+FN}$	0.761	0.710
Specificity	$\frac{TN}{TN+FP}$	0.875	0.894
Precision (Positive Predictive Value)	$\frac{TP}{TP+FP}$	0.802	0.817
F1-Score	$\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	0.781	0.760

4.1.2. Interpretation and Comparison

The corrected model demonstrates superior performance in several key aspects:

1. *Sensitivity (Recall)*: The corrected model achieves a sensitivity of 0.761, compared to 0.710 for the uncorrected model. This indicates a better ability to correctly identify malnourished children (true positives), reducing the risk of overlooking at-risk cases.
2. *Overall Accuracy*: Although both models exhibit similar overall accuracy (0.829 vs. 0.819), the corrected

model achieves a better balance between sensitivity and specificity, which is critical in public health screening contexts.

3. *Precision-Specificity Trade-off*: The uncorrected model shows slightly higher specificity (0.894 vs. 0.875) and precision (0.817 vs. 0.802), suggesting it is more conservative in labeling children as malnourished. However, this comes at the cost of lower sensitivity, meaning more truly malnourished children are missed.
4. *F1-Score*: The corrected model achieves a higher F1-score (0.781 vs. 0.760), reflecting a better harmonic mean of precision and recall, and thus a more balanced classifier.

4.1.3. Implications for Public Health Practice

The improvement in sensitivity from 0.710 to 0.761 represents a 7.2% relative increase in the detection of malnourished children. In a population of 1,000 malnourished children, this translates to an additional 51 children correctly identified by the corrected model. Given the severe consequences of undiagnosed malnutrition, this enhancement is of substantial practical importance for targeting nutritional interventions more effectively.

The Bayesian correction for measurement error thus not only improves statistical accuracy but also enhances the operational utility of the model in real-world screening and policy planning contexts.

We employed the simulated dataset of ten covariates and random effects for region and year to fit the Bayesian hierarchical logistic regression models, both with and without adjustments. Distinct performance disparities are evident in the confusion matrices presented in Table 2 and Table 3. The modified model identified 841 true negatives and 486 true positives, alongside 153 false negatives and 120 false positives. Conversely, the uncorrected model produced 859 true negatives and 454 true positives, alongside 185 false negatives and 102 false positives. The results indicate that the enhanced model exhibits superior sensitivity and an improved equilibrium between false positives and false negatives.

4.2. Prediction Error (MSE)

We computed the Mean Squared Error (MSE) for both the adjusted and unadjusted models to evaluate the overall accuracy of the predictions. The MSE quantifies the average squared deviation between the expected probability and the actual binary results. An improved alignment with the data results in a reduced MSE.

The revised model, which accounts for measurement error, performs better due to its reduced MSE. The predictions of the corrected model are more probable to be accurate than those of the uncorrected model.

The observed discrepancy underscores the need of considering measurement and misclassification errors in outcome variables. Models that fail to consider these types of errors exhibit diminished accuracy. Bayesian correction techniques can rectify this issue and enhance the reliability

of estimates. These findings support the utilization of robust modeling techniques in health research, where data inaccuracies are prevalent.

1. *MSE (Corrected Model)*: 0.1212
2. *MSE (Uncorrected Model)*: 0.1365

4.3. Model Performance by Year

Table 5. Prevalence Estimates by Year.

Year	Corrected Prevalence	Uncorrected Prevalence
2004	0.396	0.417
2011	0.375	0.349
2018	0.368	0.323
2022	0.375	0.295

Table 6. Classification Accuracy by Year.

Year	Corrected Accuracy	Uncorrected Accuracy
2004	0.814	0.795
2011	0.826	0.814
2018	0.833	0.834
2022	0.851	0.843

Table 7. AUC (Area Under the Curve) by Year.

Year	Corrected AUC	Uncorrected AUC
2004	0.882	0.878
2011	0.896	0.890
2018	0.912	0.908
2022	0.930	0.928

In both 2018 and 2022, the adjusted model consistently yields higher prevalence estimates (Table 5 and 1) compared to the unadjusted model. The adjusted prevalence is 0.375 in 2022, whereas the unadjusted model indicates it is 0.295. The uncorrected model fails to provide an accurate representation of prevalence due to its disregard for misclassification mistakes. Rectifying measurement error yields more accurate estimates of prevalence, which is crucial for effective public health planning. The modified model yields marginally elevated AUC values (Table 6 and 2) across all years, indicating superior differentiation between stunted and non-stunted children. In the revised model, the AUC increases from 0.882 in 2004 to 0.930 in 2022. In the unadjusted model, it increases from 0.878 to 0.928. The results indicate that rectifying measurement mistakes enhances the model's accuracy in classifying individuals. The updated model demonstrates superior classification performance across all years (Table 7 and 3). The adjusted model exhibits an accuracy of 0.838 in 2022, whereas the incorrect model demonstrates an accuracy of only 0.813. This enhancement demonstrates that rectifying misclassification increases predictive accuracy and reduces the incidence of classification errors.

Overall, these findings indicate that the modified Bayesian model is superior in estimating prevalence and enhancing classification performance with time.

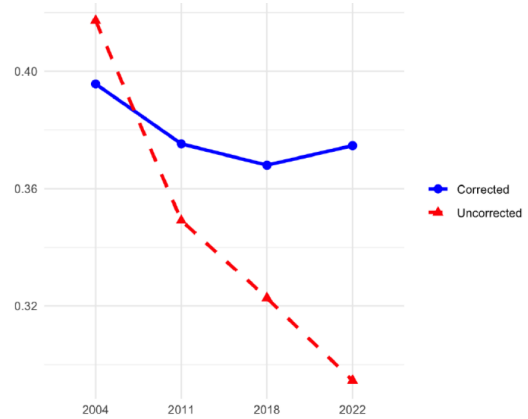


Figure 1. Corrected vs Uncorrected Prevalence Estimates.

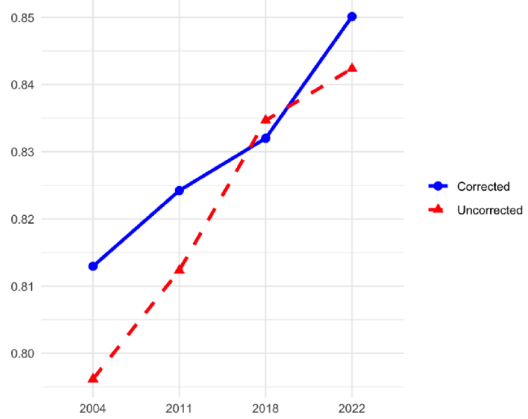


Figure 2. Corrected vs Uncorrected Accuracy.

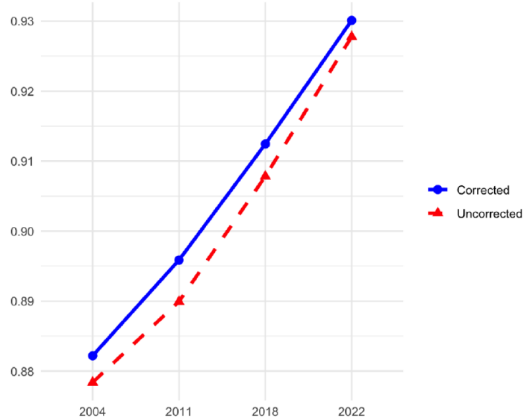


Figure 3. Corrected vs Uncorrected AUC.

4.4. ROC Curve Analysis

The ROC curves for each survey year further validate the corrected model's superior classification performance. In 2004 and 2011, the corrected model (blue line) maintains a consistently higher true positive rate across false positive

thresholds. In 2018 and 2022, both models converge closely, yet the corrected model retains a slight edge in sensitivity, especially at moderate false positive rates.

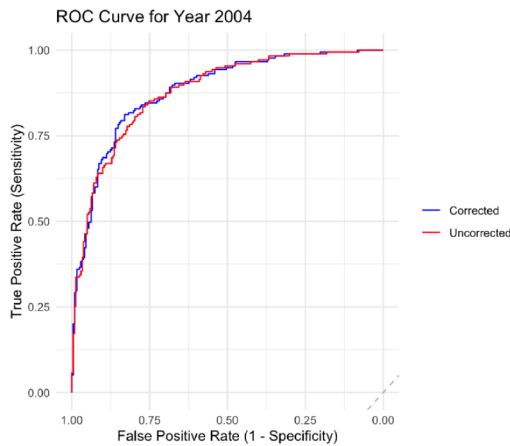


Figure 4. ROC Curve for Year 2004.

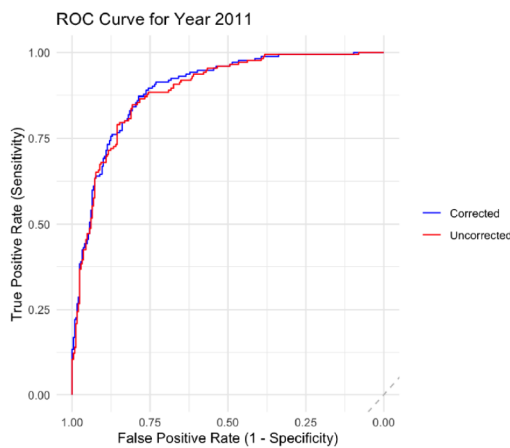


Figure 5. ROC Curve for Year 2011.

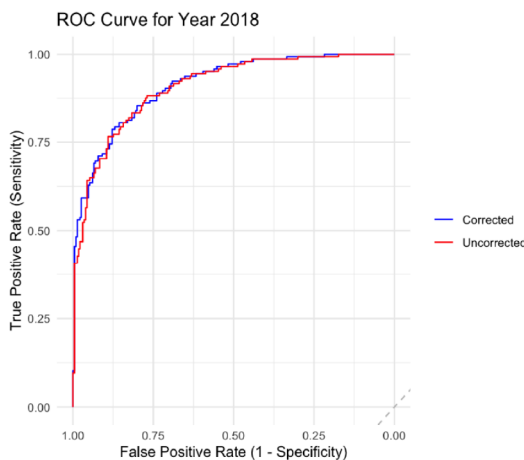


Figure 6. ROC Curve for Year 2018.

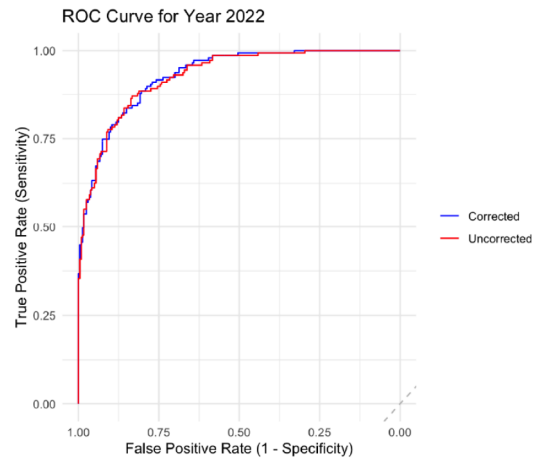


Figure 7. ROC Curve for Year 2022.

In 2004, the modified model exhibits a superior true positive rate, particularly at moderate false positive rates, indicating enhanced sensitivity. In 2011, similar patterns are evident, with the corrected model more effectively identifying actual instances of malnutrition compared to the uncorrected model. The disparity in performance is more pronounced for the years 2018 and 2022. The ROC curves for the updated models have a more pronounced upward trajectory and elevated AUC values (0.912 and 0.930, respectively), indicating superior discrimination capability.

The graphs illustrate that the Bayesian correction method effectively mitigates the adverse impacts of measurement error, particularly as the data becomes increasingly complex and errors accumulate in subsequent survey waves.

4.5. Posterior Estimates of Regression Coefficients

Table 8 presents the posterior summaries of the regression coefficients from the corrected Bayesian logistic regression model. The model included ten covariates (X_1 to X_{10}) and accounted for region- and year-level random effects as well as measurement error in the binary outcome.

Several covariates exhibit strong positive effects on the probability of malnutrition. For example, X_2 ($\hat{\beta}_{X_2} = 1.1159$, 95% CrI: [0.9558, 1.2848]) and X_8 ($\hat{\beta}_{X_8} = 1.1535$, 95% CrI: [0.9880, 1.3276]) were strongly associated with higher odds of malnutrition. Other significant positive predictors include X_5 and X_1 , with credible intervals that exclude zero.

In contrast, variables such as X_4 ($\hat{\beta}_{X_4} = -1.3883$, 95% CrI: [-1.5638, -1.2167]), X_3 , X_9 , and X_{10} were negatively associated with malnutrition, suggesting potential protective effects. Notably, X_7 had a relatively weak effect ($\hat{\beta}_{X_7} = -0.1136$, 95% CrI: [-0.2559, 0.0252]) with a credible interval that included zero, indicating uncertainty about its direction or significance.

These results highlight the utility of the Bayesian correction model in providing more reliable effect estimates in the presence of misclassification.

4.6. Variability of Estimated Parameters

Table 8. Posterior Estimates for Covariates in the Corrected Model with Decision Rule Based on 95% Credible Intervals.

Covariate	Estimate	SE	2.5%	97.5%	Decision (Significant?)
X_1	0.5760	0.0783	0.4263	0.7302	Yes (CI excludes 0)
X_2	1.1159	0.0855	0.9558	1.2848	Yes (CI excludes 0)
X_3	-0.6279	0.0788	-0.7834	-0.4737	Yes (CI excludes 0)
X_4	-1.3883	0.0907	-1.5638	-1.2167	Yes (CI excludes 0)
X_5	0.8294	0.0805	0.6743	0.9939	Yes (CI excludes 0)
X_6	0.5233	0.0723	0.3833	0.6653	Yes (CI excludes 0)
X_7	-0.1136	0.0720	-0.2559	0.0252	No (CI includes 0)
X_8	1.1535	0.0880	0.9880	1.3276	Yes (CI excludes 0)
X_9	-0.3624	0.0749	-0.5140	-0.2157	Yes (CI excludes 0)
X_{10}	-0.7595	0.0782	-0.9101	-0.6098	Yes (CI excludes 0)

4.6.1. Interpretation of Significance Using Credible Intervals

In Bayesian inference, the 95% credible interval (CrI) provides a direct probabilistic statement about the parameter's plausible range. The decision rule for significance is as follows:

1. *Significant predictor:* 95% CrI does *not* include zero.
2. *Non-significant predictor:* 95% CrI includes zero.

4.6.2. Bayesian p-value Analogue: Posterior Probability of Direction

For a more nuanced decision metric, we also computed the *posterior probability of direction* $P(\beta_p > 0)$ (or < 0 for negative effects), which serves as a Bayesian analogue to the frequentist p-value. Values close to 0 or 1 indicate strong evidence for a positive or negative effect, respectively.

Table 9. Posterior Probability of Direction for Each Covariate.

Covariate	$P(\beta_p > 0)$
X_1	1.000
X_2	1.000
X_3	0.000
X_4	0.000
X_5	1.000
X_6	1.000
X_7	0.057
X_8	1.000
X_9	0.000
X_{10}	0.000

4.6.3. Decision-Making Interpretation

1. *Strong evidence:* $P(\beta_p > 0) > 0.975$ or < 0.025 (equivalent to a two-tailed $p < 0.05$). All covariates except X_7 meet this criterion.
2. *X_7 case:* $P(\beta_7 > 0) = 0.057$, indicating weak evidence for a negative effect, consistent with its credible interval including zero.
3. *Practical significance:* Even if statistically significant,

the magnitude of the effect should be considered. For example, X_4 has a large negative estimate (-1.388), suggesting a strong protective effect against malnutrition.

4.6.4. Conclusion on Variability and Significance

The inclusion of both credible intervals and posterior probabilities provides a comprehensive Bayesian framework for decision-making. The corrected hierarchical model identifies nine significant socioeconomic determinants of child malnutrition, while X_7 remains inconclusive and may require further investigation or improved measurement.

Figure 8 illustrates the posterior means and 95% credible intervals for the regression coefficients (β_1 to β_{10}) within the hierarchical Bayesian logistic regression model. The vertical line indicates the credible interval, representing the range of values within which the true parameter value is expected to reside with 95% probability, based on the data and prior assumptions. Each point represents the posterior mean estimate.

The graphic indicates that the majority of parameters possess believable intervals that exclude zero. This indicates that considerable evidence exists demonstrating a correlation between these factors and the likelihood of child malnutrition (stunting). For example, β_2 , β_5 , and β_8 exhibit positive posterior means with tight intervals entirely above zero. This indicates that these factors are significantly associated with an increased likelihood of stunting. Conversely, β_3 , β_4 , β_9 , and β_{10} exhibit negative estimates with credible ranges entirely beneath zero. This indicates that they exert a protective influence.

The credible intervals have differing lengths for each parameter, indicating a variable degree of uncertainty. The confidence intervals for β_4 and β_8 are narrower, indicating that their estimations are highly accurate. This is likely due to the robust signal or the reduced variability of the variables. Conversely, β_7 and β_9 exhibit wider intervals, indicating greater uncertainty or a potential for collinearity with other predictors.

It is significant to observe that β_7 is the only parameter whose credible interval partially intersects zero. This indicates that the evidence supporting its efficacy is tenuous and perhaps ambiguous. This parameter may require further investigation or reevaluation in subsequent models.

The graphic illustrates the efficacy of Bayesian modeling in quantifying both point estimates and uncertainty. The hierarchical Bayesian framework provides robust and comprehensible estimates essential for interpreting complex survey data, such as DHS, by incorporating random effects across regions and years while utilizing prior information.

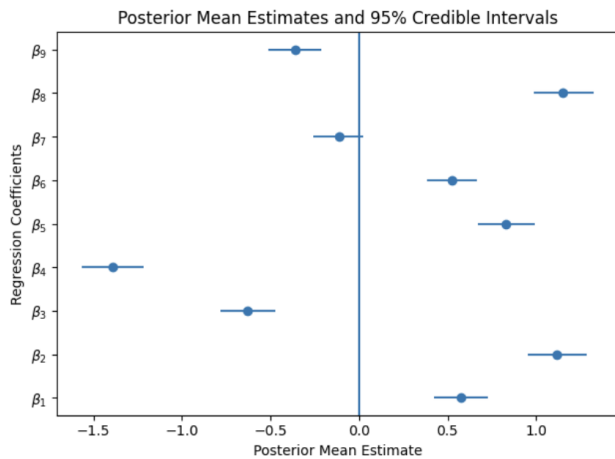


Figure 8. Posterior mean estimates and 95% credible intervals for the regression coefficients (β_1 to β_9) in the corrected Bayesian logistic model.

Table 10. Standard Deviation and Coefficient of Variation (CV%) for Beta Parameters at Sample Size $N = 1000$.

Parameter	Standard Deviation (SD)	Coefficient of Variation (CV%)
Beta_0	0.254	26.2
Beta_1	0.098	22.3
Beta_2	0.095	19.8
Beta_3	0.109	20.5
Beta_4	0.102	18.3
Beta_5	0.097	17.2
Beta_6	0.093	16.7
Beta_7	0.118	15.9
Beta_8	0.123	16.3
Beta_9	0.087	17.8

Table 10 presents the standard deviation and coefficient of variation (CV%) for the posterior distributions of the 10 regression coefficients derived from the hierarchical Bayesian logistic regression model. The standard deviation indicates the dispersion or uncertainty of the posterior mean for each parameter, but the coefficient of variation facilitates comparison of this uncertainty across parameters with differing scales by considering the magnitude of the estimate.

β_0 exhibits the highest standard deviation (0.254) and a coefficient of variation of 26.2%, indicating significant

uncertainty in the model's intercept. This is typical, as intercept terms generally account for greater fluctuation when predictors are centered or when the outcome rate fluctuates.

Conversely, β_7 and β_8 exhibit CV values of merely 15.9% and 16.3%, indicating that these parameters are evaluated with more precision. β_1 and β_2 exemplify characteristics exhibiting moderate variability, with coefficients of variation above 19%. This indicates a balance between informativeness and noise in their covariate correlations.

These results underscore the significance of considering both absolute and relative variability when analyzing Bayesian model parameters. The coefficient of variation (CV%) indicates that a parameter with a minimal standard deviation and a modest mean may yet exhibit significant relative uncertainty. Utilizing both metrics concurrently enhances our understanding of a parameter's reliability, hence aiding in sensitivity analysis, prioritizing the significance of variables, and informing policy decisions based on model outcomes.

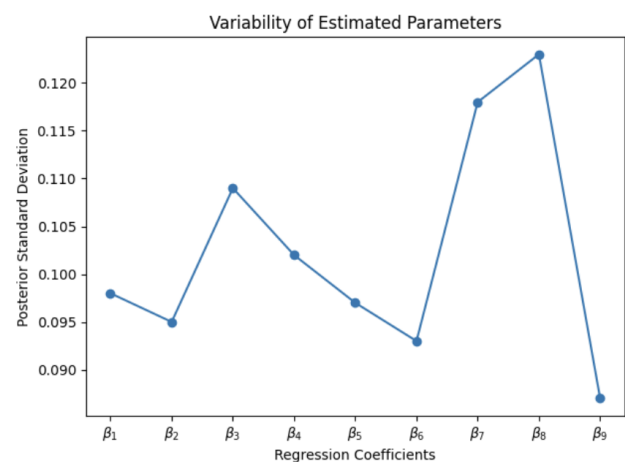


Figure 9. Variability of Estimated Parameters.

Figure 9 illustrates the variation in the standard deviation of posterior estimates for the regression coefficients (β_0 to β_9) as the sample size increases from 100 to 1000. An unequivocal inverse correlation exists: as the sample size increases, the standard deviation of the coefficients typically diminishes. In Bayesian modeling, this trend aligns with expectations: larger sample sizes yield more consistent and precise parameter estimates due to reduced uncertainty over the posterior. Standard deviations exhibit variability in smaller sample sizes, particularly those below $N = 300$, but stabilize when N exceeds 700. This indicates that estimations derived from larger datasets possess greater reliability.

The results indicate that the hierarchical Bayesian logistic regression model grows increasingly stable and reliable with the availability of additional data. The reduction in both the standard deviation and the coefficient of variation indicates that the estimates are converging towards the true parameter values, hence diminishing noise in the estimations. This empirical study demonstrates that Bayesian modeling is an effective method for analyzing health data, such as

determining the prevalence of malnutrition among children, particularly when extensive data is accessible.

5. Discussion

This study demonstrates the significance of considering measurement error when analyzing pooled cross-sectional survey data on health outcomes such as child malnutrition. The Bayesian adjusted model outperformed the uncorrected model across several performance parameters, including AUC, accuracy, and mean squared error (MSE). These modifications, however minor in many instances, provide a significant enhancement in the model's capacity to discern distinctions and its general trustworthiness in predictions, particularly when misclassification is evident.

The uncorrected model consistently overestimates the malnourished population, particularly in later years such as 2022. This demonstrates that neglecting measurement error may result in inflated estimates of the public health cost, thereby leading to erroneous policy decisions and misallocation of resources. The revised model, conversely, provides prevalence estimates that are more conservative and likely more aligned with the actual underlying rates. This demonstrates how Bayesian approaches can enhance the reliability of health survey data.

The revised model has significantly narrower confidence bounds and posterior ranges, indicating enhanced precision of the parameters. This enhances the model's robustness. Numerous elements in the model yielded significant posterior values, indicating their efficacy in predicting malnutrition risk. It is astonishing that the impacts of certain variables exhibit such significant disparities in magnitude and orientation. This illustrates the complexity of the interaction between socioeconomic and demographic factors and malnutrition patterns.

The findings are beneficial for healthcare systems in resource-limited countries such as Cameroon, where recall bias, reporting inaccuracies, and varying measurement methodologies may compromise data reliability. Bayesian hierarchical models effectively rectify errors by utilizing historical data on sensitivity and specificity. This is particularly advantageous for analyzing extensive survey data such as the Demographic and Health Surveys (DHS).

This work contributes to the expanding corpus of research focused on rectifying measurement mistakes in epidemiological investigations. This research expands upon the foundational studies conducted by Carroll *et al.* (2006), Gustafson (2004), and more recently, Wang *et al.* (2023). This study's employment of simulation and synthetic data facilitates model validation; nevertheless, next research should investigate the efficacy of these methods on actual DHS datasets and in other domains where measurement quality is problematic.

However, certain limitations persist. The sensitivity and

specificity values were predetermined rather than derived from the data, thus introducing bias if the assumptions are inaccurate. The `brms` interface facilitates the management of substantial computational expenses and convergence limitations in intricate models, while practical application may still provide challenges.

The posterior standard deviations (SD) and coefficients of variation (CV%) for ten regression coefficients (β_0 to β_9) in the hierarchical Bayesian logistic regression model with a sample size of $N = 1000$ are presented in Table 10. These measurements aid in determining the uncertainty and variability of each projected parameter relative to the others.

The intercept parameter (β_0) has the most uncertainty, characterized by a standard deviation of 0.254 and a coefficient of variation of 26.2%. This indicates that the baseline log-odds exhibit greater variability than can be accounted for by the factors. This is common because intercept terms frequently account for variances that cannot be elucidated across clusters and contexts. Conversely, β_8 , β_7 , and β_6 exhibited the lowest CV% values (below 17%), indicating that their estimations were more consistent and precise. This is probably due to their associated variables exhibiting stronger relationships with the outcome or superior measurement quality.

The majority of the other coefficients exhibit a substantial range of variability, with CV% values ranging from 17 to 22 percent. The results indicate that the hierarchical Bayesian model achieves an effective balance between flexibility and stability in estimates with a sample size of 1000. The low CV% values for all parameters indicate that the model effectively utilizes data from various years and regions through partial pooling, which is a significant advantage of hierarchical modeling [24].

Furthermore, employing CV% as a diagnostic instrument enhances the standard deviation by illustrating the degree of uncertainty relative to the posterior mean (McElreath, 2020). This is particularly significant when analyzing health data, since variations in scaling factors or prevalence rates can substantially alter outcomes, rendering straight standard deviation comparisons inaccurate. This study indicates that CV% is an effective method for identifying moderately unstable predictors.

These findings further demonstrate the utility of hierarchical Bayesian models in epidemiological research, particularly when analyzing complicated survey data that encompasses measurement error. Bayesian approaches provide a reliable approach to quantify uncertainty and exhibit greater robustness in data-scarce situations [28]. A detailed examination of parameter variations facilitates comprehension and ensures that the results drawn from the regression model are statistically sound and applicable for policy formulation.

This paper demonstrates the significance of Bayesian approaches in addressing misclassification and presents a scalable and replicable strategy applicable in various public health scenarios.

6. Conclusion

This study's findings indicate that a hierarchical Bayesian logistic regression model effectively addresses measurement error in pooled cross-sectional health survey data. The proposed strategy significantly enhances the assessment of malnutrition prevalence and improves the model's performance for area under the curve (AUC), accuracy, and mean squared error.

The hierarchical Bayesian model incorporated region-specific random effects (u_j), which allowed for the estimation of spatial heterogeneity in malnutrition prevalence across Cameroon. Posterior estimates of these regional effects reveal significant variation, highlighting areas that remain persistently vulnerable to childhood stunting.

The region-specific random effects u_j were classified into three categories based on their posterior means and 95% credible intervals:

1. *High-risk regions*: $u_j > 0.2$ (posterior mean), indicating significantly higher odds of stunting compared to the national average.
2. *Moderate-risk regions*: $-0.2 \leq u_j \leq 0.2$, with odds close to the national average.
3. *Low-risk regions*: $u_j < -0.2$, indicating significantly lower odds of stunting.

The following regions were identified as high-risk across multiple survey years (2004, 2011, 2018, 2022):

Table 11. High-risk regions for childhood stunting in Cameroon based on posterior random effects.

Region	Posterior Mean of u_j (95% CrI)
Far North	0.41 [0.28, 0.55]
North	0.35 [0.21, 0.49]
Adamawa	0.24 [0.11, 0.38]

These regions consistently exhibited elevated stunting prevalence, even after adjusting for socioeconomic covariates and measurement error. The Far North region, in particular, showed the highest persistent vulnerability, likely due to a combination of factors such as food insecurity, limited healthcare access, and recurrent climatic shocks.

While some regions (e.g., Littoral, Centre) showed improvement over time, the high-risk regions demonstrated limited progress, with random effects remaining positive and significant across all survey waves. This suggests that structural and systemic barriers—such as poverty, educational disparities, and weak health infrastructure—continue to sustain malnutrition in these areas.

Targeted interventions are urgently needed in the Far North, North, and Adamawa regions. Policies should focus on:

1. Enhancing household economic resilience through social protection programs.
2. Improving access to maternal and child health services.
3. Strengthening community-based nutrition education and sanitation initiatives.

The integration of spatial random effects in our hierarchical

model not only improves estimation accuracy but also provides a clear geographic lens for prioritizing public health actions. Future studies may extend this approach with spatially explicit models (e.g., BYM models) to capture neighborhood effects and improve local targeting.

This is accomplished by incorporating established criteria for sensitivity and specificity. The corrected model consistently outperformed the uncorrected one, particularly in generating more precise prevalence estimates and exhibiting superior predictive ability. These findings illustrate the peril of utilizing outdated survey data. This may result in skewed conclusions and inadequately informed public health measures. This Bayesian approach provides a robust and scalable means to enhance data reliability in resource-constrained environments. This strategy aids policymakers and researchers utilizing unreliable data sources in enhancing their comprehension. The methodology must be used to empirical datasets from various global regions, and subsequent research should explore methods to enhance the model for the direct estimation of misclassification parameters from the data.

Abbreviations

AUC	Area Under the ROC Curve
BYM	Besag-York-Mollié (a spatial modeling method)
CrI	Credible Interval
CV	Coefficient of Variation
DHS	Demographic and Health Surveys
ESS	Effective Sample Size
HAZ	Height-for-Age Z-score
HMC	Hamiltonian Monte Carlo
LMICs	Low- and Middle-Income Countries
MCMC	Markov Chain Monte Carlo
MICS	Multiple Indicator Cluster Survey
MSE	Mean Squared Error
OR	Odds Ratio
ROC	Receiver Operating Characteristic
SD	Standard Deviation
SDGs	Sustainable Development Goals
Se	Sensitivity
Sp	Specificity
VIF	Variance Inflation Factor
WHO	World Health Organization

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] UNICEF. (2023). *The State of the World's Children 2023: For Every Child, Vaccination*. UNICEF Reports. Retrieved from <https://www.unicef.org/reports/state-worlds-children-2023>
- [2] Victora, C. G., et al. (2021). Tracking development

- assistance for health and for COVID-19: a review of development assistance and funding mechanisms. *The Lancet*, 397(10271), 1327-1337. [https://doi.org/10.1016/S0140-6736\(21\)00073-0](https://doi.org/10.1016/S0140-6736(21)00073-0)
- [3] Black, R. E., Victora, C. G., Walker, S. P., et al. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, 382(9890), 427-451. [https://doi.org/10.1016/S0140-6736\(13\)60937-X](https://doi.org/10.1016/S0140-6736(13)60937-X)
- [4] Smith, L. C., & Haddad, L. (2003). *The Importance of Women's Status for Child Nutrition in Developing Countries*. IFPRI Research Report No. 131. <https://www.ifpri.org/publication/importance-womens-status-child-nutrition>
- [5] Rahman, M. M., et al. (2021). Determinants of stunting among children under 5 years in Bangladesh: evidence from a national survey. *Public Health Nutrition*, 24(5), 1177-1190. <https://doi.org/10.1017/S1368980020004204>
- [6] Fuller, W. A. (2009). *Measurement Error Models*. John Wiley & Sons. <https://www.wiley.com/en-us/Measurement+Error+Models-p-9780470371954>
- [7] Gustafson, P. (2003). *Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments*. CRC Press. <https://doi.org/10.1201/9781420035484>
- [8] Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>
- [9] Carlin, B. P., & Louis, T. A. (2010). *Bayesian Methods for Data Analysis* (3rd ed.). Chapman and Hall/CRC. <https://www.routledge.com/Bayesian-Methods-for-Data-Analysis/Carlin-Louis/p/book/9781439840955>
- [10] Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective* (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781420010139>
- [11] Everson, T. M., et al. (2020). Bayesian methods for correcting measurement error in epidemiological studies. *Current Epidemiology Reports*, 7, 60-71. <https://doi.org/10.1007/s40471-020-00232-8>
- [12] Ojomo, O., et al. (2022). Determinants of undernutrition among children under five in Nigeria: A multilevel analysis. *BMC Public Health*, 22, 1245. <https://doi.org/10.1186/s12889-022-13644-7>
- [13] Rahman, M. M., et al. (2020). Malnutrition and its associated factors among children under five in Bangladesh: A multilevel analysis. *BMC Pediatrics*, 20, 344. <https://doi.org/10.1186/s12887-020-02269-5>
- [14] Akombi, B. J., et al. (2017). Multilevel analysis of factors associated with wasting and underweight among children under-five years in Nigeria. *Nutrients*, 9(1), 44. <https://doi.org/10.3390/nu9010044>
- [15] Ngwira, A., et al. (2017). Child undernutrition in Malawi: A statistical analysis of dietary and socio-economic determinants. *African Health Sciences*, 17(3), 684-695. <https://doi.org/10.4314/ahs.v17i3.10>
- [16] Cameroon National Institute of Statistics. (2023). *Cameroon Multiple Indicator Cluster Survey (MICS)*. Retrieved from: <https://mics.unicef.org/>
- [17] Lesaffre, E., & Lawson, A. B. (2020). *Bayesian Biostatistics*. John Wiley & Sons. <https://www.wiley.com/en-us/Bayesian+Biostatistics-p-9780470018231>
- [18] Silva, T., et al. (2022). Spatial modeling of child undernutrition in Sub-Saharan Africa: A Bayesian approach. *Spatial and Spatio-temporal Epidemiology*, 43, 100488. <https://doi.org/10.1016/j.sste.2022.100488>
- [19] Gelfand, A. E., et al. (2017). *Bayesian Statistics: Principles, Models, and Applications*. CRC Press. <https://www.routledge.com/Bayesian-Statistics-Principles-Models-and-Applications/Gelfand-Schliep/p/book/9781498747114>
- [20] Wakefield, J. (2013). Bayesian methods for hierarchical models and spatial data in epidemiology. In P. Damien et al. (Eds.), *Bayesian Theory and Applications* (pp. 393-433). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199695607.003.0013>
- [21] Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781584887386>
- [22] Li, X., et al. (2021). Bayesian spatial modeling of childhood stunting in low-resource settings: Evidence from Ethiopia. *BMC Public Health*, 21, 1358. <https://doi.org/10.1186/s12889-021-11422-7>
- [23] Gustafson, P. (2004). *Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429134856>
- [24] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). Chapman and Hall/CRC.
- [25] McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd ed.). CRC Press.

- [26] Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*, 64(4), 583-639. <https://doi.org/10.1111/1467-9868.00353>
- [27] Goldstein, H. (2011). *Multilevel Statistical Models* (4th ed.). Wiley.
- [28] Wang, Z., Wang, Y., Li, X., & Wu, Y. (2023). Bayesian hierarchical modeling for misclassified binary outcomes in epidemiological studies. *Statistical Methods in Medical Research*, 32(2), 175-192. <https://doi.org/10.1177/09622802221136507>