International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print) 2222-5234 (Online)
http://www.innspub.net
Vol. 25, No. 6, p. 299-312, 2024

INNSPUB

RESEARCH PAPER OPEN ACCESS

Prevalence of anemia and spatial correlations among reproductive-age women in northeast India: Investigating spatially associated factors across districts

Sabrina Sultana Rahman[1], Mukesh Vishwakarma[2], Pawan Kumar Dubey[3], Ruma Talukdar*[4]

[1] *Department of Statistics, Cotton University, Guwahati, Assam, India*

[2] *Department of Mathematics and Statistics, Banasthali Vidyapith, Rajasthan, India*

[3] *Monitoring and Evaluation, IHAT, Uttar Pradesh, India*

Key words: Iron deficiency anemia, Spatial analysis, Socio-demographic factors, Dietary habits, Northeast India

http://dx.doi.org/10.12692/ijb/25.6.299-312 Article published on December 08, 2024

Abstract

Anemia, particularly iron deficiency anemia (IDA), is a critical public health issue in Northeast India, a region comprising eight states with unique geographical and socio-economic traits. To analyze the spatial distribution and factors influencing anemia among women of reproductive age in the districts of northeastern India. Descriptive statistics assessed independent variables' distribution. Spatial patterns of anemia were analyzed using Moran's I, LISA, and a LISA cluster map. Spatial Lag Model (SLM) and spatial regression models identified regional factors influencing anemia in women. The analysis reveals significant variability in anemia prevalence among women in Northeast India, influenced by socio-demographic factors and dietary habits. Higher anemia risk is linked to being Hindu, poorer, uneducated, or widowed/divorced/separated. Regular vegetable consumption reduces anemia rates, while high fried food intake increases them. High anemia rates are found in 35 districts, notably in Assam, Tripura, parts of Arunachal Pradesh, and small areas of Meghalaya. Clusters of high prevalence are present in 24 districts, including Assam and Tripura, with 22 districts showing slightly elevated rates at a 0.05 significance level. The research highlights regional disparities in anemia prevalence across Northeast India, identifying distinct clusters with high and low rates. It underscores the importance of region-specific strategies to address these variations effectively. By pinpointing critical areas and examining the key factors driving elevated anemia rates, the study lays the foundation for targeted interventions aimed at enhancing women's health outcomes in the region.

* **Corresponding Author:** Ruma Talukdar ✉ rumatalukdar@cottonuniversity.ac.in

Introduction

Anemia, as defined by the World Health Organization (WHO), is a condition characterized by a deficiency of healthy red blood cells or hemoglobin in the blood, resulting in fatigue, weakness, tiredness, pale skin, irregular heartbeat, shortness of breath, chest pain, dizziness, and ringing in ears (Mantadakis *et al.*, 2020; Safiri *et al.*, 2021; Karami *et al.*, 2022; Ali *et al.*, 2020; Owais *et al.*, 2021; Bathla and Arora, 2022). About one-third of the world's population is affected, with at least one-third of the estimated 2 billion people experiencing iron-deficiency anemia (IDA). It is the most common type of anemia and is estimated to be responsible for half of all anemia cases (Noreen *et al.*, 2020; Alkdede *et al.*, 2020). Anemia is a widespread public health issue impacting individuals across all demographics globally, leading to severe health complications such as postpartum hemorrhage, maternal deaths, perinatal and neonatal mortality, low birth weight, poor cognitive development in children, and an increased risk of chronic diseases in adults (Bogdanova *et al.*, 2020; Thiagarajan *et al.*, 2021; Addo *et al.*, 2021; Balarajan *et al.*, 2013).

Anemia is particularly prevalent in low- and middle-income countries (LMICs), where it is the most common micronutrient and nutritional deficiency, primarily caused by insufficient intake of iron, protein, and folic acid. Among all age groups, women of reproductive age are especially vulnerable to anemia due to the loss of iron through menstruation and increased need for iron during pregnancy. Its adverse effect on maternal and child health is widespread, affecting women from poor and developing countries (Sharif *et al.*, 2023; Chaudhary *et al.*, 2022; Ghosh *et al.*, 2020). Nearly 50% of the population in developing regions suffers from nutritional anemia, with a higher prevalence among females (64.6%) compared to males (33.3%) (Bathla and Arora, 2022; Ismail *et al.*, 2017; Bhadra and Deb, 2020). Globally, pregnant women have the highest prevalence of anemia at 41.8%, followed by nonpregnant, non-lactating women at 35.1% (Osborn *et al.*, 2021).

In India, anemia is a significant health problem, particularly affecting women of reproductive age (15-49 years). There are approximately 600 million women in this age group worldwide, representing 16% of the global population. In India, about 54.5% of women have experienced anemia over the past 25 years. Adolescent girls and women of childbearing age, especially during pregnancy, are most affected, with an estimated 270 to 338 million women suffering from anemia (Sharif *et al.*, 2023; Sinha *et al.*, 2021). The National Family Health Survey-4 (2015-2016) reported a prevalence of 44.6% among women aged 15-49 years, which increased to 50.3% according to the National Family Health Survey-5 (2019-2020) (Osborn *et al.*, 2021; Mog and Ghosh, 2021; Sappani *et al.*, 2023; Maji *et al.*, 2023; Let *et al.*, 2024; Chakrabarty *et al.*, 2023). Anemia is often regarded as a "female disease" and poses significant health concerns for women in India, especially among adolescent girls (Mahanta *et al.*, 2015).

The northeastern region of India, comprising eight states—Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura—has distinct geographical and socio-economic characteristics, different from the rest of India. Anemia is a significant health concern in this region, particularly among women, and the high prevalence of iron deficiency anemia (IDA) is believed to be linked to distinct dietary patterns that differ considerably from the other parts of the country (Bathla and Arora, 2022; Chakrabarty *et al.*, 2023; Kumar *et al.*, 2022; Alarcon Basurto, 2020; Hess *et al.*, 2023). Approximately 54.9% of women in the northeastern states are anemic, with nearly two out of every five women suffering from mild anemia (Ghosh *et al.*, 2021; Ragunanthanan *et al.*, 2023; Omkar, 2020; Dasgupta *et al.*, 2016).

Despite the recognition of anemia as a significant health issue, efforts to address the problem have been inadequate. While numerous studies have explored the prevalence and contributing factors of nutritional anemia, there has been insufficient investigation into the specific causes of its high occurrence among the reproductive population in Northeast India. Additionally, the geographical variability of anemia prevalence in this region has not been thoroughly examined. This study aims to analyze the spatial distribution and factors influencing anemia among women of reproductive age in the districts of the northeastern states of India.

Materials and methods

Data source

This study has utilised data from the fifth round of the National Family Health Survey (NFHS-5) conducted in 2019–20 in India. This nationally representative cross-sectional survey was overseen by the Ministry of Health *and* Family Welfare, Government of India, in collaboration with the International Institute for Population Sciences (IIPS) serving as the central agency (International Institute for Population Sciences, 2018), covered all 28 states, 8 union territories and 707 districts. The survey aimed to offer both national and sub-national level estimates pertaining to population, health, nutrition, and various demographic indicators. As the data used in this study is openly accessible through a public repository, ethical clearance was not required for its utilization. After excluding women with missing data, the final analytical sample comprises of 100,435 aged 15-49 across the northeastern 104 districts and 8 states.

Outcome variable

For the study, a dichotomous variable was formed by categorizing all women of reproductive age group (15-49) into 'anemic' and non-anemic' with code '1' and '0' respectively.

Predictor variable

After extensive literature review, explanatory variables were identified, including socio-economic factors like women's educational level, religion, wealth status, and the number of children ever born. Marital status is considered under demographic factors. Dietary factors involved the frequency of intake of various nutritional foods such as vegetables, fried foods, as well as the consumption of milk or curd.

Statistical analysis

Specification of regression models

Descriptive statistics provide insights into the distribution and variability of the independent variables within the dataset. Grasping these statistics is essential for evaluating the distribution and variability of these independent variables. Spatial patterns and clustering of anemic women were analyzed using Moran's I, the univariate local indicator of spatial association (LISA), and a LISA cluster map. To quantify spatial proximity between each possible pair of observational entities in the dataset, a spatial weight matrix (w) of order 1 was created using the Queen's contiguity method (Getis and Ord, 1992; Anselin and Rey, 2010). Global measures, such as the Global Moran's I index, were employed to estimate spatial autocorrelation across the entire study area. This index is commonly used to assess spatial patterns and clustering within data and to determine their statistical significance. This assessment is typically carried out using the following formula:

(i) $$\text{Global Moran's I} = \frac{n}{\Sigma_i \Sigma_j \omega_{ij}} \frac{\Sigma_i \Sigma_j \omega_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\Sigma_i (x_i - \bar{x})^2}$$

Here in (i), n represents the number of observations, xi and xjdenote the values of the variable being studied at locations i and j, respectively, $\bar{x}$ is the mean of the variable, and ω_{ij} represents spatial weights between locations i and j.

Moran's, I range from -1 (perfect dispersion) to 1 (perfect clustering). A positive Moran's I value signifies positive spatial autocorrelation, indicating that similar attribute values are clustered together on the map, while negative values imply dissimilar values clustering together. A value close to zero suggests no significant spatial autocorrelation, indicating a random or uniform distribution of values across the study area (Getis and Ord, 1992; Anselin, 1995).

Global measures like Moran's I indicates clustering but don't specify cluster locations on a map. For this, Local Indicators of Spatial Association (LISA), such as "Local" Moran's I, are used, known as cluster and outlier analysis (Getis and Ord, 1992). Univariate LISA measures the correlation of neighbourhood values around a specific spatial location. It determines the extent of spatial randomness and clustering present in the data (Cliff and Ord, 1970; Anselin, 1995; Clark and Evans, 1954).

Four types of spatial autocorrelation were generated:
1. Hot spots: Areas with high values surrounded by high-value neighbours (High-High).
2. Cold spots: Areas with low values surrounded by low-value neighbours (Low-Low).
3. Spatial outliers: High-prevalence surrounded by low-prevalence neighbours (High-Low).
4. Spatial outliers: Low-prevalence surrounded by high-prevalence neighbours (Low-High).

Global spatial regression model

The Spatial Lag Model (SLM) is an extension of Ordinary Least Squares (OLS) regression. by including a spatially lagged variable (spatial lag term) alongside existing independent variables. This spatially lagged variable is computed from the dependent variable, representing the weighted sum of neighbouring values for each location. By incorporating this term, the SLM accounts for spatial autocorrelation (SA) in the dependent variable, providing a more accurate analysis. The spatially lagged variable is calculated from the dependent variable and reflects the weighted sum of neighbouring values of the dependent variable for each location within its neighbourhood. This model controls for SA in the dependent variable. This model can be expressed as:

(ii) $y = pWy + Xb + u$

Where in (ii), y is the dependent variable (anaemia prevalence), Wy is the spatial lag term (spatially lagged anaemia prevalence), p is the coefficient associated with spatially lagged variable, X is the independent variable(s), b is the coefficient associated with X, and u represents the error term in the model.

Spatial regression modelling framework and model calibration

Our suggested spatial modelling framework for examining the prevalence of anemic women in Northeast India, integrating spatial autocorrelation and spatial heterogeneity, is depicted in Fig. 1. This framework was developed with the help of previous literature (Anselin, 2005; Comber *et al.*, 2023; Chien *et al.*, 2020).

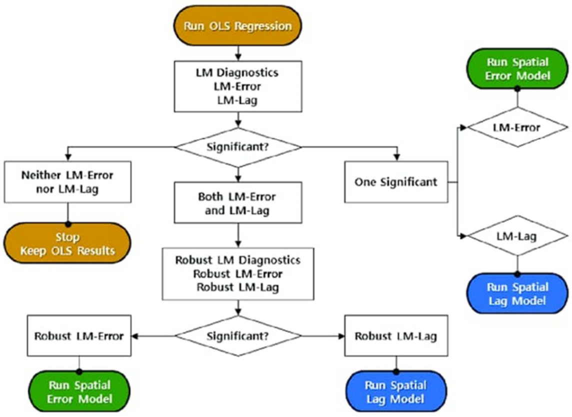

Fig. 1. Spatial regression modelling framework

For all spatial regression analysis, the OLS model is the appropriate starting point. In the first step of this study, we developed an OLS model. The OLS model is a global non-spatial model. This model is not appropriate if the residuals of the model are spatially correlated (spatial autocorrelation) or spatially heterogeneous (spatial non-stationarity) (Huang *et al.*, 2018).

After developing the OLS model, we used Moran's I tool to assess the presence of spatial autocorrelation in the residuals. To examine geographical clustering of anemic women outcomes, univariate local Moran's I and Local indicator of Spatial Association (LISA) cluster maps were created. We created the spatial weight matrix (w) of order one using the Queen's contiguity method (neighbours sharing a common boundary of non-zero length) to quantify the spatial proximity between each possible pair of observational entities (Anselin, 1995; Clark and Evans, 1954). A significant Moran's I value indicates the presence of spatial autocorrelation, which emphasizes the need to develop global spatial regression models such as SLM and SEM (Cliff and Ord, 1970; Comber *et al.*, 2023).

To determine the preferable modelling approach between SLM and SEM, it's crucial to examine the outcomes of the Lagrange Multiplier (LM)-lag and Lagrange Multiplier (LM)-error tests (Anselin, 2005). If LM-lag yields a significant result while LM-error does not, developing the SLM model would be appropriate. Conversely, if LM-error is significant while LM-lag is not, the SEM model should be pursued. In cases where both LM-lag and LM-error are significant, further examination of Robust LM-lag and Robust LM-error results are necessary. If Robust LM-error is significant while Robust LM-lag is not, the SEM model is preferable. Conversely, if Robust LM-lag is significant while Robust LM-error is not, the SLM model is favourable. In instances where both Robust LM-lag and Robust LM-error are significant, the model with the lower p-value should be selected (Anselin, 2005).

To identify potential regional factors associated with anemia in women, a spatial regression analysis was conducted. This analysis included the use of both an Ordinary Least Squares (OLS) model and a Spatial Error Model (SEM).

Software used

Non-spatial analysis was conducted using the trial version of Statistical Package for Social Sciences (SPSS) software v. 26; (IBM Inc, Chicago, IL), while spatial analysis was carried out using trial versions of ArcGIS 10.7 and GeoDa 1.20.

Results

Descriptive statistics

Descriptive statistics of all the dependent and independent variables are presented in this Table 1. On average, 49.2% of women is affected by anemia, indicating a notable public health issue. The wide range (14.9% to 81.5%) and high standard deviation (16.9%) suggest significant variability in anemia prevalence among different districts of Northeast India, emphasizing the necessity for focused interventions. About 36.9% of the population identifies as Hindu on average, but this varies greatly (SD of 34.1%), with some districts having no Hindus and others having almost the entire population identifying as Hindu (0% to 99%). This highlights the diverse religious composition and regional differences within the districts. Economic disparities are evident, with an average of 27.5% of women classified as the poorest, exhibiting notable variability (SD of 15.0%) across different districts of Northeast India. Similarly, educational attainment varies widely, with an average of 15.5% of the population having no formal education and substantial differences (SD of 8.2%) ranging from 0.2% to 40.9%, reflecting differing levels of access to education across districts. Approximately 4.9% of women are widowed, divorced, or separated, with relatively low variability (SD of 2.3%) but noticeable differences (1.7% to 15.1%) among districts. Family size also varies, with an average of 14.0% of women having four or more children, indicating moderate variability (SD of 7.8%) across different districts within the region. Dietary habits show significant diversity, with 11.1% of the population never consuming milk on average and 56.6% consuming vegetables daily, reflecting notable variability influenced by cultural, economic, or regional factors. Approximately 31.4% of the population consumes fried food daily, with wide variability (SD of 22.3%, range: 0.5% to 92.9%) possibly reflecting cultural preferences and socioeconomic factors.

Overall, the high prevalence of anemia and significant dietary disparities underscore potential nutritional and health challenges among the reproductive women of Northeast India. Meanwhile, the diversity in socio-economic factors and demographic characteristics underscores the intricate social dynamics and regional variations evident within the population.

SLM

This table presents the parameter estimation results for the "Anemic women" variable, indicating the coefficients' values and their associated p-values (Table 2).

Table 1. Descriptive statistics of variables, NE India, 2019-21

Variables	Mean (%)	SD (%)	Minimum (%)	Maximum (%)
Anemia prevalence (%)	49.2	16.9	14.9	81.5
Hindu	36.9	34.1	0	99
Poorest	27.5	15.0	1	54.5
No education	15.5	8.2	0.2	40.9
Widowed/Divorce/Separated	4.9	2.3	1.7	15.1
Four or more children	14.0	7.8	2.6	42.4
Never milk	11.1	9.4	0.3	44.9
Daily vegetable	56.6	16.0	13.1	86.9
Daily fried food	31.4	22.3	0.5	92.9

SD: Standard deviation

Table 2. Estimation of parameters for women with anemia

Variable	Anemic women	
Fixed part	Coefficient	P-value
Constant	21.5102	0.00052*
Hindu	0.181699	0.00000*
Poorest	0.235407	0.00733*
No education	0.261896	0.01655*
Widowed/Divorce/Separated	1.15222	0.03090*
Four or more children	-0.262124	0.20801
Never milk	-0.145676	0.18817
Daily vegetable	-0.16824	0.00741*
Daily fried food	0.11709	0.00855*
Lag coefficient (Rho)	0.322109	0.00003
R^2	0.789806	
Log likelihood	-361.204	
AIC	742.408	

*Significant at 5% level, AIC: Akaike Information Criterion

Fixed part

Constant

The intercept term has a coefficient of 21.5102, with a highly significant p-value of 0.00052. This indicates that when all other variables are zero, the expected value of anemia among women is 21.5102 (coefficient).

The variables "Hindu," "Poorest socioeconomic status," "No education," and "Widowed/Divorced/Separated" exhibit positive coefficients with extremely low p-values, indicating their significant association with anemia among women. Being Hindu is associated with a higher likelihood of anemia prevalence, supported by a coefficient of 0.181699 and an extremely low p-value of 0.00000. Similarly, belonging to the poorest socioeconomic category poses an elevated risk of anemia, with a coefficient of 0.235407 and a significant p-value of 0.00733. Lack of education is linked to a higher likelihood of anemia, as indicated by a coefficient of 0.261896 and a significant p-value of 0.01655. Additionally, being widowed, divorced, or separated is strongly associated with anemia prevalence, with a coefficient of 1.15222 and a significant p-value of 0.03090. Overall, factors such as religious affiliation, socioeconomic status, educational attainment, and marital status serve as significant predictors of increased anemia prevalence among women.

The variables "Four or more children" and "Never milk" have coefficients with p-values higher than the typical significance level of 0.05, suggesting they may not be significantly associated with anemia among women. The coefficient for "Four or more children" (-0.262124) implies a possible negative relationship with anemia prevalence, although it does not reach statistical significance with a p-value of 0.20801. Similarly, the coefficient for "Never milk" (-0.145676) suggests a possible negative association with anemia prevalence, but it also lacks statistical significance with a p-value of 0.18817. In summary, while there's a hint of lower anemia prevalence among women with four or more children and those who never consume milk, these findings do not reach statistical significance based on the provided p-values.

The coefficient for daily vegetable consumption (-0.16824) suggests that higher intake of vegetables daily is correlated with a reduced prevalence of anemia among women.

Conversely, the coefficient for daily consumption of fried foods (0.11709) indicates that increased intake of fried foods is associated with a higher prevalence of anemia in women.

Both coefficients have p-values below 0.05, indicating statistical significance. In essence, daily vegetable consumption and daily fried food consumption significantly influence the prevalence of anemia in women, with opposing effects.

Overall, the results suggest that variables such as being Hindu, belonging to the poorest wealth quintile, having no education, being widowed/divorced/separated, consuming daily vegetables, and avoiding daily fried food are significantly associated with anemia among women. However, the number of children and milk consumption may not have a statistically significant relationship with anemia. Additionally, the model includes a spatial lag term (Rho) indicating positive spatial autocorrelation in the occurrence of anemia among women.

Lag coefficient (Rho)

The lag coefficient (Rho) is 0.322109, indicating a positive spatial autocorrelation in the occurrence of anemia among women. A positive coefficient implies that an increase in the lagged variable is associated with an increase in anemia prevalence. The low p-value (0.00003) suggests that this relationship is statistically significant, meaning that changes in the lagged variable are related to changes in anemia prevalence, as per the model.

Model fit

R-squared (R2): The model explains approximately 78.98% of the variance in anemia among women, indicating a good fit to the data.

Log likelihood: The log likelihood value is -361.204, which measures the goodness of fit of the model. Lower values indicate better fit.

Akaike Information Criterion (AIC): The AIC value is 742.408, which is commonly used for model selection. Lower AIC values suggest better models.

Table 3. Spatial dependency tests using the lagrange multiplier diagnostic

Test	Anemic women	
	Value	Probability
Moran's I (error)	3.6741	0.00024*
Lagrange multiplier (lag)	17.1795	0.00003*
Robust LM (lag)	9.2888	0.00231*
Lagrange multiplier (error)	7.9497	0.00481*
Robust LM (error)	0.0591	0.80800
Lagrange multiplier (SARMA)	17.2386	0.00018*

LM diagnostic result

The spatial dependency tests, conducted using the Lagrange Multiplier Diagnostic, provide insights into spatial autocorrelation in the dataset (Table 3). The Moran's I test revealed significant spatial autocorrelation in anemia prevalence, with a test statistic of 3.6741 and a p-value of 0.00024, indicating clustering tendencies. Both the Lagrange Multiplier (lag) and Robust LM (lag) tests confirmed spatial autocorrelation in the dependent variable, with high test statistics (17.1795 and 9.2888) and very low p-values (0.00003 and 0.00231). The Lagrange Multiplier (error) test also indicated significant spatial autocorrelation in the residuals (test statistic of 7.9497, p-value 0.00481), though the Robust LM (error) test did not. The Lagrange Multiplier (SARMA) test further confirmed spatial autocorrelation in the residuals, with a test statistic of 17.2386 and a p-value of 0.00018. These results emphasize the need to consider spatial effects when analyzing anemia prevalence among women.

Spatial clustering

District-level quantile maps were generated to understand the spatial variation of the selected anemic women outcomes in the Northeast India.

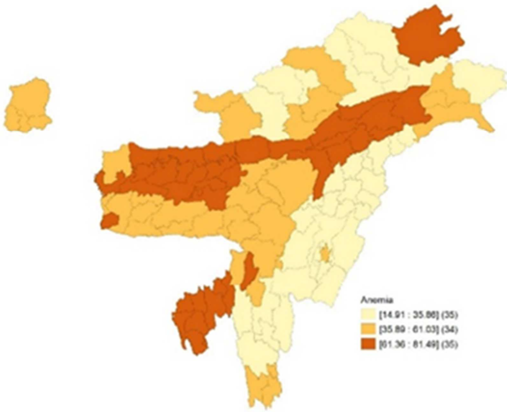

Fig. 2. Spatial distribution of prevelance of anemic women in the districts of Northeast India, NFHS-5, (2019-21)

Anemia prevalence among women varies greatly across districts. In 35 districts, prevalence ranges from 14.91% to 35.86%, including areas of Nagaland, Manipur, Arunachal Pradesh, and Mizoram. Another 34 districts have prevalence rates between 35.89% and 61.03%, with higher rates in Sikkim, Meghalaya, parts of Arunachal Pradesh and Assam, and some regions of Mizoram and Manipur. The most critical situation is found in 35 districts, where anemia prevalence is between 61.36% and 81.49%, affecting the majority of women, particularly in Assam, Tripura, parts of Arunachal Pradesh, and parts of Meghalaya (Fig. 2).

Moran's, I value of 0.645 indicates strong positive spatial autocorrelation, meaning that similar values—high or low—tend to cluster together. This significant autocorrelation necessitates further spatial analysis using techniques like local indicators of spatial association (LISA) to identify specific clusters and outliers (Fig. 3).

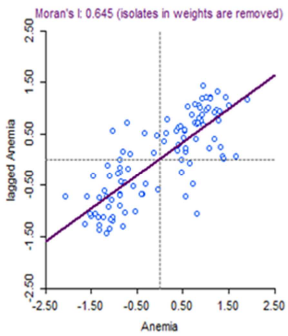

Fig. 3. Illustrates the Moran's I value, indicating the degree to which similar values are clustered together or dispersed across a geographic area

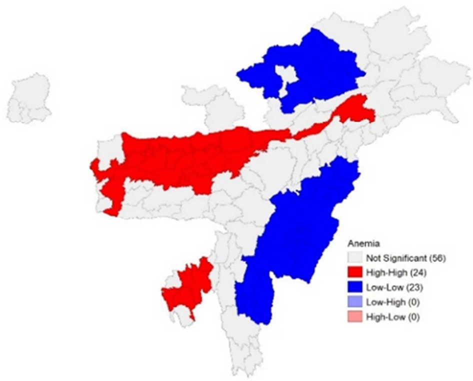

Fig. 4. LISA cluster map showing spatial correlation of anemic women in the districts of Northeast India, NFHS-5, (2019-21):light grey: not significant (56); dark red: high-high (24); dark blue: low-low (23); light blue: low-high (0); light red: high-low (0)

The univariate LISA cluster map highlighting significant local Moran's I value, illustrates the spatial autocorrelation of anemia prevalence among women aged 15-49 in Northeast India's districts. The map uses red for "hot spots" (areas with high prevalence), blue for "cold spots" (low prevalence), and light red/light blue for spatial outliers (Anselin, 1995; Clark and Evans, 1954). Anemia prevalence is statistically significant in 47 out of 104 districts. In 24 districts, mainly in Assam, Tripura, and parts of Meghalaya, high anemia prevalence clusters together

("High-High"). Conversely, 23 districts in Manipur, Nagaland, Arunachal Pradesh, and parts of Mizoram exhibit low prevalence clusters ("Low-Low"). No districts fall into the "Low-High" or "High-Low" categories, indicating areas with mixed prevalence. These findings underscore the importance of addressing spatial trends in anemia prevalence across the region (Fig. 4).

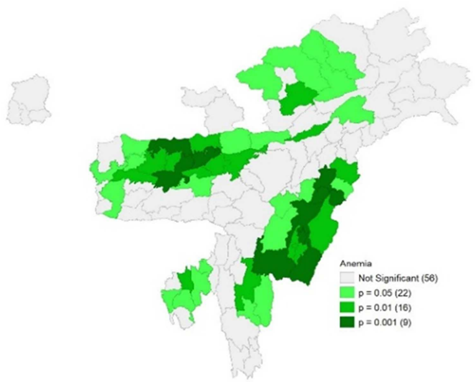

Fig. 5. LISA significance map showing spatial correlation of anemic women in the districts of India, NFHS-5, (2019-21): light grey: not significant (56); light green: p = 0.05 (22); green: p = 0.01 (16); dark green: p = 0.001 (9)

The LISA significance map classifies anemia prevalence among women in Northeast India's districts into varying levels of statistical significance. In 56 districts, the prevalence is not statistically significant, suggesting it is within an acceptable range. In 22 districts, the prevalence reaches a significance level of 0.05, indicating moderate concern. These districts are located in parts of Arunachal Pradesh, Tripura, Mizoram, Manipur, Nagaland, Assam, and Meghalaya. In 16 districts, the significance level rises to 0.01, indicating higher prevalence, while 9 districts exhibit a significance level of 0.001, representing severe anemia in parts of Manipur, Nagaland, and Assam. These findings reveal substantial variations in anemia prevalence across districts, with some areas showing more pronounced disparities (Fig. 5).

Discussion

The findings indicate a significant persistence of anemia among women of reproductive age in the districts of Northeast India. Variations in anemia prevalence rates across these districts underscore the pressing need for targeted interventions to tackle this public health issue.

Based on the National Family Health Survey-5 (NFHS5) data, certain districts primarily located in Assam, parts of Tripura, and minor sections of Meghalaya, identified as "High-High" areas, exhibit notably high rates of anemia. These districts are spatially clustered, with neighbouring districts also showing similarly elevated rates. Previous research utilizing NFHS 4 data has revealed spatial correlations between anemia and various factors across India, including the northeast region. Moran's I values indicate significant autocorrelation, highlighting hotspots in both Eastern and Western India. Northeastern states such as Manipur, Meghalaya, Mizoram, Nagaland, and Sikkim demonstrate varying rates of anemia, with notable disparities observed, particularly in Manipur and Nagaland (Sharma *et al.*, 2018).

Likewise, a prior investigation, anemia prevalence among women of reproductive age was examined. Across both NFHS4 and NFHS5 rounds, the Manipur district consistently displayed the highest rates of anemia. Conversely, Assam witnessed the most significant surge in anemia prevalence (Let *et al.*, 2024). Over time, Assam has consistently exhibited the highest prevalence of anemia among women of reproductive age.

Prior studies employing spatial analysis using Moran's I statistic have revealed a correlation between anemia among women and various factors, including wealth, geographic region, demographic characteristics, nutritional status, parity, and dietary habits. Additionally, factors such as education level and wealth status at the district level have been associated with anemia among women. Furthermore, anemia among women of reproductive age has been linked to having three to four children and belonging to the economically disadvantaged demographic (Let *et al.*, 2024; Sharma *et al.*, 2018).

In the Northeast region, Hindu women are at the highest risk of anemia, consistent with earlier research conducted in India. Areas with a larger proportion of Hindu women tend to have a higher prevalence of anemia among females (Bharati *et al.*, 2008; Sharma *et al.*, 2018). Women of reproductive age in the northeast region face an increased risk of anemia when associated with a lower wealth index or categorized as the poorest. Conversely, as the wealth index increases, the prevalence of anemia tends to decrease. Thus, the level of wealth or standard of living emerges as a critical factor in reducing the prevalence of anemia among women in India. Low socio-economic status and limited access to nutritious food are associated with a heightened risk of anemia (Sappani *et al.*, 2023; Bharati *et al.*, 2008; Sharma *et al.*, 2018).

Additionally, studies indicate that higher levels of education are linked to a reduced risk of anemia among women, with illiteracy being correlated with a higher prevalence of the condition. This connection may be attributed to education's role in promoting awareness about proper nutrition and dietary habits, which are crucial for preventing anemia (Gogoi and Prusty, 2013; Sappani *et al.*, 2023; Choudhury and Kumar, 2022). In essence, the lack of education increases the chances of anemia, while having education acts as a shield against it. Educated women are notably less likely to experience anemia compared to those with less education (Rammohan *et al.*, 2012; Sappani *et al.*, 2023; Sharma *et al.*, 2018).

Moreover, the status of being widowed, divorced, or separated significantly increases the risk of anemia among women, as observed notably in northeast India and in studies encompassing India's Empowered Action Group (EAG) states (Barman, 2024). In our study, the data concerning having four or more children did not reveal a significant outcome. However, earlier research indicated a higher prevalence of anemia among women who had given birth to two or more children. This pattern might be associated with the continuous cycle of childbirth, potentially heightening the risk of iron deficiency anemia (IDA) among women (Sharif *et al.*, 2023).

Furthermore, our study did not find any statistically significant association between consumption of milk and prevalence of anemia. However, previous research conducted in the Northeast region suggests that women who regularly consume milk or curd are less likely to experience anemia, consistent with findings from studies involving Indian women. This is attributed to milk's richness in calcium, which facilitates the absorption of iron from other dietary sources, potentially aiding in the prevention of anemia (Rammohan *et al.*, 2012).

Regarding dietary habits, daily consumption of vegetables is associated with a reduced occurrence of anemia among women. Iron-rich foods, such as pulses and green leafy vegetables, are essential for women to include in their diets, as low iron intake is recognized as a risk factor for anemia. Vegetarian diets are associated with a higher risk of iron deficiency anemia. Several studies have indicated that vegetarians tend to have lower hemoglobin levels compared to non-vegetarians, potentially contributing to the development of anemia (Rammohan *et al.*, 2012; Mog *et al.*, 2023; Elorinne *et al.*, 2016; Li *et al.*, 2000).

Lastly, our study found that daily consumption of fried foods is linked to a greater occurrence of anemia among women. However, the consumption of fried foods by women in the northeast region did not show a significant correlation with anemia prevalence, which was another conflicting finding. This aligns with the results of a prior study conducted in Northeast India, which reported similar outcomes (Mog *et al.*, 2023).

Conclusion

This analysis highlights the key predictors of anemia among women in Northeast India, including socio-demographic factors and dietary habits. Women identifying as Hindu, those in the poorest socioeconomic group, with no education, and widowed, divorced, or separated individuals are more likely to be anemic. Daily vegetable consumption reduces anemia risk, while frequent fried food consumption increases it. However, milk intake and the number of children were not statistically significant factors. Spatial analysis reveals significant clustering, with "High-High" areas in districts of Assam, Tripura, parts of Arunachal Pradesh, and Meghalaya showing high anemia prevalence. Conversely, lower rates are observed in districts of Manipur, parts of Nagaland, Arunachal Pradesh, and Mizoram. These findings demonstrate the variable nature of anemia prevalence and the urgent need for targeted public health interventions to address disparities, particularly in high-risk regions, and to improve overall health outcomes for women in the region.

Acknowledgment

We would like to thank the measure DHS program for providing the dataset.

References

Addo OY, Emma XY, Williams AM, Young MF, Sharma AJ, Mei Z, Kassebaum NJ, Jefferds ME, Suchdev PS. 2021. Evaluation of hemoglobin cutoff levels to define anemia among healthy individuals. JAMA Network Open **4**(8), e2119123.

Alarcon Basurto SG. 2020. Iron-deficiency anaemia (IDA): Socio-cultural misconceptions intersect the health of vulnerable populations in developing countries. University Honors Theses, 840.

Ali SA, Khan US, Feroz AS. 2020. Prevalence and determinants of anemia among women of reproductive age in developing countries. Journal of the College of Physicians and Surgeons Pakistan **30**(2), 177–186.

Alkdede MJ, Binsaeed AA, Alameer WH, Alotaibi AA, Alosaimi AS, Alsugair MM, Alharbi RA, Alkhulaif MA, Alanazi RS, Ghannam SA, Alshehri SA. 2020. Iron deficiency anemia, diagnosis, and treatment in primary health care centres. Archives of Pharmacy Practice **11**(3-2020), 122–126.

Anselin L, Rey SJ. 2010. Perspectives on spatial data analysis. Springer Berlin Heidelberg.

Anselin L. 1995. Local indicators of spatial association—LISA. Geographical Analysis **27**(2), 93–115.

Anselin L. 2005. Exploring spatial data with GeoDaTM: A workbook. Accessed 10 August 2021. Available from: https://www.geos.ed.ac.uk/~gisteac/fspat/geodaworkbook.pdf.

Balarajan YS, Fawzi WW, Subramanian SV. 2013. Changing patterns of social inequalities in anaemia among women in India: Cross-sectional study using nationally representative data. BMJ Open **3**(3), e002233.

Barman M. 2024. Anaemia prevalence and socio-demographic factors among women of reproductive age (WRA): A geospatial analysis of empowered action group (EAG) states in India. Spatial and Spatio-temporal Epidemiology **49**, 100644.

Bathla S, Arora S. 2022. Prevalence and approaches to manage iron deficiency anemia (IDA). Critical Reviews in Food Science and Nutrition **62**(32), 8815–8828.

Bhadra P, Deb A. 2020. A review on nutritional anemia. Indian Journal of Natural Sciences **10**(59), 18466–18474.

Bharati P, Som S, Chakrabarty S, Bharati S, Pal M. 2008. Prevalence of anemia and its determinants among nonpregnant and pregnant women in India. Asia Pacific Journal of Public Health **20**(4), 347–359.

Bogdanova A, Kaestner L, Simionato G, Wickrema A, Makhro A. 2020. Heterogeneity of red blood cells: Causes and consequences. Frontiers in Physiology **11**, 392.

Chakrabarty M, Singh A, Singh S, Chowdhury S. 2023. Is the burden of anaemia among Indian adolescent women increasing? Evidence from Indian Demographic and Health Surveys (2015–21). PLOS Global Public Health **3**(9), e0002117.

Chaudhary E, Dey S, Ghosh S, Sharma S, Singh N, Agarwal S, Tibrewal K, Venkataraman C, Kurpad AV, Cohen AJ, Wang S. 2022. Reducing the burden of anaemia in Indian women of reproductive age with clean-air targets. Nature Sustainability **5**(11), 939–946.

Chien YM, Carver S, Comber A. 2020. Using geographically weighted models to explore how crowdsourced landscape perceptions relate to landscape physical characteristics. Landscape and Urban Planning **203**, 103904.

Choudhury S, Kumar S. 2022. Gender discrimination and marginalization of women in North-East India. Journal of Public Affairs **22**(4), e2625.

Clark PJ, Evans FC. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology **35**(4), 445–453.

Cliff AD, Ord K. 1970. Spatial autocorrelation: A review of existing and new measures with applications. Economic Geography **46**(sup1), 269–292.

Comber A, Brunsdon C, Charlton M, Dong G, Harris R, Lu B, Lü Y, Murakami D, Nakaya T, Wang Y, Harris P. 2023. A route map for successful applications of geographically weighted regression. Geographical Analysis **55**(1), 155–178.

Dasgupta A, Sarkar K, Chowdhury R, Ray A, Shahbabu B. 2016. Anemia and its determinants among women of reproductive age of a slum in Kolkata: A focus group discussion among health workers in a slum of Kolkata. Journal of Family Medicine and Primary Care **5**(2), 276–280.

Elorinne AL, Alfthan G, Erlund I, Kivimäki H, Paju A, Salminen I, Turpeinen U, Voutilainen S, Laakso J. 2016. Food and nutrient intake and nutritional status of Finnish vegans and non-vegetarians. PLoS One **11**(2), e0148235.

Getis A, Ord JK. 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis **24**(3), 189–206.

Ghosh D, Panja A, Saha D, Banerjee U, Datta AK, Basu A. 2021. Drug repurposing: Hydroxyurea therapy improves the transfusion-free interval in HbE/Beta-thalassemia—Major patients with the Xmn I polymorphism. Genetic Testing and Molecular Biomarkers **25**(8), 563–570.

Ghosh P, Dasgupta A, Paul B, Roy S, Biswas A, Yadav A. 2020. A cross-sectional study on prevalence and determinants of anemia among women of reproductive age in a rural community of West Bengal. Journal of Family Medicine and Primary Care **9**(11), 5547–5553.

Gogoi M, Prusty RK. 2013. Maternal anaemia, pregnancy complications and birth outcome: Evidence from north-east India. Journal of North East India Studies **3**(1), 74–85.

Hess SY, Owais A, Jefferds ME, Young MF, Cahill A, Rogers LM. 2023. Accelerating action to reduce anemia: Review of causes and risk factors and related data needs. Annals of the New York Academy of Sciences **1523**(1), 11–23.

Huang Y, Wang X, Patton D. 2018. Examining spatial relationships between crashes and the built environment: A geographically weighted regression approach. Journal of Transport Geography **69**, 221–233.

International Institute for Population Sciences. 2018. The DHS Program - India: Standard DHS, 2015-16. Available at: https://dhsprogram.com/data/dataset/India_Standard-DHS_2015.cfm?flag=1. Accessed: 20th November 2018.

Ismail IM, Kahkashan A, Antony A, Sobhith V. 2017. Role of socio-demographic and cultural factors on anemia in a tribal population of North Kerala, India. International Journal of Community Medicine and Public Health **3**(5), 1183–1188.

Jackson Chornenki NL, James TE, Barty R, Liu Y, Rochwerg B, Heddle NM, Siegal DM. 2020. Blood loss from laboratory testing, anemia, and red blood cell transfusion in the intensive care unit: A retrospective study. Transfusion **60**(2), 256–261.

Karami M, Chaleshgar M, Salari N, Akbari H, Mohammadi M. 2022. Global prevalence of anemia in pregnant women: A comprehensive systematic review and meta-analysis. Maternal and Child Health Journal **26**(7), 1473–1487.

Kumar SB, Arnipalli SR, Mehta P, Carrau S, Ziouzenkova O. 2022. Iron deficiency anemia: Efficacy and limitations of nutritional and comprehensive mitigation strategies. Nutrients **14**(14), 2976.

Let S, Tiwari S, Singh A, Chakrabarty M. 2024. Prevalence and determinants of anaemia among women of reproductive age in aspirational districts of India: An analysis of NFHS 4 and NFHS 5 data. BMC Public Health **24**(1), 437.

Li D, Sinclair AJ, Mann NJ, Turner A, Ball MJ. 2000. Selected micronutrient intake and status in men with differing meat intakes, vegetarians and vegans. Asia Pacific Journal of Clinical Nutrition **9**(1), 18–23.

Mahanta TG, Mahanta BN, Gogoi P, Dixit P, Joshi V, Ghosh S. 2015. Prevalence and determinants of anaemia and effect of different interventions amongst tea tribe adolescent girls living in Dibrugarh district of Assam. Clinical Epidemiology and Global Health **3**(2), 85–93.

Maji I, Randhawa JK, Bakshi D, Gautam D, Mishra SS. 2023. Status of anaemia amongst women in India: Trend analysis of NFHS data. Indian Journal of Community Health **35**(3), 354–358.

Mantadakis E, Chatzimichael E, Zikidou P. 2020. Iron deficiency anemia in children residing in high and low-income countries: risk factors, prevention, diagnosis and therapy. Mediterranean Journal of Hematology and Infectious Diseases **12**(1).

Mog M, Ghosh K. 2021. Prevalence of anaemia among women of reproductive age (15–49): A spatial-temporal comprehensive study of Maharashtra districts. Clinical Epidemiology and Global Health **11**, 100712.

Mog M, Neogi D, Bharadwaz MP, Panda BK, Sil A. 2023. Prevalence and factors associated with anaemia in married women of reproductive age group: Evidence from North East India. Journal of Biosocial Science **55**(3), 425–437.

Noreen S, Bashir S, Bano S, Fatima T, Sani A, Imran S, Saeed Z, Naseer A, Ijaaz R, Riaz K, Khan M. 2020. Anemia and its consequences on the human body: A comprehensive overview. NUST Journal of Natural Sciences **5**(2).

Omkar AC. Changing pattern of tribal livelihoods: A study of Adilabad district (A case study of Narnoor and Jainoor Mandals).

Osborn AJ, Muhammad GM, Ravishankar SL, Mathew AC. 2021. Prevalence and correlates of anemia among women in the reproductive age (15–49 years) in a rural area of Tamil Nadu: An exploratory study. Journal of Education and Health Promotion **10**(1).

Owais A, Merritt C, Lee C, Bhutta ZA. 2021. Anemia among women of reproductive age: An overview of global burden, trends, determinants, and drivers of progress in low-and middle-income countries. Nutrients **13**(8), 2745.

Ragunanthanan S, Vaiyakkani G, Ramachandran A, Rajendran K, Krishnasamy N, Panneerselvam P, Pamarthi J, Rajaram M. 2023. Analysis of blood lead level and its clinical significance among occupationally exposed painters in Chennai-based population: A cross-sectional study. Journal of Trace Elements in Medicine and Biology **79**, 127257.

Rammohan A, Awofeso N, Robitaille MC. 2012. Addressing female iron-deficiency anaemia in India: Is vegetarianism the major obstacle? International Scholarly Research Notices **2012**(1), 765476.

Safiri S, Kolahi AA, Noori M, Nejadghaderi SA, Karamzad N, Bragazzi NL, Sullman MJ, Abdollahi M, Collins GS, Kaufman JS, Grieger JA. 2021. Burden of anemia and its underlying causes in 204 countries and territories, 1990–2019: Results from the Global Burden of Disease Study 2019. Journal of Hematology and Oncology **14**, 1–6.

Sappani M, Mani T, Asirvatham ES, Joy M, Babu M, Jeyaseelan L. 2023. Trends in prevalence and determinants of severe and moderate anaemia among women of reproductive age during the last 15 years in India. PLoS One **18**(6), e0286464.

Sharif N, Das B, Alam A. 2023. Prevalence of anemia among reproductive women in different social groups in India: Cross-sectional study using nationally representative data. PLoS One **18**(2), e0281015.

Sharma H, Singh SK, Srivastava S. 2018. Major correlates of anemia among women (age 15–49) in India and spatial variation, evidence from National Family Health Survey-4. Journal of Women's Health Care **7**(440), 2167.

Sinha A, Adhikary M, Phukan JP, Kedia S, Sinha T. 2021. A study on anemia and its risk factors among pregnant women attending antenatal clinic of a rural medical college of West Bengal. Journal of Family Medicine and Primary Care **10**(3), 1327–1331.

Sunuwar DR, Singh DR, Chaudhary NK, Pradhan PM, Rai P, Tiwari K. 2020. Prevalence and factors associated with anemia among women of reproductive age in seven South and Southeast Asian countries: Evidence from nationally representative surveys. PLoS One **15**(8), e0236449.

Thiagarajan P, Parker CJ, Prchal JT. 2021. How do red blood cells die? Frontiers in Physiology **12**, 655393.

Tvedten H. 2022. Classification and laboratory evaluation of anemia. Schalm's Veterinary Hematology, pp. 198–208.