भारतीय प्रबंध संस्थान अहमदाबाद
INDIAN INSTITUTE *of* MANAGEMENT AHMEDABAD

IIMA
Working Paper

Neonatal Mortality Rate (NMR) in India: A study using one-way ANOVA and multiple linear regression (MLR)

Rohan Kar
Sourav Bikash Borah

Research & Publications

Research &
Publications
W. P. No. 2023-02-02

Neonatal Mortality Rate (NMR) in India: A study using one-way ANOVA and multiple linear regression (MLR)

Rohan Kar
Sourav Bikash Borah

February 2023

IIMA
विद्याविनियोगाद्विकास:
भारतीय प्रबंध संस्थान अहमदाबाद
INDIAN INSTITUTE of MANAGEMENT AHMEDABAD

Neonatal Mortality Rate (NMR) in India: A study using one-way ANOVA and multiple linear regression (MLR)

Rohan Kar[1] & Sourav Bikash Borah[2]

Abstract

Neonatal Mortality Rate (NMR) is of grave concern for India and other low-income and middle-income countries aspiring to meet the Sustainability Development Goals by 2030 (SDG30). As per government estimates, the NMR in India was 30 per 1000 live births in 2019. Achieving the target of 12 deaths per 1000 live births by 2030 remains a considerable challenge.

This study was conducted using indicators from the State Health Index Round 4 (SHI-R4), covering 34 states and union territories (N=34). One-way ANOVA was performed to identify significant differences in mean NMR, if any, between states and union territories (UTs). Later, a model was built using multiple linear regression techniques to predict the NMR in India using indicators available in the SHI-R4.

The model obtained had an R^2 value of 0.37. Among the significant predictors that most influenced the NMR were the average occupancy of a district Chief Medical Officer (CMO), the number of caesarean sections performed at First Referral Units (FRUs), and the Kayakalp score of public health facilities.

The study findings add to the existing scholarship on NMR in India. The results are significant both in terms of future research and policymaking decisions.

Keywords: neonatal mortality, sustainability development goals, health index, ANOVA, multiple linear regression, public health

[1]Doctoral Student, Marketing Area, Indian Institute of Management Ahmedabad

[2]Assistant Professor, Marketing Area, Indian Institute of Management Ahmedabad

1. Introduction

Neonatal mortality, i.e., death within the first four weeks of a newborn's life, account for almost 40 percent of under-five child mortality globally. Further, approximately, about 4 million die in the neonatal period. Unfortunately, about 99% of neonatal deaths are reported in developing nations alone (Titaley et al., 2008)**.**

Neonatal mortality rate (NMR) represents the number of deaths (in the first 28 completed days of a newborn's life) per 1000 live births. NMR varies between nations, and in the context of India, significant variability exists at both district and state levels. Incidentally, the countries with less data available to track NMR have the highest NMRs (Blencowe & Cousens, 2013). In India, approximately one million babies die every year before completing their first month. Quite astonishingly, this is about 25% of the global burden. Common causes of neonatal mortality in India include infections, preterm birth, intrapartum-related complications (birth asphyxia or inability to breathe at birth), and congenital disabilities. The NMR is on the higher side in the northern states compared to the southern states (Patel & Kumar, 2021). As per the government estimates derived from the Sample Registration System (SRS) of India, there has been a decline in the NMR from 37 per 1000 live births in 2015 to 30 per 1,000 live births in 2019 at a national level (Press Information Bureau, n.d., sec. "Status of IMR and MMR in India"). This improvement is significant in terms of India's National Health Policy (NHP, 2017) and the sustainable development goal (SDG30), and the targets set thereof of limiting neonatal mortality to 12 per 1000 live births and under-5 mortality rate (U5MR) to 25 per 1000 live births.

High NMR is a significant concern because it reflects the availability and quality of India's prenatal, intrapartum, and neonatal care services. However, states like Kerala and Tamil Nadu have now accomplished the target (NMR of 12 per 1000 live births), and Punjab is close to achieving the same. On the other end, Odisha, Uttar Pradesh, and Uttarakhand had the highest NMR of 32, 30, and 30, respectively. Further, district-level estimates show that only 9% of the 640 districts are likely to achieve the NMR target of 12, while the remaining are most likely to fail. Most of these high-risk districts are located in the poorer states of north-central and eastern India. Nevertheless, a few high-risk districts are also from the rich and advanced states.

The central and the state governments have launched schemes such as Janani-Shishu Suraksha Karyakaram (JSSK, June 2011) and Mother and Child Tracking (MCTs, December 2019) to facilitate antenatal care and neonatal and infant care (Bora & Saikia, 2018; Kapur, 2020)**.** Despite some of these interventions, the NMR has continued to stray on the higher side for India. Therefore, much more must be done to understand district-level neonatal mortality variations and patterns. In this case, specific trends in each of the districts can serve as a valuable reference for the policymakers aiming to decentralize health planning to improve neonatal survival in India, especially in areas with persistently high NMR and low reduction rates (Dandona et al., 2020). The underlying system-based rationales of neonatal deaths need to be better comprehended. One size cannot fit all, especially in a diverse and large country like India. Based on accurate data on deaths and the underlying causes, contextual microplanning at the lowest administrative unit level needs to be taken up on priority. This would require constant monitoring regarding the availability and accessibility of quality care, knowledge translation to caregivers, and local implementation of healthcare interventions targeting neonatal deaths. Focusing on local changes using local data may lead to better outcomes, as has been shown in the case of Canada (Das et al., 2021; Kumar & Singhal, 2020).

National Institution for Transforming India (NITI) Aayog, in collaboration with the Ministry of Health & Family Welfare (MoHFW) and the World Bank, launched the State Health Index (SHI) to measure the performance of States and Union Territories (U.T.s). SHI is a composite score derived using data collected on 23 indicators, grouped into health outcomes, governance and information, and critical inputs or processes. For the generation of ranks, the states have been classified into three groups viz larger states, smaller states, and union territories (U.T.s) to ensure comparability among similar entities. The first round of the SHI was released in 2018, highlighting the performance of the states and U.T.s between 2014-15 and 2015-16. Subsequently, the second and third rounds of the SHI were released in 2018 and 2019, respectively. Most recently, round four of the SHI (hereafter addressed as SHI-R4) was released in 2020, which measured the performance of the states and U.T.s for 2019-20 (Health Index, n.d., sec. "Executive Summary").

In this study, we have constructed a predictive model of NMR in India to determine the significant predictors accounting for the variations in NMR across Indian states and U.T.s. The model used in this study is the multiple linear regression (MLR) model. In addition, I have conducted a one-way ANOVA analysis to predict the difference in mean NMR between the groups of states and U.T.s. The study adds to the existing scholarship on NMR in low and middle-income countries like India.

2. Methods

The dataset used for this study was primarily extracted from the SHI-R4 data. In addition, data from the National Family Health Survey (NFHS-5), the SRS, and the Health Management Information System (HMIS) were used as a reference to obtain the final dataset (hereafter referred to as the NMR_dataset). The dataset included observations from larger states (19), smaller states (8), and U.T.s (7) for 2019-20 [N=34]. West Bengal and Ladakh were not included in the NMR_dataset due lack of availability of reliable data points on indicators in the SHI-R4.

One-way ANOVA and Tukey's HSD test

One-way ANOVA (analysis of variance) is an F-test that generalizes the two-sample t-test to three or more samples by comparing the mean responses of two or more groups. This method is standard in research, especially when the response variable is quantitative (Agresti and Finlay, 1997; Heiberger & Neuwirth, 2009; Ross & Willson, 2017)**.** The null hypothesis tested under this study is that there is no significant difference in mean between three or more groups. Because the null hypothesis is rejected if the means are equal, it becomes necessary to identify where the significant differences in the mean arise. To do this, post hoc tests are conducted as a follow-up test to ANOVA. In addition, post hoc tests help to control for the type-I error rate and the ambiguity surrounding ANOVA (Mahajan, 2016)**.** Duncan's new multiple range test (MRT), Scheffe, Student–Newman–Keuls (SNK), least significant difference (LSD), Tukey's HSD, and Bonferroni Procedure are some of the post hoc tests available to researchers. Tukey's honestly significant difference test (Tukey's HSD) is the most used among post hoc tests. The test evaluates whether the relationship between group means is statistically significant. Tukey's HSD is robust and can be conducted in case of equal or unequal samples per group (Allen, 2017; Nanda et al., 2021)**.**

Preliminary check for linearity between predictors and the response variable

Before constructing any linear regression model(s), a preliminary check must be performed to ensure that potential predictors are linearly associated with the response variable. This check is often necessary to ascertain whether linear regression should be the method of choice or whether an alternate model needs to be chosen. Scatter plots can be used in this case to visualize the relationship between two continuous variables as linear (or nonlinear) (Sainani, 2016; Schneider et al., 2010). However, if a significant departure from linearity is observed, the study may, depending upon the research question(s), opt to exclude the predictor altogether from the study or chose to conduct a data transformation to confirm to the assumption of linearity (Lee, 2020; Marudachalam, 2017). Table 1 shows the predictors used in this study to model NMR. Logarithmic, cube root, and square root transformations were performed on predictors whose association with NMR was not found to be linear. None of the predictors were excluded from the study. Further, all subsequent analyses were performed using the transformed predictors, wherever applicable.

Table 1. List of Predictors.

Category	**Predictor**
Primary Outcome	MMR
	Number of institutional deliveries (Public & Private)
	Number of pregnant women who received 4 or more ANCs
	Number of women registered for ANC
Infrastructure	Number of DHs certified under the LaQshya initiative - labour room
	Number of DHs certified under LaQshya- maternity OT
	Number of PHCs functional as Health and Wellness Centres
	Percentage shortfall in staff nurses at PHCs, CHCs, UPHCs and UCHCs
	Number of FRUs conducting specified number of C-sections per year
	Number of DHs with Kayakalp score of >70%
	Number of SDHs/CHCs with Kayakalp score of >70%
	Number of PHCs with Kayakalp score of >70%
Governance	Level of registration of births (%)
	Average occupancy (in months) of district CMOs
	Average occupancy (in months) of PS, MD (NHM), and Director (Health Services)

Source: State Health Index (Round 4); see https://social.niti.gov.in/

Abbreviations: MMR, Maternal Mortality Ratio; ANC, Antenatal Care; DHs, District Hospitals; PHCs, OT, Operation Theatre; Primary Health Centres; CHC, Community Health Centres; UPHC, Urban Primary Health Centres; UCHC, Urban Community Health Centres; FRUs, First Referral Units; CMO, Chief Medical Officer; PS, Principal Secretary; MD, Mission Director; NHM, National Health Mission.

Multiple linear regression (MLR) and model diagnostics

One of the most frequently used methods to determine the influence of several predictors on the response variable is the multiple linear regression (MLR). The equation for a typical MLR-based model takes the below form (Ernst & Albers, 2017).

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \cdots + \beta_i x_i + \varepsilon$$

Y represents the response variable, and x_i $(i = 1, 2, 3, ..., i)$ represents the predictors. β_0 is the intercept value, i.e., the value of Y without the predictors accounted for in the model, β_i $(i = 1, 2, 3,...., i)$ is the estimated regression coefficients of individual predictors, and ε is the model error, i.e., the variation in the estimate of Y compared to the actual value.

The regression model that uses one single predictor is called a univariate regression model, while the one that explains the effect of multiple predictors on the response variable is called a multivariate model. In multivariate regression analysis, an attempt is made to account for variation in the response variable that can be explained by the model predictors (Uyanık & Güler, 2013). Nevertheless, for each of the models, the goodness of fit is what makes a difference. Several methods are available to estimate a regression model's goodness of fit. Many of these methods rely on visually inspecting the data and the residuals. Examining the 'standardized residuals' is, therefore, a crucial step in the development of any MLR model and related analyses (Casson & Farmer, 2014; Trunfio et al., 2022).

Model diagnostics involve validating the below six assumptions to determine the goodness of fit. These assumptions must hold for any linear regression model to ensure the reliability of the model's prediction.

1. *Linearity of Residuals*: Mandates the linearity of the standardized residuals when plotted against the model-fitted values. Scatter plots can used to validate this assumption.

2. *Normal Distribution of Residuals*: The normality of the residuals can be corroborated using the Shapiro-Wilk test statistic (W). In addition, Quantile-quantile (Q-Q) plots and histograms can be used to visualize the distribution of standardized residuals. For small sample sizes, W greater or lesser than 1.96 is adequate to establish the normality of the standardized residuals (Ghasemi & Zahediasl, 2012).

3. *Homoscedasticity*: implies 'constant variance'. Scale-location plots can be used to determine the homoscedasticity of standardized residuals.

4. *Independence of Residuals*: The Durbin-Watson (D.W.) statistical test can be used to verify the independence of residuals. D.W. test is a commonly used method for detecting lag-1 autocorrelation and verifying independence (Turner et al., 2021). The D.W. statistics must fall in the range of [1.4; 2.5] to establish the absence of any significant autocorrelation (Trunfio et al., 2022).

5. *Absence of Multicollinearity*: Multicollinearity illustrates a high degree of linear intercorrelation between predictors in an MLR model. This intercorrelation can significantly reduce the goodness of fit of models. Multicollinearity can be determined by examining the Variance Inflation factor (VIF) and Tolerance for each predictor in the model. VIF < 5 and Tolerance > 0.1 can be used as a cut-off to determine the absence of multicollinearity (Kim, 2019; Trunfio et al., 2022).

6. *Absence of Outliers*: Cook's distance (D) value (< 1) suggests the absence of any significant outliers in an MLR model (Trunfio et al., 2022). In addition, Cook's D plot for standardized residuals can also be visualised to confirm the absence of any significant outlier.

The coefficient of determination (R^2) measures the goodness of fit of a regression model. R^2 and adjusted R^2 are statistical values that represent the proportion of variance in the response variable, explained by the predictors in the sample (R^2) and an estimate in the population (adjusted R^2) (Miles, 2005). Adjusted R^2 denotes how well predictor values fit a curve or line by adjusting for the number of predictors in each model**.** Therefore, adjusted R^2 is normally used in place of R^2 for any MLR-based analysis. (Draper, 2011; Trunfio et al., 2022).

Cohen's f^2 test

Cohen's f^2 is an appropriate test for calculating the effect size of a regression model in which both the response and predictor variables are continuous **(Selya et al., 2012)**. Based on the f^2 value, the effect size of a model can be categorized as small (0.10 – < 0.30), medium (0.30 – < 0.50), or large (≥ 0.50) **(Cohen, 2013)**.

Software environment

RStudio ver. 2022.07.02, Build 576 was used for carrying out all the statistical analyses, as discussed in the above sections.

3. Results

Table 2 below presents the descriptive statistics of the response variable (NMR) and the model predictors. The mean of the response variable is 17.39 (± 8.44 SD), indicating high variance. The data is positively skewed (S = 0.14) and Platykurtic (K = 2.23), indicating a thin-tailed uniform distribution.

Table 2. Descriptive Statistics. M, Mean; SD, Standard Deviation; K, Kurtosis, S, Skewness. N=34.

Response Variable	M	SD	K	S
Neonatal Mortality Rate (NMR)	17.39	8.44	2.23	0.14
List of Predictors				
MMR	134.40	77.96	2.35	0.41
Number of institutional deliveries (Public & Private)	558079	755934	8.95	2.24
Number of pregnant women who received 4 or more ANCs	639103	991357	13.96	3.07
Number of women registered for ANC	805790	1291453	14.59	3.21
Number of DHs certified under the LaQshya initiative - Labour Room	4.44	5.23	5.05	1.50
Number of DHs certified under LaQshya-Maternity OT	3.79	5.18	6.01	1.78
Number of PHCs functional as Health and Wellness Centres	481.38	559.94	2.88	1.05
Percentage shortfall in staff nurses at PHCs, CHCs, UPHCs and UCHCs	38.84	29.18	1.64	0.06
Number of FRUs conducting specified number of C-sections per year	46.09	54.57	4.74	1.54
Number of DHs with Kayakalp score of >70%	10.47	14.70	13.52	2.98
Number of SDHs/CHCs with Kayakalp score of >70%	46.82	73.94	10.27	2.57
Number of PHCs with Kayakalp score of >70%	116.15	176.33	11.35	2.69
Level of registration of births (%)	90.34	10.84	3.46	(1.08)
Average occupancy (in months) of district CMOs	15.94	6.25	2.94	(0.29)
Average occupancy (in months) of PS, MD (NHM), and Director (Health Services)	13.95	5.17	3.07	1.00

Source: State Health Index (Round 4); see https://social.niti.gov.in/

One-way analysis of variance (ANOVA)

One-way ANOVA was used to determine if there were any significant differences in mean NMR between at least two groups of states. The results obtained were significant ($F (2, 31) = 3.959$, $p = 0.0294$), suggesting a difference in the mean NMR between at least two groups (Table 3 and Figure 1).

Table 3. Summary statistics (one-way ANOVA).

	Df	Sum Sq.	Mean Sq.	F value	Pr(>F)
States	2	477.7	238.83	3.959	**0.0294***
Residuals	31	1870.3	60.33		

Signif. codes: *** p < 0.001; ** p < 0.01; * p < 0.05; · p < 0.1

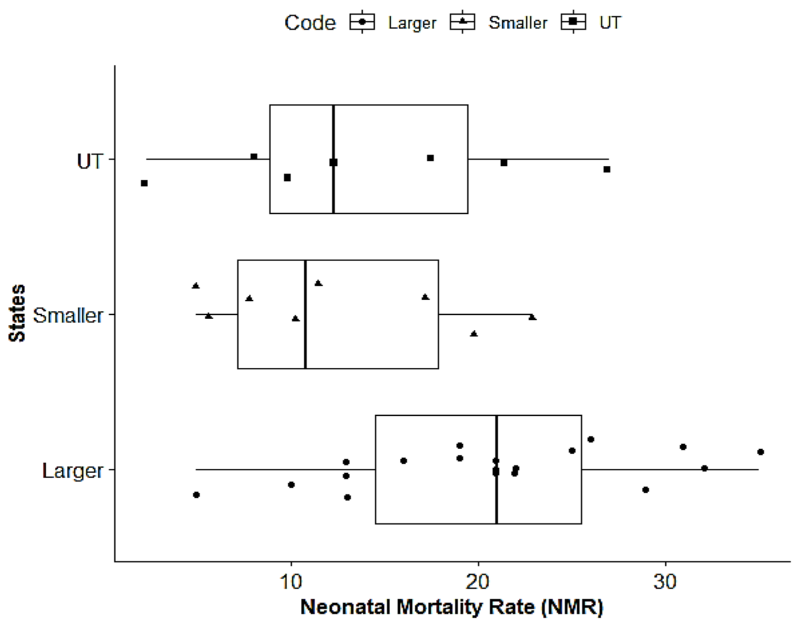

Figure 1. Box Plot Showing the Spread of NMR observations for Larger States, Smaller States and U.T.s.

Tukey's HSD post hoc analysis demonstrated that the difference in mean NMR was significant in the case of larger states (A) and smaller states (B) ($p = 0.045$, 95% C.I. = [-16.26, -0.15]) (Table 4). However, no statistically significant difference in mean NMR was reported between U.T.s (C) and larger states ($p = 0.146$) or between U.T.s and smaller states ($p = 0.92$).

Table 4. Summary statistics (Tukey's HSD). A, Larger States; B, Smaller States; C, Union Territories (U.T.s).

	diff	lower	upper	p value
B - A	-8.21	-16.26	-0.15	**0.045**
C - A	-6.64	-15.09	1.81	0.146
C - B	1.57	-8.33	11.46	0.92

α = 0.05

Model diagnostics

Two models were constructed in this study using the predictors listed in Table 5. Backward elimination technique was used derive Model 2. The results of the p-value (cut-off = $p > 0.05$) for individual predictors were examined in each iteration. The least significant predictor that did not meet the cut-off was removed after every iteration. Once the predictor was removed from the model, it remained excluded (Bursac et al., 2008)**.** After every iteration, regression was performed with the remaining predictors. These iterations were continued until no significant change is R^2 was observed by removing additional predictors based on the cut-off.

Table 5. Model-wise List of Predictors.

Predictor	Model 1	Model 2
MMR *****	✓	–
Percentage shortfall in staff nurses at PHCs, CHCs, UPHCs and UCHCs	✓	–
Number of FRUs conducting specified number of C-sections per year	✓	✓
Number of DHs with Kayakalp score of >70%	✓	✓
Number of SDHs/CHCs with Kayakalp score of >70%	✓	–
Number of PHCs with Kayakalp score of >70% *****	✓	✓
Level of registration of births (%)	✓	–
Average occupancy (in months) of district CMOs	✓	✓
Average occupancy (in months) of PS, MD (NHM), and Director (Health Services)	✓	–

** Linear Data Transformation*

Figure 2 shows the scatter plot of 'standardized residuals' against the 'fitted values.' For both models, the standardized residuals were found to fluctuate randomly, about 0, with no noticeable trend or pattern. These observations are sufficient to validate the assumption of linearity in both models. Q-Q plots and Histograms for both models are shown in Figure 3. As seen in the Q-Q plots, the points are very close to the line, indicating a normal distribution of standardized residuals in both models. The normal distribution of standardized residuals can also be visualized in the histograms, although a minor deviation from normality can be seen in the case of Model 2. Further, the W test statistics were found to be 0.99 ($p > 0.1$) and 0.97 ($p > 0.1$) for Models 1 and 2, respectively, thereby confirming the normality of standardized residual.

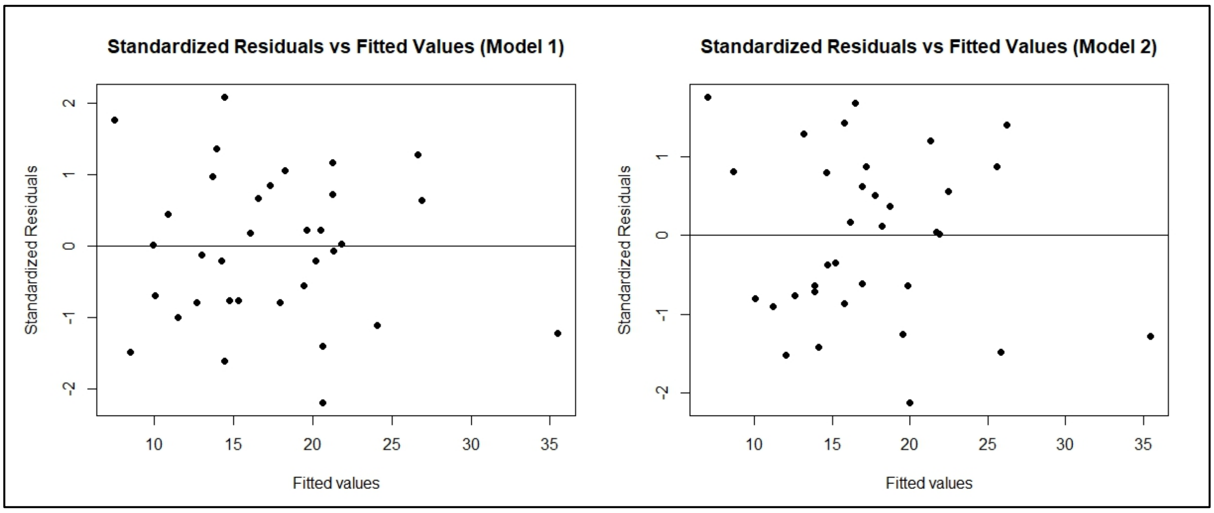

Figure 2. Scatter Plot of Standardized Residuals.

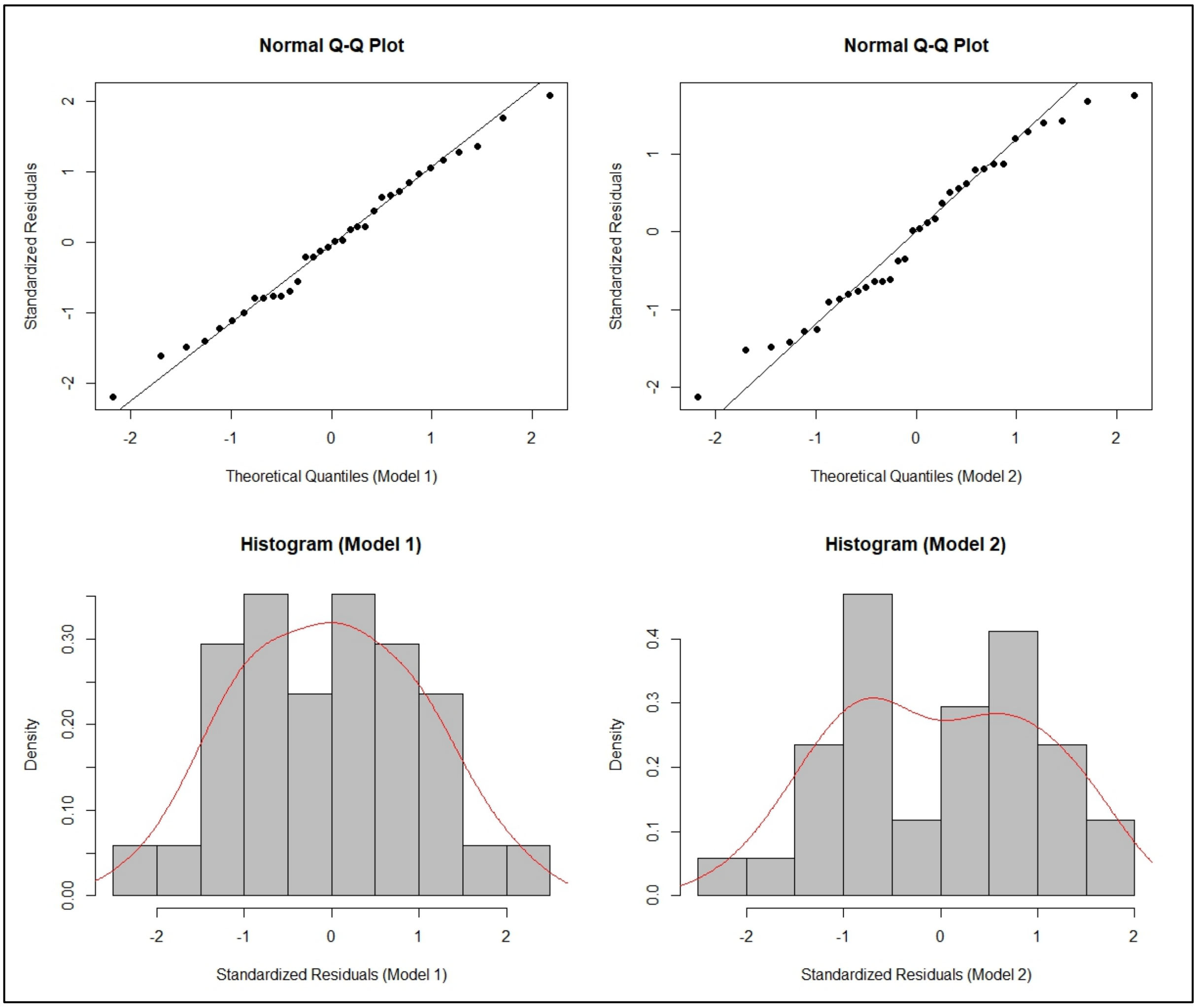

Figure 3. Upper Panel: Normal Q-Q Plot of Standardized Residuals; Lower Panel: Histogram Showing Distribution of Standardized Residuals.

Figure 4. below shows the scale location plots used to verify the homoscedasticity assumptions. The variance of residuals was not constant across fitted values for both models. So, a moderate violation of the homoscedasticity assumption was observed. Nevertheless, MLR models are robust to handle low to moderate violations of homoscedasticity (Ernst & Albers, 2017). The D.W. test statistic for Model 1 was 1.63 ($p = 0.198$), which was within the acceptable range of [1.5; 2.5], thereby demonstrating the independence of residuals. A mild deviation was observed in the case of Model 2 with a D.W. test statistic of 1.46 ($p = 0.116$). Cook's distance for each observation in Model 1 was less than 1, so there were no outliers that could negatively affect the estimate of the coefficients. However, in Model 2, one of the observations reported a Cook's distance >1 (Figure 5). Regardless, due to the small sample size, the outlier was not found to be influential and hence, was not excluded from the study.

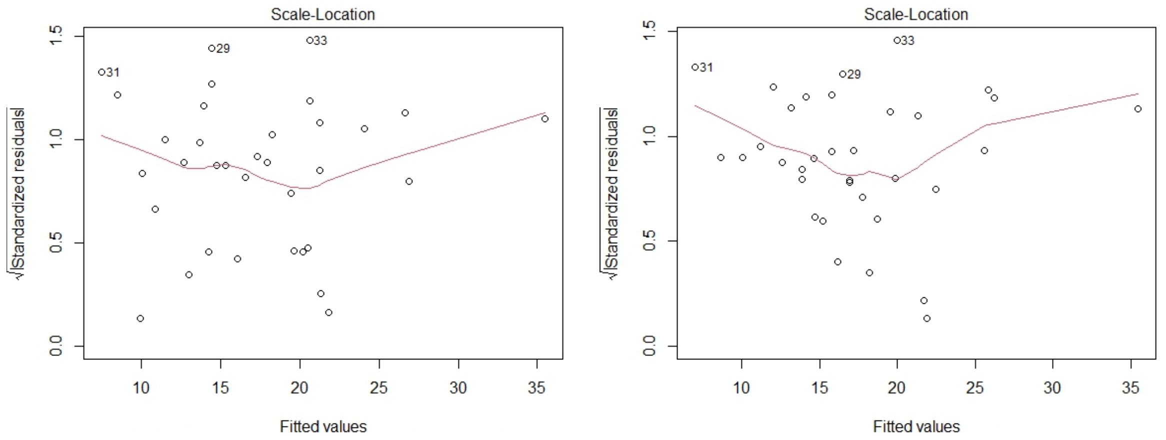

Figure 4. Scale Location Plot of Standardized Residuals. L: Model 1; R: Model 2.

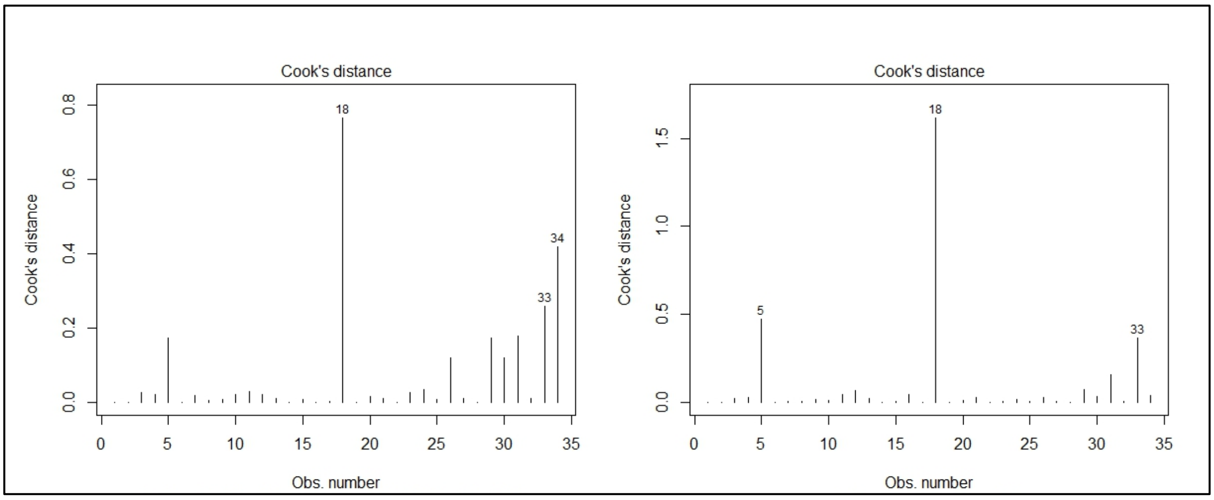

Figure 5. Cook's Distance Plot. L: Model 1; R: Model 2.

The multicollinearity of predictors was validated using the VIF values and Tolerance. The VIF for individual predictors was always less than 5, and the Tolerance was always greater than 0.1 (Table 6). These results are sufficient to establish the absence of multicollinearity in both models.

Table 6. Collinearity Statistics.

Predictor	Model 1		Model 2	
	VIF	Tolerance	VIF	Tolerance
MMR *	1.16	0.86	–	–
Percentage shortfall in staff nurses at PHCs, CHCs, UPHCs and UCHCs	1.16	0.86	–	–
Number of FRUs conducting specified number of C-sections per year	3.64	0.27	2.55	0.39
Number of DHs with Kayakalp score of >70%	1.97	0.51	1.70	0.59
Number of SDHs/CHCs with Kayakalp score of >70%	4.31	0.23	–	–
Number of PHCs with Kayakalp score of >70% *	3.20	0.31	2.12	0.47
Level of registration of births (%)	1.06	0.94	–	–
Average occupancy (in months) of district CMOs	1.06	0.95	1.01	0.99
Average occupancy (in months) of PS, MD (NHM), and Director (Health Services)	1.13	0.89	–	–

** Linear Data Transformation*

Best-fit model

Adjusted R^2 values were found to be 0.29 ($p < 0.05$) for Model 1 and 0.37 ($p < 0.01$) for Model 2, which were statistically significant. The F-values were 2.50 and 5.90 for Models 1 and 2, respectively. Table 7 below shows the regression summary of both models. Based on the observations, Model 2 was deemed a better fit than Model 1 in predicting NMR. Notably, all the predictors in Model 2 were found to be statistically significant at different significance levels; the number of FRUs conducting a specified number of C-sections per year and the number of D.H.s with a Kayakalp score of >70% ($p < 0.01$). Average occupancy (in months) of district CMOs and number of PHCs with a Kayakalp score of >70% ($p < 0.1$). The standardized regression coefficients for the predictors were found to be in the range [-0.27; 0.61].

Table 7. Summary statistics of the MLR models predicting NMR (N = 34). B: Unstandardized Coefficient; S.E.: Standard Error; β: Standardized Coefficient.

Predictor	Ordinary Least Square Regression Estimates					
	Model 1			Model 2		
	B	SE	β	B	SE	β
Intercept	13.49	14.95	–	**20.06** ***	3.59	–
MMR *****	1.30	3.22	0.06	–	–	–
Percentage shortfall in staff nurses at PHCs, CHCs, UPHCs and UCHCs	0.01	0.04	0.06	–	–	–
Number of FRUs conducting specified number of C-sections per year	**(0.08)** ·	0.04	(0.50)	**(0.09)** **	0.03	(0.63)
Number of DHs with Kayakalp score of >70%	**0.37** **	0.12	0.64	**0.35** **	0.35	0.61
Number of SDHs/CHCs with Kayakalp score of >70%	(0.03)	0.03	(0.27)	–	–	–
Number of PHCs with Kayakalp score of >70% *****	**(0.60)** ·	0.31	0.50	**0.49** ·	0.49	0.41
Level of registration of births (%)	0.00	0.11	(0.01)	–	–	–
Average occupancy (in months) of district CMOs	(0.34)	0.20	(0.26)	**(0.37)** ·	0.18	(0.27)
Average occupancy (in months) of PS, MD (NHM), and Director (Health Services)	0.19	0.25	0.12	–	–	–
Median (Residuals)	(0.15)			0.21		
Residual standard error	7.11			6.69		
Multiple R^2	0.48			0.45		
Adjusted R^2	**0.29*****			**0.37****		
F-statistic	2.50			5.90		
Δ Adjusted R^2	–			0.08		

Signif. codes: *** p < 0.001; ** p < 0.01; * p < 0.05; · p < 0.1

4. Discussion

The study's primary objective was to examine the two research questions in the context of NMR in India; First, is there any significant difference in NMR between the three groups of states and U.T.s? This inquiry is necessary because if there exist substantial differences, future rounds of health index must incorporate indicators tracking the probable factors at the level of larger states, smaller states, and U.T.s that may be driving such a phenomenon. Second, identify the best fit model that can explain the variance in NMR using predictors from the SHI-R4.

One-way ANOVA and post hoc tests were used to answer the first research question, and multiple linear regression was used to construct the best-fit model. A simple model (Model 2) was obtained with a value of adjusted R2 equal to 0.37 (Table 7). Cohen's f 2 for this model was estimated to be 0.59, indicating a medium effect size (Cohen, 2013). In addition, all four predictors were found to be significant (at different p-values), accounting for 37% of the variation in NMR.

Differences in mean NMR between larger and smaller states is significant

The findings align with another study conducted in India, citing the vast disparities in NMR that exists between and even within the states (Sankar et al., 2016). Several factors, such as the rural-urban divide, the gap between rich and poor, and gender differentials, have been found to drive inequity in healthcare services and delivery in India. Moreover, the burden of high neonatal mortality in some states and U.T.s compared to others can be attributed to various local demographic, educational, socioeconomic, and biological factors. Therefore, equitable distribution of healthcare services by improving the quality of government healthcare facilities remains a continuing challenge in the Indian context. Necessary cross-cutting interventions are needed to blur this 'great divide. Further, a panel of these interventions must be curated explicitly to reduce newborn mortality to a level defined in SDG30 goals. Policymakers must help design frameworks that can gather data from a more granular level and implement efficient tracking to improve the reliability of the collected data. In short, interventions must be designed using a data-driven approach, and efforts must be continued to improve the affordability and accessibility of these novel interventions so that all sections of society can benefit from them.

The average occupancy of district CMOs negatively correlates with NMR

A CMO is a technical or background official who advises ministers, disseminates information to the public, and functions as a senior executive responsible for providing leadership to the government healthcare institutions such as district hospitals (D.H.s). Depending on the context, a district CMO's role may be operational, strategic, or marketing-based, in addition to being head of local healthcare development. (MacAulay et al., 2022). They are considered physician leaders who play a prominent role in providing high-quality care for patients in the district and can significantly influence the overall performance of the healthcare institutions they manage. In India, this role becomes even more critical, given the shortage in the quality of healthcare services available to the citizens (Angood & Birk, 2014; MacAulay et al., 2022)**.** So, the occupancy rate of district CMOs matters. Observer Research Foundation's (ORF's) Health Initiative report (2020) articulates the fact that the short average occupancy of district CMOs is detrimental to the efficacious implementation of critical public health programs at public healthcare facilities in respective districts (Kapur, 2020).

As per the SHI-R4 estimates, the average occupancy of a district CMO in months at the pan-India level was about 16 in 2019-20. The average occupancy among the larger states was 14.20; for smaller states, it was 20.20, and in the case of U.T.s, this value was close to 15.80. Delhi (28.89) and Goa (27.02) had the highest occupancy rates among all states and U.T.s. Assam (21.85) months) and Kerela (21.92 months) were the larger states with the highest occupancy rates. In contrast, Odisha (5.19), followed by Andhra Pradesh, Haryana, Punjab, and Uttarakhand (< 12 months), had occupancy rates way below the national average of 16 months. A critical observation to be made here is that larger states, with a higher number of districts under their belt, tend to function worse than smaller states and U.T.s. A logical argument could be that tracking and managing a smaller number of districts is much easier, as with smaller states and U.T.s. A counter-intuitive argument would be that; shouldn't larger states do much better, given the population density and the varied healthcare needs of the large proportion of the Indian population that falls under these states?

Model 2 generated in this study shows a negative correlation between the average occupancy of a district CMO and NMR. The findings are novel in terms that there is no previous study available that has attempted to look at this correlation. So, this brings us to the vital question: what are the probable variables driving the low occupancy rates of district CMOs in states and U.T.s? Several factors can be considered, ranging from psychological, social, cultural, administrative, and political. These factors are incredibly dynamic, and their interplay varies considerably between states and districts within individual states. Therefore, context-based research is needed to establish causality at the district level. Some of this research may choose to investigate the association between 'transition periods' (between an outgoing CMO and an incoming CMO) and neonatal deaths occurring during this period.

Number of FRUs performing the specified number of C-sections per year negatively correlates with NMR

C-section deliveries are rising globally, leading to a proportionate rise in short-term and long-term maternal and neonatal complications (Khasawneh et al., 2020)**.** C-section delivery is a lifesaving procedure performed during obstructed labor and other emergency obstetrical conditions. Accordingly, ensuring access to caesarean delivery in a public healthcare facility is essential to meet the SDG30 goals for reducing maternal and neonatal mortality. However, just like any other surgical procedure, there involve risks of complications, and overuse can be detrimental to both mothers and newborns. According to the World Health Organization (WHO), C-section delivery rates for a country must not exceed over 10 to 15 per 100 live births. C-section delivery rates in many countries remain substantially higher (Molina et al., 2015). India National Family Health Survey (NFHS-5) estimates suggest that C-section births in India increased to 21.5% from 17.2% observed during NFHS-4 (2015-16). Notably, only 22.7% of C-section births took place in public health facilities in urban areas and 11.9 % in rural areas (Gondwe et al., 2020; *National Family Health Survey (NFHS-5)*, n.d.).

First Referral Unit (FRU) refers to an existing facility (District Hospital, Sub-divisional Hospital, Community Health Centre, etc.) that is provisioned to supply round-the-clock services pertaining to obstetric and newborn care and all other emergencies. A public health facility can be categorized as an FRU on the fulfilment of three mandatory requirements: emergency obstetric care, including surgical interventions like caesarean sections (C-sections), newborn care, and the availability of a 24-hour blood storage facility. In addition, FRUs are required to provide a whole range of other care services, including emergency care of sick children, family planning, safe abortion, treatment of STIs/RTIs, and referral transport services. As per the NHM's most updated estimates, the nationwide number of FRUs have risen from 940 in 2005 to 2996 in March 2020 (Infrastructure: National Health Mission, n.d.).

Strong negative correlation was observed between the number of FRUs conducted the specified number of C-sections and NMR. The finding is significant and may serve as the foundation for further research that focuses both on quantity (the number of C-sections performed in FRUs) and quality (the actual outcome of the procedure). It is crucial to weigh in the 'quality' factor as it is an indicator of the standard of health services being provided in FRUs. These include round-the-clock availability of obstetricians and support staff. In India, the situation of FRUs, especially in rural areas, is alarming regarding infrastructure and staffing. Consequently, much of the adverse health effects are observable in the rural population, where access to private healthcare facilities is equally limited. This situation does not bode well for a nation where the NMR is currently 30 per 1000 live births (Press Information Bureau, n.d., sec. "Status of IMR and MMR in India). Policymakers must aim to set specific guidelines and regulations as per international standards that focus more on improving the quality of care at FRUs. Further, the National Rural Health Mission (NRHM) must look at expanding the coverage of Indian Public Health Standards (IPHS) to ensure that FRUs meet the minimum acceptable quality standards in terms of maternal and neonatal care (Pandve & Giri, 2015).

"Kayakalp" scheme is significant in the context of NMR in India

The Ministry of Health & Family Welfare (MoHFW) launched the "Kayakalp" initiative to award public health facilities with high cleanliness, hygiene, and infection control levels. The scheme was initiated in D.H.s in 2015 and expanded to the PHCs in 2016 and UHCs in 2017. Under this scheme, cash rewards are provided to facilities that score at least 70% or more in each level of assessment defined in the Kayakalp assessment tool formed by the MoHFW. The scheme's primary objective is to develop and disseminate sustainable practices associated with improved cleanliness in public health facilities (linked to positive health outcomes). The Kayakalp scheme primarily applies to secondary care public facilities meeting the Indian Public Health Standards (IPHS) guidelines, albeit with some discretion, and may also be used for primary and tertiary healthcare facilities. Kayakalp aims to follow standard protocols and practices to achieve the highest cleanliness, hygiene, and infection control standards at public healthcare facilities (Quality Assurance Program Initiatives: National Health Mission, n.d.). Despite the scheme being launched in 2015, indicators tracking the Kayakalp performance for states and U.T.s were unavailable in the previous three rounds of the SHI. The indicator *"Proportion of Public Health facilities with Kayakalp score >70% against the total number of Public Health facilities"* was included only in the SHI-R4.

The discussion in the preceding section leads to the question; how is the Kayakalp scheme relevant in the context of NMR? The mother and the newborn often need to access the nearest health facility due to various health complications arising from home deliveries. These complications may occur even in the case of institutional deliveries at public health facilities. It is a proven fact that a large proportion of the Indian population belonging to rural areas has no or limited access to private facilities due to various constraints. Therefore, government health institutions remain the first and only choice for many. Risks of maternal and newborn deaths are highest during the first 24 to 48 hours after birth due to untoward infection happening to the mother or the baby. It should not come as a surprise to the readers that just by maintaining a clean and hygienic standard, some of these life-threatening infections can be prevented in public health facilities. This is precisely where the Kayakalp scheme holds significance in the NMR context. Neonatal infections contribute primarily to the high NMR burden in India, and the rate is higher in poorer states than in richer ones (Million Death Study Collaborators et al., 2010).

While a negative correlation was expected, quite surprisingly, the current study found a positive correlation between the number of D.H.s and PHCs with Kayakalp score (> 70%) and NMR. This requires further investigations (as more accurate data become available).

5. Study Limitations

Although this study is novel in terms of modeling India's NMR using the SHI-R4 data, it is not without limitations. First, the small sample size (N = 34) used for this study. Data tracking of health indicators has improved over the years and continue to evolve. Therefore, obtaining reliable estimates from the previous three rounds of SHI was challenging. With an updated set of indicators, the SHI-R4 provides a more reliable data when compared to the earlier rounds. Hence, the current study was conducted using only SHI-R4 observation leading to a small sample size. However, past literature on regression analysis have shown that reliable inferences can be drawn even with small sample sized studies (cut-off of $N \geq 25$). In addition, there should not be any observable clusters at one end of a data cloud, and the six regression assumptions must be met (Jenkins & Quintana-Ascencio, 2020). The current study meets all the criteria above and can serve as a primer to building more robust predictive models as more reliable datasets becomes available. Second, as per the SHI-R4 guidelines released by NITI Aayog, the NMR indicator applies only to larger states. As a result, the data for smaller states and U.T.s were not readily available in the SHI-R4. NMR data for the smaller states and U.T.s were obtained from SRS and HMIS. This may well have introduced mild approximation errors into the developed model. Nevertheless, the moderate effect size specifies the model's predictive ability in explaining the variance in NMR (Cohen, 2013).

6. Conclusion

In this study, an attempt was made to model the Indian NMR using MLR. NMR is significantly influenced by the average occupancy of a district CMO, the number of C-sections conducted at FRUs, and the number of D.H.s and PHCs maintaining a Kayakalp score of greater than 70%. The predictors explain about 37% of the variance in NMR. Moreover, key findings from the study align with what can be found in the existing scientific literature on India's NMR. The model, in addition, has good performance that validates it as a prediction tool for use by a range of stakeholders. The MLR model, although very simple in its interpretation, could not be robust enough owing to the limitations discussed in the preceding section. Therefore, future developments will include validation of the model after managing some of these limitations.

Future studies may look at critically examining the Kayakalp scheme, its potential merits, and its demerits and association with the NMR in India. This needs to be done as there are monetary rewards associated with this scheme, and therefore, the scheme's substantial benefits must be reflected in one of the key outcome variables, such as NMR. Lastly, policymakers can look at implementing robust mechanisms and frameworks to improve the data collection in terms of granularity at different levels, viz., community, district, and state. Stakeholders must embrace the current challenges and ensure the accessibility and reliability of future research data on NMR in India.

Abbreviations

ANOVA, Analysis of Variance; CMO, Chief Medical Officer; D.H., District Hospital; FRU, First Referral Unit; HMIS, Health Management Information System; MMR, Maternal Mortality Ratio; MoHFW, Ministry of Health & Family Welfare; NFHS, National Family Health Survey; NHM, National Health Mission; NHP, National Health Policy; NITI, National Institute for Transforming India; NMR, Neonatal Mortality Rate; PHC, Primary Health Center; CHC, Community Health Center; UPHC, Urban Primary Health Center; CPHC, Community Primary Health Center; P.S., Principal Secretary; SDG, Sustainability Development Goal; SDH, Sub District Hospital; SHI, State Health Index; U5MR, Under-five Mortality Rate; U.T., Union Territory; WHO, World Health Organization.

Author contributions

The authors confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Availability of data and materials

The datasets generated and/or analyzed during the current study are publicly available and can be downloaded.(https://social.niti.gov.in/health-index)

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Copyright statement

The copyright of the paper rests with the authors. And they are transferring the copyright to publish the article in the IIMA working paper series for public use with due reference to published matter in the name of concerned authors.

The authors reserve all proprietary rights such as patent rights and the right to use all or part of the article in future works of their own such as lectures, press releases, and reviews of textbooks.

In the case of republication of the whole, part, or parts thereof, in periodicals or reprint publications by a third party, written permission must be obtained from the authors.

The authors declare that the material being presented by them in this paper is their original work and does not contain or include material taken from other copyrighted sources. Wherever such material has been included, it has been clearly indented or/and identified by quotation marks and due and proper acknowledgements given by citing the source at appropriate places

References

Agresti, A., & Finlay B. (1997). *Statistical Methods for the Social Sciences* (3rd Edition). Prentice Hall, Upper Saddle River.

Allen, M. (2017). *The SAGE Encyclopedia of Communication Research Methods*. SAGE Publications, Inc. https://doi.org/10.4135/9781483381411

Angood, P., & Birk, S. (2014). The value of physician leadership. *Physician Executive*, *40*(3), 6–20.

Blencowe, H., & Cousens, S. (2013). Review: addressing the challenge of neonatal mortality. *Tropical Medicine & International Health*, n/a-n/a. https://doi.org/10.1111/tmi.12048

Bora, J. K., & Saikia, N. (2018). Neonatal and under-five mortality rate in Indian districts with reference to Sustainable Development Goal 3: An analysis of the National Family Health Survey of India (NFHS), 2015–2016. *PLOS ONE*, *13*(7), e0201125. https://doi.org/10.1371/journal.pone.0201125

Bursac, Z., Gauss, C. H., Williams, D. K., & Hosmer, D. W. (2008). Purposeful selection of variables in logistic regression. *Source Code for Biology and Medicine*, *3*(1), 17. https://doi.org/10.1186/1751-0473-3-17

Casson, R. J., & Farmer, L. D. (2014). Understanding and checking the assumptions of linear regression: a primer for medical researchers. *Clinical & Experimental Ophthalmology*, *42*(6), 590–596. https://doi.org/10.1111/ceo.12358

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. https://doi.org/10.4324/9780203771587

Dandona, R., Kumar, G. A., Henry, N. J., Joshua, V., Ramji, S., Gupta, S. S., Agrawal, D., Kumar, R., Lodha, R., Mathai, M., Kassebaum, N. J., Pandey, A., Wang, H., Sinha, A., Hemalatha, R., Abdulkader, R. S., Agarwal, V., Albert, S., Biswas, A., … Dandona, L. (2020). Subnational mapping of under-5 and neonatal mortality trends in India: the Global Burden of Disease Study 2000–17. *The Lancet*, *395*(10237), 1640–1658. https://doi.org/10.1016/S0140-6736(20)30471-2

Das, U., Chaplot, B., & Azamathulla, H. M. (2021). The Role of Place of Delivery in Preventing Neonatal and Infant Mortality Rate in India. *Geographies*, *1*(1), 47–62. https://doi.org/10.3390/geographies1010004

Draper, N. R. (2011). The Cambridge Dictionary of Statistics, Fourth Edition by B. S. Everitt, A. Skrondal. *International Statistical Review*, *79*(2), 273–274. https://doi.org/10.1111/j.1751-5823.2011.00149_2.x

Ernst, A. F., & Albers, C. J. (2017). Regression assumptions in clinical psychology research practice-a systematic review of common misconceptions. *PeerJ*, *5*, e3323. https://doi.org/10.7717/peerj.3323

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, *10*(2), 486–489. https://doi.org/10.5812/ijem.3505

Gondwe, T., Betha, K., Kusneniwar, G. N., Bunker, C. H., Tang, G., Simhan, H., & Haggerty, C. L. (2020). Adverse infant outcomes associated with caesarean section delivery in India. *International Health*, *12*(5), 411–416. https://doi.org/10.1093/inthealth/ihz111

Health Index. (n.d.). Retrieved December 8, 2022, from https://social.niti.gov.in/health-index

Heiberger, R. M., & Neuwirth, E. (2009). One-Way ANOVA. In *R Through Excel* (pp. 165–191). Springer New York. https://doi.org/10.1007/978-1-4419-0052-4_7

Infrastructure :: National Health Mission. (n.d.). Retrieved December 10, 2022, from https://nhm.gov.in/index1.php?lang=1&level=2&sublinkid=1220&lid=190

Jenkins, D. G., & Quintana-Ascencio, P. F. (2020). A solution to minimum sample size for regressions. *PLOS ONE*, *15*(2), e0229345. https://doi.org/10.1371/journal.pone.0229345

Kapur, K. (2020). *How Fares India in Healthcare? A Sub-National Analysis*.

Khasawneh, W., Obeidat, N., Yusef, D., & Alsulaiman, J. W. (2020). The impact of cesarean section on neonatal outcomes at a university-based tertiary hospital in Jordan. *BMC Pregnancy and Childbirth*, *20*(1), 335. https://doi.org/10.1186/s12884-020-03027-2

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, *72*(6), 558–569. https://doi.org/10.4097/kja.19087

Kumar, P., & Singhal, N. (2020). Mapping neonatal and under-5 mortality in India. *The Lancet*, *395*(10237), 1591–1593. https://doi.org/10.1016/S0140-6736(20)31050-3

Lee, D. K. (2020). Data transformation: a focus on the interpretation. *Korean Journal of Anesthesiology*, *73*(6), 503–508. https://doi.org/10.4097/kja.20137

MacAulay, M., Macintyre, A. K., Yashadhana, A., Cassola, A., Harris, P., Woodward, C., Smith, K., de Leeuw, E., Palkovits, M., Hoffman, S. J., & Fafard, P. (2022). Under the spotlight: understanding the role of the Chief Medical Officer in a pandemic. *Journal of Epidemiology and Community Health*, *76*(1), 100–104. https://doi.org/10.1136/jech-2021-216850

Mahajan, A. (2016). Post hoc tests in analysis of variance. *Indian Journal of Occupational and Environmental Medicine*, *20*(2), 121. https://doi.org/10.4103/0019-5278.197552

Miles, J. (2005). *R* -Squared, Adjusted *R* -Squared. In *Encyclopedia of Statistics in Behavioral Science*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0470013192.bsa526

Million Death Study Collaborators, Bassani DG, Kumar R, Awasthi S, Morris SK, Paul VK, Shet A, Ram U, Gaffey MF, Black RE, Jha P. (2010). Causes of neonatal and child mortality in India: a nationally representative mortality survey. *The Lancet*, *376*(9755), 1853–1860. https://doi.org/10.1016/S0140-6736(10)61461-4

Molina, G., Weiser, T. G., Lipsitz, S. R., Esquivel, M. M., Uribe-Leitz, T., Azad, T., Shah, N., Semrau, K., Berry, W. R., Gawande, A. A., & Haynes, A. B. (2015). Relationship Between Cesarean Delivery Rate and Maternal and Neonatal Mortality. *JAMA*, *314*(21), 2263. https://doi.org/10.1001/jama.2015.15553

Nanda, A., Mohapatra, Dr. B. B., Mahapatra, A. P. K., Mahapatra, A. P. K., & Mahapatra, A. P. K. (2021). Multiple comparison test by Tukey's honestly significant difference (HSD): Do the confident level control type I error. *International Journal of Statistics and Applied Mathematics*, *6*(1), 59–65. https://doi.org/10.22271/maths.2021.v6.i1a.636

National Family Health Survey (NFHS-5). (n.d.). Retrieved December 10, 2022, from http://rchiips.org/nfhs/factsheet_NFHS-5.shtml

Marudachalam, N. M. (2017). Comparison on Different Data Transformation. *International Journal of Scientific & Engineering Research*, *8*(10). http://www.ijser.org

Pandve, H., & Giri, P. (2015). First referral units (FRU): needs focus in Indian context. *International Journal of Community Medicine and Public Health*, *2*(2), 75. https://doi.org/10.5455/2394-6040.ijcmph20150501

Patel, K. K., & Kumar, M. (2021). Differential and Determinants of Neonatal Mortality: A Comparative Study in Northern and Southern Regions of India. *Indian Journal of Community Medicine : Official Publication of Indian Association of Preventive & Social Medicine*, *46*(3), 405–410. https://doi.org/10.4103/ijcm.IJCM_425_20

Press Information Bureau. (n.d.). Retrieved December 9, 2022, from https://pib.gov.in/PressReleseDetail.aspx?PRID=1796436

Quality Assurance Program | Initiatives | NHM. (n.d.). Retrieved December 10, 2022, from https://nhm.gujarat.gov.in/sqip1.htm

Ross, A., & Willson, V. L. (2017). One-Way Anova. In *Basic and Advanced Statistical Tests* (pp. 21–24). SensePublishers. https://doi.org/10.1007/978-94-6351-086-8_5

Sainani, K. L. (2016). The Value of Scatter Plots. *PM&R*, *8*(12), 1213–1217. https://doi.org/10.1016/j.pmrj.2016.10.018

Sankar, M. J., Neogi, S. B., Sharma, J., Chauhan, M., Srivastava, R., Prabhakar, P. K., Khera, A., Kumar, R., Zodpey, S., & Paul, V. K. (2016). State of newborn health in India. *Journal of Perinatology : Official Journal of the California Perinatal Association*, *36*(s3), S3–S8. https://doi.org/10.1038/jp.2016.183

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear regression analysis: part 14 of a series on evaluation of scientific publications. *Deutsches Arzteblatt International*, *107*(44), 776–782. https://doi.org/10.3238/arztebl.2010.0776

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A Practical Guide to Calculating Cohen's f2, a Measure of Local Effect Size, from PROC MIXED. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00111

Titaley, C. R., Dibley, M. J., Agho, K., Roberts, C. L., & Hall, J. (2008). Determinants of neonatal mortality in Indonesia. *BMC Public Health*, *8*(1), 232. https://doi.org/10.1186/1471-2458-8-232

Trunfio, T. A., Scala, A., Giglio, C., Rossi, G., Borrelli, A., Romano, M., & Improta, G. (2022). Multiple regression model to analyze the total LOS for patients undergoing laparoscopic appendectomy. *BMC Medical Informatics and Decision Making*, *22*(1). https://doi.org/10.1186/s12911-022-01884-9

Turner, S. L., Forbes, A. B., Karahalios, A., Taljaard, M., & McKenzie, J. E. (2021). Evaluation of statistical methods used in the analysis of interrupted time series studies: a simulation study. *BMC Medical Research Methodology*, *21*(1), 181. https://doi.org/10.1186/s12874-021-01364-0

Uyanık, G. K., & Güler, N. (2013). A Study on Multiple Linear Regression Analysis. *Procedia - Social and Behavioral Sciences*, *106*, 234–240. https://doi.org/10.1016/j.sbspro.2013.12.027