Society of Medical Education & Research

2021, Volume 1, Page No: 144-157
ISSN: 3108-4818

International Journal of Social and Psychological Aspects of Healthcare

Factors Influencing Infant Mortality in Gezira State, Sudan: A Survival Analysis Using the Cox Proportional Hazards Model

Sofia Rinaldi[1]*, Marco De Luca[1], Giulia Ferri[1], Paolo Conti[1]

[1]Department of Social Medicine, Faculty of Medicine, Sapienza University of Rome, Rome, Italy.

***E-mail** ✉ sofia.rinaldi@outlook.com

Abstract

Infant mortality serves as a vital measure of overall population health, with sub-Saharan Africa exhibiting the highest rates globally. The objective of this research is to determine the factors linked to infant mortality in Gezira State, Sudan. From July to December 2021, a cross-sectional survey was carried out, including 332 participants chosen through simple random sampling. Information was gathered via a structured questionnaire, and significant predictors of infant mortality were identified using the Cox proportional hazards regression model.

Multiple factors showed significant associations with infant mortality. Infants of mothers with no education or only primary-level education faced an elevated risk of mortality (HR = 3.003, p = 0.0014), whereas secondary education had a protective effect (HR = 0.433, p < 0.0001). Low (HR = 2.527, p = 0.0078) and moderate (HR = 3.109, p = 0.0001) household income levels were linked to higher risks. Delivery at home (HR = 1.684, p = 0.0006), birth size smaller than average (HR = 12.975, p < 0.0001), and prior stillbirth experience (HR = 2.508, p = 0.003) emerged as potent predictors of infant mortality. Furthermore, the age of the mother at first marriage and the overall number of births influenced survival probabilities significantly. To lower infant mortality rates in Sudan, interventions should focus on enhancing maternal education, improving access to healthcare services, and implementing specific measures for the high-risk populations highlighted in this research.

Keywords: Infant mortality, Gezira State, Sudan, Cox proportional hazards model

Introduction

One of the primary objectives of the United Nations Sustainable Development Goals (SDGs) is to decrease neonatal and under-five mortality to no more than 25 deaths per 1,000 live births by 2030. This goal is consistent with Sudan's national strategies for child survival. Research indicates that under-five mortality in areas affected by conflict may be as much as 80 times greater than in stable regions [1]. Infant mortality is broadly acknowledged as a key reflector of societal health and an indicator of a nation's socioeconomic progress [2].

Infant mortality denotes the passing of an infant prior to their first birthday. The infant mortality rate (IMR) is calculated as the number of such deaths per 1,000 live births [3]. On a global scale, the typical IMR stands at around 28 per 1,000 live births. Nonetheless, substantial regional disparities exist owing to variations in socioeconomic circumstances, healthcare infrastructure, and availability of high-quality maternal and child health services [4]. Sub-Saharan Africa records the world's highest IMR, with the World Bank indicating an average of 47 deaths per 1,000 live births in 2020, and certain nations reaching levels up to 100 per 1,000 live births [5]. In Sudan, the IMR stood at 49 deaths per 1,000 live births in 2020, exceeding the regional average for sub-Saharan Africa. Despite improvements from prior figures (such as 66 deaths per 1,000 in previous assessments), infant

Access this article online https://smerpub.com/

Received: 07 March 2021; **Accepted:** 03 June 2021

How to cite this article: Rinaldi S, De Luca M, Ferri G, Conti P. Factors Influencing Infant Mortality in Gezira State, Sudan: A Survival Analysis Using the Cox Proportional Hazards Model. Int J Soc Psychol Asp Healthc. 2021;1:144-57. https://doi.org/10.51847/uUItLTVAWY

mortality continues to pose a major challenge to public health.

A variety of studies have pinpointed numerous contributors to infant mortality, including maternal age, educational attainment, family size, breastfeeding habits, birth weight, birth sequence, use of contraception, paternal education, and birth intervals [6–8]. Further investigations in Ethiopia and additional African nations emphasize the roles of rural living, premature delivery, twin or multiple births, and infant gender in elevating mortality risks [9, 10]. Insufficient antenatal care (ANC), limited maternal schooling, and low family income have been demonstrated to markedly heighten the likelihood of infant death [11, 12]. Moreover, traditional customs, young maternal age at initial childbirth, and the site and standard of delivery care have been associated with child survival results [13–15].

Socioeconomic position is repeatedly identified as a fundamental influencer of infant mortality. There is a clear inverse correlation between maternal education and infant fatalities—mothers with higher education are more inclined to utilize prenatal services, practice beneficial health behaviors, and pursue prompt medical intervention. Likewise, greater household wealth and urban dwelling are tied to improved child health due to enhanced healthcare and sanitation access. In contrast, poverty and rural environments correlate with increased mortality, primarily from restricted availability of trained birth attendants, immunizations, and clean water [16–18].

Gaining insight into the drivers of infant mortality is essential for shaping effective public health policies. Pinpointing vulnerable populations and key contributing elements enables authorities and decision-makers to develop focused initiatives aimed at diminishing infant deaths and advancing maternal and child well-being. In this regard, survival analysis offers an effective approach for analyzing time-to-event outcomes and assessing the effects of various predictors on infant survival.

Accordingly, the present study employs the Cox proportional hazards model to examine the socioeconomic and demographic influences on infant mortality in Gezira State, Sudan.

Materials and Methods

Study design

This research utilized data gathered between July and December 2021 via a standardized and structured questionnaire. The target population consisted of inhabitants from different areas in Gezira State, Sudan. A total of 332 individuals were chosen through simple random sampling, a method deemed suitable given the presumed uniformity of the population [19]. The outcome variable was determined from the answers provided via this approach. All study protocols complied with applicable ethical standards and requirements. Participants, or their legal representatives where necessary, gave informed consent after being thoroughly briefed on the research aims, processes, possible risks, and advantages.

The investigation was an analytical examination that relied on available data to explore elements associated with infant mortality rates in Gezira State, Sudan. Data collection primarily involved a questionnaire. A cluster-based household survey was conducted in Gezira during the period from July to December 2021. The questionnaire was divided into two parts: the first collected details on the respondents, while the second addressed demographic, socioeconomic, and maternal healthcare aspects influencing infant mortality in the region. The final sample comprised 322 respondents. The sampling framework encompassed the whole population of Gezira State. Information was obtained via face-to-face interviews with a questionnaire tailored to meet the research goals, with a particular emphasis on Gezira.

In simple random sampling, the outcome variable can be estimated for any population. When the population is considered homogeneous, sample size determination employs the standard formula for simple random sampling:

$$n_0 = \frac{t^2 pq}{d^2} \qquad (1)$$

Where:

n_0: Preliminary sample size.

t: Standard value associated with the selected confidence level (95%).

d: Allowable margin of error for the population estimate (5%).

p: Anticipated proportion in the population (50%).

Using this calculation, a sample size of 400 was derived. Questionnaires were then administered to potential respondents. Ultimately, 322 completed responses were received, yielding a response rate of 81%. Ethical clearance was granted by the Research Deanship at the University of Gezira following review of the proposal from the Faculty of Economics and Rural Development's Department of Applied Statistics and Demography, Sudan (approval dated 9/2/2021, reference GG/KITR/GITD/21).

Study variables

The variables were categorized to support the study's objectives and selected analytical approach. The dependent variable was defined as the child's survival outcome, classified as either alive or deceased. The independent variables were organized into five primary groups: health-related, maternal, demographic, socioeconomic, and environmental factors. This categorization follows frameworks used in numerous prior studies investigating comparable drivers of child mortality [1–10].

Statistical model

The primary outcome of interest in this study is infant mortality prior to reaching one year of age. Reducing infant mortality was a core target of the Millennium Development Goals (MDGs). It is essential to investigate variations in infant mortality rates across different contexts and track disparities between the richest and poorest populations as progress toward the MDGs is evaluated, in order to identify key socioeconomic, demographic, and geographic influences. A central Sustainable Development Goal (SDG) aim is to decrease neonatal and under-five mortality rates to 25 or fewer deaths per 1,000 live births by 2030, which directly supports Sudan's priorities for child survival. In areas affected by conflict, under-five mortality rates may be as much as 80 times greater than in stable regions. Infant mortality also acts as a robust indicator of overall population health and socioeconomic progress. Consequently, greater emphasis has been placed on exploring the links between child health outcomes and socioeconomic or geographic conditions [20–22].

The study incorporated variables related to poverty as part of broader socioeconomic and geographic influences. The independent variables encompassed socioeconomic, demographic, and geographic elements, such as place of residence, educational attainment of the mother and father, parental occupations, household income, and child characteristics (e.g., birth order or type).

Infants were the unit of analysis. The effects of these contributing factors on the hazard rate can be assessed using Cox proportional hazards regression models. These models are particularly appropriate for survival (time-to-event) data analysis. In the Cox model, the hazard function for an individual is expressed in terms of their survival time t and a vector of associated explanatory variables X. Let $x_1, x_2, \ldots, x_p$ denote the values of the p covariates $X_1, X_2, \ldots, X_p$. The hazard function according to the Cox proportional hazards model is:

$$h(t, X) = h_0\,(t)\,\psi\,(X) \quad (2)$$

where $\psi\,(X) = \exp\left(\sum_{i=1}^{p} \beta_i x_i\right), \beta = (\beta_1, \beta_2, \ldots, \beta_p)$ is $1XP$ vector of regression parameters and $h_0\,(t)$ Is the baseline function at that time? Comprehensive descriptions of the Cox proportional hazards regression model are available in a wide range of sources. In-depth explanations can be found in numerous books and scholarly articles [11–15, 18, 19, 23–26].

The Cox proportional hazards model can be assessed through a variety of diagnostic techniques. Most of these diagnostics are based on different types of residuals. The first approach involves the use of martingale residuals, which serve as the primary method for evaluating the functional form of predictor variables and represent the default diagnostic tool in many implementations. Another key method relies on deviance residuals, which are a normalized transformation of the martingale residuals. This technique is particularly useful for detecting individual observations that are poorly predicted by the model. A third diagnostic option is the Cox-Snell residual, which follows a similar principle to residuals in parametric survival models. Plotting the cumulative hazard function of the Cox-Snell residuals against the residuals themselves provides a visual assessment of overall model fit.

When examining multiple risk factors simultaneously, the SAS procedure PROC PHREG is commonly employed to identify determinants of infant mortality. Before interpreting results, it is critical to evaluate the model's goodness-of-fit. Diagnostic approaches include analysis of residuals, influence diagnostics, testing of distributional assumptions, and identification of outliers [26–28]. Assessing the adequacy of the Cox regression model using martingale residuals is particularly important. In data analysis, choosing the best-fitting model among alternatives is essential for effectively controlling confounding.

Accordingly, a two-stage model-building process was adopted. In the first stage, each predictor variable was entered into the model individually (univariable analysis). Significant covariates were retained and not removed in subsequent steps. Additionally, potential interactions between variables were explored.

The primary outcome of interest was infant mortality occurring before one year of age, treated as a time-to-event outcome. Given its importance as a public health indicator and its direct relevance to global targets such as the Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs), examining socioeconomic and demographic disparities in infant survival is vital, particularly in resource-limited settings like Gezira State, Sudan [20, 21, 23, 24].

The study included a range of explanatory variables encompassing socioeconomic, demographic, and geographic domains, such as urban/rural residence, maternal and paternal education levels, parental occupations, household income, and birth type (singleton or multiple). These covariates were chosen for their established theoretical and empirical associations with poverty and child health outcomes.

The analysis utilized the Cox proportional hazards regression model, which is well-suited for survival data subject to censoring. The model estimates the hazard rate of the event (infant death) at time t, given a set of covariates.

Under the proportional hazards assumption, the hazard function for an individual with survival time t and covariate vector X is specified in the Cox model as follows. Let $x_1, x_2, \ldots, x_p$ denote the values of the p covariates $X_1, X_2, \ldots, X_p$. The hazard function is:

$$h(t, X) = h_0(t)\, \psi(X) \tag{3}$$

where $\psi(X) = \exp\left(\sum_{i=1}^{p} \beta_i x_i\right), \beta = (\beta_1, \beta_2, \ldots, \beta_p)$ is $1XP$ vector of regression parameters and $h_0(t)$ Is the baseline function at that time? The Cox proportional hazards regression model is well-documented across numerous sources, with comprehensive explanations available in a variety of books and scholarly articles [11–26].

Various diagnostic methods were applied to assess the fitted model, relying on different types of residuals. The martingale residual served as the main tool for diagnostics, proving particularly useful in evaluating how well the included covariates fit the data and acting as the standard approach for checking predictor performance. Additionally, the deviance residual—a standardized version derived from the martingale residual—was utilized to pinpoint cases where the model poorly forecasted individual outcomes.

The Cox-Snell residual was another important diagnostic applied, operating on similar principles to those in parametric survival analyses. A visual evaluation of the model's overall fit can be achieved by graphing the cumulative hazard of these Cox-Snell residuals against time. These diagnostic strategies contribute to confirming the fitted model's strength and dependability.

To examine the factors driving infant mortality while controlling for several covariates, a Cox proportional hazards regression was implemented via the SAS PROC PHREG procedure. Prior to result interpretation, the model's suitability was thoroughly checked using techniques like residual checks, influence assessments, verification of underlying assumptions, and outlier identification [22, 25, 26]. Checking the Cox model's assumptions through martingale residuals is vital for confirming its validity and effectiveness.

The modeling process involved two phases. Initially, univariate screening was performed on each potential predictor. Covariates showing significance were included to maintain the model's explanatory strength. Potential interactions were also tested and incorporated where relevant, further improving the model's ability to explain variations. Careful variable selection is key to reducing bias from confounding and enhancing the reliability of conclusions.

The main endpoint of interest was death occurring in the first year of life. A major Sustainable Development Goal focuses on reducing global infant mortality rates. Examining how socioeconomic, demographic, and regional variables contribute to disparities in infant mortality between higher- and lower-income groups is essential for tracking advancement toward these goals. Consequently, increasing attention is being paid to the connections between children's health outcomes and their families' economic and locational circumstances.

Results and Discussion

This research posits that various demographic, biologic, and health service elements—including levels of maternal schooling, feeding methods, gender of the infant, nutrition levels, prior illnesses, and health service availability—play a substantial role in shaping infant mortality risks in Gezira State, Sudan. Through Cox proportional hazards survival modeling, the analysis seeks to pinpoint critical influences on infant survival and quantify the contribution of each to the timing of deaths. These results are anticipated to guide precise interventions and policy decisions aimed at enhancing child health in the area.

Analysis of the dataset provides valuable information on elements affecting survival probabilities, drawn from comparisons between deceased and surviving infants across categories **(Table 1)**. Several aspects demonstrate clear links to improved survival: elevated educational attainment for both parents, occupational status of the father and mother, and successful breastfeeding. For instance, infants born to parents with higher education or to working mothers exhibited greater survival chances. Furthermore, better outcomes were observed among infants free of congenital disorders, boys, those without sibling deaths, and those from pregnancies without prior stillbirths. Strikingly, every infant diagnosed with a congenital condition in the sample succumbed, underscoring the profound impact of such disorders.

In contrast, certain elements—like delivery location, household earnings, mode of birth, and singleton versus multiple birth—did not reach statistical significance in relation to survival within this sample. While household income approached borderline significance, it lacked strong predictive value. This implies that these aspects might exert indirect effects or be moderated by other unobserved influences.

Table 1. Association between infant mortality and demographic variables

Variables	Child Survival status					Alive	Total	Chi-square
	Dead							
	Number	%	Number	%				
Residence	Urban	85	58.2%	61	41.8%	146	0.046
	Rural	128	68.8%	58	31.2%	186	
Total		213	64.2%	119	35.8%	332	
Mother's educational level	Illiterate/Primary	49	65.3%	26	34.7%	75	0.000
	Secondary	92	76.7%	28	23.3%	120	
	University and above	72	52.6%	65	47.4%	137	
Total		213	64.2%	119	35.8%	332	
Father's educational level	Illiterate/Primary	63	75.9%	20	24.1%	83	0.000
	Secondary	96	70.6%	40	29.4%	136	
	University and above	54	47.8%	59	52.2%	113	
Total		213	64.2%	119	35.8%	332	
Father's Occupation	Farmer	16	84.2%	3	15.8%	19	0.006
	Employee	88	56.4%	68	43.6%	156	
	Driver	11	52.4%	10	47.6%	21	
	Free Businesses	98	72.1%	38	27.9%	136	
Total		213	64.2%	119	35.8%	332	
Place of child delivery	Hospital/Health care	145	61.4%	91	38.6%	236	**0.106**
	Home	68	70.8%	28	29.2%	96	
Total		213	64.2%	119	35.8%	332	
Mother Occupation	Housewife	191	67.3%	93	32.7%	284	0.000
	Teacher	10	76.9%	3	23.1%	13	
	Employee	12	34.3%	23	65.7%	35	
Total		213	64.2%	119	35.8%	332	
Family Income	Low Income	71	74.0%	25	26.0%	96	**0.053**
	Average Income	92	61.3%	58	38.7%	150	
	High Income	50	58.1%	36	41.9%	86	
Total		213	64.2%	119	35.8%	332	
Nature of the child	Single	188	62.9%	111	37.1%	299	**0.143**
	Twins	25	75.8%	8	24.2%	33	
Total		213	64.2%	119	35.8%	332	
Genetic diseases	Yes	26	100.0%	0	0.0%	26	0.000
	No	187	61.1%	119	38.9%	306	
Total		213	64.2%	119	35.8%	332	
Sex of the child	Male	126	72.4%	48	27.6%	174	0.001
	Female	87	55.1%	71	44.9%	158	
Total		213	64.2%	119	35.8%	332	
Dead siblings	Yes	68	87.2%	10	12.8%	78	0.000
	No	145	57.1%	109	42.9%	254	

Total		213	64.2%	119	35.8%	332	
Stillbirth	Yes	58	75.3%	19	24.7%	77	0.020
	No	155	60.8%	100	39.2%	255	
Total		213	64.2%	119	35.8%	332	
Kind of delivery	Normal/ vaginal	150	62.8%	89	37.2%	239	**0.279**
	Elective Caesarean delivery	22	59.5%	15	40.5%	37	
	Emergency cesarean delivery	41	73.2%	15	26.8%	56	
Total		213	64.2%	119	35.8%	332	
Child Size	Lower than natural	25	75.8%	8	24.2%	33	0.002
	Natural	171	60.6%	111	39.4%	282	
	Bigger than natural	17	100.0%	0	0.0%	17	
Total		213	64.2%	119	35.8%	332	
The ability of the child to breast or bottle feed	Yes	133	53.6%	115	46.4%	248	0.000
	No	80	95.2%	4	4.8%	84	
Total		213	64.2%	119	35.8%	332	

Selection of variables for the cox proportional hazards model

A two-stage process was employed to select variables for inclusion in the final Cox proportional hazards model. In the initial stage, univariable (bivariate) analyses were conducted to evaluate each potential predictor against infant mortality. Covariates demonstrating a p-value below 0.25 were advanced to the multivariable stage, in line with standard recommendations that use a more liberal threshold to prevent overlooking variables that may gain importance when adjusted for others. In the multivariable model, final retention of variables was determined by statistical significance at $p < 0.05$, their impact on overall model fit, and the evaluation of possible interaction terms. This strategy promotes a parsimonious yet comprehensive model, reliably highlighting key risk factors for infant death.

The dataset originated from an infant mortality survey carried out in Gezira State, Sudan, from July to December 2021. These data were analyzed with a Cox proportional hazards regression model to determine factors linked to infant mortality. The modeling was conducted in SAS version 9.4 via the PROC PHREG procedure, with supplementary options activated for thorough diagnostic evaluation. The outcome variables consisted of the infant's age at death (or censoring time) and survival status.

An extensive array of covariates was examined for their association with infant mortality risk. The predictors entered into consideration included: the infant's capacity to breastfeed or bottle-feed, birth size relative to average, number of previously deceased siblings, household income level, father's education, father's occupation, presence of congenital/genetic disorders, mode of delivery, mother's occupation, mother's education, singleton versus multiple birth, location of delivery, child's sex, prior history of stillbirth, mother's age at first marriage, mother's current age, birth order of the child, and total number of pregnancies/deliveries. Selection of these variables was guided by established theoretical frameworks and empirical findings from earlier research on child survival determinants.

Figure 1 displays the diagnostic plot of Martingale residuals plotted against Deviance residuals. Martingale residuals (shown on the Y-axis) assess model fit at the individual level and characteristically display asymmetry and skewness. Deviance residuals (X-axis), being a symmetric transformation of Martingale residuals, facilitate the identification of outlying observations. The nonlinear, curved pattern evident in the scatter is anticipated given the mathematical relationship between the two residual types. The majority of observations cluster tightly around the smoothed lowess curve (solid line), signifying adequate model fit. No extreme outliers or systematic deviations are apparent, supporting the conclusion that the Cox model's assumptions are appropriately met and the overall fit is satisfactory.

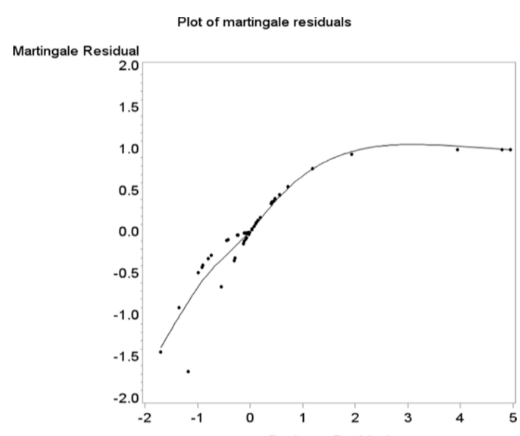

Figure 1. Martingale residual and deviance residual plot

The comprehensive test of the null hypothesis that all coefficients are zero ($\beta = 0$) establishes the model's strong overall significance. Each of the three evaluation methods—Likelihood Ratio ($\chi^2 = 147.30$), Score ($\chi^2 = 116.38$), and Wald ($\chi^2 = 85.06$)—returned p-values under 0.0001. These results firmly reject the idea that none of the predictors matter, confirming that one or more variables meaningfully influence the dependent variable. Across 28 degrees of freedom, this underscores the model's solid explanatory strength and supports detailed review of individual effects.

Findings from **Table 2** relate to the Supremum Test assessing functional form. In the Cox proportional hazards regression, this test—drawing on cumulative martingale residuals—was applied to each basis function for the spline modeling of mother's age at first marriage (Q21). Results indicated that four of the six terms (splineQ212 with p = 0.1440, splineQ213 with p = 0.1390, splineQ214 with p = 0.2000, and splineQ215 with p = 0.0980) displayed no noteworthy departure from expected residual trends, consistent with adequate specification. That said, two terms—splineQ211 (p = 0.0470) and splineQ216 (p = 0.0220)—achieved significance at the 0.05 level, hinting at limited mismatches in the assumed shape. Even so, given the modest scale of these issues, the broader residual behavior, and the clear superiority over a basic linear approach, the spline parameterization of Q21 is judged to capture the true association with the hazard much more effectively.

Table 2. Supremum test for functional form

Variable	Replications	Maximum absolute value	Pr > MaxAbsVal	Seed
splineQ211	1000	2.6861	0.0470	1,275,827,554
splineQ212	1000	2.9379	0.1440	1,275,827,554
splineQ213	1000	3.7371	0.1390	1,275,827,554
splineQ214	1000	2.9847	0.2000	1,275,827,554
splineQ215	1000	3.2364	0.0980	1,275,827,554
splineQ216	1000	2.9379	0.0220	1,275,827,554

The Type III Wald Chi-Square tests indicate that multiple socio-demographic and clinical variables exert a statistically significant influence on the hazard rate of the outcome under investigation. Key predictors demonstrating robust effects include the educational attainment of both parents, the father's occupational category, and household income, with evidence suggesting that elevated parental education and improved socioeconomic position are linked to lower risk. Similarly, clinical variables—such as prior stillbirth experience, delivery location and mode, infant birth size, and gestational duration—emerge as significant, highlighting the critical role of perinatal care and maternal health status. Furthermore, the mother's age at first marriage and her parity (number of prior births) prove influential, pointing to the relevance of reproductive patterns in shaping outcomes.

Conversely, certain factors fail to reach statistical significance in affecting the hazard. Residence type (urban versus rural), multiple birth status (singleton versus twin), and the occurrence of genetic disorders show no independent predictive value for the outcome. The mother's occupation and the child's birth order display only borderline significance, implying potentially limited or conditional roles. These patterns suggest that not every routinely examined socio-demographic characteristic operates as an independent risk modifier, and their contributions may be contingent upon interactions with other elements.

Of particular note are the detected interaction terms, which reveal that the effect of certain predictors is moderated by others. For example, the association involving child sex differs according to residential setting and maternal occupation, illustrating how environmental and social contexts can shape sex-specific risks. Likewise, interactions involving family income with maternal occupation and birth order indicate that financial disadvantage can exacerbate vulnerabilities tied to these factors. Overall, these results underscore the importance of accounting for intertwined risk profiles in designing public health initiatives, since approaches targeting single factors might fail to address the compounded challenges faced by specific vulnerable subgroups.

Table 3. Type III tests for covariates and interaction effects

Effect	P-value	Wald Chi-Square	DF

Residence	0.4910	0.474	1
Mother's educational level	< 0.0001	38.722	2
Father's educational level	0.0190	7.922	2
Father's Occupation	< 0.0001	21.485	3
Mother Occupation	0.0648	5.473	2
Family Income	0.0003	16.164	2
Nature of the child	0.6745	0.176	1
Genetic diseases	0.9945	0.000	1
Sex of the child	< 0.0001	17.078	1
Dead siblings	0.0268	4.903	1
Still Birth	0.0030	8.784	1
Place of child delivery	0.0006	11.717	1
Kind of delivery	< 0.0001	40.019	2
Child Size	< 0.0001	32.417	1
The ability of the child to breast or bottle feed	0.0207	5.355	1
Age of mother at first marriage	< 0.0001	28.160	1
Mother current age	0.0258	4.968	1
The rank of the child	0.0535	3.729	1
Number of deliveries for the mother	0.0086	6.914	1
Duration of pregnancy of the child	0.0049	10.644	2
Residence & sex of the child	0.0101	6.6189	1
Mother's Occupation & sex of the child	0.0090	9.4311	2
The rank of the child & sex of the child	< 0.0001	20.2717	1
Residence & stillbirth	< 0.0001	18.8877	1
Residence & family income	0.0366	6.6159	2
Mother's Occupation & Family Income	0.0003	16.4870	2
Rank of the child & family income	< 0.0001	45.7173	2

This Cox proportional hazards analysis explores how a range of socio-demographic and clinical characteristics affect the risk of the studied event (likely related to child survival) across time. Multiple predictors exhibit clear statistical links to the hazard rate. For example, mothers whose highest education was illiterate or primary level faced substantially elevated risk relative to university-educated mothers (HR = 3.003, $p = 0.0014$), whereas secondary education conferred protection (HR = 0.433, $p < 0.0001$). Fathers with primary education only showed a lower hazard (HR = 0.503, $p = 0.0384$). Household income exerted a notable influence, with low income linked to markedly higher risk (HR = 2.527, $p = 0.0078$) and average income similarly so (HR = 3.109, $p = 0.0001$) when compared against high-income families, underscoring the key role of socioeconomic conditions.

Several clinical and perinatal factors emerged as important drivers. A prior stillbirth markedly raised the hazard (HR = 2.508, $p = 0.003$), and delivery outside a medical facility (e.g., at home) was tied to increased risk (HR = 1.684, $p = 0.0006$). Delivery mode had pronounced associations: planned vaginal birth (HR = 2.249, $p < 0.0001$) and scheduled cesarean (HR = 1.297, $p < 0.0001$) both carried significantly greater hazards than emergency cesarean. Small size at birth stood out as a particularly strong risk factor (HR = 12.975, $p < 0.0001$), highlighting the vital importance of newborn condition. Gestational length also mattered, with pregnancies lasting eight months showing a notably lower hazard (HR = 0.343, $p = 0.0011$), although the specific implication requires careful contextual consideration.

Additional maternal and family history elements played meaningful roles. Later age at first marriage corresponded to higher hazard (HR = 1.708, p < 0.0001), whereas greater current maternal age was linked to decreased risk (HR = 0.894, p = 0.0258). The presence of previously deceased siblings was associated with a substantial reduction in hazard (HR = 0.304, p = 0.0268), possibly due to heightened medical vigilance in affected families. Higher maternal parity indicated elevated risk (HR = 5.076, p = 0.0086), pointing to potential challenges with multiparity. Certain variables, including urban versus rural residence and the infant's capacity for breastfeeding or bottle-feeding, failed to reach statistical significance, suggesting they exert little independent influence in this analysis.

Table 4. Main effects of covariates on infant mortality

Factors	**Category (Reference)**	**Estimate**	**P-Value**	**Hazard Ratio (HR)**	**95% Confidence Interval for HR**
Place of residence	Urban (Ref: Rural)	0.260	0.491	1.297	0.627 – 2.643
Mother's educational level	Illiterate/Primary (Ref: University and above)	1.100	0.0014	3.003	2.095 – 11.594
	Secondary (Ref: University and above)	−0.837	<0.0001	0.433	0.214 – 0.879
Father's educational level	Illiterate/Primary (Ref: University and above)	−0.686	0.0384	0.503	0.068 – 0.929
Father's occupation	Farmer (Ref: Free businesses)	−1.754	<0.0001	0.173	0.162 – 1.563
	Employee (Ref: Free businesses)	0.298	0.1439	1.347	0.304 – 1.190
	Driver (Ref: Free businesses)	0.168	0.2519	1.183	0.130 – 1.709
Mother's occupation	Housewife (Ref: Employee)	−0.191	0.0544	0.826	0.979 – 9.771
	Teacher (Ref: Employee)	−1.325	0.6058	0.266	0.096 – 3.917
Family income	Low income (Ref: Higher income)	0.927	0.0078	2.527	1.589 – 7.574
	Average income (Ref: Higher income)	1.134	0.0001	3.109	2.613 – 8.494
Nature of the child	Single (Ref: Twins)	−0.400	0.6745	0.670	0.336 – 5.405
Sex of the child	Male (Ref: Female)	0.099	<0.0001	1.104	0.124 – 0.475
Dead siblings	Yes (Ref: No)	−1.191	0.0268	0.304	0.059 – 0.897
Stillbirth history	Yes (Ref: No)	0.920	0.003	2.508	2.363 – 67.850
Place of delivery	Home (Ref: Hospital/Health care)	0.521	0.0006	1.684	2.342 – 22.911
Type of delivery	Normal/vaginal (Ref: Emergency cesarean)	0.810	<0.0001	2.249	1.003 – 5.150
	Elective cesarean (Ref: Emergency cesarean)	2.563	<0.0001	1.297	40.544 – 1223.916
Child size at birth	Smaller than average (Ref: Larger than average)	2.563	<0.0001	12.975	1.134 – 48.409
Ability to breastfeed or bottle-feed	Yes (Ref: No)	0.476	0.5465	12.975	0.343 – 7.559
Gestational age	7 months (Ref: 9 months)	−4.050	0.9699	0.017	0.271 – 2.398
	8 months (Ref: 9 months)	−1.070	0.0011	0.343	0.087 – 3975
Age of mother at first marriage	(Continuous)	0.085	<0.0001	1.708	1.402 – 2.082
Mother's current age	(Continuous)	−0.043	0.0258	0.894	0.810 – 0.987
Birth rank of the child	(Continuous)	0.317	0.0535	0.366	0.132 – 1.015
Number of deliveries for the mother	(Continuous)	−0.215	0.0086	5.076	1.512 – 17.038

Various individual socioeconomic and demographic characteristics at the personal level played a key role in determining child survival outcomes. Offspring of mothers with little or no formal schooling faced dramatically increased risks of dying compared to those whose mothers had completed higher education (OR =

25.17, 95% CI: 3.41–185.9). Likewise, kids from lower- or middle-income families showed substantially greater mortality risks than those from wealthier homes (low income: OR = 5.53, 95% CI: 1.59–19.24; average income: OR = 6.95, 95% CI: 2.30–20.98). Home deliveries carried a markedly higher danger (OR = 7.33, 95% CI: 2.18–24.62), while vaginal births (OR = 43.95, 95% CI: 10.93–176.71) and planned cesarean sections (OR = 222.76, 95% CI: 28.70–1728.4) were linked to sharply elevated mortality odds. Infants perceived as small at birth displayed an extraordinarily high risk (OR = 70,401, 95% CI: 209.8 to over 1 million). These patterns emphasize how strongly maternal schooling, family wealth, and childbirth circumstances affect a child's chances of survival.

Analyses of combined factors uncovered important vulnerabilities that illustrate the multifaceted drivers of child deaths **(Table 5)**. In particular, boys in rural communities had much greater mortality risks than girls in the same areas (OR = 6.34, 95% CI: 1.38–29.14), pointing to notable differences tied to gender and location. Such disparities could arise from varying healthcare-seeking patterns, preferential resource distribution within families, or heightened environmental hazards more common in countryside settings.

Family economic status interacted powerfully with place of residence. Young children in city-based poor households experienced an extraordinarily high mortality threat (OR = 100.56, 95% CI: 18.85–536.60), highlighting how urban poverty can block effective use of nearby services due to financial barriers, dense living conditions, or marginalization. Rural poor children also faced elevated dangers (OR = 13.94, 95% CI: 1.44–135.40), though less extreme than in cities. This evidence shows how location and economic disadvantage amplify each other, suggesting that effective strategies must address both aspects simultaneously.

Mothers' work situation combined with household earnings further shaped outcomes. Kids of non-working mothers in mid-level income homes had considerably higher death risks (OR = 15.88, 95% CI: 2.06–122.40) versus those with working mothers in affluent families, possibly due to reduced decision-making power, health awareness, or prompt medical access. Surprisingly, boys with stay-at-home mothers showed dramatically reduced mortality odds (OR = 0.01, 95% CI: 0.001–0.08). Although this might imply intensive maternal care, the extreme estimate and broad confidence range suggest potential issues like limited cases in certain categories, unstable modeling, or specification problems—underscoring the value of robustness checks and methods such as shrinkage estimation.

Certain child traits and family reproductive patterns also mattered. Notably, boys overall had lower mortality risks than girls (OR = 0.24, 95% CI: 0.13–0.42), a pattern that runs counter to typical evidence of greater male biological fragility in infancy. This reversal might reflect societal preferences in care allocation or unaccounted influences in the dataset.

A mother's past childbearing experiences proved influential too. Surprisingly, kids following a sibling's death had much lower mortality odds (OR = 0.06, 95% CI: 0.01–0.76), perhaps because families became more cautious afterward or due to reporting biases in historical data. By contrast, those born after a stillbirth faced far higher risks (OR = 12.66, 95% CI: 2.30–69.86), likely tied to persistent maternal health issues or ongoing barriers to quality delivery care.

Finally, birth sequence interacted meaningfully with economic status. In poorer households, later-born children enjoyed better survival chances (OR = 0.34, 95% CI: 0.20–0.57), possibly from growing parental expertise or adaptive practices—or because only stronger families continued having more children.

Overall, these insights reveal the intricate connections among economic conditions, demographics, biology, and setting in driving child mortality. They call for comprehensive, locally tailored health initiatives tackling wealth gaps, gender influences, maternal well-being, and service access across urban and rural divides. Unusual patterns also stress the need for careful model verification, reliable data gathering, and supportive in-depth community research to better understand local practices and cultural factors shaping child health.

Table 5. Interaction effects of covariates on infant mortality

Interaction	Category	Estimate	Standard Error (SE)	Hazard Ratio (HR)	P-value
Place of residence × Sex of child (Ref: Rural & Female)	Urban, Male	0.843	1.098	2.324	0.4426
	Urban, Female	0.428	0.979	1.534	0.6621

	Rural, Male	1.847	0.818	6.338	0.0240
Mother's Occupation × Sex of child (Ref: Employee & Female)	Housewife, Male	−4.604	1.238	0.010	0.0002
	Housewife, Female	−0.463	1.019	0.630	0.6500
	Teacher, Male	−21.275	965.656	0.000	0.9824
	Teacher, Female	1.156	0.821	3.178	0.1588
Rank of child × Sex of child (Ref: Female)	Male	1.165	0.216	3.207	<0.0001
	Female	0.220	0.210	1.246	0.2946
Residence × Stillbirth (Ref: Urban & No)	Urban, Yes	−0.463	0.602	0.629	0.4415
	Rural, Yes	0.742	0.390	2.100	0.0570
Residence × Family Income (Ref: Urban & High Income)	Urban, Low Income	4.611	0.879	100.558	<0.0001
	Urban, Average Income	1.537	0.900	4.652	0.0876
	Rural, Low Income	2.635	1.272	13.942	0.0383
	Rural, Average Income	1.227	0.866	3.411	0.1566
Mother's Occupation × Family Income (Ref: Employee)	Housewife, Average Income	2.045	1.079	7.729	0.0104
	Teacher, Average Income	−2.737	11.380	0.065	0.9932
Rank of child × Family Income (Ref: High Income)	Low Income	−1.094	0.276	0.335	<0.0001
	Average Income	−1.026	0.229	0.358	<0.0001

Figure 2 displays a Kaplan-Meier survival curve accompanied by its 95% confidence bands. The vertical axis represents the probability of survival, whereas the horizontal axis indicates time elapsed. This plot depicts the estimated proportion of individuals remaining event-free across the follow-up period. The curve starts at 1.0 (100%) and descends in a step-like fashion as events occur, with each vertical drop corresponding to one or more events at that specific time point and flat sections reflecting intervals without any events. The decline appears fairly continuous without prolonged flat periods, implying a relatively steady hazard rate over the observation window.

Plus marks (+) positioned on the curve denote censored cases—participants who left the study, were lost to follow-up, or remained event-free until the study's conclusion. These censored observations do not trigger drops in the curve but do decrease the number of subjects still at risk, influencing subsequent estimates. The Kaplan-Meier estimator properly accounts for such censoring, yielding unbiased survival probabilities even with incomplete data for some participants.

The shaded region around the curve illustrates the 95% confidence interval, indicating the plausible range for the true survival probability at each time point. This band is typically narrower early on, when larger numbers of subjects are still at risk, and widens later as the at-risk population diminishes due to events and censoring. Toward the end of the follow-up period (approximately 12.5 units), the estimated survival probability drops below 0.1 (10%), meaning the majority of subjects had experienced the endpoint by then. Overall, this graph offers a concise yet comprehensive visual summary of the event-time distribution, effectively integrating both event timings and censoring information for reliable inference.

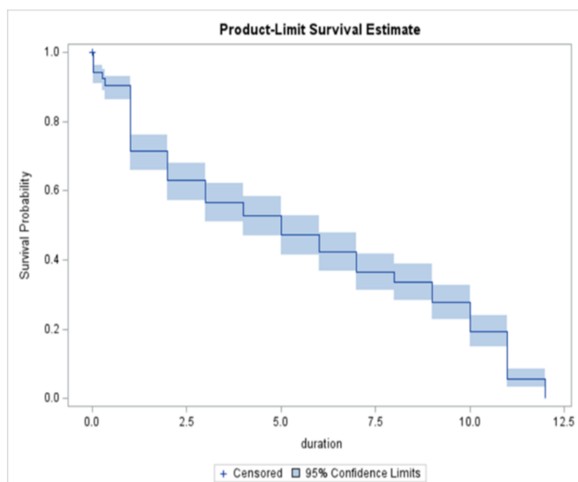

Figure 2. Kaplan-meier survival curve with 95% confidence intervals

Sub-Saharan Africa records the highest probability of infants not surviving their first year, with rates exceeding 70 deaths per 1,000 live births in several countries, reflecting elevated infant mortality levels [19]. In 2023, low-income countries experienced 72 deaths per 1,000 live births among infants. This figure is over 15 times greater than the typical rate of 3.5 deaths per 1,000 live births observed in high-income nations. Addressing preventable child deaths and reducing inequalities across countries remains essential to improve child survival outcomes [23].

In Sudan, infant mortality has shown notable improvement over time. By 2023, the rate stood at 38.04 deaths per 1,000 live births, marking a 2.47% reduction compared to the previous year.

Mortality among children under five in Eastern Africa has decreased over the last quarter-century, with an accelerated decline evident in the most recent decade [24]. This progress in lowering under-five deaths aligns with findings from a 2015 investigation by Ayele, Zewotir, and Mwambi [20]. Nonetheless, rates in Ethiopia continue to be elevated, prompting the application of survival analysis techniques to pinpoint key socioeconomic and demographic influences on child mortality under five.

Data from an infant mortality survey spanning July to December 2021 in Gezira State, Sudan, formed the basis of this research. The analysis explored socioeconomic, demographic, and geographic elements affecting infant mortality in the region. Key determinants were identified through Cox proportional hazards regression, which assesses hazard ratios across different categories of these variables. Although overall trends indicate progress in reducing mortality over recent decades, sub-Saharan Africa maintains the world's highest infant and under-five death rates. Progress has been documented, yet pinpointing the precise socioeconomic, demographic, and geographic contributors to these improvements remains difficult.

The Cox proportional hazards model revealed several factors linked to infant mortality. This approach involves estimating the hazard of an event (such as death) over time, conditional on covariate values. Here, 28 predictors—mixing continuous and categorical types—were examined to predict infant death (a binary outcome variable), with age at death serving as the time metric.

Analysis indicated that certain social, demographic, and economic conditions markedly affect an infant's chances of reaching one year of age. Applying this framework to data from Gezira State highlighted the value of survey-derived insights. Results suggested that infants of mothers with primary-level education or no schooling face greater mortality risks than those born to mothers with university education or above. In contrast, infants of fathers with primary education or less, or those working as farmers, exhibited lower risks compared to cases where fathers had higher education or engaged in independent businesses. Infants with below-average birth weight also showed elevated mortality relative to those of normal or larger size. Additionally, the hazard of infant death rose with increasing maternal age at first marriage.

The analysis also uncovered several counterintuitive patterns requiring further explanation. For instance, lower paternal education and farming occupations were associated with reduced mortality risks, diverging from common expectations linking advanced education and professional roles to improved child health. A prior record of deceased siblings correlated with lower infant death rates, potentially indicating heightened parental awareness or medical engagement. Normal vaginal births carried higher risks than emergency cesareans, possibly due to variations in care quality or facility access. Notably large hazard ratios for planned cesareans and breastfeeding capacity point to potential issues like limited data in certain categories or classification inaccuracies. Such inconsistencies underscore the importance of interpretive caution, while pointing to opportunities for additional research and refinement of the modeling approach.

Examining interaction terms introduced additional layers of complexity to the results. For example, male infants in rural settings emerged as particularly at risk. Similarly, infants born to homemaker mothers in households with moderate income faced elevated mortality hazards, potentially reflecting limited maternal knowledge about health practices or restricted decision-making power regarding medical care. Certain intersections of parental occupation, household income, and infant sex produced unexpected patterns, such as markedly reduced mortality risks for male children of homemakers, pointing to dynamics that could benefit from in-depth qualitative exploration or mixed-methods approaches.

These observations highlight the necessity for comprehensive public health strategies to lower infant mortality rates. Effective measures might include enhancing maternal education levels, increasing access to facility-based deliveries, improving prenatal and postpartum services, and alleviating economic hardship via targeted financial assistance or incentive-based programs. Such efforts must also consider regional inequalities and overlapping vulnerabilities related to gender, residence, and socioeconomic position, through bolstered rural healthcare infrastructure and context-specific initiatives.

A key strength of this research is its application of survival analysis techniques, which effectively manage time-dependent outcomes and incomplete observations, yielding reliable and statistically robust conclusions. Validation through diagnostic graphics and goodness-of-fit evaluations further supports the integrity of the results. However, constraints involve dependence on respondent-provided information, risks of categorization errors, and the omission of system-wide factors like healthcare facility standards or proximity to services.

The outcomes carry substantial relevance for public health planning in Sudan, particularly for efforts to curb infant deaths. Prominent modifiable risks—including limited maternal schooling, low family earnings, non-institutional births, and suboptimal birth characteristics—pinpoint priority domains for action. Decision-makers should focus on broadening educational opportunities for women and fostering health awareness, especially in marginalized areas. Bolstering maternal and child health provisions, with emphasis on professional delivery assistance and hospital-based childbirth, is vital. Moreover, focused aid for vulnerable households and pregnancies at higher risk—such as those with low birth weight or preterm delivery—could markedly enhance survival prospects. Tackling these alterable social and medical determinants offers a pathway for Sudan to achieve meaningful reductions in infant mortality and better overall early-life health.

Conclusion

Ultimately, this investigation sheds critical light on the drivers of infant mortality within Gezira State, with broader applicability to comparable resource-constrained environments. By delineating actionable risk elements and vulnerable populations, the evidence supports the development of fairer, data-driven policies for maternal and child welfare that align with the Sustainable Development Goals (SDGs). The results particularly emphasize the value of professional and facility-based delivery services. Non-institutional births were strongly linked to increased mortality, reinforcing the urgency of encouraging hospital deliveries and guaranteeing prompt, quality obstetric care. Although direct indicators of service access were not captured, patterns tied to delivery practices underscore the need to boost uptake of maternal healthcare.

Strengths and limitations

This research benefits from a solid survival analysis approach centered on Cox proportional hazards regression. This method suitably accommodates time-to-event outcomes and censoring mechanisms, providing accurate hazard estimates across the infant's first year. The inclusion of an extensive array of socioeconomic, demographic, and obstetric predictors facilitates a thorough examination of the multiple influences on infant mortality. Additionally, the substantial sample contributes to greater statistical reliability and applicability of the results to the target population.

That said, certain shortcomings must be acknowledged. The study's reliance on retrospective, observational data restricts the ability to establish causality. Potential inaccuracies in variable classification or memory-related biases may affect self-reported items. Sparse observations in some subgroups could produce imprecise or overly variable hazard estimates. Unaccounted confounders, including variations in service quality or external environmental risks, were also absent from the model. Lastly, the proportional hazards assumption was not comprehensively verified across all predictors, which could compromise the model's appropriateness in isolated instances.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Moucheraud C, Worku A, Molla M, Finlay JE, Leaning J, Yamin AE. Consequences of maternal mortality on infant and child survival: a 25-year longitudinal analysis in Butajira, Ethiopia (1987–2011). Reproductive Health, pp. 1–8, 2015.
2. Finlay JE, Moucheraud C, Goshev S, Levira F, Mrema S, Canning D, Masanja H, Yamin AE. The Effects of Maternal Mortality on Infant and Child Survival in Rural Tanzania: A Cohort Study, *Matern Child Health J*, 2015.

3. Kiross GT, Chojenta C, Barker D, Tiruye TY, Loxton D. The effect of maternal education on infant mortality in ethiopia: A systematic review and meta-analysis. PLoS ONE, 2015.
4. Schellekens J. Maternal education and infant mortality decline: the evidence from Indonesia, 1980–2015. Demographic Res. 2021;45(24):807–24.
5. Tymicki K. Correlates of infant and childhood mortality: A theoretical overview and new evidence from the analysis of longitudinal data of the Bejsce (Poland) parish register reconstitution study of the 18th-20th centuries. Demographic Res. 2009;20(23):559–94.
6. Satti MI, Ali MW, Irshad A, Shah MA. Studying infant mortality: A demographic analysis based on data mining models. Open Life Sci. 2023;18(1):1–10.
7. Wellington O. Determinant of infant mortality rate: A panel data analysis of African countries. Developing Ctry Stud. 2014;4(18):111–5.
8. Fasina F, Oni G, Azuh D, Oduaran A. Impact of mothers ' socio-demographic factors and antenatal clinic attendance on neonatal mortality in Nigeria. Cogent Social Sci. 2020;6(1):1–15.
9. Kumar S, Sahu D. Socio-economic, demographic and environmental factors effects on under-five mortality in empowered action group States of india: an evidence from NFHS-4. Global Health. 2019;9(2):23–9.
10. Wu H. The effect of maternal education on child mortality in Bangladesh. Popul Dev Rev. 2022;48(2):475–503.
11. Hosmer DW, Lemeshow S. Applied survival analysis: regression modeling of time to event data. New York: Wiley; 1999.
12. Kleinbaum DG, Klein M. Survival analysis: A Self-Learning text. New York: Springer; 2012.
13. Emmert-Streib F, Dehmer M. Introduction to survival analysis in practice. Mach Learn Knowl Extr, pp. 1013–38, 2019.
14. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival Analysis Part I: Basic concepts and first analyses, *Br J Cancer, vol. 89, no. 2, pp. 232–238*, 2003.
15. Leung KM, Elashoff RM, Afifi AA. Censoring issues in survival analysis. Annu Rev Public Health. 1997;18:83–104.
16. United Nations. The Millennium Development Goals Report, 2015.
17. Kurinczuk JJ, Hollowell J, Brocklehurst P, Gray R. Infant mortality: overview and context. Natl Perinat Epidemiol Unit Univ Oxf, 2009.
18. Ayele DG, Zewotir TT, Mwambi H. Using Rasch modeling to Re-Evaluate rapid malaria diagnosis test analyses. Int J Environ Res Public Health. 2014;11(7):6681–91.
19. WHO. Infant mortality. World Health Organization; 2021.
20. Mulugeta SS, Muluneh MW, Belay AT, Moyehodie YA, Agegn SB, Masresha BM, Wassihun SG. Multilevel log linear model to estimate the risk factors associated with infant mortality in ethiopia: further analysis of 2016 EDHS. BMC Pregnancy Childbirth, 22, 597, 2022.
21. Baraki AG, Akalu TY, Wolde HF, Lakew AM, Gonete KA. Factors affecting infant mortality in the general population: evidence from the 2016 Ethiopian demographic and health survey (EDHS); a multilevel analysis. BMC Pregnancy Childbirth, 20, 299, 2020.
22. Islam MA, Bassum TRT, Moni MA. Exploring the influencing factors for infant mortality: a mixed-method study of 24 developing countries based on demographic and health survey data. Family Med Prim Care Rev. 2022;24(3):227–36.
23. Bank W. Mortality rate, infant (per 1,000 live births) - Sub-Saharan Africa. The World Bank; 2021.
24. WHO. Newborn mortality. World Health Organization; 2022.
25. QAYYUM A. Comparative analysis of factors affecting. Res J Social Sci. 2015;4(2):1–17.
26. Shobiye DM, Omotola A, Zhao Y, Zhang J, Ekawati FM, Shobiye HO. Infant mortality and risk factors in Nigeria in 2013 – 2017: A population-level study, *eClinicalMedicine*, 2022.
27. Bank W. Mortality rate, infant (per 1,000 live births). World Bank Group; 2021.
28. Bank W. Mortality rate, infant (per 1,000 live births) - Sudan. The World Bank; 2021.