Fenta *et al. BMC Public Health* (2025) 25:2884
https://doi.org/10.1186/s12889-025-24186-x

BMC Public Health

RESEARCH

Open Access

Comparisons of cox semi-parametric and parametric shared frailty models: application for under-five children survival in sub-Saharan Africa

Haile Mekonnen Fenta[1,2,6*], Ding-Geng Chen[1,3], Temesgen T. Zewotir[4], Najmeh Nakhaei Rad[1], Deneke Bitew Belay[1,2] and Seyifemickael Amare Yilema[1,5]

Abstract

Background The under-five child mortality in sub-Saharan African (sSA) countries is a persistent problem with limited effort being made to explore the determinants of disparities across countries and their lower administrative districts. A child's survival may depend on several known and unknown covariates and vary across the study areas. The main objective of this study is to assess the time to death of under-five children and its associated risk factors by comparing the performance of semiparametric and parametric frailty models across sSA regions.

Methods We used a dataset from the Demographic and Health Survey (DHS) across 33 sSA countries. The semiparametric and parametric models with different frailty distributions were used to model the under-five survival time of children across the administrative districts of 33 sSA countries.

Results A total of 330,373 under-five children were included in the study, of whom 19,893 (6.02%) died before reaching their 5th birthday. Unobserved country-level variance (0.421) and district-level variance (0.183) effects considerably impacted the survival time of under-five children in sSA countries. Under-five children born to mothers aged 25–29 and 30–49 were 16% and 20% less likely to die compared to children born to mothers younger than 24 years. Moreover, children born in rural areas were 8.3% more likely to die than those who were born in urban areas. Children who were born from mothers with better access to improved water sources and clean fuel were 9% and 11% less likely to die than their counterparts, respectively.

Conclusions The exponential shared frailty hazard model with lognormal frailty distribution demonstrated better performance compared to the Cox semiparametric model for identifying risk factors for under-five children across sSA countries. Place of residence, wealth index, media exposure, birth order, birth interval, access to improved water, and use of clean fuels for cooking were the significant risk factors on time to death of under-five children in sSA.

Keywords Demographic and health survey, Frailty models, Under-five mortality, Parametric models, Random effects, Unobserved heterogeneity

*Correspondence:
Haile Mekonnen Fenta
hailemekonnen@gmail.com

Full list of author information is available at the end of the article

Introduction

The estimation of the under-five mortality rate (U5MR) is one of the basic priorities for Sustainable Development Goals (SDG) 3.2: reducing the U5MR to at least 25 per 1000 live births by 2030 [1]. To achieve this goal, the United Nations (UN) gives high priority to intervention for low- and middle-income countries (LMICs) over high-income countries. Nationally acceptable estimates can be found in the previous literatures on country level [2–4], but the estimates at the sub-nation (district) level is vital since this is where any potential interventions occur [2, 5]. Although promising progress has been made in improving survival rates of children in Africa through the reduction of diseases and improved access to health facilities, sub-Saharan Africa (sSA) carries about half of the burden of the world's under-five deaths in 2015 [6]. Children's survival prospects continue to be influenced by broad geographical and socioeconomic differences [7–10]. The sSA suffers the greatest burden of child mortality due to limited access to quality healthcare, inadequate nutrition, high prevalence of infectious diseases, and insufficient infrastructure [7, 8].

Survival (time to event) data are usually modelled with Cox proportional hazards (CPH), sometimes called the semi-parametric CPHs regression model, which estimates the effects of the covariates as log hazard ratios [11]. However, alternative parametric approach including exponential, Weibull, log-normal, and log-logistic models [11–13] can also be used and a parametric model sometimes can give more efficient estimates [14, 15] than the semiparametric models. Various studies have been done to compare different survival regression methods, of which some authors proposed parametric models as the most appropriate models [16–24] than others semi-parametric methods like Cox regression [25, 26]. A researcher applied both nonparametric and parametric methods in a survival analysis of patients with gastric cancer and the result of Cox regression and parametric models were almost consistent [17]. The other researchers suggested that log-logistic model best fitted in the analyses of patients with gastrointestinal cancer [23] and Weibull model was selected as the best fitted model in survival time of diabetic nephropathy [20] and acute myocardial infection [21]. In addition, log-normal model revealed an excellent fit to the event of time of retinopathy [18, 22, 24] and regression model showed best fit for the risk factors for lower-limb assumption due to neuropathy [26, 27].

The distribution of the survival model in a homogeneous population, described by the hazard function is the same for each subject. However, the data may present extra variation due to unmeasured and unobservable factors and the inclusion of the frailty term in the hazard distribution function accounts this unmeasured heterogeneity [11–13, 28–30].The significant variance of frailty implies a large heterogeneity across the strata and hence greater correlation among individuals from the same community, but a variance close to zero in the frailty term indicates that there is minimal unobserved heterogeneity between clusters, but this does not imply complete independence of observations within the same cluster [12, 29].

Most of the previous studies had tried to investigate the risk factors of under-five mortality using DHS datasets in different sSA countries [7–10]. Those studies [7–10, 31–33] focused on the risk factors of U5C at country level, and did not address the unobserved heterogeneity at district levels. However, ignoring the frailty term which account such unobserved variability at the district level in the survival model may lead to biased survival estimates of the parameters [29, 30]. Previous studies had a methodological gap in identifying the determinant factors of under-five survival times among children across the sSA countries by incorporating the unobserved heterogeneity at district level. Therefore, this study aims to fill this gap to assess the associated factors for time to death of U5C by incorporating district and country level frailty terms in the survival model by comparing the performance of semi-parametric and parametric shared frailty models.

Data sources and variables

We used the Demographic and Health Survey (DHS), which consists of nationally representative surveys mainly conducted in low-income countries (https://dhsprogram.com). The DHS employs a multistage sampling method to select participants for each survey across various countries. The initial step in this sampling process involves choosing clusters, known as enumeration areas (EAs), followed by a systematic selection of households within those EAs. The clusters are selected from a list of EAs created during the latest population census in each country, and households are randomly chosen within each EA. Women aged 15–49 years from these selected households are then interviewed in detail [34]. Many recent DHS waves incorporate global positioning systems (GPS) coordinates (latitude and longitude) for household clusters, enabling the collection of geospatial data to pinpoint the central location of each EA. The GPS urban and rural locations have been masked for confidentiality reasons as a result, utilizing DHS geo-masking, urban clusters are moved up to 2 km, rural clusters up to 5 km, and 1% of rural clusters up to 10 km [35]. Furthermore, the DHS geo-referenced data displacement process and the geographical variability of the produced data are described in further depth [36, 37]. Additionally, birth record files from under-five children across 33 sSA countries provide comprehensive birth history data for all reproductive women interviewed, along with health indicators related to fertility and mortality rates (Fig. 1; Table 1).

Study variables

Outcome variable

The outcome variable for this study is the time-to-under-five children mortality (from birth up to 59 months of age), quantified by the number of months a child survives after birth. To create the outcome variable for the survival "time to event" analysis [38–40], the children's survival status and their age at death in months or the last month they were confirmed to be alive were combined. Children who died under the age of five years were deemed to have experienced the event and were assigned the number 1, while children who did not die within the specified period were censored (right censored) and assigned the number 0.

Independent variables

The independent variables were extracted based on a review of the earlier studies [31, 38–48]. The variables included in the analysis are summarized in Fig. 2.

Methodology

Suppose that the covariates are represented by $\boldsymbol{X}$ and let T be the non-negative random variable representing an individual's survival time, with t being a realization of random variable T. Then the Cox Proportional Hazard (CPH) model is:

$$h(t, \boldsymbol{X}) = h_0(t)\, exp\left(\boldsymbol{X}^T\boldsymbol{\beta}\right) \quad (1)$$

where $\boldsymbol{X} = [X_1, X_2, \ldots, X_p]$ and $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_p]$ are the covariate and regression parameters respectively, $h_0(t)$ is the baseline hazard function, when all the covariates are equal to zero [11, 29, 33].

The estimation methods for this model include the partial Likelihood Eq. 2 and Log- partial Likelihood Eq. 3 respectively [28]. Assume that there are n individuals and D uncensored (observed) events at $t_1 < t_2 < \ldots < t_D$..

$$L(\beta) = \prod_{\rangle=1}^{D} \frac{exp\left(\S_{\rangle}^{T}\beta\right)}{\sum_{| \in R(\sqcup_{\rangle})} exp\left(\S_{|}^{T}\beta\right)} \quad (2)$$

$$\ell(\beta) = \sum_{i=1}^{D}\left[\boldsymbol{x}_i^T\boldsymbol{\beta} - log\left(\sum_{j\in R(t_i)} \exp\left(\boldsymbol{x}_j^T\boldsymbol{\beta}\right)\right)\right] \quad (3)$$

Where i index the individual with an event at time t_i and $R(t_i)$ be the risk at time t_i that individuals still under observation just before t_i. This maximizes with respect to $\boldsymbol{\beta}$ using numerical optimization (e.g., Newton-Raphson) [49].

The cumulative hazard function is the integral of the hazard function up to time t and different approaches are used to estimate the cumulative hazard functions [33]. In the CPH model, the unknown hazard function h_0 is the nonparametric part whereas the unknown $\boldsymbol{\beta}$ is the parametric part, which together makes a semi-parametric model and the baseline hazard function does not assume any specific distributions [11, 12, 29,

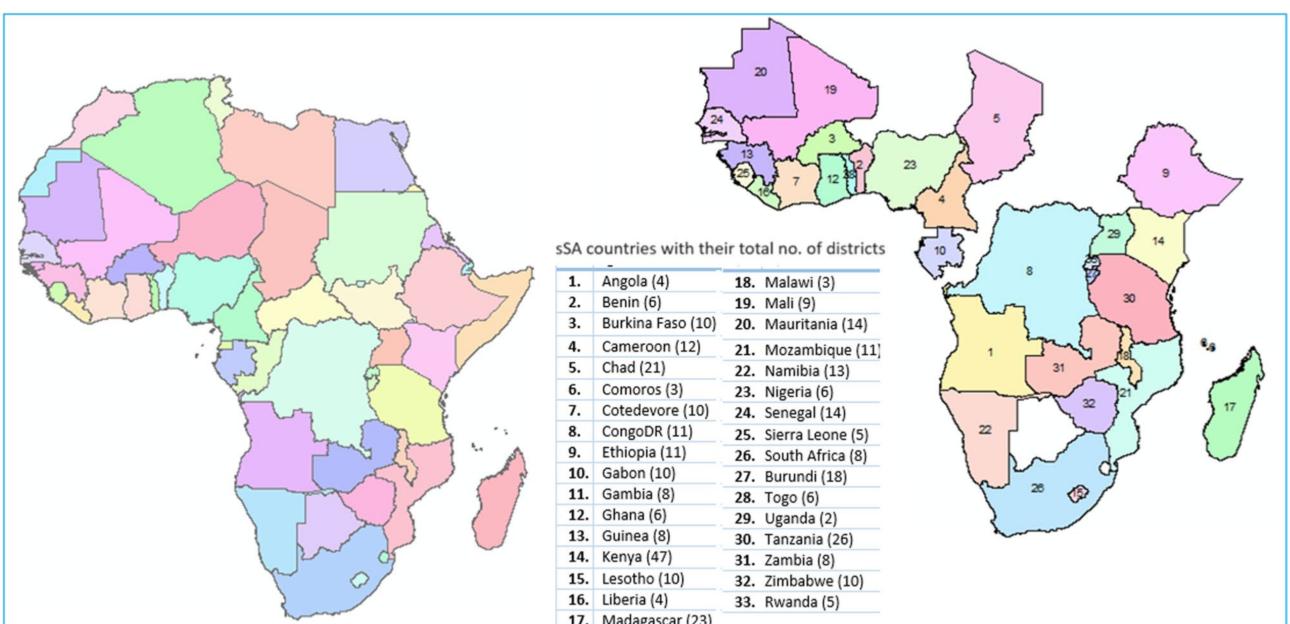


Fig. 1 Eligible sub-Saharan African countries and their total number of lower-level administrative areas included in the study. Source: Authors drawings

Table 1 Selection of study participants from 33 sSA countries with recent DHS reports

A total of 49 countries are located in Sub-Saharan Africa			
East African regions 18 countries	West African regions 17 countries	Central Africa regions 9 countries	Southern Africa regions 5 countries
A total of 16 countries were excluded for the following reasons			
6 countries were excluded. • 3 countries no DHS report • 3 countries no GPS is available	4 countries were excluded. • 1 country with no DHS report • 3 countries no GPS available	4 countries were excluded. • 1 country with no DHS report • 3 countries no GPS available	2 countries were excluded. • 2 countries where no GPS is available
A total of 33 countries included			
East African regions • 12 countries (Burundi, Comoros, Ethiopia, Madagascar, Malawi, Mozambique, Rwanda, Tanzania, Uganda, Zambia, Zimbabwe, and Kenya) • 167 Districts • 7,148 PSU • 119,515 under-five children	West African regions • 13 countries (Benin, Burkina Faso, Gambia, Ghana, Guinea, Ivery Coast, Liberia, Mali, Mauritania. Nigeria, Senegal, Sierra Leone, and Togo • 106 Districts • 6,453 PSU • 132,032 under-five children	Central African regions • 5 countries (Angola, Cameroon, Chad, Democratic Republic Congo, and Gabon) • 58 Districts • 2,565 PSU • 67,094 under-five children	Southern African regions • 3 countries (South Africa, Lesotho, and Namibia) • 31 Districts • 1,605 PSU • 11,732 under-five children
A total of • 33 countries • 362 Districts • 17,771 PSU • 330,373 under-five children			

PSU primary sampling unit (Enumeration areas) where the GPS location is collected

33]. However, for the parametric approach, the baseline hazard is defined as a parametric function and the vector parameters, say φ, are estimated together with the regression parameter(s) in the model [11, 15, 29, 33]. The common parametric distributions for the hazard baseline include Weibull, exponential, Gompertz, log-normal, and log-logistic distributions [13, 14, 33] which are summarized in Table 2.

The frailty models are equivalent to mixed effects (random effect) models in survival analysis [13, 14, 33]. The frailty model includes the unobserved effects which are represented by ,and included in equation 1 as:

$$h\left(t,\,\boldsymbol{x}_{ij},\,u_i\right) = h_0\left(t\right) exp\left(\boldsymbol{x}_{ij}{}^T\boldsymbol{\beta}\ + u_i\right). \tag{4}$$

$$h\left(t,\,\boldsymbol{x}_{ij},\,u_i\right) = h_0\left(t\right) exp\left(\boldsymbol{x}_{ij}{}^T\boldsymbol{\beta}\ + u_i\right). \tag{4}$$

Let $Z = exp\left(u_i\right)$ then

$$h\left(t,\,\boldsymbol{x}_{ij},\,u_i\right) = \boldsymbol{Z}h_0\left(t\right) exp\left(\boldsymbol{x}_{ij}{}^T\boldsymbol{\beta}\ \right), \tag{5}$$

where u_i is the frailty term of all subjects in the group, and $\boldsymbol{x}_{ij}$ the vector of covariates for subject $j = \{1,\ .\ .\ .\ ,\ n_i\}$, in group $i = \{1,\ .\ .\ .\ ,\ G\}$.

The frailty, $\boldsymbol{U}$ has different distributions [50] and in this paper, we focus on gamma, log-normal, positive stable, and inverse Gaussian frailty distributions. The frailty models are classified in two different forms (univariate and multivariate). In the univariate context, the frailty model introduces an unobservable multiplicative effect of $\boldsymbol{z}$ on the hazard, so that conditional on the frailty as:

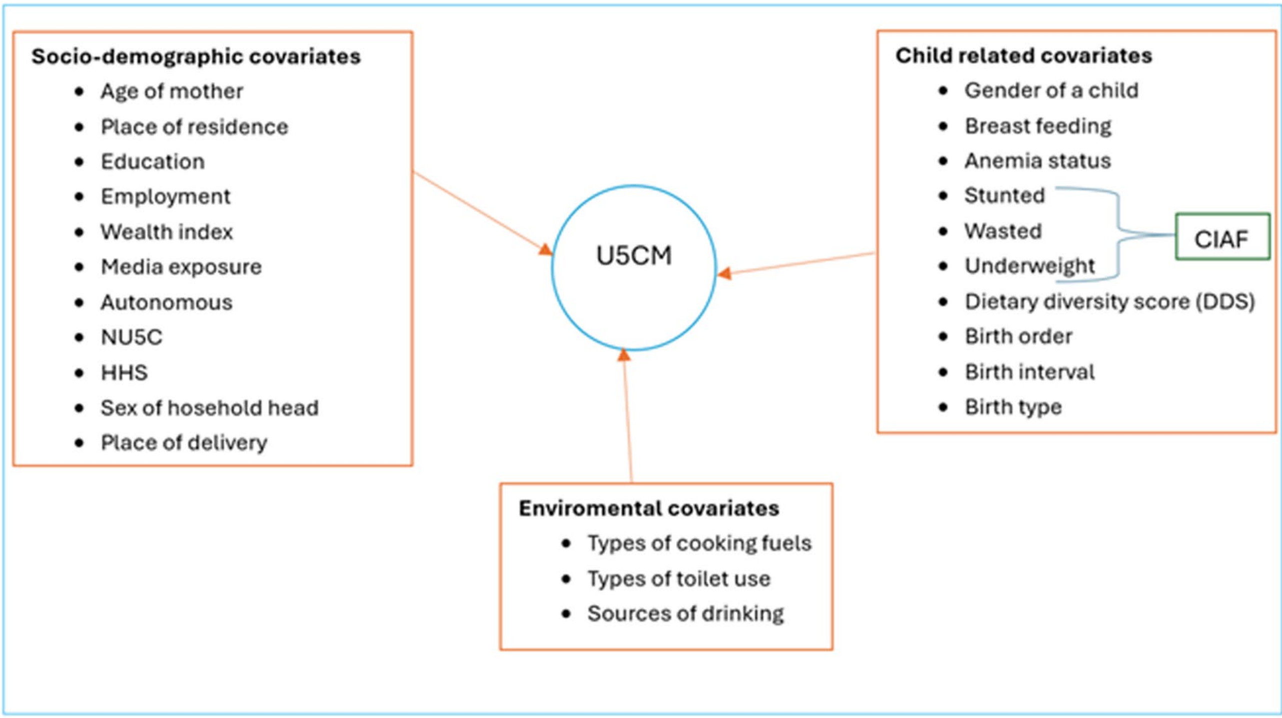
$$h\left(t|z\right) = zh\left(t\right), \quad z > 0 \tag{6}$$

,

where $\boldsymbol{z}$ is some random quantity assumed to have a unit mean and a constant variance θ. However, in the multivariate survival model, which is an extension of univariate case, individuals are allowed to share the same frailty value. This indicates that a frailty value also generates dependence between those individuals who share frailties, whereas conditional on frailty, those individuals are independent [13, 14, 50]. Let the data consist of n stratum with the i^{th} strata comprised of n_i individuals $(i = 1,\ .\ .\ .\ ,G)$ given as:

$$h_{ij}\left(t|z_i\right) = z_i h_{ij}\left(t\right), \tag{7}$$

where $h_{ij}\left(t\right) = h\left(t|x_{ij}\right)$. This means, that any member of the i^{th} strata, the standard hazard function is multiplied by the shared frailty term z_i. Before we fit the standard CPH model, we have to check the proportional hazard (PH) assumptions [11, 50]. The model assumes that the hazard of the different clusters formed by the level of the covariates is proportional [11, 12, 50]. If the PH assumption does not hold, the results from a CPH model are misleading, and alternative approach strategies should be carried out [11, 12, 50, 51]. To check the CPH assumption, we can use the Kaplan-Meier plot but this graphical approach may be insufficient in cases where the violation of the assumption is marginal [52]. Grambsch and Therneau [53] presented a goodness of fit testing



CIAF: composite index for anthropometric failure [49] and DDS: dietary diversity score [49, 50]

Fig. 2 Conceptual framework for variables description

Table 2 The hazard and cumulative hazard functions of the common parametric distributions

Distribution	$h_0(t)$	$H_0(t) = \int_0^t h_0(s)\,ds$	Parameter space
Exponential	λ	λt	$\lambda > 0$
Weibull	$\lambda \rho t^{\rho-1}$	λt^{ρ}	$\lambda, \rho, > 0$
Gompertz	$\lambda \exp(\gamma t)$	$\frac{\lambda}{\gamma}(\exp(\gamma t) - 1)$	$\lambda, \gamma > 0$
log-normal	$\frac{\varphi(\log(t)-\mu/\sigma)}{\sigma t[1-\varphi(\log(t)-\mu/\sigma)]}$	$-log[1 - \Phi(\log(t) - \mu/\sigma)]$	$\mu \in R,\ \sigma > 0$
log-logistic	$\frac{\exp(\alpha) k t^{k-1}}{1+\exp(\alpha) t^k}$	$\log(1 + \exp(\alpha) t^k)$	$\alpha \in R,\ k > 0$

φ and Φ respectively denote the probability density and cumulative hazard functions for each of these distributions

approach, which gives a test statistic for checking the PH assumption.

The estimation method includes the Log-Likelihood approach presented in Eq. 8) Assume $i = 1, \ldots, G$ index clusters and $j = 1, \ldots, n_i$ index individuals in group i, observed time (T_{ij})event indicator (δ_{ij}) $1 = event,\ 0 = censored)$, hence the individual likelihood contribution given the frailty z_i [29, 49] is:

$$L_{ij}(z_i) = \left[z_i h_0(t_{ij})\ exp\left(\boldsymbol{x}_{ij}^T\boldsymbol{\beta}\right)\right]^{\delta_{ij}} .exp\left[-z_i H_0(t_{ij})\, exp\ \left(\boldsymbol{x}_{ij}^T\boldsymbol{\beta}\right)\right] \tag{8}$$

,

the group level contribution:

$$L_i = \int_0^{\infty}\left[\prod\nolimits_{j=1}^{n_i} L_{ij}(z_i)\right] f(z_i;\boldsymbol{\theta})\, dz_i. \tag{9}$$

where $f(z_i;\boldsymbol{\theta})$ is the density of the frailty distribution and $\boldsymbol{\theta}$ is the dispersion parameter.

Moreover the flexible parametric survival models such as Royston and Lambert (ref) and the Exponentiated Exponential Sinh Cauchy [54, 55]can be used. These models are flexible parametric model (FPSM) which extends beyond the Cox proportional hazards models by modelling the baseline hazard or log cumulative hazard directly, using the restricted cubic splines (natural splines).

Identifiability of frailty models

It is a foundational and subtle issue in the analysis of survival data, particularly when the unobserved heterogeneity is incorporated via random effects (frailty term) [29, 55, 56]. This model uniquely estimates model parameters from the observed data and a model is identifiable if different values of the parameters lead to different distributions of the observed data.

Dependence of shared frailty model

The correlation between any two event times from the same cluster is measured using Kendall's tau (τ). This is measured from the frailty term (θ) [57] given as.

Kendall's tau $(\tau) = \frac{\theta}{\theta + 2}$,

where $\tau \in (0, 1)$ and θ is the frailty distribution parameter which provides information on the heterogeneity (variability) of the population of clusters (strata). The larger the value of θ , the higher the degree of

Model selection

To compare the group of parametric frailty and semiparametric frailty models, Akaike Information Criteria (AIC) and Bayesian Information Criterion (BIC) were used [58]. The model with the smaller AIC fits the data better than the model with the larger AIC [51, 52]. All the statistical analysis were carried out with R software and the statistical significance level was set at p-value < 0.05.

Ethical consideration

The study conducted was a secondary analysis of publicly available survey data obtained from the MEASURE DHS program. The MEASURE DHS Program is a global initiative that collects and analyzes data on population, health, and nutrition in developing countries. Funded by USAID, it provides key information on topics such as maternal and child health, family planning, and HIV, supporting evidence-based policymaking and health program development. Since no original data collection was involved, the research was exempted from obtaining ethical approval and participant consent. Permission to access and utilize the data was granted by [www.dhsprogram.com]. It is essential to mention that the datasets contained no personal information, such as names or household addresses, to safeguard the privacy of the individuals involved.

Results

This study included a total of 330,373 under-five children (U5C), with time-to-death as the primary outcome of interest. A total of 19,893 (6.02%) U5C died before celebrating their fifth birthday, of whom 14,719 (74%) resided in rural areas. Most women fell within the 30–49 age group, and a significant proportion had completed their primary education or higher, 201,718 (61.1%>50%). A notable number (227,695) of these women were employed, and the majority of children came from poor households. Moreover, a significant percentage of dead children (6.40%) were from women who had low levels of

Table 3 Background characteristics and percentage distribution of under-five mortality by survival determinants of sampled children

Variables	Categories	*N* (Total)	Child if dead *n* (%)	Variables	Categories	*N* (Total)	Child if dead *n* (%)
Woman's age	15–24	94,061	5,883 (6.25)	Sex of child	Male	167,124	10,893 (6.52)
	25–29	87,859	4,870 (5.54)		Female	163,249	9,000 (5.51)
	30–49	148,453	9,140 (6.16)	NU5C	1–2	244,113	16,773 (6.86)
Mother's educ.	Illiterate	128,405	9,035 (7.04)		3–4	71,422	2,536 (3.55)
	Primary	106,288	6,333 (5.96)		5+	14,838	584 (3.94)
	Secondary	83,288	4,133 (4.95)	Breastfeed	No	146,633	13,119 (8.95)
	Higher	12,142	392 (3.23)		Yes	183,740	6,774 (3.69)
Occupation	Yes	227,695	14,719 (6.46)	DDS	< minimum	165, 589	6,582 (3.97)
	No	102,678	5,174 (5.04)		Minimum	164,784	13,311 (8.08)
Place residence	Urban	104,212	5,53 (5.04)	Birth order	First	74,685	4,440 (5.94)
	Rural	226,161	14,640 (6.47)		2nd –3rd	116,820	6,017 (5.15)
Wealth index	Poorest	84,234	5,759 (6.84)		>=4th	138,868	9,436 (6.79)
	Poorer	71,638	4,816 (6.72)	Birth-type	Single	318,839	17,666 (5.54)
	Middle	67,066	4,024 (6.00)		Multiple	11,534	2,227 (19.31)
	Richer	58,960	3,160 (5.36)	Birth interval	< 18 months	95,808	6,941 (7.24)
	Richest	48,475	(2,134 (4.40)		18–23 months	32,570	2,832 (8.70)
Media exposure	Yes	194,549	10,709 (5.50)		$\geq$ 24 months	201, 995	10,120 (5.01
	No	135,824	9,184 (6.76)	Toilet	Unimproved	101,478	6,772 (6.67)
Autonomous	Low	123,834	7924 (6.40)		Improved	228,895	13,121 (5.73)
	Medium	112,753	6,853 (6.08)	Water	Unimproved	99,551	7,014 (7.05)
	High	93,786	5,117 (5.46)		Improved	230,822	12,879 (5.58)
Sex of household head	Male	258,931	16,071 (6.21)	Fuel use	Unclean	290,688	18,301 (6.30)
	Female	71,442	3,822 (5.35)		Clean	39,685	1,592 (4.01)
Place of delivery	Home	101,151	8,052 (7.96)				
	Health fa.	229,221	11,841 (5.17)				

N number of children in the analysis, *DDS* dietary diversity score, *NU5C* number of under five children

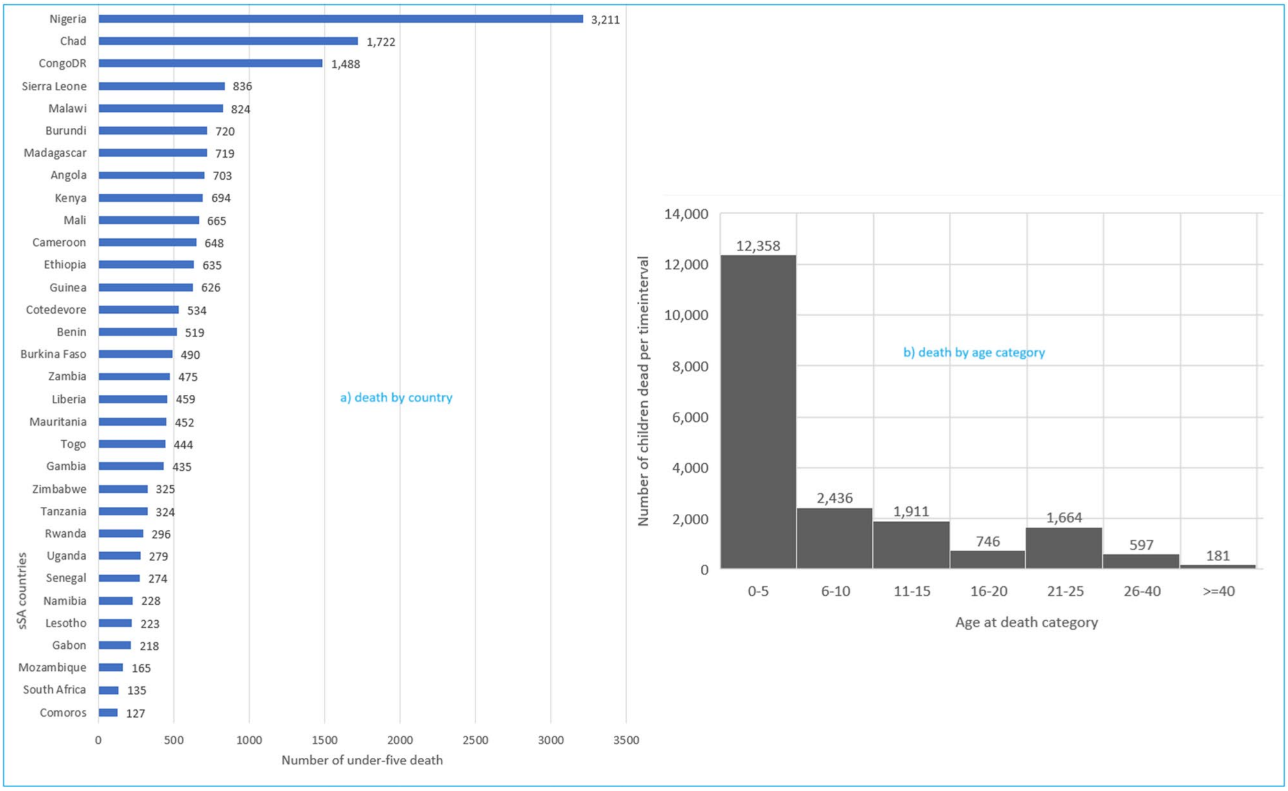


Fig. 3 Distribution of deaths by (**a**) country (**b**) by age groups (in months)

Table 4 Testing the proportional hazard assumption using the scaled Schoenfeld residuals

Variables	Global		Variables	Global	
	X^2	*p*-value		X^2	*p*-value
Mother's age	4.02	0.13	Breastfeed	4.44	0.035
Mother's educ.	2.52	0.47	DDS	7.24	0.007
Occupation	4.06	0.044	CIAF	8.45	< 0.001
Place residence	2.09	0.15	Birth order	2.04	0.36
Wealth index	5.37	0.25	Birth-type	3.0	0.083
Media exposure	0.417	0.52	Toilet	5.03	0.025
Autonomous	3.78	0.15	Water	1.71	0.19
Sex of household	6.5	0.011	Fuel use	3.37	0.066
Place of delivery	8.52	0.004	NU5C	17.7	< 0.001
Sex of child	9.17	0.003	Birth interval	19.09	< 0.001

CIAF Composite index for anthropometric failure

decision-making power regarding their healthcare, and a small proportion (5.50%) were from women who had media exposure. The majority of the under-five deaths were associated with unimproved sources of drinking water, unimproved toilet facilities, and unclean use of fuels (Table 3).

Figure 3 displays the number of under-five deaths among children by country (Fig. 3a) and by age category (Fig. 3b). Nigeria had the highest number of under-five deaths, with 3,211, followed by Chad (1,722), and the Democratic Congo Republic (1,488), suggesting a need for targeted interventions to increase awareness and promote strategies to reduce under-five mortality in these areas. In contrast, Comoros (127), South Africa (135), and Mozambique (165) had a lower number of under-five deaths (Fig. 3a). Moreover, the distribution of the number of under-five deaths among children by age group is also summarized in Fig. 3b. The number of deaths was greater in the lower age group than in the others.

Table 4 shows the scaled Schoenfeld residual test results. The variables, including the mother's age, mother's education, place of residence, wealth index of the household, media exposure of the household, number of under-five children in the house, birth order of the child, birth type of the child, use of improved water and use of cleaned fuels were satisfied the proportional hazard assumptions (p-values were greater than 0.05) and they were considered in the multiple covariates semiparametric CPH model and parametric models.

The frailty model (both Cox semiparametric and parametric) was found to fit better than the PH (without frailty), which indicates that district level unobserved random effects had a considerable impact on time to death of U5C across sSA countries. The model selection criteria (AIC and BIC) values for the CPH model were the highest, and the exponential hazard baseline distribution with log-normal frailty distribution was the best model since it has the lowest BIC and AIC values (Table 5).

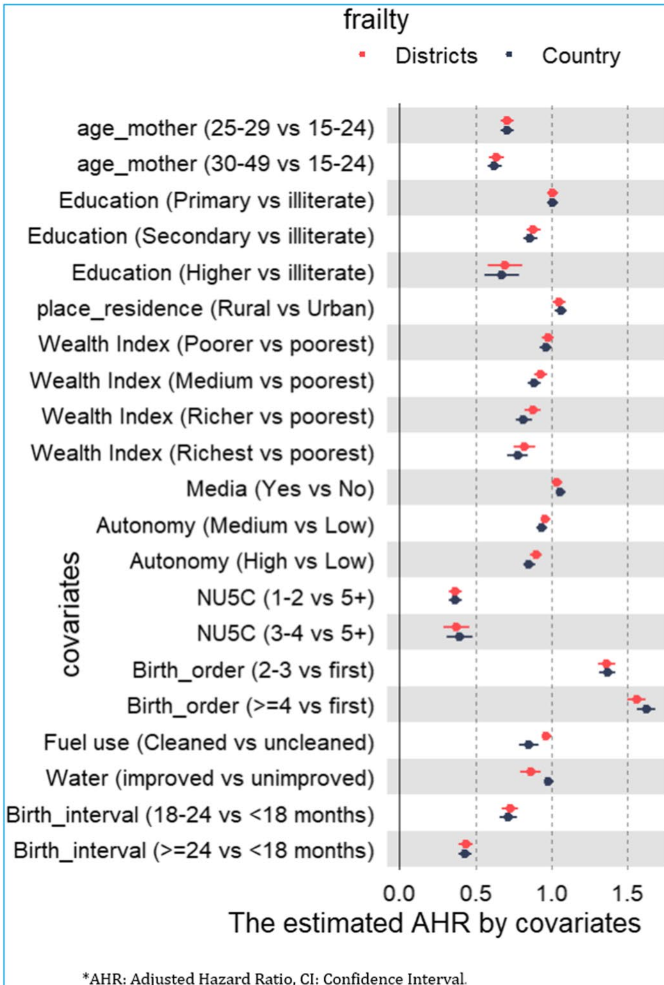
Table 5 The comparison of the AIC between the Cox proportional hazard model and shared frailty parametric model

Model types	Baseline hazard distribution	Frailty distribution	AIC	BIC	LR
Semiparametric	Cox	NA	34166.15	34167.50	−17080.075
		Gamma	31189.4	31194.75	−15593.7
Flexible parametric	Roston and Mambert	NA	165689.3	165753.4	−82838.7
Flexible parametric	Roston and Mambert	Gamma	161,296	161541.5	−80,625
Parametric	Exponential	Gamma	14426.89	14451.44	−7210.45
		Lognormal	14423.66	14448.2	−7208.83
		I. gaussian	14424.69	14449.23	−7209.34
		Stable positive	14445.91	14470.46	−7219.96
		NA (UAPHM)	14574.21	14590.58	−7285.11

NA no specific distribution (without frailty term), *AIC* Akaike Information Criteria, *BIC* Bayesian Information Criteria, *LR* Loglikelihood Ratio, *I.gaussian* inverse gaussian

The exponential baseline hazard model with log-normal frailty was found to fit the dataset better than the CPH model and other exponential parametric models with different frailty distributions, indicating that country (district) level unobserved random effects impact on the survival of under-five children. The heterogeneity (unobserved) in the population of the country (district), which was used as a cluster, was estimated by the exponential baseline hazard model with log-normal frailty parameters $(\sigma^2 = 0.421 \,\&\, 0.183 \quad \text{and} \quad \lambda = 0.033 \,\&\, 0.032)$ within the countries (districts) was measured by Kendall's tau at $\tau = 0.163 \,\&\, 0.082$, respectively. The result revealed that the frailty component had a significant contribution to modelling the survival of U5C across the sSA countries. After controlling for the district-level frailty term, the results from the exponential parametric baseline hazard distribution revealed that maternal characteristics such as age, educational status, and decision-making ability were statistically significant covariates for under-five death. Moreover, household characteristics such as place of residence, wealth index, media exposure, access to improved water, use of clean fuel for cooking, and the number of under-five children in the household were also important covariates for predicting under-five child death. Finally, child-level characteristics such as birth order, and birth intervals were also significant variables for under-five deaths. Specifically, the estimated hazard ratios of under-five death for children born to mothers aged 25–29 and 30–49 were (HR:0.84, 95% CI: 0.81–0.88, HR=0.8, 95% CI: 0.76–0.84) higher respectively compared with the baseline age (15–24 years). This revealed that lower-age pregnancy was at a risk for under-five mortality implying that children born to mothers aged 25–29 and 30–49 were 16% and 20% less likely to die compared with children born from mothers younger than 24 years. Moreover, children born from women with high autonomy were 8.5% (AHR=0.941, 95% CI: 0.91–0.98) less likely to be died than those from low autonomy levels. Children who were born in rural areas were 8.3% (AHR=1.083,95% CI: 1.04–1.13) more likely to die than those born in urban areas. Women who had long birth intervals (spacing) such as 18–23 months and > = 24 months had a lower AHR [28%, 95% CI:0.689, 0.769 and 56%, 95% CI: 0.417, 0.456], respectively of under-five mortality compared to women with birth interval of less than 18 months. Moreover, under-five children born to mothers with 1–2 and 3–4 parities are around [63.5%, 95% CI:0.349, 0.381] and [62.4%, 95% CI: 0.0.345, 0.410], respectively were less likely to face the risk of mortality than those mothers of 5+ parities. Children born in households that used improved water sources and uncleaned fuel use were 9% (AHR=0.90, 95% CI: 0.87–0.93, 11% (AHR = 0.895:95% CI: 84-0.96) less likely to die than their counterparts, respectively. The risk of death among under-five children with birth intervals of more than two years (> 24 months) was 54% (AHR=0.464, 95% CI: 0.396- 0,544) less likely to die than those with birth intervals of less than 18 months, but children with a birth interval of 18–23 months had a similar risk compared with those who had less than 18 months of birth interval (Fig. 4). Moreover, a frailty model that considered country or districts as a frailty term is summarized in Figs. 5 and 6. The result showed that estimates of the frailty variance are 0.42 [95% CI:0.17, 0.68] (country level) and 0.18 [95% CI: 0.15, 0.22] (district level), indicating that under-five survival time varies more across countries than across districts. In addition, estimates of Kendall's tau are 0.163 (country level) and 0.082 (district level) for frailty models.

Predicted frailty plots Fig. 5 for countries and Fig. 6 (Additional file 1) for districts indicate an unobserved term (frailty) exists. Predicted values above one indicate, an increasing risk while those below one indicate a decreasing risk. Figure 5 reveals that countries like Kenya, Ghana, Cote d'Ivoire, Guinea, and South Africa had the lowest risk of death, respectively. However, countries such as Togo, Lesotho, Mauritania, Senegal, and Benin had the highest risk of under-five mortality in sSA, respectively.



District-level AHR (95% CI) ●	District-level AHR (95% CI) ●
0.712 (0.682, 0.743)	0.707 (0.677, 0.738)
0.636 (0.606, 0.668)	0.627 (0.597, 0.658)
1.008 (0.971, 1.046)	1.010 (0.973, 1.048)
0.882 (0.841, 0.924)	0.862 (0.824, 0.901)
0.696 (0.622, 0.778)	0.673 (0.602, 0.753)
1.051 (1.009, 1.095)	1.065 (1.024, 1.108)
0.979 (0.942, 1.018)	0.964 (0.927, 1.002)
0.930 (0.889, 0.972)	0.888 (0.850, 0.927)
0.880 (0.835, 0.928)	0.820 (0.779, 0.862)
0.823 (0.768, 0.881)	0.778 (0.728, 0.832)
1.038 (1.004, 1.073)	1.061 (1.026, 1.097)
0.961 (0.929, 0.993)	0.937 (0.906, 0.969)
0.899 (0.867, 0.934)	0.856 (0.824, 0.888)
0.365 (0.349, 0.381)	0.368 (0.352, 0.384)
0.376 (0.345, 0.410)	0.395 (0.363, 0.430)
1.365 (1.292, 1.442)	1.369 (1.296, 1.446)
1.562 (1.467, 1.663)	1.626 (1.527, 1.731)
0.969 (0.937, 1.001)	0.850 (0.797, 0.907)
0.865 (0.809, 0.925)	0.983 (0.951, 1.016)
0.728 (0.689, 0.769)	0.715 (0.677, 0.756)
0.436 (0.417, 0.456)	0.430 (0.411, 0.450)

	Districts	Countries
Sigma2 (σ^2)	0.183 (0.148, 0.218)	0.421 (0.166, 0.676)
Lambda (λ)	0.032 (0.03, 0.034)	0.033 (0.025, 0.041)
Kendall's Tau (τ)	0.082	0.163

Fig. 4 Comparisons of cox PH frailty model with exponential parametric frailty models among under-five children across sSA countries

Moreover, the frailty estimate for each of the local districts is mapped in Fig. 6. The green colour corresponds to the lowest risk of death among children and the red colour towards the highest risk of death. As a result districts in Chad (*N'Djamena* and Logone Occidental), Nigeria (northwest and northeast), Guinea (Konkan), and the Democratic Congo (Katanga and Sud-Kivu) had the highest risk of under-five child death, respectively, while regions in Mali (Kidal), Kenya (Kwale and *Marsabit*), and Tanzania (Mtwara) had the lowest risk of death compared to other districts across the sSA countries (Fig. 6 and Additional file 1).

Discussions

The present study was conducted to determine the timing of death of under-five children in sSA countries. The choice of the random effects (frailty) distribution is very critical and different frailty models have been used in literature [13, 59]. For modelling the time-to-event datasets, the semiparametric CPH is more common rather than the shared frailty parametric models [11, 12]. The frailty model accounts for unobserved heterogeneity (random effect) in survival analysis rather than models without frailty implicitly assuming that populations are homogeneous [13, 14], meaning that all subjects have the same risk of an event. However, in reality, the unobserved heterogeneity is assumed to represent different strata, and the strata are assumed to be independent and considered proportional hazard structures conditional on random effects [11, 29]. Therefore, the present study aimed to identify the risk factors of time-to-death among under-five children across the countries (districts) of sSA and determine the most efficient model for the analysis among semi-parametric and parametric models.

There are different views on the most efficient model in prediction of the time to death of U5C including the semi-parametric CPH and the five possible parametric models with four frailty distributions by considering countries (districts) as frailty components. The frailty term was statistically significant for semiparametric CPH and all the

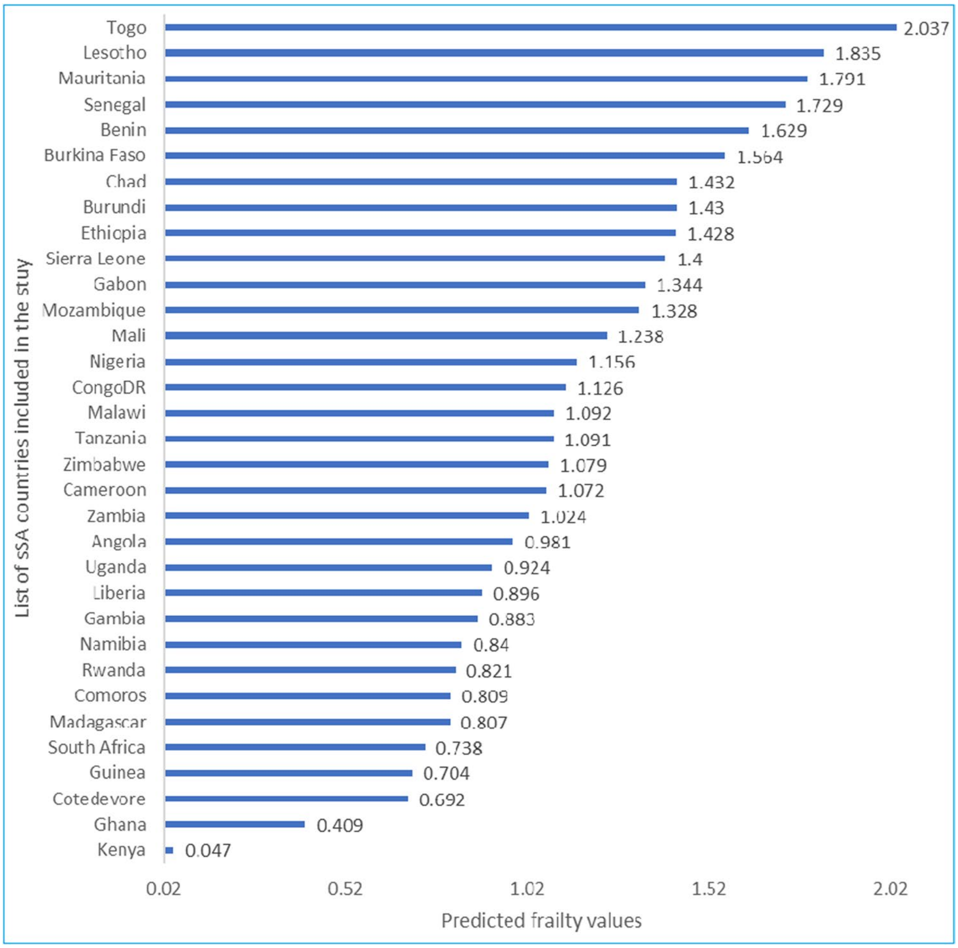


Fig. 5 The plot of predicted frailty values for the lognormal frailty model with exponential baseline

parametric models. The shared frailty model is the common model used to model the survival data with frailty terms, however, it is not valid in all cases, especially when the PH assumption does not hold or the survival data follow a parametric frailty model [11, 12, 15, 29, 31, 53]. Based on the Akaike's information criteria, the parametric shared frailty model (exponential with lognormal frailty) had better performance in predicting the survival time of under-five children across districts in sSA countries. However, the log-logistic, the log-normal, Weibull and Gompertz baseline distributions with different frailty terms did not converge in the algorithm used [51, 57, 60]. This might be due to the number of parameters included in the models. From the exponential baseline distribution with the log-normal frailty model, the estimated heterogeneity (variability) in the countries and districts was 0.163 and 0.082, respectively, which implies that a significant variation in survival of under-five children was accounted for unobservable children/district-level effects. The inclusion of the frailty in the model thereby minimizes both overestimation and underestimation of the model parameters and also correctly measure the effect of the covariates on the outcome variable (s) [33, 51, 57, 60, 61]. This is consistent with several previous studies that compared parametric models with the Cox proportional hazards model. For example, a study conducted on 197 children with acute leukaemia in Iran found that parametric models were preferred, with the Weibull model being the most efficient among them [23]. Similarly, another study involving 484 patients with gastrointestinal cancer reported that the logistic-log parametric model outperformed the Cox model in terms of efficiency [62]. Additional studies involving 3,421 [63] and 408 [64] participants, respectively, also demonstrated that the parametric gamma and parametric Weibull models were the most efficient compared to other models respectively. Furthermore, various studies [65–67] that compared Cox regression with parametric models concluded that parametric models were generally more effective in predicting survival outcomes across different conditions. These findings align well with the results of our study.

Children from rural areas have a higher risk of death than those who are from urban counterparts, which is in line with the studies conducted in different countries [31,

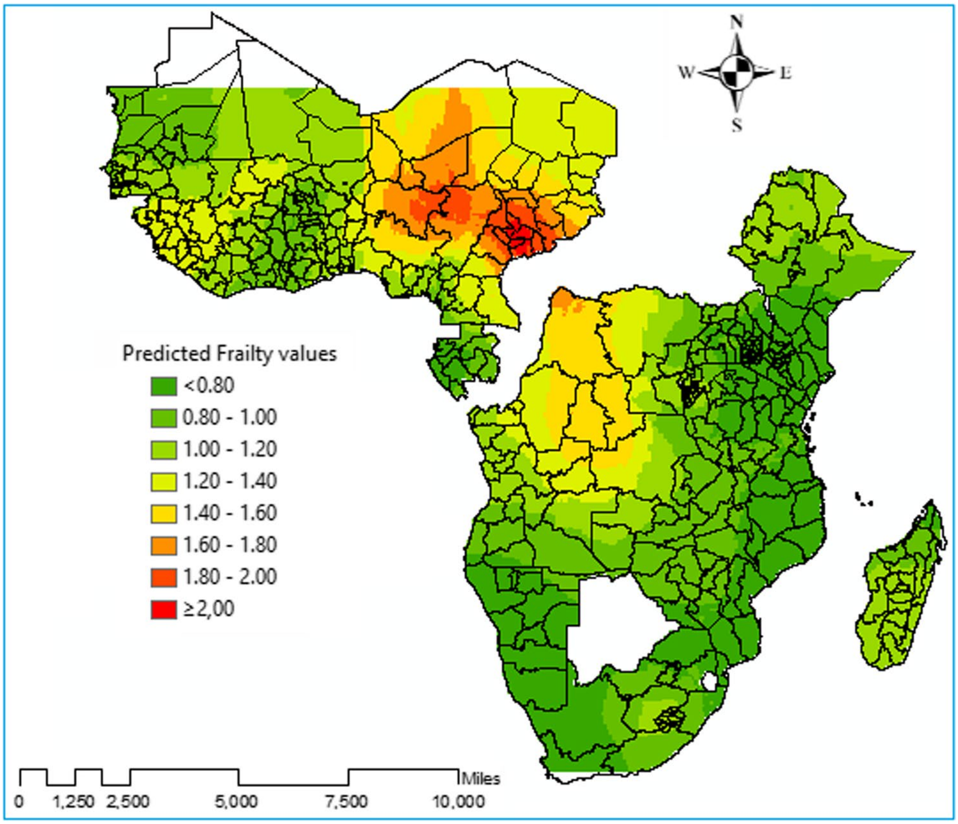


Fig. 6 Estimated predicted frailty values among districts across the sSA countries

46, 47, 68]. This might be the reason that there are differences in the distribution of healthcare facilities in rural and urban communities. Our study revealed that children born to younger mothers have a higher risk of mortality, which is consistent with the previous studies [42, 68–72]. This is because children born to adolescent mothers experience fragile health outcomes and have a higher risk of death [73]. Our finding revealed that the educational status of mothers was negatively associated with under-five mortality, which is in agreement with the previous studies [31, 42, 68, 71–73]. This might be the reason that education increases awareness of maternal-related healthcare services and practices, which ensures the survival of their children [31, 73, 74]. Moreover, short birth spacing is highly associated with the high risk of under-five child mortality. This finding is in line with the studies that indicate an association between child mortality and preceding birth interval [31, 42, 47, 68, 69]. This may be due to the fact that short birth intervals have been linked to negative health outcomes, such as maternal, newborn, and child mortality, and they are in violation of WHO birth spacing recommendations [61]. On the other hand, improved toilet facilities are also negatively associated with under five mortalities [75]. This is the reason that an intervention to improve existing sanitations facilities may reduce under five mortality rates.

This is study is used to identify the importance of access to improved water and the use of clean fuel for cooking for under five mortalities. Access to improved water and the use of clean fuel for cooking are negatively associated with under-five mortality risk. These findings are consistent with other studies conducted [31, 42, 73, 74]. This might be the reason that utilizing clean cooking fuels results in significantly lower emissions of pollutants than using unclean fuels, which emit toxic smoke into the air. In the same way, children who live in homes with better water facilities are more likely to survive than those who live in homes with unimproved water [76–78]. Therefore, the usage of clean cooking fuels and improved water has a negative relationship with the consequences of under-five mortality.

This study has several strengths. Firstly, we employed nationally representative cross-sectional data providing a comprehensive view of the population in sSA countries. Moreover, the selection of respondents for the study was random and covered a large population in sSA allowing us to make accurate conclusions about under-five children mortality in the region. Another strength of this study is that the DHS program used is transparent and has a high response rate which helps us to obtain accurate estimates when analyzing the data. Despite these strengths, the study has some limitations. The study did not include several important factors that significantly contribute to under-five mortality. Additionally, the cross-sectional nature of the DHS program datasets allows us to conclude associations but does not enable researchers to establish causal links. The study also couldn't see all possible baseline distributions (Weibull, Gompertz, lognormal, and logistic distributions) to fully leverage the performance of the parametric model due to the convergence problem. This might be the reason for the rigidly defined assumptions of each model under survival frailty analysis. This gap may be filled with the use of different machine learning and deep learning algorithm approaches for all frailty models.

Conclusions

Understanding the survival probability of under-five children and the associated risk factors is crucial for both practitioners and researchers in sub-Saharan Africa (sSA). This study emphasizes the significant impact of various factors on under-five mortality across sSA. By comparing semiparametric and parametric frailty models, we found that the exponential CPH model with log-normal frailty provided the best fit, outperforming the Cox CPH model, which did not account for the frailty effects adequately. The inclusion of a frailty component revealed unobserved heterogeneity in under-five mortality risks across different districts. This approach highlighted that children in some districts were at a significantly higher or lower risk of dying, necessitating district-specific interventions. Key risk factors identified include maternal characteristics (age, educational status, decision-making ability), household factors (place of residence, wealth index, media exposure, access to improved water, clean cooking fuel, number of under-five children), and child-level variables (birth order, birth interval). These findings underscore the multifaceted nature of under-five mortality and the importance of tailored public health strategies. Interventions should prioritize improving maternal education, enhancing women's autonomy, and ensuring access to clean water and cooking fuel. Additionally, focusing resources on high-risk districts identified through frailty modelling can optimize the effectiveness of public health strategies and reduce regional disparities in child survival.

Finally, our findings suggest that addressing district-level unobserved heterogeneity is crucial for understanding under-five mortality disparities across the region. We recommend that targeted interventions in high-risk districts and countries, such as those identified in our frailty models, can help reduce under-five mortality rates. Further research should incorporate infrastructure and other variables to provide a more comprehensive understanding of the determinants of under-five mortality.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-025-24186-x.

Supplementary Material 1.

Acknowledgements

The datasets used in this study were obtained from the DHS program, thanks to the authorization received to download the dataset on the website. This work is partially based upon research supported by the South Africa Department of Research and Innovation (DSI), National Research Foundation (NRF), South Africa Medical Research Council (SAMRC), South Africa DSI-NRF-SAMRC SARChI Research Chair in Biostatistics, Grant number 114613). Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the NRF and SAMRC.

Authors' contributions

HMF*1,2,6 was involved in this study from data management, data analysis, drafting, and revising the final manuscript. DGC1,3, TTZ 4, NNR1, DBB1,2, and SAY1,5 contributed to the conception, design, and interpretation of data, as well as to manuscript reviews and revisions. All authors have read and approved the manuscript.

Funding

Not applicable.

Data availability

The dataset used for the current study is available at the DHS program repository and the shapefile of the map of countries was accessed as an open-source without restriction from open Africa 2016 https://dhsprogram.com/data/available-datasets.cfm.

Declarations

Ethics approval and consent to participate

All methods were carried out following relevant guidelines of the Demographic and Health Surveys (DHS) program. All experimental protocols were approved by the Institutional Review Board (IRB) of Bahir Dar University. Informed consent was waived from the International Review Board of Demographic and Health Surveys (DHS) program data archivists after the consent paper was submitted to the DHS Program, a letter of permission to download the dataset for this study. The dataset was not shared or passed on to other bodies and was anonymized to maintain its confidentiality. All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

We, the authors, declare that we have no competing interests.

Author details

[1]Department of Statistics, University of Pretoria, Pretoria, South Africa
[2]Department of Statistics, Bahir Dar University, Bahir Dar, Ethiopia
[3]College of Health Solutions, Arizona State University, Phoenix, AZ, USA
[4]School of Mathematics, Statistics and Computer Science, College of Agriculture Engineering and Science, University of KwaZulu-Natal, Durban, South Africa
[5]Department of Statistics, College of Natural and Computational Science, Debre Tabor University, Debre Tabor, Ethiopia
[6]Center for Environmental and Respiratory Health Research (CERH), Research Unit of Population Health, University of Oulu, Oulu, Finland

Received: 16 August 2024 / Accepted: 28 July 2025
Published online: 22 August 2025

References

1. Sharrow D, et al. Global, regional, and National trends in under-5 mortality between 1990 and 2019 with scenario-based projections until 2030: a systematic analysis by the UN Inter-agency group for child mortality Estimation. Lancet Global Health. 2022;10(2):e195–206.
2. Wang H, et al. Global, regional, national, and selected subnational levels of stillbirths, neonatal, infant, and under-5 mortality, 1980–2015: a systematic analysis for the global burden of disease study 2015. Lancet. 2016;388(10053):1725–74.
3. Reidpath DD, Allotey P. Infant mortality rate as an indicator of population health. J Epidemiol Community Health. 2003;57(5):344–6.
4. Macharia PM, et al. Sub national variation and inequalities in under-five mortality in Kenya since 1965. BMC Public Health. 2019;19(1):1–12.
5. Song P, et al. Causes of death in children younger than five years in China in 2015: an updated analysis. J Glob Health. 2016. https://doi.org/10.7189/jogh.06.020802.
6. Nations U. The millennium development goals report 2012. Millennium Development Goals Report; 2012.
7. Fenta SM, Fenta HM. Risk factors of child mortality in Ethiopia: application of multilevel two-part model. PLoS One. 2020;15(8):e0237640.
8. MacDonald N, et al. Global vaccine action plan lessons learned I: recommendations for the next decade. Vaccine. 2020;38(33):5364–71.
9. Bobo FT, et al. Child vaccination in sub-Saharan Africa: increasing coverage addresses inequalities. Vaccine. 2022;40(1):141–50.
10. Fisher CB, et al. COVID-19 vaccine hesitancy among parents of children under five years in the United States. Vaccines. 2022;10(8): 1313.
11. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. Wiley; 2011.
12. Kleinbaum DG, Klein M. Survival analysis a self-learning text. Springer; 1996.
13. Legrand C. Advanced survival models. Chapman and Hall/CRC; 2021.
14. Lawless JF. Parametric models in survival analysis. Wiley StatsRef: statistics reference online; 2014.
15. Efron B. The efficiency of cox's likelihood function for censored data. J Am Stat Assoc. 1977;72(359):557–65.
16. Giolo SR, et al. Survival analysis of patients with heart failure: implications of time-varying regression effects in modeling mortality. PLoS One. 2012;7(6): e37392.
17. Rajaeefard A, et al. Applying parametric models for survival analysis of gastric cancer. Feyz Med Sci J. 2009;13(2):83–8.
18. Orbe J, Ferreira E, Núñez-Antón V. Comparing proportional hazards and accelerated failure time models for survival analysis. Stat Med. 2002;21(22):3493–510.
19. Pourhoseingholi MA, et al. Comparing Cox regression and parametric models for survival of patients with gastric carcinoma. Asian Pac J Cancer Prev. 2007;8(3):412.
20. Grover G, Sabharwal A. A parametric approach to estimate survival time of diabetic nephropathy with left truncated and right censored data. Int J Stat Probab. 2012;1(1):128.
21. Roshany D et al. Application of parametric, semiparametric and nonparametric approaches in survival analysis of patients with acute myocardial. 2011.
22. Baghestani A, Hajizadeh E, Fatemi S. To evaluate the prognostic factors in using Bayesian interval censoring analysis on survival rate of gastric cancer in Iran. Iran J Epidemiol. 2010;6(3):18–21.
23. Ghadimi M, et al. Family history of the cancer on the survival of the patients with Gastrointestinal cancer in Northern Iran, using frailty models. BMC Gastroenterol. 2011;11:1–9.
24. Askarishahi M, et al. Using parametric and Cox models in analysis of factors influencing the diagnosis of retinopathy in type II diabetes. J Mazandaran Univ Med Sci. 2014;24(113):27–35.
25. Teshnizi SH, Zare S, Tazhibi M. The evaluation of Cox and Weibull proportional hazards models and their applications to identify factors influencing survival time in acute leukemia. Hormozgan Univ Med Sci. 2010;15:269–78.
26. Laclé A, Valero-Juan LF. Diabetes-related lower-extremity amputation incidence and risk factors: a prospective seven-year study in Costa Rica. Rev Panam Salud Publica. 2012;32:192–8.
27. Teshnizi SH, Ayatollahi SMT. Comparison of Cox regression and parametric models: application for assessment of survival of pediatric cases of acute leukemia in Southern Iran. Asian Pac J Cancer Prevention: APJCP. 2017;18(4):981.
28. Cox DR. Regression models and life-tables. J Roy Stat Soc: Ser B (Methodol). 1972;34(2):187–202.
29. Duchateau L, Janssen P. The frailty model. Springer; 2008.
30. Schervish MJ, Carlin BP. On the convergence of successive substitution sampling. J Comput Graphical Stat. 1992;1(2):111–27.
31. Getachew Y, Bekele S. Survival analysis of under-five mortality of children and its associated risk factors in Ethiopia. J Biosens Bioelectron. 2016;7(213):2.
32. Argawu A. Multilevel modelling of under-Five time to death, and risk factors. Stat Ukraine. 2021;92(1):34–46.
33. Wegbom AI, Kiri VA, Essi ID. Comparison between semi-parametric Cox and parametric survival models in estimating the determinants of under-five mortality in Nigeria: application in Nigerian demographic and health survey. Afr J Maths Stat Stud. 2019;2(2):1–12.
34. Aliaga A, Ren R. The optimal sample sizes for two-stage cluster sampling in demographic and health surveys. ORC Macro; 2006.
35. Yilema SA, et al. Spatial small area estimates of undernutrition for under five children in Ethiopia via combining survey and census data. Spatial and Spatio-temporal Epidemiology. 2022;42: 100509.
36. Burgert CR, et al. Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys. Icf International; 2013.
37. Perez-Heydrich C, et al. Guidelines on the use of DHS GPS data. ICF International; 2013.
38. Adebowale AS, et al. Parental educational homogamy and under-five mortality in sub-Saharan Africa: clarifying the association's intricacy. Sci Afr. 2020;7: e00255.
39. Van Malderen C, et al. Socioeconomic factors contributing to under-five mortality in sub-Saharan africa: a decomposition analysis. BMC Public Health. 2019;19:1–19.
40. Zike DT, et al. Determinants of under-five mortality in Ethiopia: an application of Cox proportional hazard and frailty models. Turk Klin J Biostat. 2018. https://doi.org/10.5336/biostatic.2018-60550.
41. Bryce J, Victora CG, Black RE. The unfinished agenda in child survival. Lancet. 2013;382(9897):1049–59.
42. Fenta SM, Fenta HM, Ayenew GM. The best statistical model to estimate predictors of under-five mortality in Ethiopia. J Big Data. 2020;7(1):1–14.
43. Fenta SM, et al. Community and individual level determinants of infant mortality in rural Ethiopia using data from 2016 Ethiopian demographic and health survey. Sci Rep. 2022;12(1): 16879.
44. Brockerhoff M, Hewett P. Inequality of child mortality among ethnic groups in sub-Saharan Africa. Bull World Health Organ. 2000;78:30–41.
45. Rutherford ME, Mulholland K, Hill PC. How access to health care relates to under-five mortality in sub-Saharan Africa: systematic review. Trop Med Int Health. 2010;15(5):508–19.
46. Sear R, et al. The effects of kin on child mortality in rural Gambia. Demography. 2002;39(1):43–63.
47. Antai D. Regional inequalities in under-5 mortality in Nigeria: a population-based analysis of individual-and community-level determinants. Popul Health Metrics. 2011;9:1–10.
48. Singh A, et al. Do antenatal care interventions improve neonatal survival in India?? Health Policy Plann. 2014;29(7):842–8.
49. Therneau TM, et al. The cox model. Springer; 2000.
50. Hougaard P, Hougaard P. Analysis of multivariate survival data, vol. 564. Springer; 2000.
51. Klein JP. Handbook of survival analysis. Boca Raton, FL: CRC; 2014.
52. Kleinbaum DG, et al. Logistic regression. Springer; 2002.
53. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994;81(3):515–26.
54. Royston P, Lambert PC. Flexible parametric survival analysis using Stata: beyond the Cox model, vol. 347. Stata press College Station, TX; 2011.
55. Martins BA, et al. Frailty prevalence using frailty index, associated factors and level of agreement among frailty tools in a cohort of Japanese older adults. Arch Gerontol Geriatr. 2019;84: 103908.
56. Wienke A. Frailty models in survival analysis. Chapman and Hall/CRC; 2010.
57. Balan TA, Putter H. A tutorial on frailty models. Stat Methods Med Res. 2020;29(11):3424–54.
58. Akaike H. Factor analysis and AIC. Psychometrika. 1987;52:317–32.
59. Hougaard P. Frailty models for survival data. Lifetime Data Anal. 1995;1:255–73.
60. Duchateau L, Janssen P. Frailty distributions. The Frailty Model, 2008: pp. 117–197.
61. DR C. Regression models and life tables. JR Stat Soc. 1972;34:248–75.
62. Ravangard R et al. Comparison of the results of Cox proportional hazards model and parametric models in the study of length of stay in a tertiary teaching hospital in Tehran, Iran. 2011.

63. Adelian R, et al. Comparison of cox's regression model and parametric models in evaluating the prognostic factors for survival after liver transplantation in Shiraz during 2000–2012. Int J Organ Transplantation Med. 2015;6(3):119.
64. Hsu C-L, et al. Advanced non-small cell lung cancer in patients aged 45 years or younger: outcomes and prognostic factors. BMC Cancer. 2012;12:1–7.
65. Dai H, et al. Postoperative radiotherapy for resected pathological stage IIIA–N2 non-small cell lung cancer: a retrospective study of 221 cases from a single institution. Oncologist. 2011;16(5):641–50.
66. Santoro IL, et al. Non-small cell lung cancer in never smokers: a clinical entity to be identified. Clinics (Sao Paulo). 2011;66:1873–7.
67. Wang W-l, Shen-Tu Y, Wang Z-q. Prognostic factors for survival of stage IB upper lobe non-small cell lung cancer patients: a retrospective study in Shanghai, China. Chin J Cancer Res. 2011;23:265–70.
68. Li Z, et al. Changes in the spatial distribution of the under-five mortality rate: small-area analysis of 122 DHS surveys in 262 subregions of 35 countries in Africa. PLoS One. 2019;14(1): e0210645.
69. Tiruneh SA, Zeleke EG, Animut Y. Time to death and its associated factors among infants in sub-Saharan Africa using the recent demographic and health surveys: shared frailty survival analysis. BMC Pediatr. 2021;21:1–13.
70. Naz L, Patel KK, Uzoma IE. Crucial predicting factors of under-five mortality in Sierra Leone. Clin Epidemiol Glob Health. 2020;8(4):1121–6.
71. Rahman M. Factors affecting on child survival in Bangladesh: cox proportional hazards model analysis. Internet J Trop Med. 2008;6(1):1–5.
72. Amir-ud-Din R, et al. Impact of high-risk fertility behaviours on underfive mortality in Asia and Africa: evidence from demographic and health surveys. BMC Pregnancy Childbirth. 2021;21(1): 344.
73. Andriano L, Monden CW. The causal effect of maternal education on child mortality: evidence from a quasi-experiment in Malawi and Uganda. Demography. 2019;56(5):1765–90.
74. Adedini SA, et al. Regional variations in infant and child mortality in Nigeria: a multilevel analysis. J Biosoc Sci. 2015;47(2):165–87.
75. Fink G, Günther I, Hill K. The effect of water and sanitation on child health: evidence from the demographic and health surveys 1986–2007. Int J Epidemiol. 2011;40(5):1196–204.
76. Gaffan N, et al. Effects of household access to water, sanitation, and hygiene services on under-five mortality in sub-Saharan Africa. Front Public Health. 2023;11: 1136299.
77. Owili PO, et al. Cooking fuel and risk of under-five mortality in 23 sub-Saharan African countries: a population-based study. Int J Environ Health Res. 2017;27(3):191–204.
78. Fenta HM, Chen D-G, Zewotir TT. *Geostatistical Analysis of Under-Five Children Mortality and Associated Factors Across Sub-Saharan African Countries*, in *Biostatistics Modeling and Public Health Applications*. Springer; 2024. p. 231–56.

Publisher's Note