SciencePG
Science Publishing Group

On Multi-Level Modeling of Child Mortality with Application to KDHS Data 2014

Otieno Otieno[1, *], Mathew Kosgei[1], Nelson Onyango Owuor[2]

[1]School of Sciences and Aerospace Studies, Moi University, Eldoret, Kenya

[2]School of Mathematics, University of Nairobi, Nairobi, Kenya

Email address:

otienochriso@yahoo.com (Otieno Otieno), mkosgei12@gmail.com (Mathew Kosgei), nelsonowuor@gmail.com (Nelson Onyango Owuor)

[*]Corresponding author

Abstract: One of the dominant challenges affecting low and middle countries is the regard of child mortality. It had been a millennium development goal to reduce infant and child mortality by two-thirds in 1990 mortality levels by the year 2015. Therefore, the aspiration to recognize the causal factors of under five child mortality poses a crucial aspect of research. In principal, remarkable progress has been made in bringing down mortality in children under 5 years of age. The global under five mortality rate declined by 59 per cent, from 93 deaths per 1,000 live births in 1990 to 38 in 2019. In Kenya, the infant mortality rate in 2021 is 32.913 deaths per 1000 live births, a 3.36 per cent decline from the year 2020. It was 34.056 deaths per 1,000 live births in 2020, a decline of 3.24 per cent from the year 2019. In Nyanza Province, Kenya, has the highest infant mortality rate (133 deaths per 1,000 live births) while the lowest in Central Province (44 deaths per 1,000 live births). Despite all that improvement, the world is still doubtful to achieve that millennium development goal target number four, of diminishing child mortality. Our study aims to scrutinize on vital covariates affecting child mortality in Nyanza, Kenya. The principal purpose of this paper is to scrutinize the effect of demographic and socioeconomic variables on child mortality. We carried out a series of model evaluations to ascertain the best model under various scenarios bearing in mind the presence of dependencies due to Clusters and households. Then, performed a linear mixed effects model with the best fit based on data from Kenya Demographic and Health Survey (KDHS 2014) which was collected by use of questionnaires. Child mortality from the, KDHS 2014 data, was analyzed in an age period: mortality from the age of 12 months to the age of 60 months. The study reveals that, number of children under 5 in household, number of births in last 5 years, modern family planning and contraceptive use had an exceptional impact on child mortality.

Keywords: Mixed Models, Correlated Data, Lmer, Glmer, Best Linear Unbiased Predictors (BLUP)

1. Introduction

Child mortality refers to the death of children under age of five. The under-5 mortality rate (U5MR), the probability of dying before 5 years of age (per 1000 live births), is a key global indicator of child health [1], and one of the most important measures of global health [2].

Child mortality is a essential measure of child health and overall national development [3]. It also reflects a country's level of socio-economic develop-ment, quality of life and are used for monitoring and evaluating population, health programs and policies. In the past few decades there is decline in under-five mortality in almost all countries of the world, regardless of initial levels, socio-economic circumstances and development strategies [4].

A high mortality rate generally signify unmet human health needs in education, medical care, nutrition and sanitation. The desire to understand the determinants of Under 5 Child mortality (U5CM) poses a very important aspect of research. Intents to reduce under-5 mortality to at least as low as 25 deaths per 1000 live births in all countries by 2030, was previously targeted in the fourth Millennium Development Goal (MDG). Today, it appears in the third Sustainable Development Goal (SDG3) [5].

The Demographic and Health Surveys (DHS) program has been very instrumental for obtaining and disseminating authentic, national representative data on family planning, fertility, maternal and child health, among other health issues. The most recent DHS survey conducted in Kenya was KDHS 2014 [6].

This study aims at identifying the determinants of Child mortality in Nyanza, Kenya. We chose a range of covariates gotten from three different publications based on Demographic health survey data in those three countries [7-9]. Those covariates were exceptional in determining child mortality. In the study, we then test for normality assumption using residual plots and analysis for purposes of diagnostics on the best fit from those specified covariates. We ultimately, construct a Linear Mixed Effects model that handles the dependence within clusters and households.

A linear mixed effects model has two parts, that of fixed effects and random effects factors. A fixed effects factor is a factor whose levels are the only possible levels in the population being studied. This is opposed to a random effect factor whose levels in the study are just a sample of all the possible choices.

2. Methods

2.1. Data Description and Ethical Approval

In this paper, we use the Kenya Demographic Health Survey data KDHS 2014. It is the sixth Demographic and Health Survey (DHS) conducted in Kenya since 1989. KDHS is a national research undertaking conducted every five years with an intention of collecting a wide range of data with a strong interest on indicators of reproductive health, fertility, mortality, maternal and child health, nutrition and self-

reported health habits among adults [10]. It is a household sample survey data with a national representation where households are selected at random from Kenya National Bureau of Statistics (KNBS) sampling frame. The survey procedures, instruments and sampling methods used in the KDHS 2014 acquired ethical recommendation from the Institutional Review Board of Opinion Research Corporation (ORC) Macro International Incorporated, a health, demographic, market research and consulting company situated in New Jersey, USA. We sought official registration on the DHS website and got permission to use the KDHS 2014 data. The data was downloaded in SPSS format and constituted 1,099 variables and 20,964 observations. Using package foreign, the data was imported to R software version 4.2.2 for analysis. KDHS data is a national survey data that is classified into 8 regions, constituting former provinces in Kenya. For this work, we analyzed data only for Nyanza region, being a region with the highest child mortality in Kenya. A set of dependent variables are chosen from literature given that they were profound in explaining child mortality. Survival time and status variables which are important considerations when analyzing survival data were calculated and included in the dataset.

2.2. Ethical Approval

The study did not need any approvals because it was secondary data.

3. Model Formulation

In our study, we explored a series of models. First, is the null model, without any covariate.

>null model < −glm(factor(Child is alive) ∼ 1, family = binomial(link = "logit"), data = Child Mortality)

The second model is one with Cluster number as a random effect.

> fit1 < −glmer(factor(Child is alive) ∼1 + (1 | Cluster number),

family = binomial(link = "logit"), data = Child Mortality)

Model with Household number as a random effect.

> fit2 < −glmer(factor(Child is alive) ∼1 + (1 | Household number),

family = binomial(link = "logit"), data = Child Mortality)

Interaction model of Cluster and Household as a random effects.

> fit3 < −glmer(factor(Child is alive) ∼1 + (1| Cluster number/Household number),

family = binomial(link = "logit"), data = Child Mortality)

Model with fixed effects plus Cluster number as the random effect.

> f it4 < −glmer(Child is alive∼ as.factor(Sex of child) + as.factor(Education program)

+as.factor(Number of children under 5 in HH) + as.f actor(Number of births in last 5 years) + Maternal age at birth + as.factor(FP contraceptive use) + (1 | Cluster number),

family = binomial(link = "logit"), data = Child Mortality)

A model with fixed effects plus Household number as the random effect.

fit5 < −glmer(Child is alive∼ as.factor(Sex of child) + as.factor(Education program) +

as.factor(Number of children under 5 in HH) + as.factor(Number of births in last 5 years) + Maternal age at birth + as.f actor(F P contraceptive use) + (1|Household number),

family = binomial(link = "logit"), data = Child Mortality)

A model with fixed effects and an interaction between random effects.

fit6 < −glmer(Child is alive∼ as.factor(Sex of child) + as.factor(Education program) +

as.factor(Number of children under 5 in HH) + as.factor(Number of births in last 5 years) + Maternal age at birth + as.f actor(F P contraceptive use) +

(1|Cluster number/Household number), family = binomial(link = "logit"), data = Child Mortality)

Since Household number is a random effect nested in Cluster number, we consider other model formulations bearing that in mind, prior to performing an ANOVA for comparison of the models. Accordingly, choose the optimal one.

A nested model without any fixed effect covariate.

f it7 < −glmer(Child is alive∼ Cluster number | Household number), family = binomial(link = "logit"), data = Child Mortality)

A nested model with a few fixed effects covariates

fit8 < −glmer(Child is alive∼ as.factor(Sex of child) + as.factor(Education program) +

as.factor(Number of children under 5 in HH) + as.factor(Number of births in last 5 years) + (Cluster number|Household number), family = binomial(link = "logit"),

data = Child Mortality)

A nested model with all the fixed and random effects.

f it9 < −glmer(Child is alive∼ as.factor(Sex of child) + as.factor(Education program) +

as.factor(Number of children under 5 in HH) + as.factor(Number of births in last 5 years) + Maternal age at birth + as.factor(FP contraceptive use) + (Cluster number|Household number), family = binomial(link = "logit"), data = Child Mortality)

The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) information for model fit4 are AIC = 940.89, BIC = 1042.6.

Log Likelihood = - 453.45 and a p-value < 2e-16. This is reasonably a good model comparing the criterion values to those of other models. Un-fortunately, we cannot compare fit4 directly with other the model fits since Bayesian Information Criterion for instance, dictates that we compare nested models [11].

It is recognised that smaller values for AIC, BIC define the best models.

anova(fit7, fit8, fit9)

Table 1. *Comparison of the nested models.*

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
fit7	4	1296.22	1320.2	-644.11	1288.22			
fit8	16	945.44	1041.2	-456.72	913.44	374.7781	12	<2.00E-16***
fit9	19	944.12	1057.8	-453.06	906.12	7.3271	3	0.06217
Signif.codes:	0.001	***						

Even though fit9 is not statistically significantly, it has the lowest AIC value compared to the other model fit7 and fit8.

We recollect that an AIC is given by AIC = $−2\log L + 2k$, where logL is the maximum log-likelihood and k is the number of parameters in the model. Then BIC is given by BIC = $−2\log(L) + k\log(n)$ and for normally distributed errors, we have BIC = $\log(\sigma_\varepsilon^2) + _{k/n} \log(n)$, where logL is the maximum log-likelihood, k is the number of parameters in the model, n is the number of observations and σ_ε^2 is the error variance. BIC is therefore an increasing function of error variance and the number of parameters [12].

By the "principle of parsimony" we choose the model

fit9 < −glmer(Child is alive∼ as.factor(Sex of child) + as.factor(Education program) +

as.factor(Number of children under 5 in HH) + as.factor(Number of births in last 5 years) + Maternal age at birth + as.factor(FP contraceptive use) + (Cluster number|Household number), family = binomial(link = "logit"), data = Child Mortality)

fit10 < −lmer(Child is alive∼ as.factor(Sex of child) + as.factor(Education program)+

as.factor(Number of children under 5 in HH) + as.factor(Number of births in last 5 years) + Maternal age at birth + as.factor(FP contraceptive use) + (Cluster number|Household number), data = Child Mortality, method = "REML")

anova(fit9,fit10)

Table 2. Comparison of the models.

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)	
fit9	19	944.12	1057.76	-453.06	906.12				
fit10	20	-689.15	-569.53	364.58	-729.15	1635.3	1	<2.20E-16	***
Signif.	codes:	0.001	***						

In table 2 above, we made a comparison of the same covariates but under different criterions to understand their performance. Using the same rule of thumb that the best fit is the one with the lowest AIC, fit10 is statistically significant with a p-value <2.20E-16. Implying that it is a better model that a null model or even fit9.

3.1. Parameter Estimation

A linear mixed effects model is fitted of the form below

$$Y_{ij} = \mu + \alpha_{ij} + \beta_{ij} + \varepsilon_{ij}$$

The matrix form for the mixed effects model is

$$Y = X\beta + Z\alpha + \varepsilon \qquad (1)$$

where the vector β represents the fixed effect parameters, usually estimated by Generalised Least Squares (GLS) approach, and the vector α represents the random effects and are estimated as Best Linear Unbiased Predictors (BLUPs). The mathematical theory of Mixed Model Analysis based on model (1) and well illustrated in literature requires that we estimate the parameters β and α, Searle et. al., [13]. To estimate β, the fixed effects parameters, we use the method of GLS that maximizes the log-likelihood function with respect to β. We obtain Best Linear Unbiased Estimators (BLUEs) β via,

$$X'V^{-1} Y = X' V^{-1}X\beta \qquad (2)$$

The GLS function (2) depends on the variance components via the matrix, V, and one has to obtain an estimate of a matrix as a first step. This is done by REML (Restricted or Residual Maximum Likelihood) estimation. Then one needs the estimates of random effects. The estimates are referred to as "Predictors" to distinguish them from fixed effects for which the word "Estimates" has been used. The BLUPs of α are obtained from the equation,

$$\alpha = \left(Z'Z + \hat{\Gamma}^{-1} \right)^{-1} Z' \left(Y - X\hat{\beta} \right) \qquad (3)$$

In (3), the parameters to be estimated include Γ and β. The

BLUE of β and the REML estimate of the variance components contained now in Γ, are substituted in the equation to finally obtain the BLUPs.

3.2. Restricted Maximum Likelihood Estimation

The variance components are two main parameters σ_α^2 and σ_ε^2 contained in the matrix V. The variance components can be estimated by a number of methods, including ANOVA, Maximum Likelihood, Bayesian Estimation and Method of Moments. Nevertheless, REML, developed by is more attractive since it offers unbiased and non-negative estimates of variance components [14]. Maximum likelihood estimates (MLE) of variance components may turn out to be negative [15]. Searle et al., mentioned that such variance components can be set to zero [13].

4. Data Analysis

4.1. Descriptive Statistics

In Table 3, children who died in Nyanza were still high even though comparatively those who survive account for 2757 (94.2%). This imply that 169 (5.8%) died and that should be a matter of concern and should be mitigated to re-duce it further.

Table 3. Child is alive in Nyanza.

Child alive	Frequency	Percentage
No	169	0.058
Yes	2757	0.942
Total	2926	

In Table 4, majority in the survey resided in the rural areas of Nyanza province with 2014 (69%) While the rest are in the urban areas of Nyanza 912 (31%).

Table 4. Place of residence.

	Frequency	Percentage
Urban	912	0.311
Rural	2014	0.689
Total	2926	

In Table 5, even though modern techniques of family planning in Nyanza seem to have been preferred or were being used with 1583 (54%). There was still a big portion who either did not use any method 1284 (43.8%). Traditional methods were also still being used though with a smaller number of 54 (1%).

Table 5. Contraceptives/Family planing methods.

	Frequency	Percentage
No Method	1284	0.438
Traditional Method	54	0.018
Modern Method	1588	0.542
Total	2926	

In Table 6, the distribution of children born in Nyanza province by gender.

The numbers of males were slightly higher with 1485 (50.75%) as compared to their female counterparts with 1441 (49.25%).

Table 6. Gender of children.

	Frequency	Percentage
Male	1485	50.75
Female	1441	49.25
Total	2926	

In table 7, shows the duration it took before a child is born, and on average the period was 41.57 months. The maximum time was 225 and minimum of 9 months before a child birth.

Table 7. Preceding birth interval (months).

min	1st Qu	median	mean	3rdQu	Max	NA's
9	25	34	41.57	51	225	634

In table 8, showed the duration it took after a child is born and on average the period was 27.17 months. The maximum time was 58 months and minimum of 9 months before a child birth.

Table 8. Succeding birth interval (months).

min	1st Qu	median	mean	3rdQu	Max	NA's
9	20	26	27.17	33	58	2120

In Table 9, almost all the families interviewed had 2274 (99.7%) of no a live birth in between the births. Only 6 (0.2%) had all alive children which is a very small number.

Table 9. Live birth in between births.

	Frequency	Percentage
No	2274	0.997
Yes	6	0.002
Total	2280	

In Table 10, the table showed that 111 households did not have children below 5 years, 1188 households had 2 children each and that 5 households had 5 children each.

Table 10. Number of children under 5 years in household.

Children	0	1	2	3	4	5
Households	111	1083	1188	438	101	5

In Table 11, in the last 5 years, 1358 households in Nyanza had 1 child born. 1238 households had 2 and 20 households had 4 children living with them.

Table 11. Number of births in last 5 years.

Children	1	2	3	4
Households	1358	1238	312	20

In Table 12, the lowest portion 33 (0.011%) had no education at all while 1958 (0.669%) had primary level of education.

Table 12. Education level.

Highest education level	Frequency	Percentage
No education	33	0.011
Primary	1958	0.669
Secondary	759	0.259
Higher	176	0.06
Total	2926	

4.2. Results of Fitting Linear Mixed Model Using Lmer

We consider the output for model from fit10. In our output below, shows that REML was the tool that produced the variance estimates. It is the default tool under the lmer function, and if specified to be false, then MLE is used. The other information includes criterion for model choice, include AIC, BIC and the other criteria.

Linear mixed model fit by REML. t-tests use Satterthwaite's method [lmerModLmerTest].

fit10 < −lmer(Child is alive∼ as.factor(Sex of child) + as.factor(Education program)+

as.factor(Number of children under 5 in HH) + as.factor(N umber of births in last 5 years) + Maternal age at birth + as.factor(FP contraceptive use) + (Cluster number|Household number), data = Child Mortality, method = "REML")

Npar AIC BIC logLik deviance

20 -689.15 -569.53 364.58 -729.15

Table 13. Random effects.

Groups	Name	Variance	Std.Dev.	Corr
Household number	(Intercept)	4.57E-02	0.2136471	-1
	Cluster number	5.41E-08	0.0002325	

Groups	Name	Variance	Std.Dev.	Corr
Residual		4.57E-02	0.2137744	
Number of obs: 2926	groups:	Household number	148	

The output includes estimates of the variance components and standard deviations. The standard deviation of the variance component of Household number is equal to that of residual error but greater than that of the Cluster number justifying its inclusion as a random effect. These variance estimates are elements of matrix V are then used to compute the fixed effect parameters via the formula (2).

***Table 14.** Fixed effects.*

	Estimate	Std.Error	t value	Pr (>| t |)
(Intercept)	6.35E-01	5.03E-02	12.619	<2.00E-16***
as.factor (Sex of child)2	1.14E-02	7.93E-03	1.442	0.14948
as.factor (Education program)1	7.54E-03	3.83E-02	0.197	0.8438
as.factor (Education program)2	2.45E-02	3.88E-02	0.631	0.5279
as.factor (Education program)3	6.01E-02	4.19E-02	1.433	0.15183
as.factor (Number of children under 5 in HH)1	2.87E-01	2.15E-02	13.387	<2.00E-16***
as.factor (Number of children under 5 in HH)2	4.12E-01	2.22E-02	18.608	<2.00E-16***
as.factor (Number of children under 5 in HH)3	4.98E-01	2.44E-02	20.397	<2.00E-16***
as.factor (Number of children under 5 in HH)4	5.03E-01	3.19E-02	15.767	<2.00E-16***
as.factor (Number of children under 5 in HH)5	5.52E-01	9.86E-02	5.6	2.7E-08***
as.factor (Number of births in last 5 years)2	-1.34E-01	1.09E-02	-12.305	<2.00E-16***
as.factor (Number of births in last 5 years)3	-2.38E-01	1.67E-02	-14.233	<2.00E-16***
as.factor (Number of births in last 5 years)4	-2.52E-01	5.37E-02	-4.692	2.99E-06***
Maternal age at birth	-5.03E-04	1.40E-03	-0.36	0.71878
as.factor (FP contraceptive use)2	2.89E-02	3.03E-02	0.956	0.33892
as.factor (FP contraceptive use)3	2.20E-02	8.15E-03	2.699	0.00699**

For fixed effects, only three variables are statistically significant in explaining child survival. They are Number of children under 5 in a household, Number of births in last 5 years and Family Planning contraceptive use (modern method use).

We could also infer that since the absolute t values for the above mentioned predictors are (greater than 2), that is an indication of higher reliability of the predictive power of that coefficient.

The variable Sex of child had two levels namely, Male and Female. The male is the reference category. Consequently, female children had higher odds of surviving by a factor of 1.0114 than their male counterparts.

Those children born in a household with mothers having higher education had higher odds of surviving by a factor of 1.061 than those who did not have education at all.

The chances of child survival is higher by a factor of 1.7367 for children born in households with 5 children under 5 years as compared to those who did not have any child or children.

Children born in houses with already 4 births in the last five years, were 22% less likely to survive until the fifth birth day compared to those with just 1 birth in the last five years.

When a mother's age increases by one year, the odds of having a child alive to the fifth year, becomes smaller by a factor of 0.9994.

Those who used modern family planning or contraceptives methods had a higher odds by a factor of 1.0222 of having their children being alive as compared to those who never used any family planning or contraceptives.

4.3. Residual Analysis

Apart from the information on the exploratory data analysis, it is as useful to consider the residual plots and analysis for purposes of diagnostics. Residual versus fitted value plots, normal QQ-plots, factor versus residual plots are some of the common tools in understanding the residual structure in the data. By using our fit10, we obtain the residual plot. Moreover, we employ a statistical inference method which is a more liberal test of normality done on the residuals of model fit10, which is the Shapiro Wilk test,

> shapiro.test(residuals(fit10))data: Child Mortality

residuals($f it10$) W = 0.57995, $p-value < 2.2e-16$

From our result above, the normality assumption is violated by the results of Shapiro-Wilk test, i.e., the null hypothesis of normally distributed data is rejected as the p-value is significant at 5% level, (p −value = 2.2e −16). The stem and leaf diagram gotten by

> stem(residuals(fit9)) also confirms the non-normality assumption.

To establish the independence of the residuals and the homoskedasticity, we use a plot of the observed values versus residuals predicted values.

In figure 1, shows how normality assumption does not

quite hold for the residuals. The points do not fall on the straight line violating that assump-tion. On the other plot, even though the points fall above and below the zero line. They are not randomly scattered with constant spread on that space for independence and homoskedasticity to hold. Secondly, the points are conjested at the zero line [12].
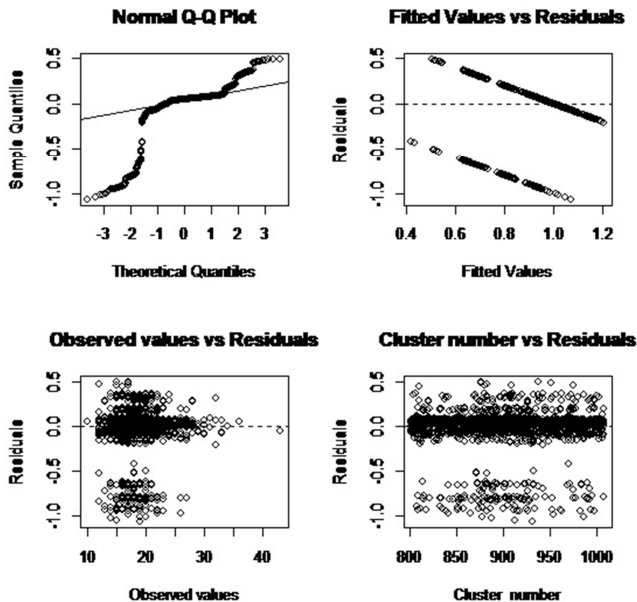

Figure 1. Residual plots of Child mortality dataset.

5. Discussion

The study attempts to understand the determinants of child mortality using survey data from KDHS. In this case, Kenya DHS survey 2014 dataset was used for the analysis. For this work, we analysed data for Nyanza province only being a region with the highest child mortality in Kenya.

A high mortality rate generally signify unmet human health needs in education, medical care, nutrition and sanitation. Therefore, aspiration to understand the determinants of Child mortality gives rise to a very impor-tant aspect of research. Many studies have employed regression techniques to explore the determinants of U5CM, these include the Cox PH regression among others [16-18].

In our findings using Linear Mixed Effects model we realised that the fol-lowing covariates;, Number of children under five years in household, Number of births in the last five years, Modern methods of family planning and con-traceptive use were statistically significant in determining child mortality.

In comparison with other studies that were published, there was no big mismatch in that findings showed that child mortality was associated with variables related to: child characteristics at birth (such as age at birth), reproduction factors of the mother (such as number of siblings born before), feeding characteristics and anthropometric measurements.

This was in line with other findings which used Cox PH regression and established that region of residence, sex of the

child, type of birth (multiple), birth interval (less than 24 months after the preceding birth), and mother's education were related with an increased risk of children mortal-ity before their fifth birthday [16]. Searle et. al., also established that factors related to mother characteristics and previous births such as sex of the child, sex of the head of the household and the number of births in the past one year was found to be significant [17]. Patterson et al., also explored the effect of mother's education, child's sex, rural/urban residence, household wealth index, regions ecological zones and development [18].

In our work, we had a large sample of 2926 observations and 1132 variables which was statistically sufficient. Nevertheless, there was a limitation of missing cases which we ignored and just used complete cases in our analysis.

Another limitation was that we did not use any scientific method on choosing the covariates we worked on. That, was just choosen from literature.

The study is of significance in that it will help the government, Non-governmental organisations to get information to guide them on how to intervene in order to ultimately achieve the global target on Child mortality.

Furthermore, students who are actively engaged in doing research in child mortality may not only fill the gap but also add on the already existing bank of knowledge.

6. Conclusion

Correlated data arises from numerous health fields especially where participants are in a cluster or household sharing the same prognostic factors. In this research, we have managed to present a framework for determination of child mortality using the 2014 KDHS data from Nyanza Province in Kenya.

Majorly, the framework involved checking which covariates were notable in explaining child survival. Moreover, based on this research, child mortality is associated with variables related to reproduction factors of the mother (such as number of children under five years in the household, number of births in the last five years and Family planning and contraceptive use (modern).

These results are significant in that they can inform policy from the national government. The national government will have known which covariates to give more priority in its endeavours to further reduce child mortality.

References

[1]   UNICEF. (2016). The under-five mortality rate: The indispensable gauge of child health.www.unicef.org/sowc08/docs/sowc08_panels. pdf.AccessedMar2016.

[2]   Black, R. E., Levin, C., Walker, N., Chou, D., Liu, L., Temmerman, M., Group, D. R. A. (2016). Reproductive, maternal, newborn, and child health: key messages from disease control priorities 3rd edition. The Lancet, 388 (10061), 2811-2824.

[3]   McGuire, J. W. (2006). Basic health care provision and under-5 mortality: a cross-national study of developing countries. World Development, 34 (3), 405-425.

[4]   Child Mortality. (2022, January 20). UNICEF DATA. https://data. unicef.org/topic/child-survival/under-five-mortality/

[5]   Nations, U. (2016). The Sustainable Development Goals 2016. eS-ocialSciences. http://www.un.org/sustainabledevelopment/health/.AccessedM ar2016.

[6]   Wanjiru, W. H. (2021). Improved balanced random survival forest for the analysis of right censored data: application in determining under five child mortality (Doctoral dissertation, Moi University).

[7]   Bereka, S. G., Habtewold, F. G., Nebi, T. D. (2017). Under-five mor-tality of children and its determinants in Ethiopian Somali regional state, Eastern Ethiopia. Health Science Journal, 11 (3), 1.

[8]   Khan, J. R., Awan, N. (2017). A comprehensive analysis on child mor-tality and its determinants in Bangladesh using frailty models. Archives of Public Health, 75 (1), 1-10.

[9]   Sharma, A. K., Dutta, R. (2020). Determinants of child survival at the household level: An insight of the method of factor analysis. In Contem-porary Issues in Sustainable Development (pp. 253-271). Routledge India.

[10]   Corsi, D. J., Neuman, M., Finlay, J. E., Subramanian, S. V. (2012). Demographic and health surveys: a profile. International journal of epi-demiology, 41 (6), 1602-1613.

[11]   Burnham, K. P. (2000). Model selection and multimodel inference. A practical information-theoretic approach.

[12]   Nelson Owuor Onyango (2009). On the Linear Mixed Effects Regression (lmer) R Function for Nested Animal Breeding Data. CS-BIGS 4 (1): 44-58. http://erepository.uonbi.ac.ke:8080/xmlui/ handle/123456789/38574

[13]   Searle, S. R., Casella, G., McCulloch, C. E. (2009). Variance compo-nents. John Wiley Sons.

[14]   Patterson, H. D. (1971). Thompson R. Recovery of inter-block informa-tion when block sizes are unequal. Biometrika, 58, 545-554.

[15]   Duchateau, L., Janssen, P., Rowlands, J. (1998). Linear mixed models. An introduction with applications in veterinary research. ILRI (aka ILCA and ILRAD).

[16]   Ayiko, R., Antai, D., Kulane, A. (2009). Trends and determinants of under-five mortality in Uganda. East African journal of public health, 6 (2), 136-140.

[17]   Nasejje, J. B., Mwambi, H. G., Achia, T. N. (2015). Understanding the determinants of under-five child mortality in Uganda including the estimation of unobserved household and community effects using both fre-quentist and Bayesian survival analysis approaches. BMC public health, 15 (1), 1003.

[18]   Sreeramareddy, C. T., Kumar, H. N., Sathian, B. (2013). Time Trends and Inequalities of Under-Five Mortality in Nepal: A Secondary Data Analysis of Four Demographic and Health Surveys between 1996 and 2011. PLoS ONE, 8 (11): e79818. doi: 10.1371/journal.pone.0079818.