Analysis of Determinants of Birth Interval in Four Disadvantaged Regions of Ethiopia

Etiyopya'daki Dört Dezavantajlı Bölgede Doğum Aralıklarının Belirleyici Etkenlerinin Analizi

Dereje TESFAYE,[a]
Shibru TEMESGEN,[b]
Zinabu TEKA[c]

[a]Department of Statistics,
College of Science, Bahir Dar University,
[b]Department of Statistics,
College of Natural Science,
Addis Ababa University
[c]Department of Statistics,
College of Natural and Computational
Science, University of Gondar

Geliş Tarihi/*Received:* 08.04.2015
Kabul Tarihi/*Accepted:* 03.07.2015

Yazışma Adresi/*Correspondence:*
Zinabu TEKA
College of Natural and Computational
Science, University of Gondar
Department of Statistics,
ETHIOPIA/ETİYOPYA
take.z70@gmail.com

ABSTRACT Objective: Studying the dynamics of spacing of births is important for several reasons including an understanding of completed family size. In this paper the length of birth interval between successive children that is called inter-birth interval in four disadvantaged regions of Ethiopia was considered. To identify and analyze socioeconomic and demographic factors that may have a significant influence on birth interval length in four disadvantaged regions of Ethiopia. **Material and Methods:** The data for the study was obtained from Ethiopian Demography and Health Survey data conducted in 2011. The data contains 3340 women aged from 15-49 years. The Kaplan-Meier and Cox's proportional hazards model were employed for the analysis of birth interval data using SPSS 16 and STATA 11 software. The Kaplan Meier median length was used to examine birth interval differentials by socio economic and demographic characteristics of women. **Results:** The results indicated that almost in all birth intervals, educated women, rich women, orthodox, urban women, women belonging to Benishangul-gumuz region and women whose index child has survived have longer birth interval lengths. The Cox regression analyses revealed that region, place of residence and survival status of the index child were consistently significant while women educational level and wealth index were significant in some birth intervals. **Conclusion:** Almost in all birth intervals, educated women, rich women, orthodox, urban women, women belonging to Benishangul-gumuz region and women whose index child has survived have longer birth interval lengths.

Key Words: Birth intervals; disadvantaged regions; survival analysis; proportional hazards model

ÖZET Amaç: Doğum aralığı dinamiklerinin çalışılması, tamamlanmış aile büyüklüğünü de içeren birçok neden için önemlidir. Bu makalede birbirini izleyen çocuklar arasındaki doğum aralığının uzunluğu Etiyopya'da ki dört dezavantajlı bölgede dikkate alınan doğum aralığı olarak adlandırılmıştır. Sosyoekonomik ve demografik faktörleri belirlemede ve analiz etmede Etiyopya'daki dört dezavantajlı bölgede doğum aralığının anlamlı bir etkisi olabilir. **Gereç ve Yöntemler:** Çalışmadaki veri seti, 2011'de yapılan Etiyopya Demografi ve Sağlık Çalışması'ndan elde edilmiştir. Veri, 15-49 yaş aralığında 3340 kadını içermektedir. Doğum aralığı verisinin analizi için SPSS 16 ve STATA 11 programları kullanılarak Kaplan-Meier ve Cox oransal hazard modeli kullanılmıştır. Kaplan-Meier medyan süresi, kadınların sosyoekonomik ve demografik karakterlerine göre doğum aralığındaki farklılıkları incelemek için kullanılmıştır. **Bulgular:** Sonuçlar, yaklaşık olarak tüm doğum aralıklarında eğitimli kadınlar, zengin kadınlar, ortodoks, şehirde yaşayan kadınlar, Benishangul-gumuz bölgesine mensup kadınlar ve çocukları sağ kalan kadınların doğum aralıklarından daha uzun olduğunu belirtmektedir. Cox regresyon analiz sonuçları, kadınların eğitim düzeyinin ve sağlık indekslerinin bazı doğum aralıklarında anlamlı olmasına rağmen bölgenin, indekslenen çocukların sağkalım durumunun ve ikamet yerinin tutarlı olarak anlamlı olduğunu açığa çıkarmıştır. **Sonuç:** Yaklaşık olarak tüm doğum aralıkları, eğitimli kadınlar, zengin kadınlar, ortodoks, şehirde yaşayan kadınlar, Benishangul-gumuz bölgesine mensup kadınlar ve çocukları sağ kalan kadınlar daha uzun doğum aralıklarına sahiptir.

Anahtar Kelimeler: Doğum aralıkları; dezavantajlı bölgeler; yaşam analizi; oransal hazard modeli

doi: 10.5336/biostatic.2015-44961

Turkiye Klinikleri J Biostat 2015;7(2):63-76

A birth interval, defined as the length of time between two successive live births, indicates the pace of childbearing. Information on birth intervals provides insight into birth spacing patterns, which have far-reaching impact on both fertility and child mortality levels. The analysis of birth intervals explicitly recognizes the distinct renewable nature of the fertility process. It considers the progression from one birth to the next during a woman's reproductive life. This study look at the socio-economic, proximate (biological) and demographic factors associated with women's preferred birth intervals across four disadvantaged regions of Ethiopia: namely Affar, Somali, BenishangulGumuz and Gambela.

Research findings show also that births occurring within 2 years are riskier and their intervals are considered to be too short.[1] Recent findings show that intervals of 3 to 5 years are safer for both mother and infant compared to ≤ 2 years.[2-4] However, too long inter-birth intervals (>5 years) are associated with increased risk of complications such as preeclampsia because the mother loses protective effect from previous pregnancy.[2] Poorly spaced pregnancies have been documented worldwide to result into unwanted maternal and child health outcomes.

The World Health Organization (WHO) recommends an inter-birth interval length of at least 33 months between two live births in order to reduce the risk of adverse maternal and child health outcomes. However, birth spacing practices in many developing countries, including Ethiopia remain scantly addressed. According to EDHS 2011the median birth interval is 34 months in Ethiopia which is slightly greater than reported in 2005 that is 33.3 months. Empirical evidence from many different cultural settings has identified several correlates of birth intervals including breast feeding, contraceptive use, and maternal education.[5,6] However, to the investigators knowledge there is limited evidence and there is no study conducted so far to assess factors that determine birth intervals among four disadvantaged regions of Ethiopia. Therefore, analysis of factors that can influence birth interval among women will provide regional planners useful information that could encourage optimal intervals.

METHODOLOGY

The study used the data collected in the Ethiopian Demographic and Health Survey (EDHS) conducted in 2011. The 2011 EDHS used a two-stage stratified nationally representative sample of households for the data collection. In the first stage, the sample included 624 EAs, 187 in urban areas and 437 in rural areas. In the second stage, a fixed number of 30 households were selected for each EA. Of all the selected 18,720 households, 5,610 are in urban areas and 13,110 are in rural areas. In 2011 EDHS, a total of 17,385 eligible women were identified and interviews were completed for 16515 women, yielding a response rate of 95 percent. Details of the reproductive history of women were collected using the individual women's questionnaire together with background information. The study didn't consider those women who had twin or multiple births and women whose birth intervals are less than 9 months are excluded from the analysis. With these restrictions, the study has finally found 3340women for first birth interval that had at least one live birth in four disadvantaged region of Ethiopia. Therefore the study has found 3340 women for first birth interval analysis. Accordingly, the study considers 2805, 2285 and 1864 women for second, third and fourth birth intervals respectively.

SURVIVAL DATA ANALYSIS

THE SURVIVAL FUNCTION

Let T be a random variable, which can take any non-negative value, associated with the actual survival times, t (time of having birth). When the random variable T has a probability distribution with underlying probability density function $f(t)$, the distribution function (cumulative distribution function) of T is given by:

$$F(t) = P(T<t) = \int_0^t f(u)du, \quad t \geq 0 \qquad (1)$$

Which represents the probability that a subject selected at random will have a survival time less than some stated value t. Then, the survival function $S(t)$ is defined as:

$$S(t) = P(T \geq t) = 1 - F(t) \quad (2)$$

The survivor function can therefore be used to represent the probability that an individual survives (without having birth) from the time origin to sometime beyond t. The survival function is the probability that an individual will survive at time tor beyond t, and then relationship between the probability density function f (t) and S (t) will be:

$$f(t) = \frac{d(1-S(t))}{dt} = \frac{-dS(t)}{dt} \quad (3)$$

THE HAZARD FUNCTION

The hazard function is widely used to express the risk or hazard of experiencing the event (having birth) at some time t, and is obtained from the probability that an individual experiencing the event at time t, conditional on he or she having survived (without having birth) to that time. That is, the function represents the instantaneous failure rate for an individual surviving to time t.

The hazard function h(t) is defined by:

$$h(t) = \lim_{\Delta t \to 0} \frac{p\{\text{an individual fails in the time interval}(t, t+\Delta t) \backslash \text{it survives until time } t\}}{\Delta t}$$

$$h(t) = \lim_{\Delta t \to 0} \frac{P[t < T < t + \Delta t / T \geq t]}{\Delta t} \quad (4)$$

By applying the theory of conditional probability and the relationship in equation (3), the hazard function can be expressed in terms of the underlying probability density function and the survivor function as follows.[7]

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt}\{\log S(t)\} \quad (5)$$

The corresponding integrated or cumulative hazard function H(t) is defined by:

$$H(t) = \int_0^t h(u)du = -\log S(t) \quad (6)$$

Hence the survival function can be rewritten as

$$S(t) = \exp\{-H(t)\}, \quad (7)$$

The hazard rate is not a probability, it is a probability rate. Therefore it is possible that a hazard rate can exceed one in the same fashion as a density function f (t) may exceed one. Survival data are summarized through estimates of the survival and hazard function. The Kaplan-Meier, Nelson-Aalen and Life Tables are the three commonly used methods for estimating survival and hazard functions.

KAPLAN-MEIER ESTIMATOR

The first step in the analysis of ungrouped censored survival data is normally to obtain the Kaplan-Meier estimate of the survivor function. The Kaplan-Meier estimator is the standard estimator of the survival function (Collett, 2003). This estimator is also known as the Product-Limit estimator of the survivor function. Chakraborty et al. (1996) had used Product Limit Survivorship Function to study the differential pattern of birth intervals in Bangladesh. This method is non-parametric or distribution-free since it does not require specific assumption to be made about the underlying distribution of the survival times. Suppose the data consist of n survival times $t_1, t_2 \ldots\ldots\ldots t_n$ i.e. time since the previous birth of a child to that woman and some of these observations are right-censored times, i.e. for some of the t_j, it is only know that individual j was still without having births at time t_j. Let r be the number of distinct failure times, $r \leq n$, and $t_{(1)} < t_{(2)} < \ldots\ldots < t_{(r)}$ be the ordered failure times. And assume that n_j number of women just prior to time t_j exposed to the risk of having birth. And d_j is the number of women having births at time t_j.

Then the Kaplan-Meier estimator of the survival function at time t is given by:

$$\hat{S}(t) = \prod_{j=1}^{k} \left\{\frac{n_j - d_j}{n_j}\right\} \quad (8)$$

for $t_{(k)} \leq t_{(k+1)}$, k=1,2,…,r, with $\hat{S}(t)=1$ for $t < t_{(1)}$.

The standard error of the Kaplan-Meier survival estimator which is also known as the Greenwood's formula (Collett, 2003) is given as:

$$Se\{\hat{S}(t)\} = \hat{S}(t)\left\{\sum_{j=1}^{k} \frac{d_j}{n_j(n_j - d_j)}\right\}^{1/2} \quad (9)$$

Comparing Survival Distributions

The Kaplan-Meier estimator of the survival and hazard functions is the basic quantity in describing the overall survival experience. Many types of survival curves can be shown, but the important point to note is that they all have the same basic properties. They are monotonic, non increasing functions equal to one at zero and zero as the time approaches infinity. Their rate of decline, of course, varies according to the risk of experiencing the event at time t but it is difficult to determine the essence of a failure pattern by simply looking at the survival curve. Moreover, it is difficult to provide useful information about how time to event distributions differs among groups. Hence, it is frequently of interest to compare the survival of one group of study subjects with another. There are several tests to compare survival functions between two or among several groups. Most tests can be computed from contingency tables for those at risk at each event time.

The general form of the test statistic for the comparison of survival functions between two groups can be defined as follows:

$$Q = \frac{\left[\sum_{j=0}^{r} w_j (d_{1j} - \hat{e}_{1j})\right]^2}{\sum_{j=0}^{r} w_j^2 \hat{v}_{1j}} \quad (10)$$

Where: r is the number of rank-ordered failure times (event times).

w_j is the weight for censor adjustment at time $t_{(j)}$.

$\hat{e}_{1j} = \frac{n_{1j} d_j}{n_j} =$ is the expected number of individuals who experience an event at time $t_{(j)}$ in group 1,

$\hat{v}_{1j} = \frac{n1j\ n2j\ dj(nj-dj)}{nj^2(nj-1)} =$ is the variance of the number of event occurs at time $t_{(j)}$ in group 1,

d_{1j} is the observed number of failure (event occur) at time $t_{(j)}$ in group 1,

n_{1j} is the number of individuals at risk of event occur in the first group just before time $t_{(j)}$,

n_{2j} is the number of individuals at risk in the second group just before time $t_{(j)}$,

d_j is the total number of events occurred at $t_{(j)}$,

n_j is the total number of individuals at risk before time $t_{(j)}$.

Under the null hypothesis that the two survivorship functions are the same, and assuming that the censoring experience is independent of group, and that the total number of observed events and the sum of the expected number of events is large, Q follows a chi-square distribution with one degree of freedom.

Log-rank test sometimes called the Mantel-Haenszel test and Cox Mantel log-rank test is the most frequently used test which is based on weights equal to one, i.e. $w_j = 1$. The log-rank test is a non-parametric test for comparing two or more independent survival curves. Since it is a non-parametric test, no assumptions about the distributional form of the data need to be made. For the comparison of two groups of survival data the log rank test statistic is given (Hosmer and Lemesho, 1989) by:

$$Q_{LR} = \frac{\left[\sum_{j=0}^{r} (d_{1j} - \hat{e}_{1j})\right]^2}{\sum_{j=0}^{r} \hat{v}_{1j}} \quad (11)$$

The statistic Q_{LR} follows a chi-square distribution with one degree of freedom. It tests the extent to which the observed survival times in the two groups deviate from those expected under the null hypothesis of no group differences.It is used to compare Survival Distribution of various categories of factors (Nathan, 1966). For the two groups, Hypotheses are given as

$H_o : S_1(t) = S_2(t)$

$H_1 : S_1(t) > S_2(t)$

Khan and Raeside (1998) had used Log-Rank Test to compare Survival Distribution across various categories of factors for the determinants of first and subsequent births in urban and rural areas of Bangladesh.

Breslow's test (also known as Gehan's generalized Wilcoxon test) (Collett, 2003) is applicable to data where there is progressive censoring. It is more powerful than the log-rank test when the hazard functions are not parallel and where there is little censoring.

It has low power when censoring is high. It gives more weight to early failures.

The Wilcoxon test statistic:

$$Q_{WT} = \frac{\left[\sum_{j=0}^{r} n_j\left(d_{1j} - \hat{e}_{1j}\right)\right]^2}{\sum_{j=0}^{r} n_j^2 \hat{v}_{1j}} \quad (12)$$

has a chi-square distribution with one degree of freedom when the null hypothesis is true. This test uses weights equal to the number of subjects at risk at each survival time, i.e. $w_j = n_j$.

The log-rank test and Wilcoxon test can easily be generalized to the comparison of more than two groups. The statistic for $g > 2$ groups follows an approximate χ^2 distribution with g−1 degrees of freedom (Collett, 2003).

THE COX PROPORTIONAL HAZARDS (PH) REGRESSION MODEL

The Cox proportional hazards regression model, also referred to as the Cox model or the relative risk model, is a flexible tool for assessing the relationship of multiple predictors to a right-censored, time-to-event outcome, and has much in common with linear and logistic models. Since the dependent variables, length of birth intervals are time dependent event, the proportional hazard regression is made to the relative risk of covariates of having birth among women in four disadvantaged regions of Ethiopia using 10 independent variables as predictors.

The model assumes that all women with same covariates have identical risk of having birth over the course of study but these may vary among the groups with different covariates. The cases only considered here are those women who already have one child but have not given birth to a second child until the survey time are treated as "censored" for the first birth interval because in retrospective survey it is not possible to follow them until they either have a birth or reach menopause or even no longer to conceive, those women who already have two children but not given birth to a third child are censored for the second birth interval. Similarly, women with only three and four children but not given to the fourth and fifth child on the survey date are considered as "censored" cases for the third and fourth birth intervals respectively.

Let T denote failure time (having birth) and $\mathbf{X} = (x_1, x_2, \ldots, x_m)'$ represent a vector of available covariates. We are interested in modeling and determining the relationship between T and **X**.

For the PH model, the hazard function is:[8]

$$h(t, \mathbf{X}, \boldsymbol{\beta}) = h_o(t).\exp(\mathbf{X}'\boldsymbol{\beta}) \quad (13)$$

Where

$h(t, \mathbf{X}, \boldsymbol{\beta})$ represents the hazard function at time t with covariates $\mathbf{X} = (x_1, x_2, \ldots, x_m)'$,

$h_0(t)$ is an unspecified baseline hazard function that characterizes how the hazard function changes as a function of survival time (independent of the covariates),

$\beta = (\beta_1, \beta_2, \ldots, \beta_m)'$ is a column vector of m regression parameters,

$e^{(X'\beta)}$ = Characterizes how the hazard function changes as a function of subject covariates,

t is the failure time (having birth).

For two different individuals with covariates $\mathbf{X_1} = (x_{11}, x_{12}, \ldots, x_{1m})'$ and $\mathbf{X_2} = (x_{21}, x_{22}, \ldots, x_{2m})'$, the proportion

$$\frac{h(t, X1, \beta)}{h(t, X2, \beta)} = \frac{\exp(X_1\beta)}{\exp(X_2\beta)} = \exp\left(\left(\mathbf{X}'_1 - \mathbf{X}'_2\right)\boldsymbol{\beta}\right), \quad (14)$$

called the hazards ratio (HR), is constant with respect to time t. This defines the proportional hazards property (i.e., the hazard functions for two different levels of a covariate are proportional for all values of t).As with linear and logistic regression modeling, the statistical goal of survival analysis is to obtain some measure of effect that will describe the relationship between a predictor variable of interest and time to failure, after adjusting for the other variables included in the model.

The Cox model formula has the property that if the X_i's are entirely zero, the formula reducesto the baseline hazard function. This property of the Cox model is the reason why $h_o(t)$ is called baseline function. Another appealing property of the Cox model is that, even though the baseline hazard part of the model is unspecified, it is still possible to estimate the β's in the exponential parts of the model. So, it can equally be regarded as linear model which is a linear combination of the covariates of the logarithm transformation of the hazard ratio. It is given as:

$$log\left\{\frac{h(t,X,\beta)}{h_0(t)}\right\} = log\{e^{\beta' X}\} = \beta' X = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \quad (15)$$

The quantity $\beta' X = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$ is called the linear combination of the Cox proportional hazards model.

The hazard function in the Cox model is called semi-parametric function since it does not explicitly describe the baseline hazard function, $h_0(t)$. The survival function of the proportional hazard model is estimated as:

$$S(t,X,\beta) = e^{-H(t,\beta' X)} \quad (16)$$

Where $H(t,X,\beta)$ is the cumulative hazard function at time t for a subject with covariate x.Since we have assumed that survival time is absolutely continuous, the value of the cumulative hazard function is expressed as:

$$H(t,X,\beta) = H_0(t).\exp(\beta' X) \quad (17)$$

Consequently, from the proportional hazards function, we obtained the survivor function given by:

$$S(t,X,\beta) = [So(t)]^{\exp(\beta' X)} \quad (18)$$

Where: $H_0(t)$ is the baseline cumulative hazard function and $S_0(t)$ is the baseline survival function.[7] And estimation is made through maximum likelihood function.

RESULTS AND DISCUSSION

RESULTS

The study included 3340 women who had at least one live birth during the five years preceding the date of the survey; with 84.1% having second live birth and 15.9% without having second live birth (Table 1). Therefore the study has found 3340 women for first birth interval analysis. Accordingly, the study considers 2805, 2285 and 1864 women for second, third and fourth birth intervals respectively (Table 1). From a total of 2805, 2285 and 1864 women 82.2%, 81% and 77.7% were having third, fourth and fifth live births and 17.8%, 19% and 22.32% were without having fourth, fifth and sixth live births respectively (Table 1). Of the total of 3340 women included 70.7% had no work, 81.9% live in rural part of the four disadvantaged regions of Ethiopia (Table 1). Regarding mother's religion 14.6% were Orthodox, 61.3% Muslim, 20.6% Protestant and 3.6% were others (Table 1). With regard to educational attainment, about 74% of mothers and 60.6% of fathers had no education while 22% of mothers and 27% of fathers had primary education and the remaining 4% mothers and 12.4% fathers had attained secondary and higher education level (Table 1). Among 3340 live births in first birth interval 54% were female and 20.2% died as infant (before two years). About 59.0% of the households were classified as poor, 10.8% of the households were classified as middle income, while 30.2% were rich (Table 1).

The mean and median duration of first birth interval in four disadvantaged regions of Ethiopia for surviving women were 28.364 and 26.000 respectively (Table 2). Second birth interval is the shortest among all birth intervals Length of second birth interval is approximately 27 months. Marginal difference is observed in the length of first, third and fourth birth interval length. The average length of these birth intervals is one month more than average length of second birth interval (Table 2). It means birth spacing behavior of the four disadvantaged regions of Ethiopian women is almost same for different parities.

The graph of the estimate of overall Kaplan-Meier survivor function (survival curve) Figure 1-4, showed that the probability of mothers not having successive birth is high at ninth months, which relatively decreases as follow up time (months after delivery of index child) increases.

The Log-rank test was performed to investigate the significance of the observed difference in

TABLE 1: Summary of some important socio-economic, demographic and proximate characteristics of mother and her child in four disadvantaged regions of Ethiopia.

	1st Birth interval			2nd Birth interval			3rd Birth interval		
Covariates	**Event/having birth/**	**Censored**	**Total**	**Event/having birth/**	**Censored**	**Total**	**Event /having birth**	**Censored**	**Total**
Mother's education									
No education	2197 (88.9%)	275	2472	1900 (86.1%)	307	2207	1582 (84.1%)	298	1880
Primary	534 (80.4%)	201	735	363 (69.8%)	157	520	248 (68.5%)	114	362
Secondary & Higher	77 (57.9%)	56	133	43 (55.1%)	35	78	20 () (46.5%)	23	43
Father's education									
No education	1792 (88.6%)	231	2023	1564 (86.9%)	236	1800	1314 (84.8%)	235	1549
Primary	726 (72.7%)	177	903	553 (77.1%)	164	717	414 (76%)	131	545
Secondary & Higher	290 (70%)	124	414	189 (65.6%)	99	288	122 (64.9%)	69	191
Region									
Affar	791 (83.8%)	153	944	646 (81.6%)	146	792	528 (82%)	116	644
Ben-gumuz	761 (83.7%)	148	909	626 (82.6%)	132	758	500 (80.3%)	123	623
Ethio-Somali	600 (89.7%)	69	669	529 (87.6%)	75	604	452 (85.6%)	76	528
Gambella	656 (80.2%)	162	818	505 (77.6%)	146	651	70 (75.5%)	120	490
Residence									
Rural	2329 (85.2%)	405	2734	1965 (83.9%)	376	2341	1558 (82%)	343	1901
Urban	479 (79.0%)	127	606	341 (73.5%)	123	464	292 (76%)	92	384
Sex									
Female	1521 (84.4%)	281	1802	1209 (82.5%)	256	1465	976 (81.1%)	228	1208
Male	1287 (83.7%)	251	1538	1097 (81.9%)	243	1340	874 (71.9%)	207	1081
Survival status									
Dead	634 (93.9%)	41	675	530 (93.3%)	38	568	393 (93.6%)	27	420
Survived	2174 (81.6%)	491	2665	1776 (79.4%)	461	2237	1457 (79.1%)	408	1865
Religion									
Orthodox	377 (77.6%)	109	486	287 (76.55)	88	375	219 (76.3%)	68	287
Muslim	1753 (85.7%)	293	2046	1486 (84.6%)	270	1756	1229 (87%)	252	1481
Protestant	578 (83.9%)	111	689	574 (79.6%)	117	574	345 (77.9%)	98	443
Other	100 (84%)	19	119	76 (76%)	24	100	57 (77%)	17	74
Wealth index									
Poor	1715 (87.1%)	254	1969	1443 (84.1%)	273	1716	1185 (82.9%)	245	1430
Medium	309 (85.4%)	53	362	263 (85.4%)	45	308	226 (80.4%)	55	281
Rich	784 (77.7%)	225	1009	600 (76.8%)	181	781	439 (76.5)	135	574
Work status									
Not working	1998 (84.6%)	364	2362	1652 (82.7%)	346	1998	1329 (81%)	312	1641
Working	810 (82.8%)	168	978	654 (81%)	153	807	521 (80.9%)	123	644
Access to mass media									
No access	2237 (85.4%)	382	2619	1872 (83.6%)	366	2238	1507 (81.4%)	344	1851
Have access	571 (79.2%)	150	721	434 (76.5%)	133	567	343 (79%)	91	434

TABLE 2: Mean and median for survival time of birth intervals and their 95% CI & SE.

	Mean				Median			
			95% Confidence Interval				95% Confidence Interval	
Birth intervals	**Estimate**	**Std. Er.**	**Lower Bound**	**Upper Bound**	**Estimate**	**Std. Er.**	**Lower Bound**	**Upper Bound**
1st Birth interval	28.364	0.262	27.851	28.878	26.000	0.185	25.637	26.363
2nd Birth interval	26.722	0.232	26.267	28.178	26.000	0.217	25.575	26.425
3rd Birth interval	28.815	0.303	28.222	29.408	27.000	.297	26.418	27.582
4th Birth interval	28.335	0.329	27.691	28.980	26.000	0.299	25.414	26.586

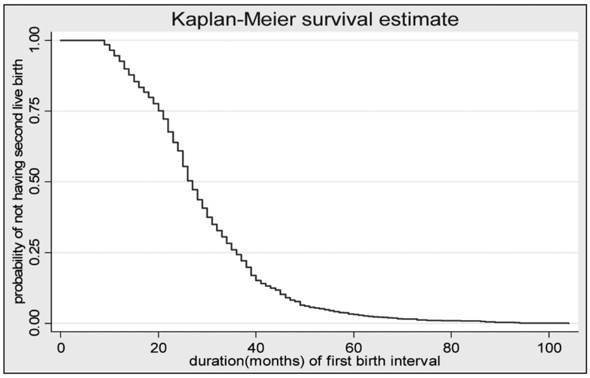

FIGURE 1: The plot of the overall estimate of Kaplan-Meier survivor function for women who were not having second live birth at different durations (months) of first birth interval.

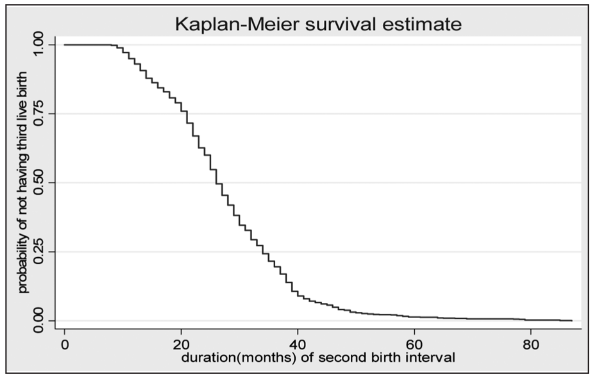

FIGURE 2: The plot of the overall estimate of Kaplan-Meier survivor function for women who were not having third live birth at different durations (months) of second birth interval.

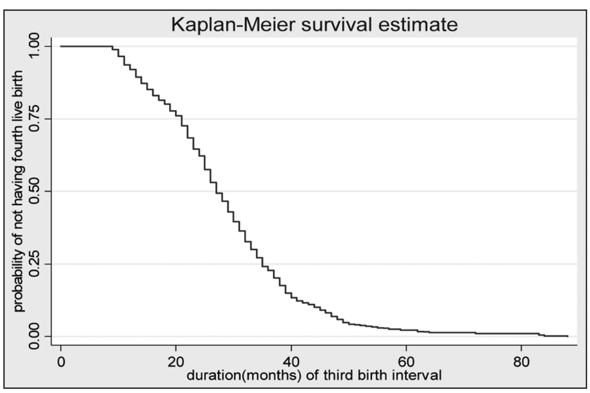

FIGURE 3: The plot of the overall estimate of Kaplan-Meier survivor function for women who were not having fourth live birth at different durations (months) of third birth interval.

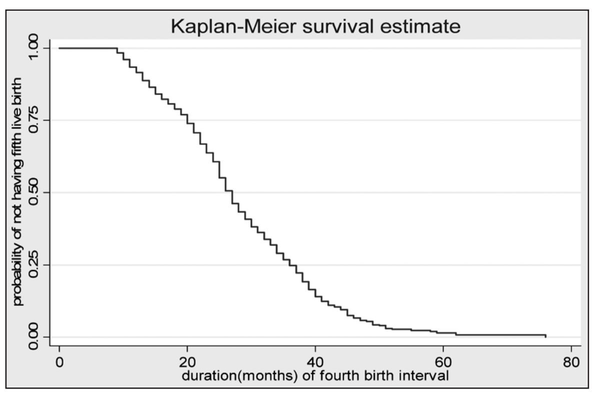

FIGURE 4: The plot of the overall estimate of Kaplan-Meier survivor function for women who were not having fifth live birth at different durations (months) of fourth birth interval.

the Kaplan-Meier estimates of the survivor functions among different categories of the covariates (Table 3). There is significant variations in the failure time (occurrence of birth) of different categories for, educational level of woman, education level of husband, wealth index, religion, survival status of index child, place of residence, region for all birth intervals (Table 3). There were also significant variations in the failure time (occurrence of birth) of different categories for sex of index child and exposure to mass media for fourth birth interval (Table 3). It means significant variability exists between birth intervals among various categories of these factors. Sex of preceding child, access to

TABLE 3: Results of log-rank test of equality of survival distribution for the different categorical covariates

		1st Birth interval		2nd Birth interval		3rd Birth interval		4th Birth interval	
Covariates	**Df**	**Chi-square**	**p-value**	**Chi-square**	**p-value**	**Chi-square**	**p-value**	**Chi- square**	**p-value**
Mother's educational level	2	52.045	0.000	33.812	0.000	50.786	0.000	17.864	0.000
Region	3	25.364	0.000	19.983	0.000	58.947	0.000	23.919	0.000
Father's educational level	2	48.352	0.000	19.096	0.000	35.937	0.000	13.987	0.001
Religion	3	43.109	0.000	10.723	0.000	38.146	0.000	26.895	0.000
Residence	1	17.827	0.000	22.616	0.000	8.948	0.003	8.154	0.004
Sex of index child	1	3.641	0.056	0.106	.744	1.780	0.182	7.280	0.007
Current Work status of mothers	1	3.035	0.082	1.244	0.265	2.194	0.139	4.149	0.042
Wealth index	2	8.071	0.018	19.696	0.000	21.208	0.000	39.130	0.000
Media access	1	2.018	.155	.995	0.318	0.450	0.503	3.013	0.083
Survival Status of index child	1	97.283	0.000	18.826	0.000	104.418	0.000	86.099	0.000

mass media and work status of woman do not have significant variation among their categories in first, second and fourth birth intervals. But sex of index child and exposure to mass media has significant variations among their categories in fourth birth interval (Table 3).

Cox Proportional Hazard Regression Model Results

Model adequacy

At this point we have a preliminary model and the next step is to assess its fit and adherence to key assumptions before we move to interpretation of the results obtained. We start here first by checking the overall goodness of fit using r-square and LR, Score and Wald tests. We then proceed to check the proportionality assumption for each covariate included in the final model.

Overall goodness of fit

The value of R_p^2 for each birth interval calculated as:

$R_p^2=1-\exp[\frac{2}{n}(L_O - L_p)] = 1-\exp[\frac{2}{3340}((-20017.810-(-19965.088))]$ = 0.0311 for first birth interval.

$R_p^2=1-\exp[\frac{2}{n}(L_O - L_p)] = 1-\exp[\frac{2}{2805}((-16118.271-(-16081.0195))]$ = 0.0262 for second birth interval.

$R_p^2=1-\exp[\frac{2}{n}(L_O - L_p)] = 1-\exp[\frac{2}{2285}((-12637.9325-(-12566.2095))]$ =0.0608 for third birth interval.

$R_p^2=1-\exp[\frac{2}{n}(L_O - L_p)] = 1-\exp[\frac{2}{1864}((-9621.500-(-9592.308))]$ = 0.0308 for fourth birth interval.

Due to the presence of high censoring the value of R_p^2 is very low for all birth intervals.

From Table 4 we can see that the likelihood ratio, Score and Wald test statistics are all significant at the 1% level of significance for all birth intervals. Thus, the fitted model is a good-fit for all birth intervals.

Testing the Proportional Hazards Assumption

Two basic assumptions of the Cox model are log-linearity and proportional hazards. Just like other regression models, these assumptions need to be examined. Since all covariates used in the final model are categorical, there is no need of checking linearity. The validity of Cox's regression analysis relies heavily on the assumption of proportionality

TABLE 4: Results of Likelihood ratio, Score and Wald tests of fit of the final model.

	First birth interval			Second birth interval			Third birth interval			Fourth birth interval		
Test	**Chi Square**	**Df**	**p value**	**Chi Square**	**Df**	**p value**	**Chi Square**	**Df**	**p value**	**Chi Square**	**Df**	**p value**
Likelihood Ratio	112.7480	6	<.0001	59.6424	4	<.0001	105.4590	5	<.0001	62.5731	5	<.0001
Score	108.2604	6	<.0001	56.4028	4	<.0001	102.0546	5	<.0001	61.6755	5	<.0001
Wald	107.6993	6	<.0001	56.0137	4	<.0001	100.6297	5	<.0001	61.1800	5	<.0001

Testing Global Null Hypothesis: BETA=0.

TABLE 5: Result of test of proportionality assumption for each covariate in the preliminary final model.

Birth intervals	Covariates	DF	Parameter estimates	Standard Error	Chi-square	p value	Hazard Ratio
1st birth interval	REGION	1	0.10283	0.13326	0.5955	0.4403	1.108
	SVST	1	-0.19537	0.34685	0.3173	0.5733	0.823
	MOEDC	1	-0.56846	0.28901	3.8688	0.0492	0.566
	RESIDENCE	1	-0.51452	0.40328	1.6278	0.2020	0.598
	RELIGION	1	-0.31275	0.20230	2.3900	0.1221	0.731
	WEALTH	1	-0.01834	0.17903	0.0105	0.9184	0.982
	REGION*log(time)		-0.04381	0.04096	1.1438	0.2848	0.957
	SVST*log(time)	1	-0.00203	0.10820	0.0004	0.9850	0.998
	MOEDC*log(time)	1	0.10211	0.08818	1.3410	0.2469	1.108
	RELIGION*log(time)	1	0.10359	0.06195	2.7965	0.0945	1.109
	RESIDENCE*log(time)	1	0.11224	0.12268	0.8371	0.3602	1.119
	WEALTH* log(time)	1	-0.02179	0.05500	0.1570	0.6919	0.978
2nd birth interval	SVST	1	0.19008	0.41174	0.2131	0.6443	1.209
	REGION	1	-0.11073	0.15433	0.5148	0.4731	0.895
	MOEDC	1	0.47753	0.38452	1.5423	0.2143	1.612
	RESIDENCE	1	-0.35232	0.48862	0.5199	0.4709	0.703
	SVST*log(time)	1	-0.12243	0.12917	0.8983	0.3432	0.885
	MOEDC*log(time)	1	-0.21554	0.12015	3.2182	0.0728	0.806
	REGION*log(time)	1	0.03807	0.04834	0.6201	0.4310	1.039
	RESIDENCE*log(time)	1	0.05006	0.15182	0.1087	0.7416	1.051
3rd birth interval	SVST	1	-0.86233	0.43145	3.9947	0.0456	0.422
	REGION	1	-0.01749	0.16486	0.0113	0.9155	0.983
	MOEDC	1	-0.34188	0.50200	0.4638	0.4958	0.710
	RESIDENCE	1	0.21926	0.52471	0.1746	0.6760	1.245
	WEALTH	1	-0.22079	0.23187	0.9067	0.3410	0.802
	SVST*log(time)	1	0.15980	0.13739	1.3529	0.2448	1.173
	MOEDC*log(time)	1	-0.00570	0.15553	0.0013	0.9708	0.994
	REGION*log(time)	1	0.00565	0.05145	0.0121	0.9125	1.006
	RESIDENCE*log(time)	1	-0.10059	0.16280	0.3818	0.5366	0.904
	WEALTH* log(time)	1	0.04561	0.07209	0.4003	0.5269	1.047
4th birth interval	SVST	1	-0.83871	0.50178	2.7938	0.0946	0.432
	RESIDENCE	1	-0.73828	0.64008	1.3304	0.2487	0.478
	REGION	1	-0.02895	0.18388	0.0248	0.8749	0.971
	WEALTH	1	0.28892	0.25167	1.3180	0.2510	1.335
	SEX	1	0.18118	0.40143	0.2037	0.6517	1.199
	SVST*log(time)	1	0.18113	0.16063	1.2715	0.2595	1.199
	REGION*log(time)	1	0.00565	0.05820	0.0094	0.9227	1.006
	RESIDENCE*log(time)	1	0.15828	0.19785	0.6400	0.4237	1.171
	WEALTH* log(time)	1	-0.13380	0.07894	2.8727	0.0901	0.875
	SEX* log(time)	1	-0.10146	0.12679	0.6403	0.4236	0.904

Linear Hypotheses Testing Results for each birth intervals

Label		Wald Chi-Square	DF	p value
test_proportionality	1st Birth interval	8.0076	6	0.2375
	2nd Birth interval	2.2296	4	0.6936
	3rd Birth interval	5.9920	5	0.3070
	4th Birth interval	5.2162	5	0.3901

of the hazard rates of individuals with distinct values of a covariate. If the proportionality assumption holds the lowest smoothing curve should be approximately horizontal line around zero and the distribution of residuals over time is random, with no particular trend with time. Alternatively, we can run a model with each covariate (individually) by introducing a time-dependent interaction term for that covariate. If the proportional hazards assumption is valid for the covariate, the time-dependent interaction term should not be significant. The following table display the SAS output of test of proportionality assumption.

From Table 5, we can see the Wald chi-square values and the corresponding p-values for each covariate in all birth intervals. Since the p-values for each interaction of covariate with logarithm of time are greater than 0.05, the proportionality assumption is satisfied. The global fit test also shows that the Wald chi-square test statistic is not significant which indicates that the proportional hazards assumption is not violated.

DISCUSSION OF THE RESULTS

The study assessed duration of birth interval and examined the demographic and socioeconomic determinants of birth interval in four disadvantaged regions of Ethiopia. The effects of the various covariates on transition to 1^{st} up to 5^{th} births are shown in Table 6.This table provide the final model results based on the determinants of length of birth interval.

From the final model (Table 6), the significant determinants of first birth interval are region, place of residence, religion, wealth, mothers' educational level and survival status of index child. Second birth interval is significantly determined by region, place of residence, mothers' educational level and survival status of index child. In the determination of third birth interval region, place of residence, wealth, mothers' educational level and survival status of index child has played significant role. While in the determination of fourth birth interval region, place of residence, sex of index child, wealth and survival status of index child contribute significantly.

In all birth orders the socio demographic factor (region) has influence in the length of all birth intervals (Table 6). The hazard ratios of having second birth(second child) for mothers of Benishangul-gumuz region and Ethiopia-somali region as compared to mothers of Affar region were 1.163(95% CI: 1.050, 1.289) and 1.120(95% CI: 1.000,1.254)respectively. This means mothers of Benishangul-gumuz and Ethiopia-somali had 16.3% and 12% more likely to have second birth than mothers of Affar region respectively (Table 6).

The hazard ratios of having third birth (third child) for mothers of Ethiopia-somali region as compared to mothers of Affar region was 1.158(95% CI:1.031, 1.301). This means mothers of Ethiopia-somali region mothers had 15.8% more likely to have third birth than mothers of Affar region. Mothers of Ethiopia-somali region as compared to mothers of Affar region had 12%, 15.8%, 33.2% and 12.9% more likely to have their second, third, fourth and fifth birth than mothers of Affar region respectively. This indicates that Ethiopia-somali region mothers had short birth interval (Table 6).

In this study education of women has shown positively related to birth spacing. The effect of education is significant in all models except for fourth birth interval (Table 6). Women's education has important significant effect on risk of having successive birth. It is observed that women with secondary and higher educational background had 32.7% ,38.8% and 35.7% less chance of having second birth (second child),third birth(third child) and fourth birth(fourth child) than non educated women respectively .From first to third birth intervals, the likelihood of having births decreases with the increasing level of education of women. The spacing between births after first birth among educated women indicates the intention to limit total children ever born. This finding is consistent with that educated women always had longer birth interval than non-educated women.[9]

Wealth index has shown consistent positive effect on birth spacing except for second birth interval. The estimated relative risks (hazard ratios)

TABLE 6: Parameter estimates of the final model.

	Covariates	B	SE	Wald	Df	Sig.	Exp(B)	95.0% CI for Exp(B)	
1st birth interval	REGION(Affar)	0.151	0.052	19.037	3	0.000			
	Ben-gumuz	0.114	0.058	8.388	1	0.004	1.163	1.050	1.289
	Ethio-Somali	-0.062	0.065	3.863	1	0.049	1.120	1.000	1.254
	Gambella			0.905	1	0.342	0.940	0.827	1.068
	RELIGION(Orthodox)			13.469	3	0.004			
	Muslim	0.236	0.065	13.205	1	0.000	1.266	1.115	1.437
	Protestant	0.177	0.072	5.953	1	0.015	1.193	1.035	1.376
	Other	0.142	0.112	1.620	1	0.203	1.153	0.926	1.434
	MOEDC(No edu)			27.074	2	0.000			
	Primary	-0.211	0.050	17.817	1	0.000	0.810	0.734	0.893
	Secondary& higher	-0.396	0.113	12.240	1	0.000	0.673	0.539	0.840
	WEALTH(poor)			12.981	2	0.002			
	Medium	-0.095	0.064	2.191	1	0.139	0.910	0.802	1.031
	Rich	-0.175	0.050	12.330	1	0.000	0.839	0.761	0.926
	SVST(Dead)	-0.191	0.048	15.925	1	0.000	0.826	0.752	0.908
	RESIDENCE(Rural)	-0.129	0.055	5.493	1	0.019	0.879	0.789	0.979
2nd birth interval	REGION(Affar)			17.489	3	0.001			
	Ben-gumuz	-0.093	0.057	2.693	1	0.101	0.911	0.815	1.018
	Ethio-Somali	0.147	0.059	6.141	1	0.013	1.158	1.031	1.301
	Gambella	-0.066	0.062	1.128	1	0.288	0.936	0.828	1.058
	MOEDC(No edu)			10.602	2	0.005			
	Primary	-0.099	0.061	2.647	1	0.104	0.906	0.804	1.020
	Secondary& higher	-0.491	0.163	9.092	1	0.003	0.612	0.444	0.842
	SVST(Dead)	-0.196	0.050	15.496	1	0.000	0.822	0.746	0.906
	RESIDENCE(Rural)	-0.224	0.063	12.668	1	0.000	0.800	0.707	0.904
3rd birth interval	REGION(Affar)			45.771	3	0.000			
	Ben-gumuz	-0.083	0.064	1.667	1	0.197	0.921	0.812	1.044
	Ethio-Somali	0.287	0.065	19.495	1	0.000	1.332	1.173	1.513
	Gambella	-0.149	0.071	4.410	1	0.036	0.861	0.749	0.990
	MOEDC(No edu)			19.241	2	0.000			
	Primary	-0.299	0.073	16.806	1	0.000	0.742	0.643	0.856
	Secondary& higher	-0.441	0.230	3.674	1	0.055	0.643	0.410	1.010
	WEALTH(poor)			6.403	2	0.041			
	Medium	-0.062	0.075	0.693	1	0.405	0.940	0.811	1.088
	Rich	-0.151	0.060	6.351	1	0.012	0.860	0.764	0.967
	SVST(Dead)	-0.366	0.057	40.665	1	0.000	0.693	0.620	0.776
	RESIDENCE(Rural)	-0.155	0.067	5.287	1	0.021	0.857	0.751	0.977
4th birth interval	REGION(Affar)			17.451	3	0.001			
	Ben-gumuz	-0.143	0.073	3.868	1	0.049	0.867	0.751	1.000
	Ethio-Somali	0.121	0.073	2.790	1	0.095	1.129	0.979	1.301
	Gambella	-0.154	0.079	3.788	1	0.052	0.857	0.733	1.001
	SVST(Dead)	-0.279	0.066	17.722	1	0.000	0.756	0.664	0.861
	WEALTH(poor)			12.046		0.002			
	Medium	-0.113	0.084	1.783		0.182	0.893	0.757	1.054
	Rich	-0.235	0.069	11.695		0.001	0.790	0.691	0.904
	RESIDENCE(Rural)	-0.261	0.082	10.204	1	0.001	0.770	0.656	0.904
	SEX(female)	-0.140	0.053	7.004	1	0.008	0.870	0.784	0.964

of having second, fourth and fifth births for rich women as compared to poor women are.839 (95% CI: .761, .926), .860(95% CI: .764, .967) and .790 (95% CI: .691, .904) respectively (Table 6). This means rich women were 16.1%, %, 14% and 21% less likely to have second, fourth and fifth births respectively than poor women. Apart from parity two women belonging to rich wealth index had longer birth interval as compared to those belonging to poor categories of wealth index. This finding is consistent with.[10]

Survival status of the index child is consistently significant in each birth intervals. The estimated relative risks (hazard ratios) of having second, third, fourth and fifth births for women whose index child survive as compared to women whose index child did not survive are .826 (95% CI: .752-.908), .822 (95% CI: .746-.906), .693 (95% CI: .620-.776) and .756 (95% CI: .664-.861) respectively (Table 6). This means women whose index child survive were 17.4%, 17.8%, 30.7% and 24.4% less likely to have second, third, fourth and fifth births respectively than women whose index child did not survive. This finding is consistent with previous research in Ghana and Kenya about the effects of the status of the index child on the risk of subsequent births.[11] Based on this, women with child loss experience are less likely to use contraception and more likely to discontinue if they are already using contraception. The reason behind this is that couples want make deliberate efforts to bear another child in the hope of replacing the lost one.

In all birth orders the socio demographic factor place of residence has influence in the length of birth interval (Table 6). The estimated relative risks (hazard ratios) of having second, third, fourth and fifth birth for urban women as compared to rural women are .879 (95% CI: .789-.979), .800(95% CI: .707-.904), .857 (95% CI: .751-.977) and .770 (95% CI: .656-.904) respectively. This indicates that women who are lived in urban area were 12.1%, 20%, 14.3% and 23% less likely to have second, third, fourth and fifth birth respectively than the women in rural area. The 95% confidence intervals also suggests that the risks of having second, third, fourth and fifth births for urban women could be as low as .789 and as high as.979, as low as .707 and as high as .904, as low as .751 and as high as .977, as low as.656 and as high as .904 respectively. Hence, urban mothers had birth interval for a relatively-longer duration than the rural Mothers did in first to fourth birth intervals. The reason may be the lack of educational facilities, working status and for lack of consciousness. This finding is consistent with.[3,12]

CONCLUSION

This study was intended to identify socio-economic and demographic determinants of birth interval in four disadvantaged regions of Ethiopia based on 2011 EDHS data. Ten covariates were selected for the study and 10 uni-variable Cox Proportional Hazards regression Model were developed for each birth interval to assess the relation between length of birth interval and the selected variables. Based on the results, the multi-variable Cox Proportional Hazards regression Model of length of birth interval was employed for each birth interval to select the most important determinants of birth interval.

The Kaplan-Meier survival estimate results showed that the probability of women not having successive birth is high in the ninth months, which relatively decreases as follow up time (months after delivery of index child) increases. The probabilities of not having second, third, fourth and fifth live births at ninth months were 98.92%, 98.97%, 98.95% and 98.5%respectively. The probability of not having second, third, fourth and fifth live births at 36 months were 20.15%, 19.51%, 22.80% and 24.88% respectively in four disadvantaged regions of Ethiopia . The mean and median duration of first birth interval in four disadvantaged regions of Ethiopia for surviving women were 28.364 and 26.000 respectively. Second birth interval is the shortest among all birth intervals Length of second birth interval is approximately 27 months. Marginal difference is observed in the length of first, third and fourth birth interval length. The average length of these birth intervals is one month more than average length of second birth interval.

The Kaplan Meier median length was used to examine birth interval differentials by socio economic and demographic characteristics of women. The results indicated that almost in all birth intervals, educated women, rich women, orthodox, urban women, women belonging to Benishangulgumuz region and women whose index child has survived have longer birth interval lengths

We recommend the following based on our findings:

- Government should motivate couples to increase the birth interval length in case of death of preceding child and also strengthen health programs. It is necessary for the maternal and child health. If long birth interval is promoted in case of death of preceding child, it will cause decline in fertility.
- Women should be assisted with the development and implementation of comprehensive family planning programs that have effective outreach services and are accessible geographically, socially and financially.
- Birth interval length is shorter than three years for all order births. There is need of effective policy for promotion of long birth space (at least 4 to 5 years) between two consecutive children. Lady health visitors can be used for this purpose.

REFERENCES

1. RamaRao S, Townsend J, Askew I. Correlates of Inter-birth Intervals: Implications of Optimal Birth Spacing Strategies in Mozambique. Population Council 2006:1-17.
2. Orji E, Shitu A, Makinde O, Sule S. Effect of prolonged birth spacing on maternal and perinatal outcome. East Afr Med J 2004;81(8): 388-91.
3. Ayanaw A. proximate determinants of birth interval length in Amhara region: the case of FagitaLekoma district, Awi-zone. Addis Ababa, Ethiopia, 2008.
4. Yohannes S, Wondafrash M, Abera M, Girma E. Duration and determinants of birth interval among women of child bearing age in Southern Ethiopia. BMC Pregnancy and Childbirth 2011;11:38.
5. Norton M. New evidence on birth spacing: promising findings for improving newborn, infant, Child, and maternal health. Int J Gynecol Obstet 2005;89(Suppl 1):S1-6.
6. Yigzaw M, Enquselassie F. Birth spacing and risk of child mortality at Kalu District South Wollo Zone of Amhara Region. North East Ethiopia. Ethiop Med J 2010;48(2):105-15.
7. Collett, D.(2003). Modeling survival data in medical research (Second edition).ces.1995; 8:2:69-72.
8. Cox DR. Regression models and life tables. Journal of the Royal Statistical Society: Series B (Methodological) 1972;34(2):187-220.
9. Ahbab M, Fazle R, Shamal C, Karmaker S, Ali M, Sayema S. (2012): Determinants of Birth Spacing and Effect of Birth Spacing on Fertility in Bangladesh.
10. Youssef RM. Duration and determinants of inter birth interval: community-based survey of women in southern Jordan. East Mediterr Health J 2005;11(4):559-72.
11. Gyimah, S. O. (2001). Childhood Mortality and Reproductive Behavior in sub-Saharan with Emphasis on Ghana and Kenya. A paper presented at the Population Association of America Meetings, Atlanta Hilton, and GA May 8-11, 2002.
12. Sharif Mahmood and Begum Zainab (2010): Analysis of Birth Interval in Bangladesh Using Product-Limit Estimate and Accelerated Failure Time Regression Model.