

## Estimation of Survival Function of Duration of Breast-feeding in the Presence of Some Censored Data

<sup>1</sup>N. Sanajaoba Singh and <sup>2</sup>N. Sharat Singh

<sup>1</sup>Directorate of Census Operations, Porompat, Manipur, India

<sup>2</sup>Thoubal College, Thoubal, Manipur, India

---

**Abstract:** Breast-feeding (BF) has direct impact not only on the health of young children but on reducing fertility through duration of Post Partum Amenorrhoea (PPA). The present study is to examine the pattern of BF and its associated factors by estimating the survival functions of duration of BF. A cross sectional study of 1225 ever married women of reproductive age having at least one live birth was conducted during 2008-09 in valley districts of Manipur under cluster sampling scheme. Survival analysis techniques using life table method and Cox's pH Model has been adopted through SPSS vs. 16. The median duration of BF is found to be 20.37 months. Place of residence ( $p < 0.01$ ), employment status ( $p < 0.01$ ), educational level ( $p < 0.05$ ), family income ( $p < 0.05$ ) and parity ( $p < 0.05$ ) have their significant impacts on the dynamics of BF duration. The current duration of BF is about 20 months which is far behind the National figure of 25 months and WHO recommendation of 24 months. Thus, program implementers are needful to promote BF at an adequate level.

**Key words:** Censored case, parity, log rank test, Cox's Model, relative risk, India

---

### INTRODUCTION

Recent research has documented the benefits of breast-feeding not only for the health of the infant but also for lowering fertility. The suckling of infant stimulates the flow of hormones within the mother that delay the return of ovulation. Extended breast-feeding lengthens the period of Post Partum Amenorrhoea (PPA) which may be considered as a component of birth interval and thus lengthens the inter live birth intervals which in turn indirectly reduces fertility (Huffman, 1984; Thapa and Williamson, 1990; Lechtig and Jelliffe, 1990; Mannan and Islam, 1995).

Extensive research has recently proven that BF reduces the morbidity and mortality of infants and children and promotes maternal health (Chhabra *et al.*, 1998; Yngve and Sjostrom, 2001; Lawrence, 2002; Koosha *et al.*, 2008). BF prevents some bacterial infections such as diarrhoeal disease, urinary tract infections and upper and lower respiratory tract as well as bacteraemia and meningitis. It also reduces certain immunological disorders such as atopic eczema, food and respiratory allergy and also the risk of chronic disorders later in life such as Crohn's and coeliac diseases, childhood cancers and diabetes mellitus (Agrawal *et al.*, 2003; Koosha *et al.*, 2008; Aryal, 2004). Some earlier studies have shown that breast-feeding might even protect mothers against breast cancer as well as some

types of ovarian cancers (Yngve and Sjostrom, 2001; Lawrence, 2002; Koosha *et al.*, 2008). Usually, BF is universal in most of the society. But its degree varies from society to society depending on the characteristics of mothers in the society such as education, occupation, availability of supplementary foods, socio-economic status, etc. (Aryal, 2004, 2005; Giashuddin and Kabir, 2004; Islam *et al.*, 2006). So the study of behaviour and differential of BF is very important in a society where the fertility, woman as well as child health is becoming a serious problem. Therefore, this study tries to touch these issues. The main objective of this study is to estimate survival function of the duration of Breast-feeding (BF) by using survival analysis techniques. This estimation is used to investigate the Breast-feeding (BF) patterns and also to identify the determinants which may have pronounced effects on the duration of BF.

### MATERIALS AND METHODS

A cross sectional study of 1225 ever married women of reproductive age having at least one live birth was conducted during the period from 1 January to 30 June, 2009 in valley districts of Manipur namely; Bishnupur, Thoubal, Imphal West and Imphal East under cluster sampling scheme. The pre-tested and semi-structural type interview schedule was used as tool for collecting the required information. The duration of BF is taken as the

response variable. The duration variable is defined as the child's age (in months) at the time of complete weaning, regardless of when consumption of other food began. For the children who have died or still being breast-fed at the time of survey, the duration is the child's age at death or at the time of survey. Such cases are treated as censored ones. The explanatory variables or so termed as covariates are socio-economic and demographic characters. The categorical variables are quantified by binary dummy variable (0, 1). The socio-economic variables include place of residence (urban = 1, rural = 0), educational level, employment status (employed = 1, unemployed = 0) and family income. The demographic variables are age at marriage, sex of previous child (male = 1, female = 0) and parity.

As the study is confined in censored data, statistical analysis is therefore carried out using survival analysis techniques.

**Non-parametric estimation of survival functions**

**The life-table analysis:** Berkson and Gage (1950) and Cutler and Ederer (1958) presented a non-parametric approach to estimate survival function using life table method.

Suppose the duration of breast-feeding (0, t) is under consideration. Let this interval be partitioned into a fixed sequence of intervals (t<sub>i</sub>, t<sub>i+1</sub>), i = 1, 2, 3....., k. Let n<sub>i</sub> the number of women who are breast-feeding at the beginning of the ith interval:

- d<sub>i</sub> = The number of women who terminate breast-feeding in the ith interval
- l<sub>i</sub> = The number of women reporting continuance of breast-feeding in the ith interval
- w<sub>i</sub> = The number of women with the death of child before the termination of breast-feeding in the ith interval
- P<sub>i</sub> = P [continuance of breast-feeding through ith interval/continuance of breast-feeding at the beginning of ith interval]
- Q<sub>i</sub> = P [termination of breast-feeding through ith interval/continuance of breast-feeding at the beginning of ith interval] = 1 - P

Life-table estimate of survival function, S(t<sub>k</sub>) is given by:

$$S(t_k) = P[T > t_k] = P[T > t_1] P\left[\frac{T > t_2}{T > t_1}\right] \dots \tag{1}$$

$$P\left[\frac{T > t_k}{T > t_{k-1}}\right] = \prod_{i=1}^k P_i$$

Where:

$$P_i = P\left[\frac{T > t_i}{T > t_{i-1}}\right]$$

If w<sub>i</sub> = 0 and l<sub>i</sub> = 0 that is if all the women terminate breast-feeding at the date of survey then:

$$\hat{P}_i = 1 - \frac{d_i}{n_i}$$

However with w<sub>i</sub> and l<sub>i</sub> non zero, researchers assume that on the average, those women with the continuance of breast-feeding during the ith interval and with the death of child during the ith interval are at risk for half of the interval. Therefore, the effective sample size is defined as:

$$n'_i = n_i - \frac{1}{2}(l_i + w_i)$$

and;

$$\hat{Q}_i = q_i = \frac{d_i}{n_i}$$

$$\hat{P}_i = p_i = 1 - \hat{Q}_i$$

**Non-parametric methods for comparing survival distributions**

**Log rank test:** In the case of censored data, log rank test is the most commonly used nonparametric test for comparing two survival functions (Peto and Peto, 1972). Suppose that there are two groups where the observations of duration of breast-feeding in group 1 are samples from a distribution with survival function S<sub>1</sub>(t) and the observations in group 2 are samples from the distribution with survival function S<sub>2</sub>(t).

The null hypothesis is H<sub>0</sub>: S<sub>1</sub>(t) = S<sub>2</sub>(t), against the alternative hypothesis, H<sub>1</sub>: S<sub>1</sub>(t) ≠ S<sub>2</sub>(t). Let t<sub>1</sub> < t<sub>2</sub> < ... < t<sub>g</sub> be distinct ordered duration of breast-feeding for combined sample of the two groups. Let d<sub>j</sub> is the number of women experiencing the termination of breast-feeding at t<sub>j</sub> in both groups (for complete data d<sub>j</sub> = 1 for every j). n<sub>j</sub> is the number of women at risk of termination of breast-feeding at t<sub>j</sub>. Researchers suppose, n<sub>1j</sub> is the number of women at risk at t<sub>j</sub> from group 1. n<sub>2j</sub> is the number of women at risk at t<sub>j</sub> from group 2. The log rank test compares the observed number of women experiencing termination of breast-feeding and expected (under H<sub>0</sub>) number of women experiencing termination of breast-feeding in group 1. Let, E is expected (under H<sub>0</sub>) number of women experiencing the termination of breast-feeding in:

$$\text{Sample 1} = \sum_{j=1}^g d_j \frac{n_{1j}}{n_j}$$

The variance of E is given by:

$$V = \sum_{j=1}^g d_j \frac{n_j n_{2j}}{n_j^2}$$

Let, O is observed number of women experiencing the termination of breast-feeding in sample (group) 1. Under  $H_0$ , the test statistics is given by:

$$Z = \frac{O - E}{\sqrt{V}}$$

has asymptotically standard normal distribution, i.e.:

$$Z = \frac{(O - E)}{\sqrt{V}} \sim \chi^2 \quad (2)$$

**Cox's Proportional Hazard Model:** In this part, Cox (1972)'s Proportional Hazard Model has been used to determine the effects of various socio-economic and demographic factors on the duration of waiting time to conception. The simple form of the model is given by:

$$\lambda(t; \underline{x}) = \lambda_0(t) \varphi(\underline{x}) \quad (3)$$

where,  $\lambda_0(t)$  is the baseline hazard function and is defined as the hazard function when all the x variables are ignored that is all x's equal zero.  $\varphi(\underline{x})$  is a parametric link function bringing in the covariates. It satisfies  $\varphi(0) = 1$  and  $\varphi(\underline{x}) \geq 0$  for all  $\underline{x}$ . The commonly used form of  $\varphi$  is:

$$\varphi(\underline{x}) = \varphi(\underline{x}, \underline{\beta}) = e^{\underline{\beta}' \underline{x}}$$

so that the ratio:

$$\frac{\lambda(t; \underline{x})}{\lambda_0(t)} = e^{\underline{\beta}' \underline{x}}$$

represents the Relative Risk of failure (RR) or so termed as the risk of termination of breast-feeding in the present study.

Suppose that  $y_{(1)} < y_{(2)} < \dots < y_{(k)}$  are the duration of breast-feeding for k women who have terminated breast-feeding before the survey data.

The rest of (n-k) study subjects are randomly censored. Suppose  $R(y_{(i)})$  to be the risk set at time  $y_{(i)}$  which is the set of women reporting the continuance of breast-feeding at time  $y_{(i)}^-$  (just prior to  $y_{(i)}$ ). The relevant likelihood function for the proportional hazards model in Eq. 3 is given by:

$$L(\underline{\beta}) = \prod_{j \in R(i)} \frac{e^{\underline{\beta}' \underline{x}_{(i)}}}{\sum_{j \in R(i)} e^{\underline{\beta}' \underline{x}_j}}$$

where,  $\underline{x}_{(i)}$  be the covariates associated with  $y_{(i)}$ . The maximum likelihood estimates of the parameters ( $\underline{\beta}$ ) in the proportional hazards model can be found by maximizing the log-likelihood function using numerical methods. The regression coefficient ( $\underline{\beta}$ ) is tested using Wald's test for its significant effect on the duration of BF and the test statistics is given by:

$$\hat{\underline{\beta}}' I^{-1} \hat{\underline{\beta}} \sim \chi_p^2$$

Where:

$$I = E \left( - \frac{\partial^2 \log L(\underline{\beta})}{\partial \underline{\beta}^2} \right)$$

The calculations in this study have been carried out using the package SPSS.

## RESULTS

**Life-table analysis:** Table 1 provides information on life table estimates of median duration of BF and the survival function of the BF with respect to socio-economic and demographic characteristics. The overall median duration of BF of the study population is 20.37 months. About 59, 35, 24, 16 and 13% of the mothers who have married during the age group of <15, 15-20, 20-25, 25-30 and  $\geq 30$  years, respectively terminate BF within 12 months. It also highlights that the median duration of BF increases with the increase in the age at marriage of mothers.

By the log rank test, the association between age at marriage and duration of BF is highly significant ( $\chi^2 = 26.181, p < 0.001$ ). It could be examined that 59% of mothers having parity zero terminate BF within 12 months. On contrary, 23% of mothers having parity one that of 22% of parity two, 20% of parity three and 19% of parity four and above terminate BF within 12 months. The median duration of BF increases with the increase in parity and this variation is highly significant in the study population irrespective of other covariates ( $p < 0.001$ ). Mothers residing in rural areas have longer duration of BF (20.85 months) than residing in urban areas (18.48 months) and relationship is found to be statistically significant ( $p < 0.001$ ). About 22% of mothers residing in rural areas terminate BF within 12 months whereas about 25% of mothers in urban areas terminate BF at the same time. Educational level of mother shows an inverse relationship with the duration of BF ( $p < 0.001$ ). About 50% college and university level mothers terminate BF within 12 months

Table1: Life table of duration of BF according to different characteristics of mother

Variables	Proportion of weaning of BF upto the months (1-S <sub>(t<sub>h</sub>)</sub> )				Median (month)	Log rank test
	6	12	18	24		
<b>Demographic</b>						
<b>Age at marriage of wife (years):</b>						
<15	0.16	0.59	0.69	0.77	10.04	$\chi^2 = 26.181,$ df = 4, p<0.001
15-20	0.13	0.35	0.45	0.68	19.38	
20-25	0.09	0.24	0.33	0.61	20.63	
25-30	0.08	0.16	0.28	0.58	21.39	
≥30	0.07	0.13	0.25	0.55	22.82	
<b>Parity:</b>						
0	0.16	0.59	0.85	0.90	10.72	$\chi^2 = 95.624,$ df = 4, p<0.001
1	0.10	0.23	0.38	0.66	20.72	
2	0.09	0.22	0.33	0.63	20.76	
3	0.08	0.20	0.32	0.61	20.84	
4+	0.05	0.19	0.30	0.52	23.28	
<b>Sex of the previous child:</b>						
Female	0.10	0.32	0.36	0.62	20.52	$\chi^2 = 0.240,$ df = 1, p>0.05
Male	0.09	0.24	0.34	0.61	20.54	
<b>Socio-economic</b>						
<b>Place of residence:</b>						
Urban	0.10	0.25	0.38	0.69	18.48	$\chi^2 = 24.970,$ df = 1, p<0.001
Rural	0.08	0.22	0.32	0.54	20.85	
<b>Educational level of mother :</b>						
No schooling	0.05	0.22	0.25	0.46	28.18	$\chi^2 = 35.405,$ df = 4, p<0.001
Primary school	0.08	0.26	0.30	0.59	20.99	
Sec. school	0.09	0.28	0.32	0.62	20.64	
Higher sec. school	0.10	0.35	0.41	0.63	20.12	
College and university	0.16	0.50	0.63	0.77	12.64	
<b>Employment status of mother:</b>						
Unemployed	0.09	0.22	0.14	0.40	20.55	$\chi^2 = 59.408,$ df = 1, p<0.001
Employed	0.12	0.41	0.22	0.53	15.40	
<b>Family income (Rs.):</b>						
<2000	0.05	0.15	0.25	0.46	26.10	$\chi^2 = 98.909,$ df = 5, p<0.001
2000-4000	0.06	0.19	0.31	0.49	25.43	
4000-6000	0.08	0.21	0.32	0.51	21.88	
6000-8000	0.09	0.25	0.33	0.58	20.68	
8000-10000	10.00	0.29	0.38	0.72	19.68	
≥10000	0.14	0.34	0.43	0.74	18.53	
Overall				20.37		

Table 2: Cox's regression analysis of covariates on the duration of BF

Explanatory variables	$\beta$	SE	Wald	p-value	e <sup><math>\beta</math></sup>	95% CI for e <sup><math>\beta</math></sup>
Place of residence	0.300	0.062	23.114	0.000	1.350	1.195-1.526
Educational level of mother	0.016	0.008	4.512	0.034	1.016	1.001-1.032
Employment status of mother	0.609	0.093	42.550	0.000	1.838	1.531-2.207
Family income	0.025	0.011	4.658	0.031	1.025	1.002-1.048
Age at marriage of mother	-0.013	0.007	3.785	0.052	0.987	0.975-1.000
Parity	-0.186	0.076	5.958	0.015	0.830	0.715-0.964
Sex of the previous child	0.010	0.060	0.026	0.872	1.010	0.897-1.136

whereas 22% of mother with no schooling and 26% of mothers with primary school level, 28% of mothers with secondary school level and 35% of mothers with higher secondary level terminate BF within the 12 months. A longer duration of BF is found for unemployed mothers (20.55 months) than employed mothers (15.40 months).

About 22% of unemployed mothers terminate BF within 12 months while 41% of employed mothers terminate BF within 12 months. Only 15% of mother having family income of <Rs. 2000 terminate BF within 12 months. But 19% of the study subjects having family income of Rs. 2000-4000, 21% in the income of

Rs. 4000-6000, 25% in Rs. 6000-8000, 29% in Rs. 8000-10000 and 34% in Rs. 10000 and above terminate BF within the same period. The variation in the median duration of BF with respect to family income is highly significant (p<0.001).

**Cox's regression analysis:** Further, the partial effects on of each of the explanatory variables on the duration of BF while controlling other covariates, Cox's Proportional Hazard Model analysis is performed. The results are shown in Table 2. After controlling other covariates, the parity is negatively associated with the risk of termination

of BF ( $\beta = -0.186$ ,  $p < 0.05$ ). By increasing a parity, the adjusted Relative Risk (RR) of termination of BF becomes 0.83 (95% CI: 0.715-0.964) indicating the risk is decreased by around 17%. The place of residence has also significant impact on the risk of termination of BF ( $\beta = 0.300$  and  $p < 0.001$ ). The risk of termination of BF of mothers residing in urban is 1.35 times higher than those residing in rural areas (RR = 1.35). By Wald's test, the educational level of mothers is positively associated with the risk of termination of BF ( $\beta = 0.016$ ,  $p < 0.05$ ) in the sense that when the educational level is increased by a 1 year standard, the risk of termination of BF is increased by 2% (RR = 1.02).

The employment status of mothers also plays a significant role in the variation of duration of BF ( $\beta = 0.609$ ) which is highly significant ( $p < 0.001$ ). Employed mothers are found to be subject to a hazard of termination of BF 1.84 higher than those of unemployed mothers. The family income is positively associated with the risk of conception ( $\beta = 0.025$ ,  $p < 0.05$ ). If the monthly family income is raised by Rs. 100, the hazard of termination of BF is increased by 3% (RR = 1.03).

#### DISCUSSION

These result shows that the median duration of BF is 20.37 months which is lower than the minimum of 24 months recommended by WHO for most children. The duration of about 20 months is also lacking behind the India's national figure of 25 months observed by National Family Health Survey-3 (NHFS-3). The findings arrived in the present study also highlight that different factors have different effects on duration BF. The duration of BF increases with the increase in parity. It might be due to the fact that mothers of high parity may be older and they produce less milk but they may be more traditional in orientation and usually lower order births occur in quick succession than higher order births and hence the chance of voluntary termination of BF at an early age of child might be higher for the lower birth order babies than for the higher birth order. This view is consistent with the other findings observed by several researchers (Mannan and Islam, 1995; Aaryal, 2005; Islam *et al.*, 2006).

The educational level of mothers show an inverse relationship with the duration of BF, owing to the fact that educated mothers probably start giving food supplements to their children earlier and so a shorter period of lactation. Similar, findings have also been obtained by other researchers in developing countries (Huffman, 1984; Koosha *et al.*, 2008; Aaryal, 2005). Place of residence has a significant impact on the duration of BF. The risk of

termination of BF of mothers residing in urban is higher than those in rural. The shorter duration of BF in urban areas is thought to be caused by higher educational level and higher income than that of rural. This view is incorporated with the past findings (Koosha *et al.*, 2008; Islam *et al.*, 2006). Besides, employed mothers have shorter duration of BF than unemployed one. The reason may be that employed mothers do not get enough time to breast-feed their children as they work outside and thus tend to lactate for shorter period and probably also provide food supplements to the children much earlier. It is supported by the earlier findings too (Huffman, 1984; Giashuddin and Kabir, 2004). Mothers of high income group have shorter duration of BF. It may perhaps be the fact that economically affluent families could afford to buy powder milk, other baby formulae, wet nurses, etc. Similar views have also been reported by other researchers based on the data from developing countries (Giashuddin and Kabir, 2004).

#### CONCLUSION

The present duration of breast-feeding that is about 20 months is lower than the WHO recommended figure of 24 months and even lacking behind the all India's figure of 25 months. This low status of breast-feeding may cause hazardous problem in reproductive and child health. Thus, the breast-feeding promotion programme in Manipur should address among the educated and employed women having high income and the women residing in urban areas through vigorous implementation of inter-sectoral operational strategies since these mothers tend to breast-feed their children for a relatively shorter period of time.

#### REFERENCES

- Aaryal, T.R., 2004. Some techniques to estimate child mortality in Nepal. *J. Inst. Sci. Technol.*, 13: 48-61.
- Aaryal, T.R., 2005. Differentials of breast-feeding among rural women of Western Nepal: A survival analysis. *J. Nepal Health Res. Council*, 3: 58-64.
- Agrawal, D.K., K.N. Agarwal, I.C. Tiwari and K.N.S. Yadava, 2003. Breast-feeding practices in urban slum and rural areas of varanasi. *J. Trop. Pediatrics*, 28: 220-224.
- Berkson, J. and R.R. Gage, 1950. Calculation of survival rates for Cancer. *Proc. Staff Meetings Mayo Clin.*, 25: 270-286.
- Chhabra, P., V.L. Grover, O.P. Aggarwal and K.K. Dubey, 1998. Breast-feeding patterns in an urban resettlement colony of Delhi. *Indian J. Pediatr.*, 65: 867-872.

- Cox, D.R., 1972. Regression models and life tables. *J. Royal Stat. Soc. Ser. B (Methodological)*, 34: 187-220.
- Cutler, S.J. and F. Ederer, 1958. Maximum utilization of the life table method in analyzing survival. *J. Chronic Dis.*, 8: 699-712.
- Giashuddin, M.S. and M. Kabir, 2004. Duration of breast-feeding in Bangladesh. *Indian J. Med. Res.*, 119: 267-272.
- Huffman, S.L., 1984. Determinants of breast-feeding in developing countries: Overview and policy implications. *Studies Family Planning*, 15: 144-154.
- Islam, S., K.N.S. Yadava and M.A. Alam, 2006. Differentials and determinants of the duration of breast-feeding in Bangladesh: A multilevel analysis. *Proc. Pakistan Acad. Sci.*, 43: 1-14.
- Koosha, A., R. Hashemifesharaki and N. Mousavinasab, 2008. Breast-feeding patterns and factors determining exclusive breast-feeding. *Singapore Med. J.*, 49: 1002-1006.
- Lawrence, R.A., 2002. Peer support: Making a difference in breast-feeding duration. *Can. Med. Assoc. J.*, 166: 42-43.
- Lechtig, J.A.D. and E. Jelliffe, 1990. The first workshop on national breast-feeding programmes in Latin America: A personal view. *J. Trop. Pediatr.*, 29: 278-282.
- Mannan, H.R. and M.N. Islam, 1995. Breast-feeding in Bangladesh: Patterns and impact on fertility. *Asia Pacific Popul. J.*, 10: 23-38.
- Peto, R. and J. Peto, 1972. Asymptotically efficient rank invariant procedures. *J. Royal Statist. Soc. Series A*, 135: 185-207.
- Thapa, S. and N.E. Williamson, 1990. Breast-feeding in Asia: An overview. *Asia-Pacific Popul. J.*, 5: 7-25.
- Yngve, A. and M. Sjostrom, 2001. Breast-feeding in countries of the European Union and EFTA: Current and proposed recommendations, rationale, prevalence, duration and trends. *Public Health Nutr.*, 4: 631-645.