Assessing Survival Time of Women with Cervical Cancer Using Various Parametric Frailty Models: A Case Study at Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia

Selamawit Endale Gurmu¹

Received: 29 November 2017 / Revised: 24 January 2018 / Accepted: 31 January 2018 /

Published online: 17 February 2018

© The Author(s) 2018. This article is an open access publication

Abstract Cervical cancer is one of the leading causes of death in the world and represents a tremendous burden on patients, families and societies. It is estimated that over one million women worldwide currently have cervical cancer; most of them have not been diagnosed or have no access to treatment that could cure them or prolong their lives. The goal of this study is to investigate potential risk factors affecting survival time of women with cervical cancer at Tikur Anbessa specialized hospital. Data were taken from patients' medical record card that enrolled during September 2011–September 2015. Kaplan–Meier estimation method, Cox proportional hazard model and parametric shared frailty model were used to analysis survival time of cervical cancer patients. Study subjects (cervical cancer patients) in this study came from clustered community and hence clustered survival data correlated at the regional level. Parametric frailty models will be explored assuming that women with in the same cluster (region for this study) shares similar risk factors. We used Exponential, Weibull, Log logistics and Log normal distributions and based on AIC criteria, all models were compared for their performance. The lognormal inverse Gaussian model has the minimum AIC value among the models compared. The results implied that not giving birth up to the study ends and married after twenty years were significantly prolong the survival time of patients while age class 51–60, 61–70, > 70, smoking cigarettes, patients with stage III and IV disease, family history of cervical cancer, history of abortion and living with HIV AIDS were significantly shorten survival time of patients. The findings of this study suggested that age, smoking cigarettes, stage, family history, abortion history, living with HIV AIDS, age at first marriage and age at first birth were major factors to survival time of patients. Heterogeneity between

---

✉ Selamawit Endale Gurmu
sendale2017@gmail.com

¹ Assosa University, Assosa, Ethiopia

the regions in the survival time of cervical cancer patients, indicating that one needs to account for this clustering variable using frailty models. The fit statistics showed that lognormal inverse-Gaussian frailty model described the survival time of cervical cancer patients dataset better than other distributions used in this study.

Keywords Survival Data Analysis · Proportional hazard · Parametric frailty · Cervical cancer · Survival time

Abbreviations

AFT	Accelerated failure time
AIC	Akaike information criterion
ANC	Antenatal care
PH	Proportional hazard
WHO	World Health Organization

1 Background of the Study

Cancer is the name given to a collection of related diseases. Cancer can start almost anywhere in the human body, which is made up of trillions of cells. It is one of the leading causes of death in the world and represents a tremendous burden on patients, families and societies [1].

There were 12.7 million new cancer cases in worldwide, of which 5.6 million occurred in developed countries and 7.1 million in developing countries. The corresponding estimates for total cancer deaths were 7.6 million 2.8 million in developed countries and 4.8 million in developing countries. There were an estimated 4.9 million new cases and 0.266 million global deaths from cervical cancer accounting for 7.5% of all female cancer deaths. Cervical cancer is one of the leading causes of death in the world and represents a tremendous burden on patients, families and societies. It is estimated that over one million women worldwide currently have cervical cancer; most of them have not been diagnosed or have no access to treatment that could cure them or prolong their lives [2,3].

Survival data is a term used for describing data that measure a time to the occurrence of a given event of interest. In this study the event of interest is survival time of cervical cancer patients from the day of diagnosis. One of the major aims of this analysis was to assess the survival of women with cervical cancer using various parametric frailty models. Kaplan and Meier obtained one important development in non-parametric methods [4]. The non-parametric methods work well for homogeneous samples; they do not determine whether certain variables are related to the survival times [5]. The Cox PH model has the restriction that proportional hazards assumption holds with time-fixed covariates; and it may not be appropriate in many situations and other modifications such as stratified Cox model or Cox model with time-dependent variables are required [6]. Study subjects (cervical cancer patients) in this study came from clustered community and hence clustered cervical cancer patients survival data may be correlated at the regional level. In this study shared frailty models were explored

assuming that cervical cancer patients with in the same cluster (region) shares similar risk factors. Frailty model is common to all individuals in a cluster and responsible for creating dependence between event times [7]. Although the Cox regression model is the most favorable employed technique in survival analysis, parametric models do have a number of benefits [8].

2 Methodology

2.1 Study Population and Setting

The population of this study was all cervical cancer patients who had been registered at Tikur Anbessa specialized hospital, Oncology center, from 2011–2015. All data were carefully reviewed from the registration log book and patients' registration card; if any inadequate information was countered it was checked from the file and excluded from analysis if proven to be inadequate. A total of 907 women with cervical cancer from whole regions of Ethiopia were considered.

The response variable for this study is survival time of the women with cervical cancer which measured in months. Death of cervical cancer patients was event of interest for this research. The patients were censored for those loss follow up and not died up to study ends. The following variables were considered for their influence on survival time of cervical cancer patients; age, smoking status, recurrence, stage of the disease, treatment taken, cycles of chemotherapy, number of sexual partner, number of children, aim of radiography, history of abortion, HIV status, family history, age at marriage and age at first birth.

2.2 Hazard Function

The hazard function $h(t)$ the instantaneous potential for failing at time t , given that the individual has survived up to time t [9].

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

where $f(t)$ and $S(t)$ are the probability density function and survival probability function at time t .

2.3 Kaplan–Meier Product Method

Kaplan–Meier estimator incorporates information from all of the observations available, both censored and uncensored, by considering any point in time as a series of steps defined by the observed survival and censored times [4]. Therefore, the Kaplan–Meier estimate of the survival function at time t is given by:

$$\hat{S}(t) = \prod_{i=1}^k \left(\frac{R(t_{(j)}) - d_{(j)}}{R(t_{(j)})} \right), \quad \text{for } t_{(j)} < t < t_{(j+1)}, \quad k = 1, 2, \dots, r \quad (2)$$

2.4 Log-rank Test

Assessing whether or not there is a real difference between groups can only be done, with any degree of confidence, by utilizing statistical tests [10].

$$Q = \frac{[\sum_i^m w_i (d_{1i} - \hat{e}_{1i})]^2}{\sum_i^m w_i^2 \hat{V}_{1i}} \sim \chi_{k-1}^2 \quad (3)$$

2.5 Cox PH Regression Model

The non-parametric method does not control for covariates and it requires categorical predictors. A broadly applicable and the most widely used method of survival analysis [11].

$$\widehat{HR} = \frac{h_0(t) \exp(\hat{\beta} \cdot X)}{h_0(t) \exp(\hat{\beta} \cdot X^*)} \quad (4)$$

2.6 Parametric Frailty Models

The concept of frailty provides a suitable way to introduce random effects in the model to account for association and unobserved heterogeneity [12]. Models constructed in terms of group-level frailties are referred to as ‘shared’ frailty models because observations within a subgroup share unmeasured ‘risk factors’ that prompt them to exit earlier than other subgroups [13]. A random effect describes excess risk or frailty for distinct categories, such as individual or families, over and above any measured covariates. It is recognized that individuals in the in the same group (cluster) are more similar than individuals in different cluster because thus, frailty or random effect model try to account for correlations within groups [14]. The inverse Gaussian distribution was introduced as a frailty distribution alternative to the gamma distribution [15].

2.7 The Gamma Frailty Distribution

The gamma distribution has been widely applied as a mixture distribution . From a computational and analytical point of view, it fits very well as a mixture distribution to failure data. The most common reason for using the gamma distribution is its mathematical convenience. This is due to the simplicity of the derivative of the Laplace transform, meaning that traditional maximum likelihood procedures can be used for parameter estimation [15]. The density of a gamma-distributed random variable with parameter θ is given by

$$f_z(z) = \frac{Z_i \exp(-Z_i/\theta)}{\theta^{\frac{1}{\theta}} \Gamma(\frac{1}{\theta})}, \quad \theta > 0 \quad (5)$$

2.8 The Inverse Gaussian Frailty Distribution

The inverse Gaussian (inverse normal) distribution was introduced as a frailty distribution alternative to the gamma distribution by Hougaard [16]. The probability density function of an inverse Gaussian distributed random variable with parameter $\theta > 0$ is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} z^{-3/2} \exp\left(-\left(\frac{1}{2\theta_z}\right)(z-1)^2\right) \quad (6)$$

In order to investigate effect of the candidate covariates on survival time of women with cervical cancer, we first did univariable parametric frailty models analysis by fitting a separate model for each candidate covariates. Covariates that were found to be significant in the univariable analysis were included in the multivariable analysis.

The multivariable parametric frailty models in the study was done by assuming the Exponential, Weibull, Log logistics and Log normal distributions for the base line hazard function. It was performed by using the twelve most significant covariates namely age, smoking history, cancer recurrence, stage, cycles of chemotherapy, sexual partner, family planning, family history, abortion history, HIV status, age at first marriage and age at first birth. However we excluded, Aim of radiotherapy, Treatment taken, Number of children, which were not significant in the univariable analysis.

3 Results

Of all 907 cervical cancer patients 349 (38.48%) were died and the overall median survival time was 26 months while the minimum observed event time was 6 months and the maximum was 60 months, respectively (Table 3).

The AIC value of Log-normal Inverse-Gaussian model 1159.95 was the minimum among all the other AIC values of the models indicating that it was the most efficient model to describe cervical cancer patients' dataset (Table 1).

Analysis based on Log-normal Inverse-Gaussian model showed that age, smoking history, stage, family history, abortion history, HIV status, age at first marriage and age at first birth were significant.

The output of the Log-normal Inverse-Gaussian showed patients with age groups 51–60, 61–70 and > 70 , women who smoking cigarettes, patients with stage III and IV, women who had family history of cervical cancer, patients with previous abortion history and living with HIV AIDS were statistically significantly shorten survival time of women with cervical cancer while patients married after arrived 20th birth day and not giving birth up to the study ends prolong the survival time of cervical cancer patients.

The acceleration factor for patients with age group 51–60 and 61–70 are estimated to be 0.56 and 0.49 respectively this indicates patients with age group 51–60 and 61–70 have less survival time than the age group of 20–30 patients. Similarly acceleration

Table 1 AIC values of the parametric frailty models

Model	Frailty distribution	AIC
Baseline Distribution	Gamma	1389.318
	Inverse-Gaussian	1389.304
Weibull	Gamma	1186.041
	Inverse-Gaussian	1185.965
Log- logistic	Gamma	1163.071
	Inverse-Gaussian	1163.027
Log- normal	Gamma	1160.008
	Inverse-Gaussian	1159.95

factor for patients with age group > 70 is estimated to be 0.67 this implied that women with age group 20–30 have longer survival time than women with age group > 70 . The acceleration factors of patients with stage III and IV were estimated to be 0.56 and 0.35. This oblique that the stage I patients have longer survival time than stage III and IV.

The acceleration factor for women that had family history of cervical cancer is 0.63 which indicates that the survival of time of patients that had family history of cervical cancer shorten than patients not have family history. The estimated acceleration factor for patients who had history of abortion is estimated to be 0.62 with very small P -values ($P = 0.00$) which indicate patients who had no history of abortion prolong the survival time than patients who had history of abortion. Patients not living with HIV AIDS prolong the survival time than patients living with HIV AIDS with acceleration factor 0.69. The acceleration factor for women married at age of 21–25 is estimated to be 5.72 and P -value is small ($P = 0.01$) this indicates women who married at late age have prolonged survival time than women that married at age of ≤ 15 years. the acceleration factor for women give birth at age of 16–20 is estimated to be 1.22 and P -value is small ($P = 0.01$) this indicates women who got first baby at age of 16–20 have prolonged survival time than women that got the first baby at age of ≤ 15 years. The acceleration factor for women not giving birth up to the study ends is estimated to be 2.95 with P -value ($P = 0.002$) This indicates women not giving birth up to the study ends have prolonged survival time than women that got the first baby at age of ≤ 15 years (Table 2).

To check the adequacy of baseline hazard, the exponential is plotted by the $-\log(S(t))$ with the time of the study; the Weibull is plotted by $\log(-\log(S(t)))$ with the logarithm of time of the study; the log-logistic is plotted by $\log(\frac{\hat{S}(t)}{1-\hat{S}(t)})$ with the logarithm of time of the study and the log-normal baseline by $\phi^{-1}[1 - S(t)]$ against $\log(t)$ of time of the study. If the plot is linear, the given baseline distribution is appropriate for this dataset (Fig. 1). The plot of log-normal was more linear than other plots.

The Cox-Snell residuals are one way to investigate how well the model fits the data. The plot for fitted model of residuals for log-normal to our data via maximum likelihood estimation with cumulative hazard functions (Fig. 2). If the model fits the

Table 2 Summary result for the final Log-normal inverse-Gaussian frailty models

Covariates	Category	$\hat{\beta}$	SE	Sig.	ϕ	95% CI for ϕ
Age	20–30	Ref				
	31–40	0.26	0.18	0.165	1.68	[0.79, 3.60]
	41–50	0.12	0.18	0.503	1.45	[0.69, 2.96]
	51–60	−0.36	0.17	0.037*	0.56	[0.28, 0.82]
	61–70	−0.45	0.22	0.011*	0.49	[0.24, 0.99]
	> 70	−0.32	0.45	0.0147*	0.67	[0.29, 1.63]
Smoking status	Nonsmoker	Ref				
	Smoker	−0.32	0.77	0.00***	0.65	[0.49, 0.86]
Recurrence	No	Ref				
	Yes	−0.11	0.79	0.147	0.82	[0.62, 1.08]
Stage	I	Ref				
	II	0.012	0.13	0.926	1.04	[0.62, 1.74]
	III	−0.52	0.13	0.00***	0.56	[0.34, 0.93]
	IV	−0.89	0.14	0.00***	0.35	[0.20, 0.62]
Cycles of chemotherapy	No chemo	Ref				
	First cycle	−0.06	0.13	0.445	0.89	[0.68, 1.17]
	Second cycle	0.04	0.15	0.767	1.02	[0.62, 1.67]
	Third cycle	0.21	0.14	0.253	1.21	[0.70, 2.10]
	Fourth cycle	−0.16	0.12	0.721	0.85	[0.43, 2.34]
Sexual partner	None	Ref				
	One	−0.26	0.27	0.924	1.10	[0.55, 1.55]
	Few	−0.22	0.28	0.416	0.75	[0.34, 3.53]
	Multiple	−0.38	0.27	0.163	0.57	[0.23, 2.41]
Family history	No	Ref				
	Yes	−0.28	0.08	0.00***	0.63	[0.48, 0.83]
Use family planning	No	Ref				
	Yes	−0.07	0.06	0.262	0.89	[0.70, 1.12]
Abortion	No	Ref				
	Yes	−0.27	0.07	0.00***	0.62	[0.48, 0.80]
HIV status	No	Ref				
	Yes	−0.32	0.06	0.00***	0.69	[0.54, 0.87]
Age at marriage	≤ 15	Ref				
	16–20	0.18	0.08	0.811	1.01	[0.79, 1.36]
	21–25	0.69	0.27	0.01**	5.72	[1.35, 24.19]
Age at first birth	≤ 15	Ref				
	16–20	0.11	0.74	0.126	1.25	[0.94, 1.36]
	21–25	−0.32	0.12	0.789	0.95	[0.63, 1.49]
	> 25	−0.27	0.24	0.267	0.55	[0.23, 1.29]
	Not give birth	0.73	0.23	0.002**	2.95	[1.18, 7.37]

ϕ , acceleration factor; *significant at 5% level; 95% CI for ϕ , 95% confidence interval for acceleration factor; SE, standard error for estimates; Ref., reference

Fig. 1 Graphical evolution of Exponential, Weibull, Log logistic and Log normal assumption

Fig. 2 Cox- Snell residuals plots of log-normal distribution for survival time of patients'

data, the plot of cumulative hazard function of residuals against Cox-Snell residuals should be approximately a straight line with slope 1. The plot makes straight lines through the origin for log-normal baseline distribution suggesting that it is appropriate for survival time of cervical cancer patients' data set.

A quantile–quantile plot is made to check if the accelerated failure time provided an adequate fit to the data by using two different groups of population. We checked the adequacy of the model by comparing the significantly different groups of patients by family history, smoking cigarettes, HIV status and abortion history. The figures appear to be approximately linear for all covariates family history, smoking status, HIV status and abortion history of patients (Fig. 3). Therefore the accelerated failure time appears to be the best to describe survival time of cervical cancer data set.

Fig. 3 Quantile–quantile plot to check the adequacy of the accelerated failure time mode

4 Discussion

The main purpose this study was assessing the determinant of survival time of cervical cancer dataset which was obtained from Tikur Anbessa Specialized Hospital. The comparison of distribution of the model was performed using the AIC criteria, where the mode with minimum AIC is accepted to be the best [10]. Accordingly, Log Log-normal Inverse-Gaussian model to describe the cervical cancer dataset. This finding is in accordance with the studies [17–22] with regards the Log-normal Inverse-Gaussian model.

The numbers of cervical cancer cases were not related to population size of the regions in Ethiopia at the study period. The number of cases in other region of the country depends on their distance from Addis Ababa; the farther the region, the fewer number of cases. Because patients may not got the opportunity to go and be diagnosed. Those patients who are relatively near to the hospital such as patients in Addis Ababa, Oromiya and Amhara are able to go the hospital and diagnosed. But, those patients who are far from the hospital are not able to go the hospital easily due to many reasons

like transportation cost, accommodation, job and family related cases. Similar finding was obtained from [23,24].

Patients treated at old age were significantly associated to shorten to time death of cervical cancer. This implied survival time for cervical cancer patients was highest in the youngest women and decreases with increasing age. The current study is consistent with other findings of [25–28] reported that increases in age have been related to poorer survival time of cervical cancer.

The results of this study suggested that smoking was significant predictive factor for survival time of the patients. Non-smokers had longer survival time than smokers. The current study is consistent with other findings of [29–31].

Accordingly the stage of the disease also had significant association to survival timing of patients. Stage I and Stage II patients had long survival time than stage III and IV patients. This is due to early to an early stage cervical cancer can simply be cured by available cancer treatments. This result is consistent with [29,30].

Family history of cervical cancer is also one factor that significantly predicts the survival of the patients. In this study, family history is significantly associated with survival time cervical cancer patients. The current study is consistent with other findings [32] reported that women who have a mother or sister diagnosed with cervical cancer have a greater risk of developing this cancer than women without any cervical cancer family history

The history of abortion also had significant association to survival time of cervical cancer patients'. Those women who had no history of abortion had long survival time than those who had abortion history. This result is consistent with [33] reported that previous abortion plays a role in cervical cancer. And also a study by Ibrahim et al. [34] in Danish found a significantly positive trend between those women had abortion and the death of cervical cancer.

The results of this study indicated that status of HIV AIDS was significant predictive factor for survival time of cervical cancer patients in TASH. Women who lived with HIV AIDS had smaller survival time than women who are free from HIV. This happens since cervical cancer makes CD4 counts lower and they are more likely to die. Somers [35] and Abel et al. [36] found a significantly positive trend between those women having had HIV and the death of cervical cancer.

The results of the study revealed that having first baby at early age (≤ 15) years has a significant association with survival time of cervical cancer patients. Women who has had first child at early age had smaller survival time. Similarly Similar finding was obtained from [37] reported having your first baby before the age of 16 gives a higher risk, compared to women who had their first baby after the age of 25. This result is also in accordance with the studies [38,39].

5 Conclusions

This study used survival time of cervical cancer patients' dataset of those patients who started their cancer treatment from 2003–2007 years with the aim of assessing the determinant of time-to-death of women with cervical cancer in Tikur Anbessa specialized hospital. Out of the total 907 women who started cancer medicine (treat-

ments), about 38.48% died at the end of the study. The estimated mean survival time of women was 26 months.

To model the determinants of survival time of cervical cancer patients, Cox Ph model was used. Then parametric frailty models were fitted because the assumption of Cox proportional model was violated. Different parametric frailty models by using different baseline distributions were applied. Among this using AIC, Log-normal Inverse-Gaussian model is better fitted survival time of cervical cancer patients' dataset among various parametric frailty models.

The result of Log-normal Inverse-Gaussian model showed that age, smoking status, stage, family history, abortion history, HIV status, age at first marriage and age at first birth were found significant predictors for survival time of patients in Tikur Anbessa specialized hospital. Of which married after 20 years and not giving birth up to study the end of time prolong the timing death of cervical cancer patients. Similarly age classes 51–60, 61–70 and > 70, smoking cigarettes, stage III and IV, family history of cervical cancer, abortion history, HIV status and having first baby at early age (≤ 15) were statistically significantly shorten timing of death of women with cervical cancer.

Acknowledgements The authors gratefully acknowledge Tikur Anbessa Specialized Hospital and College of Nursing for allowing using the data.

Author contributions SE conceived the idea, contributed in the design analyses completed all statistical analyses and interpretation, SE the corresponding author drafted the manuscript and read and approved the manuscript.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no competing interests

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

See Table 3.

Table 3 Descriptive summary of cervical cancer data (2011–2015)

Variable	Status of patient			Total (%)
	Category	Number of event (%)	Number of censored (%)	
Age	20–30	9 (2.58)	20 (3.58)	29 (3.20)
	31–40	30 (8.60)	123 (22.04)	153 (16.87)
	41–50	53 (15.19)	204 (36.56)	257 (28.34)
	51–60	155 (44.41)	135 (24.19)	290 (31.97)
	61–70	90 (25.79)	59 (10.59)	149 (16.43)
	> 70	12 (3.44)	17 (3.44)	29 (3.20)
Smoking habit	Non-smoker	268 (76.79)	516 (92.47)	784 (86.44)
	Smoker	81 (23.21)	42 (7.53)	123 (13.56)
Region	Addis Ababa	146 (41.83)	155 (27.78)	301 (33.19)
	Oromiya	97 (27.79)	200 (35.84)	297 (32.75)
	Amhara	65 (18.62)	128 (22.94)	193 (21.28)
	Tigray	9 (2.58)	22 (3.94)	31 (3.42)
	SNNPR	20 (5.73)	39 (6.99)	59 (6.5)
	Others	12 (2.87)	14 (2.51)	26 (2.87)
Recurrence	No	280 (80.23)	497 (89.07)	777 (85.67)
	Yes	69 (19.77)	61 (19.77)	130 (14.33)
Stage	I	19 (5.44)	28 (5.02)	47 (5.18)
	II	90 (25.79)	187 (33.51)	277 (30.54)
	III	189 (54.15)	293 (52.51)	482 (53.14)
	IV	51 (14.61)	50 (14.61)	101 (11.14)
Sexual partner	None	3 (0.86)	10 (1.79)	13 (1.43)
	One	91 (26.07)	91 (26.07)	390 (43.00)
	Few	81 (23.21)	81 (23.21)	217 (23.93)
	Multiple	174 (49.86)	113 (20.25)	287 (31.64)
Aim of radiography	No RT	4 (0.36)	2 (0.36)	6 (0.66)
	Palliative	240 (68.77)	354 (63.44)	594 (65.49)
	Radical	105 (30.09)	202 (36.20)	307 (33.85)
Family history	No	275 (78.80)	499 (89.43)	774 (85.34)
	Yes	74 (21.20)	59 (10.57)	133 (14.66)
Treatment taken	Surgery	4 (1.15)	1 (0.18)	5 (0.55)
	Chemotherapy	2 (0.57)	3 (0.54)	5 (0.55)
	Radiotherapy	259 (74.21)	438 (78.49)	697 (76.85)
	Combination of two or three	84 (24.07)	116 (20.79)	200 (22.05)

Table 3 continued

Variable	Status of patient			Total (%)
	Category	Number of event (%)	Number of censored (%)	
Cycles of chemotherapy	No chemo	228 (65.33)	406 (72.33)	634 (69.90)
	First cycle	71 (20.34)	78 (13.98)	149 (16.43)
	Second cycle	18 (5.16)	32 (5.73)	50 (5.16)
	Third cycle	15 (4.30)	23 (4.12)	38 (4.19)
	Fourth cycle	17 (4.87)	19 (3.41)	36 (3.97)
Number of children	No child	10 (2.87)	23 (4.12)	33 (3.64)
	1–4	24 (6.88)	91 (16.31)	115 (12.68)
	5–8	72 (20.63)	228 (40.86)	300 (33.80)
	9–12	134 (38.40)	130 (23.30)	264 (29.11)
	Above 12	109 (31.25)	86 (15.41)	195 (21.50)
Family planning	No	235 (67.34)	409 (73.30)	644 (71.00)
	Yes	114 (32.66)	149 (26.70)	263 (29.00)
Abortion	No	256 (73.35)	466 (83.51)	722 (79.60)
	Yes	93 (26.65)	92 (16.49)	185 (20.40)
HIV status	No	231 (66.19)	422 (75.63)	653 (72.00)
	Yes	118 (33.81)	136 (24.37)	254 (28.00)
Age at marriage	≥ 15	236 (67.62)	315 (56.45)	551 (60.75)
	16–20	111 (37.63)	210 (37.63)	321 (35.39)
	21–25	2 (.57)	33 (5.91)	35 (3.86)
Age at first birth	≤ 15	168 (48.14)	188 (33.69)	356 (39.25)
	16–20	136 (38.97)	271 (48.57)	407 (44.87)
Not give birth	21–25	34 (9.74)	77 (13.80)	111 (12.24)
	> 2	6 (1.72)	6 (1.08)	12 (1.32)
		5 (1.43)	16 (2.87)	21 (2.32)

References

1. WHO (2009) Human papillomavirus and related cancers in Ethiopia. In: Summary report
2. Rebecca L, Kimberly D, Ahmedin Jemal DVM (2015) Global cancer facts & figures, 2nd edn. <https://doi.org/10.3322/caac.21254>
3. World Health Organization (2015) Comprehensive cervical cancer control: a guide to essential practice. 2nd edn
4. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958(53):457–81
5. Klein JP, Moeschberger ML (1997) Survival analysis: techniques for censored and truncated data. Springer, New York
6. Collett D (2003) Modeling survival data in medical research. Chapman & Hall, London
7. Sastry N (1997) A nested frailty model for survival data, with an application to the study of child survival in Northeast Brazil. *J Am Stat Assoc* 92:426–43
8. Andersen PK, Borgan O, Gaill RD, Keiding N (1993) Statistical methods based on counting processes. Springer, New York
9. Kleinbaum DG, Klein M (2011) Survival analysis: a self-learning text, 3rd edn. Springer, New York

10. Mantel N, Haenszel W (1959) Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 22:719–748
11. Cox D (1972) Regression models and life tables (with discussions). *J R Stat Soc* 34:187–220
12. Wintrebret C (2007) Statistical modeling of repeated and multivariate survival data. PhD thesis
13. Liu X (2014) Survival models on unobserved heterogeneity and their applications in analyzing large-scale survey data. *J Biom Biostat*. <https://doi.org/10.4172/21556180.1000191>
14. Holt J, Prentice R (1974) Survival analysis in twin studies and matched pair experiments. *Biometrika* 61:17–30
15. Hougaard P (1984) Life table methods for heterogeneous populations: Distributions describing heterogeneity. *Biometrika* 71:75–83
16. Hougaard P (1986) Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73(2):387–396
17. Akaike's H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723
18. Akalu B, Dinberu S, Tefera, Belay B, Yehnew G (2015) Modeling time-to-cure from severe acute malnutrition: application of various parametric frailty models. *Bio Med Cent*. <https://doi.org/10.1186/2049-3258-73-6>
19. Chen LM, Ibrahim JG, Chu H (2014) Sample size determination in shared frailty models for multivariate time-to-event data. *J Biopharm Stat* 24(4):908–923
20. Congdon P (1995) Modeling frailty in area mortality. *Stat Med* 14:1859–1874
21. Lee ET, Wang JW (2013) Statistical methods for survival data analysis. 4th edn. Wiley, New York
22. Pourhoseingholi M, Hajizadeh E, Moghimi D, Safaee A, Abadi A, Zali MR (2007) Comparing cox regression and parametric models for survival of patients with gastric carcinoma. *Asian Pac Organ Cancer Prev* 8(3):412–416
23. Chapman JW, Trudeau ME, Pritchard KI, Sawka CA, Mobbs BG, Hanna WM, Kahn H, McCready DR, Lickley LA (1992) A comparison of all-subset Cox and accelerated failure time models with Cox step-wise regression for node-positive breast cancer. *Breast Cancer Res Treat* 22(3):263–272
24. Royston P (2001) The lognormal distribution as a model for survival time in cancer, with an emphasis on prognostic factors. *Stat Neerl* 55:89–104
25. Abate S (2015) Trends of cervical cancer in Ethiopia, Oncology Pharmacy Case Team, Tikur Anbessa specialized hospital, College of Health Sciences. Addis Ababa University, Ethiopia
26. Sule ST, Shehu MS (2007) Cervical cancer management in Zaria, Nigeria. *Afr J Health Sci* 14:149–153
27. Miura K, Mishima H, Yasunami M, Kaneuchi M, Kitajima M, Abe S, Higashijima A, Fuchi N, Miura S, Yoshiura K, Masuzaki H (2016) A significant association between gene expression variation and invasive cervical cancer originally identified by a genome-wide association study in Han Chinese is replicated in a Japanese population
28. Wright JD (2014) cervical intraepithelial neoplastic: terminology, incidence, pathogenesis, and prevention. <http://www.uptodate.com/contents/cervical-intraepithelial-neoplasia-terminology-incidence-pathogenesis-and-prevention>. Accessed 3 July 2014
29. Castillo OO, Fernandez R, Arauz R, Zarak E, Diaz J, Alcedo C, Lopez RI, Cebamanos F (2009) Prognostic factors for death in elderly Women with cervical carcinoma. *J Clin Oncol* 26:16564
30. Kumari K, Gupta S, Parida SK, Chattopadhyay D, Prasad M (2013) Genome-wide development and use of microsatellite markers for large-scale genotyping applications in foxtail millet [*Setaria italica* (L.)]. *DNA Res* 20:207
31. Tesfahun H, Bizunesh Y, Hiwot A, Engidawork M, Goye O, Mihret S (2017) Prevalence of cervical cancer and associated risk factors among women attending cervical cancer screening and diagnosis center at Yirgalem General Hospital, Southern Ethiopia. *J Cancer Sci Ther* 9(11):730–735. <https://doi.org/10.4172/1948-5956.1000500>
32. Waggoner SE et al (2010) Smoking behavior in women with locally advanced cervical cancer: a gynecologic oncology group study. *Am J Obstet Gynecol* 202(3):283.e1–283.e7
33. Bellinger JD (2013) The role of family history of cancer on cervical cancer screening behavior in a population-based survey of women in the Southeastern United States
34. Ibrahim A et al (2011) Predictors of cervical cancer being at an advanced stage at diagnosis in Sudan. *Int J Women Health* 3:385–389
35. Somers R (2001) Risk of admission to psychiatric institutions among Danish women who experienced induced abortion: an analysis on national record linkage. Dissertation Abstracts. International, Public Health 2621-B, Order No. 7926066

36. Gedefaw A, Astatkie A, Tessema GA (2013) The prevalence of precancerous cervical cancer lesion among HIV-infected women in Southern Ethiopia: a cross-sectional study. *PloS one* 8(12):e84519 <https://doi.org/10.1371/journal.pone.0084519>
37. Massad LS (2008) Squamous cervical lesions in women with human immunodeficiency virus: long-term follow-up. *Obstet Gynecol* 111(6):1388–1393
38. Brinton LA, Hamman RF, Huggins GR et al (2002) Sexual and reproductive risk factors for invasive squamous cell cervical cancer. *J Natl Cancer Inst* 79(1):23–30
39. Chelimo C, Woulde TA, Cameron LD, Elwood JM (2013) Risk factors for and prevention of human papillomaviruses (HPV), genital warts and cervical cancer. *J Infect* 66(3):207–217