Asian Journal of Probability and Statistics

Volume 26, Issue 11, Page 124-143, 2024; Article no.AJPAS.126704
ISSN: 2582-0230

Modeling the Risk Factors of Miscarriage Using Censored Quantile Regression in Kenyan Cohort

John Olives Chibayi [a*]

[a] *Department of Mathematics and Statistics, Kaimosi Friends University, Kenya.*

Author's contribution

The sole author designed, analyzed, interpreted and prepared the manuscript.

Article Information

DOI: https://doi.org/10.9734/ajpas/2024/v26i11676

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/126704

Original Research Article

Received: 10/09/2024
Accepted: 13/11/2024
Published: 19/11/2024

Abstract

Background: Numerous researches have been undertaken to delineate the significant risk factors linked to miscarriage. The etiology of miscarriage has presented challenges in its elucidation, necessitating a comprehensive understanding of the elements that are linked to this phenomenon. The examination of these aspects has been conducted utilizing diverse methodological techniques and analytical procedures. The Cox's proportional hazards model and the accelerated failure time models are commonly utilized in the field of survival analysis. While these models provide a certain level of flexibility, they do not allow for the differentiation of predictor indicators between high-risk and low-risk subpopulations. In contrast, quantile regression allows for the assessment of separate effects of covariates at different quantiles of the conditional distribution of miscarriage.
Objective: This study assessed the risk factors for miscarriage using the censored quantile regression model and compared the results to that of Cox proportional hazard model.

**Corresponding author: Email: john.chibayi@yahoo.com;*

Cite as: *Chibayi, John Olives. 2024. "Modeling the Risk Factors of Miscarriage Using Censored Quantile Regression in Kenyan Cohort". Asian Journal of Probability and Statistics 26 (11):124-43. https://doi.org/10.9734/ajpas/2024/v26i11676.*

Methodology: A secondary data from 6077 records of pregnant women with recognized pregnancy and enrolled in the pre-natal care in Kakamega County General Teaching and Referral Hospital (KCGTRH) in Kakamega county, western Kenya during the period from 1 January, 2019 up to 31 October, 2020 were recruited into the study. The study used chi-squared tests for descriptive analysis. Cox proportional hazards and censored quantile regression were used for modeling. Analysis was done using the R package and 5% was the level of significance.

Results: Censored quantile regressions were conducted to estimate the parameters for the quantiles, $\tau = 0.05, 0.1, 0.15, 0.2$. The results showed that ethnicity had a positive effect and was statistically significant at $\tau = 0.1$ ($p = 0.0023$) and number of previous stillbirths had a negative effect and was significant at $\tau = 0.15$ ($p = 0.032$) and at $\tau = 0.2$ ($p = 0.036$). Graphical comparison of censored quantile model effects to that of cox PH model revealed that the variables; prior miscarriages, prior stillbirths, place of residence, and number of ANC visits have a statistically significant adverse effect on the survival time of miscarriages in the initial phases of pregnancy, contrary to cox PH model.

Conclusion: No variable had a statistically significant influence across all quantiles. The results depicts that at a certain moment in the monitoring procedure, they will have a significant impact on the length of survival for miscarriages. A quantile regression model have the capacity to incorporate cross-over effects, whereby the influence of a covariate can manifest as either positive or negative. In contrast, Cox estimates do not allow for the potential scenario when a covariate initially increases the hazard for a specific time period and subsequently decreases it.

Keywords: Quantile regression; miscarriage; modeling; statistically.

1 Introduction

Miscarriage is a common event in women's lives. It is considered irreversible, and therefore prevention is probably the only method of intervention. According to Kline and Stein (1984), the World Health Organization (WHO) released a suggestion in 1957 to develop a precise definition of abortion as the unintentional termination of pregnancy before the conclusion of 28 weeks of gestation, particularly in instances unrelated to the human condition. In 1977, the WHO made adjustments to its definition, wherein it broadened the scope to encompass the cessation of a non-viable pregnancy weighing below 500 g prior to attaining 22 weeks of gestation (WHO, 1977).

The incidence of miscarriages has demonstrated notable fluctuations among many groups and temporal epochs. The current prevalence of miscarriage is commonly claimed to be around 10% to 15% of all pregnancies that have been clinically identified (Simpson and Carson, 1993; Simpson and Mills, 1986; Zinaman et al., 1996; Garcia-Enguidanos et al., 2002). The empirical estimation of the prevalence of the phenomenon being studied has yielded a range of reported percentages, with the lower end estimates ranging from 2% to 3%, and the top end estimates reaching as high as 30% (Cai and Feng, 2005).

The cause of miscarriage remains uncertain. Nevertheless, some studies have identified several factors that contribute to and elevate the likelihood of spontaneous abortion. Various risk factors are involved in reproductive health, including demographic parameters such as a woman's years of age, gravidity, duration of gestation gap, and pregnancy history (including the number of prior live births and fetal losses) (Andersen et al., 2002; Regan et al., 1989; Coste et al., 1991; Reagan, 1991; Parazzin et al., 1997; Ogasawara et al., 2000; Andersen, Wohlfahrt, Christens, Olsen, Melbye, Nybo, 2000; Sufian et al., 2024; De la Rochebrochard and Thonneau, 2002). Genetic factors encompass chromosomal abnormalities and mutations in genes (Garcia-Enguidanos et al., 2002; Simpson and Carson, 1993; Wu Y et al., 2024; Jackson&Watkins, 2021; Neill, 2023). Social and environmental factors encompass various determinants that can influence individuals' health outcomes. These factors include drug usage, caffeine intake, smoking and drinking habits, obesity, place of residence, and race (Frazier, Hogue, Bonney, Yount, Pearce, 2018; Maraka, Ospina, O'Keeffe, Espinosa De Ycaza, Gionfriddo, Erwin, Coddington, Stan, Murad, Montori, 2016; Practice Committee of the American Society for Reproductive Medicine, 2012; Metwally, 2008; Chatenoud, Parrazin, Di cintio, Zanconato, Benzi, Bortolus and La Vecchia, 1998; Cnattingius et al., 2000; Lindbohm et al., 2002; Rasch, 2003; Kicia, Skurzak, Wiktor, Iwanowicz-Palus, & Wiktor, 2015; Mills, Ricklesford, Cooke, Heazell, Whitworth & Lavender, 2014; Pasquali, 2006; Wong, Ray, Gao et al., 2010; Peck et al., 2010). Additional factors that can contribute to adverse pregnancy outcomes encompass uterine anomalies, fibroids, cervical insufficiency, post-operative alterations, chronic diseases, starvation, trauma, infectious diseases, social upheaval, and induced abortion (Hendriks, MacNaughton, MacKenzie, 2019; Wu Y, Yu X, Li M, Zhu J, Yue J, Wang Y, Man YZhou C, Tong R and Wu X., 2024; Carlson and Mourgova, 2003; Ellett, Buxton, and Luesley, 1992; Shapiro and Bross, 1980) as cited in (Wood, 1994).

The Cox's PH model, introduced by Cox (1972), and the AFT models are commonly utilized in the field of survival analysis. While Cox models provide a certain level of flexibility, they do not allow for the differentiation of predictor indicators between high-risk and low-risk subpopulations. The evidence offered exclusively presents a static condition, indicating that a certain variable acts as a risk factor. However, it does not indicate the specific point in the patient's survival timeline when the variable became a significant risk factor. The phenomenon can be ascribed to the sustained preservation of the hazard ratio across the whole duration. The application of quantile regression (QR) models, which were first introduced by Koenker and Basset (1978), and the advancement of censored quantile regression models by Powel (1984, 1986), have become valuable alternative methodologies for analysing the distribution of a response variable in the presence of a specified set of independent variables. The Cox proportional hazard and acceleration failure time models are commonly utilized tools in the examination of time to event data, but they possess certain inherent limitations. An inherent constraint to consider for the cox model is the underlying assumption of a consistent HR during the entire duration of observation. Furthermore, this particular model places emphasis on the modeling of the hazard rate as opposed to directly measuring the duration of survival. Furthermore, this approach demonstrates a constrained ability to effectively handle statistical anomalies and lacks adaptability in analyzing the variables that impact miscarriage, particularly in situations when the data distribution is skewed.

The utilization of the censored quantile regression model is a more flexible and easily understandable alternative. In contrast to traditional methodologies such as the Cox model and the Accelerated Failure Time model, the utilization of quantile regression (QR) allows for the assessment of separate effects of covariates at different quantiles of the conditional distribution of miscarriage duration. More so, it relaxes the inherent assumption of the AFT model that the relationship between the response variable and the covariates are constant across all levels. Quantile regression models are utilized to capture changes in the quantiles of the conditional distribution of survival time with respect to modifications in the explanatory variables. Fitzenberger and Wilke (2005) describe the distinguishing features of the models as their independence from modeling assumptions and their increased flexibility in comparison to the AFT and Cox's PH models. Additionally, it is frequently seen that the distribution of the response variable in survival analysis has substantial skewness. Quantile regression techniques demonstrate robustness in the characterization and examination of survival data distributions that possess skewness.

In their study, Koenker and Geling (2001) present a comprehensive overview of the advantages linked to duration analysis based on QR. Censored quantile regression for fixed censoring was first introduced by Powel in his influential publications in (1984, 1986). Portnoy (2003) proposed a technique for QR in the presence of random censoring, which has similarities to the Kaplan-Meier estimator. In a similar vein, the authors Peng and Huang (2008) put up a method based on QR to address the issue of random censoring, which exhibits certain resemblances to the Nelson-Aalen estimator.

In order to acquire a deeper understanding of the practical implementations of QR, it is advisable to consult the studies conducted by Fitzenberger et al (2002) or Koenker and Hallock (2001).

Quantile regression has a wide application in failure time analysis. For instance, Carey et al. (2004) conducted a study in the domain of clinical research, which unveiled a significant correlation between diminished growth velocity (particularly, falling below the 10th quantile) in individuals with AIDS and a substantially increased likelihood of dying. In the healthcare domain, a quantitative research approach was employed by Austin et al. (2005) to examine the various factors that impact waiting times for critical medical treatment. The study focused on investigating the correlation between patient and system variables and waiting times. The study's results indicated that gender was a notable factor in predicting the degree of delays encountered by patients while seeking treatment, particularly among persons who experienced the lengthiest waiting periods. Further noteworthy applications may also be discerned in the fields of clinical and biomedical research (Flemming et al., 2017; Faradmal et al., 2016), as well as in the sphere of healthcare (Ding et al., 2010; Cournane et al., 2016; Yazdani, A et al., 2021)). However, the application of the censored QR model in medical research has been significantly restricted. To date, there has been a scarcity of research undertaken on the subject of miscarriages in Kenya, with the existing studies failing to employ censored QR models.

2 Materials and Methods

2.1 Study design

A retrospective cohort study was conducted for pregnant women who were enrolled for antenatal care in Kakamega County General Teaching and Referral Hospital (KCGTRH) in Kakamega county, western Kenya, hence secondary data from records of pregnant women was collected. Pregnant women with recognized pregnancy and enrolled in the pre-natal care during the period from 1 January, 2019 up to 31 October, 2020 were recruited into the study.

2.2 Study area

The research was carried out inside Kakamega County, western Kenya and limited to pregnant women enrolled for antenatal care services in Kakamega County General and Teaching hospital (KCGTRH). The county is located on the western part of Kenya. According to Kenya Bureau of Statistics (2019), the county under discussion ranks as the fourth most populated in Kenya, boasting a population above 1.8 million individuals. The KCGTRH was chosen to host the study as it is a level 5 referral hospital receiving patients from the entire Kakamega county, with good laboratory and clinical infrastructure and therefore suitable for recruiting pregnant mothers all over the county. The study population included all pregnant women living in Kakamega county, western Kenya during the period from 1, January 2019 up to 31 October 2020, where cases of spontaneous abortion were identified.

2.3 Sampling procedure

All pregnant women enrolled in antenatal care from January 1, 2019 up to 31 October 2020 in KCGTRH in Kakamega County, western Kenya and met the inclusion criteria were included in the sample. Decision was made to consider all of them as sample to get the accurate and precise findings which will be inferred to county population as the KCGTRH serve the whole county. All pregnancy records with the results of the pregnancy outcome within the stated period was selected and reviewed for this study.

Inclusion criteria

- Subjects under study must be a pregnant woman enrolled in antenatal care in KCGTRH
- Must have a confirmed pregnancy of a gestational age of 6 to 12 weeks or earlier.
- Subjects whose documentation of medical, gynaecological history and demographical data are available
- The result of the pregnancy outcome of the woman is documented.

Exclusion criteria

- Miscarriages from unrecognized pregnancy.
- Pregnant women not under antenatal care. This because of difficulty obtaining full information about them, such as their total number.
- Pregnancy of gestational age beyond 12 weeks
- Threatened miscarriage

2.4 Study variables

The factors under consideration in this investigation were obtained from the existing standard medical records. The analysis utilized the covariates that were recorded during the initial antenatal care visit due to a scarcity of available secondary data.

The factors that are being measured or observed in a study are referred to as outcome variables. The event being examined was a spontaneous termination of pregnancy. The dependent variable will be operationalized as the time interval between the last menstrual period and either the incidence of a miscarriage or the termination of the study period, which aligns with the ending of the pregnancy. Participants who were unable to experience the event being studied before the end of the designated research period were categorized as censored, along with those who were no longer available for further observation. Miscarriage, commonly referred to as spontaneous abortion, denotes the involuntary cessation of a pregnancy before to attaining a gestational age of 28 weeks.

Independent variables included;

Socio-demographic characteristics: ethnicity, maternal age, marital status, length of pregnancy interval, history of previous pregnancy (number of previous live births and foetal losses), parity, gravidity, induced abortion and use of certain contraceptives and number of antenatal care visits.

Socio-economic characteristics: Education level, employment status and place of residence (urban/rural).

Clinical characteristics: signs and symptoms of pregnancy loss (blood discharge), chronic diseases (HIV, TB, diabetic, hypertension), infections (malaria, STDs, UTIs), haemoglobin status (HB) and folic intake.

Lifestyle characteristics: caffeine intake, exercises, stress, smoking, exposure to cigarette smoke and alcohol consumption, obesity and drug use. Pregnant mothers that were eliminated from the sample were those whose folders lacked essential information. The researcher utilized a checklist to gather data from these folders, focusing on various demographic and medical factors. These factors included ethnicity, gravidity, maternal age, parity, marital status, history of miscarriages and stillbirths, educational level, profession, HIV status, frequency of antenatal care visits, presence of malaria infection, urinary tract infection, sexually transmitted diseases, gestational age, place of residence, survival time, and survival status.

Miscarriage was the study endpoints with mothers follow up to go on as up to date of transferring out from health facility, death or until end of follow up period (end of pregnancy).

2.5 Censored Quantile Regression (CQR)

Linear censored quantile regression was introduced by (Powell, 1984). Given the linear latent variable model;

$$T_i = X^T{}_i\beta + \varepsilon_i \tag{2.1}$$

Where ε_i are identically and independently distributed with distribution function F. Powell noted if that censoring values C_i are observed for all $i = 1,2, \dots n$ and we observe $Y_i = \max(C_i, T_i)$. Then the conditional function is:

$$Q_{Y_i}(\tau/x_i) = F^{-1}{}_\tau + X^T{}_i\beta \tag{2.2}$$

which can be estimated by letting $\rho_\tau(\mu) = \mu(\tau - I(\mu < 0))$ by

$$\hat{\beta} = argmin_{b\in R^n} \sum \rho_\tau(Y_i - max\{C_i, X^T{}_i b\}) \tag{2.3}$$

So long as that the design matrix $X = (x_i)$ contain an intercept to absorb the τ dependent contribution $F^{-1}(\tau)$. Right censoring is accommodated by replacing max by min above. That is the Powell estimator is given as;

$$\mathcal{R}_\tau(b) = argmin_{b\in R^n} \sum \rho_\tau(Y_i - min\{C_i, X^T{}_i b\}) \tag{2.4}$$

For this method C_i's are known for all observations, this situation is known as fixed censoring.

Portnoy (2003) proposed a recursive reweighting method that provides a broader application of the Kaplan-Meier estimator. The proposed methodology facilitates the estimation of conditional quantile functions through the utilization of censored survival data. The revelation was made by the researcher that the quantiles of the Kaplan-Meier distribution function can be formally characterized as a solution to an optimization problem that incorporates weighted quantiles. In this particular situation, the weight assigned to observations that have been censored is divided into smaller halves. The aforementioned assertion was expanded to encompass the domain of regression analysis. The quantile regression estimator introduced by Portnoy, which incorporates censoring, can be mathematically represented as

$$Min \sum_{i \notin K(\tau)} \rho_\tau \left(y_i - x^T{}_i\beta\right) + \sum_{i \in K(\tau)} [w_i(\tau)\rho_\tau(y_i - x^T{}_i\beta) + (1 - w_i(\tau))\rho_\tau(y_\infty - x^T{}_i\beta)] \tag{2.5}$$

where $K(\tau)$ denote the index set of the censored observations encountered up to τ,

$$w_i(\tau) = \frac{\tau - \tau_i}{1 - \tau_i} \text{ ‘and } y_\infty = +\infty. \tag{2.6}$$

Peng and Huang (2008) have presented a methodology that is closely associated with censored quantile regression for the analysis of censored survival data. The approach utilized in this methodology is grounded on the Nelson-Aalen estimation of the cumulative hazard function.

3 Results and Discussion

3.1 Result

3.1.1 Distribution of study subjects and description of variables

The study utilized a secondary dataset of 6,077 records of pregnant women who initiated prenatal treatment at Kakamega County Teaching and Referral Hospital (KCGTRH) in Kakamega County. The dataset encompassed a time frame that extended from January 1, 2019, to December 31, 2020.

The primary variable under investigation in this study was the time in weeks until the occurrence of miscarriage. The independent variables, or covariates, included factors such as ethnicity, gravidity (number of pregnancies), maternal age, parity (number of previous live births), marital status, history of previous miscarriages, history of previous stillbirths, educational level, profession, HIV status, frequency of antenatal care visits, presence of malaria infection, presence of urinary tract infection, presence of sexually transmitted diseases, and place of residence. For easier analysis of the data, several covariates, including gravidity, mother’s age, number of previous miscarriages, number of previous stillbirths, ANC visits, and parity, were grouped into categories. However, despite this categorization, these variables were still utilized in the model fitting process. Table 1 displays the covariates together with their respective descriptions, categorizations, and distributions of the study individuals.

Table 1. Description and categorization of variables

	Overall (N=6077)	
Variable	**No. of mothers**	**Percentage (%)**
Ethnicity		
Kalenjin	89	1.5
Kikuyu	104	1.7
Luhya	5231	86.1
Luo	565	9.3
Others	88	1.4
Marital status		
Single	815	13.4
Married	5262	86.6
Educational level		
Primary	878	14.4
Secondary	4715	77.6
College	484	8.0
HIV Status		
Positive	111	1.8
Negative	5966	98.2
Malaria infection		
No	6073	99.9
Yes	4	0.1

Overall (N=6077)		
Variable	**No. of mothers**	**Percentage (%)**
STDs status		
No	6073	99.9
Yes	4	0.1
UTI status		
No	6073	99.9
Yes	4	0.1
Profession		
Unemployed	4407	72.5
Employed	1670	27.5
Place of residence		
Rural	1296	21.3
Urban	4781	78.7
Age of mother		
<20	863	14.2
21-25	2055	33.8
26-30	1556	25.6
31-35	1016	16.7
>35	587	9.7
Number of previous miscarriage		
≤2	6058	99.7
3-4	16	0.3
5-6	3	0.0
Number of previous stillbirths		
≤2	6069	99.9
3-4	7	0.1
>41	1	0.0
Number of ANC visits		
≤2	1009	16.6
3-4	4794	78.9
5-6	246	4.0
>6	28	0.5
Gravidity		
≤2	4003	65.9
3-4	1651	27.2
5-6	373	6.1
>6	50	0.8
Parity		
≤2	5053	83.1
3-4	858	14.1
5-6	147	2.4
>6	19	0.3

3.2 Descriptive analysis

A total of 248 female participants, constituting 4.1% of the sample, reported experiencing miscarriages. Table 2 displays the descriptive statistics for the variables of interest. From Table 2, the distribution of the number of women who experienced miscarriages, varied across different levels of some categorized variables. This was indicated by the computed Pearson chi-square statistic with P-value below 5%.

3.3 Fitting cox proportional model

The analysis commenced by including all covariates in the model (referred to as the complete model). Subsequently, a backwards stepwise regression procedure was employed, with the Akaike Information Criterion

(AIC) serving as the metric for assessing the model's fit. The analysis yielded the multivariable model, which included continuous variables in their original continuous form, as presented in Table 3.

Table 2. Descriptive statistics

Variable	(N=248) Miscarriage n(%)	(N=5829) Not miscarriage n(%)	Chi Square Test X^2	p-value
Ethnicity				
Kalenjin	10(11.2)	79(88.8)	25.7	0.0000
Kikuyu	10(9.6)	94(90.4)		
Luhya	192(3.7	5039(96.3)		
Luo	32(5.7)	533(94.3)		
Others	4(4.5)	84(95.5)		
Marital status				
Single	35(4.3)	780(95.7)	0.110	0.741
Married	213(4.0)	5049(96.0)		
Educational level				
Primary	33(3.8)	845(96.2)	6.075	0.048
Secondary	185(3.9)	4530(96.1)		
College	30(6.2)	454(93.8)		
HIV Status				
Positive	6(5.4)	105(94.6)	0.507	0.477
Negative	242(4.1)	5724(95.9)		
Malaria infection				
No	244(4.0)	5829(96.0)	94.078	0.0000
Yes	4(100.0)	0(0.0)		
STDs status				
No	248(4.1)	5825(95.9)	0.172	0.680
Yes	0(0.0)	4(100.0)		
UTI status				
No	246(4.1)	5827(95.9)	21.561	0.000
Yes	2(50.0)	2(50.0)		
Profession				
Unemployed	160(3.6)	4247(96.4)	8.310	0.004
Employed	88(5.3)	1582(94.7)		
Place of residence				
Rural	81(6.2)	1215(93.8)	19.799	0.000
Urban	167(3.5)	4614(96.5)		
Age of mother				
<20	26(3.0)	837(97.0)	3.506	0.469
21-25	83(4.0)	1972(96.0)		
26-30	71(4.6)	1485(95.4)		
31-35	43(4.2)	973(95.8)		
>35	25(4.3)	562(95.7)		
Number of previous miscarriage				
≤2	246(4.1)	5812(95.9)	3.031	0.220
3-4	2(12.5)	14(87.5)		
5-6	0(0.0)	3(100.0)		
Number of previous stillbirths				
≤2	246(4.1)	5823(95.9)	10.78	0.005
3-4	2(28.6)	5(71.4)		
>4	0(0.0)	1(100.0)		
Number of ANC visits				
≤2	85(8.4)	924(91.6)	59.466	0.000
3-4	157(3.3)	4637(96.7)		

Variable	(N=248) Miscarriage n(%)	(N=5829) Not miscarriage n(%)	Chi Square Test	
			X^2	p-value
5-6	6(2.4)	240(97.6)		
>6	0(0.0)	28(100.0)		
Gravidity				
≤2	148(3.7)	3855(96.3)	7.483	0.058
3-4	81(4.9)	1570(95.1)		
5-6	19(5.1)	354(94.9)		
>6	0(0.0)	50(100.0)		
Parity				
≤2	199(3.9)	4854(96.1)	2.970	0.396
3-4	43(5.0)	815(95.0)		
5-6	6(4.1)	141(95.9)		
>6	0(0.0)	19(100.0)		

Table 3. **Cox proportional hazard model with included covariates in the best selection of covariates**

Covariate	Coef.	Exp(coef.)	s.e(coef.)	Z-statistic	P-value	95% conf. interval Lower	Upper
Ethnicity							
Luhya(Ref.)							
Kalenjin	1.168	3.216	0.324	3.598	0.000	1.702	6.077
Kikuyu	0.905	2.472	0.325	2.788	0.005	1.308	4.672
luo	0.404	1.499	0.192	2.105	0.035	1.028	2.184
Residence (urban)	-0.482	0.617	0.139	-3.468	0.001	0.470	0.811
No. of Previous miscarriages	0.443	1.557	0.099	4.474	0.000	1.283	1.891
No. of Previous stillbirths	0.662	1.939	0.106	6.243	0.000	1.575	2.387
Malaria(+)	3.456	31.682	0.517	6.687	0.000	11.506	87.235
No. of ANC visits	-0.664	0.515	0.096	-6.936	0.000	0.427	0.621

3.4 Censored quantile regression model

A quantile regression model with censoring was employed in this section, to estimate quantiles up to the 20th percentile. Censored quantile regressions were conducted to estimate the parameters for the quantiles 0.05, 0.1, 0.15, and 0.2. The results derived from the censored quantile regression model are displayed in Table 4.

The Fig. 1 depicts the estimated treatment effects (blue curve) observed across all quantiles, with a specific focus on the outcomes presented in Tables 4. The provided Figure additionally displays the 95% (lighter blue shaded region) bootstrap confidence intervals pertaining to the coefficient effects of the CQR model and the null effect (solid horizontal black line).

The cox PH model and censored quantile regression were compared. The two models were computed and subsequently merged to facilitate the graphical evaluation of the impact of the coefficients on the distribution of survival time. In Fig. 2, the coefficient effects obtained from the Cox PH model are combined and evaluated with those obtained from the censored quantile regression model. The coefficient effects are represented by a solid red line for the Cox PH model and a dotted blue line for the censored quantile regression model.

3.5 Discussion

The main aim of this study was to construct censored quantile regression model that captures the risk factors linked to miscarriage. The Table 4 illustrates a positive association between ethnicity and survival time, suggesting that women of Luhya ethnicity had longer quantile of survival of miscarriage compared to those from other ethnic groups. The variable ethnicity was coded (Luhya=1 and Others=0). The coefficient estimates for $\tau = 0.05, 0.1, 0.15$ and 0.2 for the factor ethnicity as in the table above are 2.126, 3.038, 2.753 and 1.723 respectively, which shows that luhya ethnicity has longer quantile of miscarriage survival than other ethnicities. The coefficients estimates depict how survival time varies across luhya ethnicity and other ethnicities. Nevertheless, the effect found did not exhibit statistical significance across many of afore mentioned quantiles except at 0.1 quantile. At the 0.05 quantile, the p-value was determined to be 0.8463. Likewise, when considering the 0.1 quantile, the p-value was determined to be 0.0023. Moreover, the p-value at the 0.15 quantile was found to be 0.5515, whereas at the 0.2 quantile, the p-value was determined to be 0.7712.

Number of previous miscarriages on the other hand was associated with having a negative effect on survival time. The effect identified in the study did not exhibit statistical significance at any quantile level. Moreover, based on the data presented in the table, it is evident that the presence of prior stillbirths has a negative influence on the length of survival. The observed effect exhibited statistical significance at the 0.05 quantile level (p=0.032). Similarly, the occupation of the individual had a negative effect on their lifespan, and this effect was statistically significant at the 0.05 quantile level (p=0.0324) and at the 0.2 quantile (p=0.0357). In relation to the place of residence, there was a positive decreasing correlation observed between domicile and survival duration, with no statistical significance observed at any quantile level. In addition, it was observed that the presence of malaria had a non-significant influence on the duration of survival when compared to persons who did not suffer from malaria. The frequency of ANC visits has demonstrated a positive influence on the length of survival. However, it is important to note that this effect was not significant a cross quantiles. The significance of this discovery lies in the potential oversight of important patterns that are essential for the development of effective intervention strategies when utilizing approaches that rely on analyzing average effects.

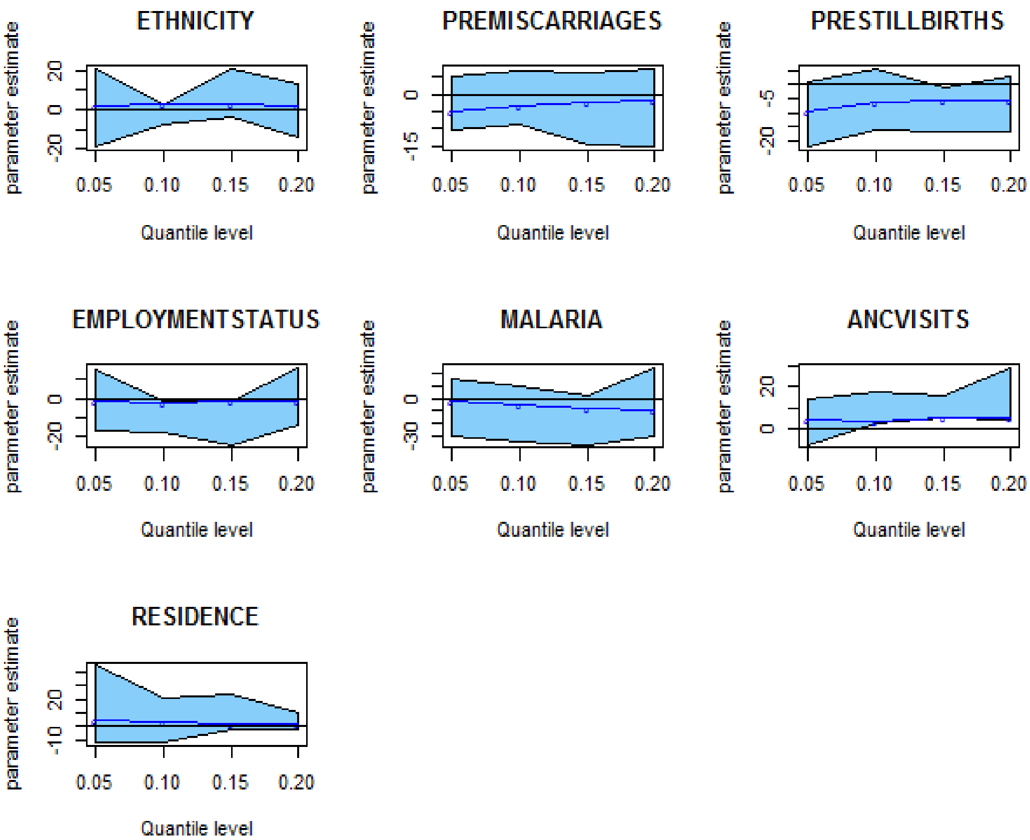

Fig. 11. The effects of ethnicity, previous number of miscarriages, previous number of stillbirths, occupation status, malaria infection, number of ANC visits and place of residence on women miscarriages' survival time

Table 4. The estimated parameters and p-values for t-statistic for variables at each quantile level

Variable	Quantile levels (τ_s)							
	0.05		0.1		0.15		0.2	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Intercept	14.00	0.154	19.62	0.091	18.00	0.073	19.11	0.003
Ethnicity	2.13	0.846	3.04	0.002	2.75	0.552	1.72	0.771
Number of Previous miscarriage	-4.87	0.187	-3.69	0.392	-2.62	0.619	-1.86	0.705
Number of Previous stillbirths	-9.75	0.032	-6.50	0.149	-5.62	0.218	-5.45	0.036
Employment status	-1.87	0.844	-2.73	0.55	-1.46	0.811	-1.56	0.839
Place of residence	4.75	0.756	3.35	0.700	2.54	0.73	2.45	0.516
Malaria status	-2.00	0.873	-5.42	0.620	-8.79	0.509	-9.58	0.443
Number of ANC visits	4.00	0.356	3.35	0.4537	4.71	0.356	4.72	0.301

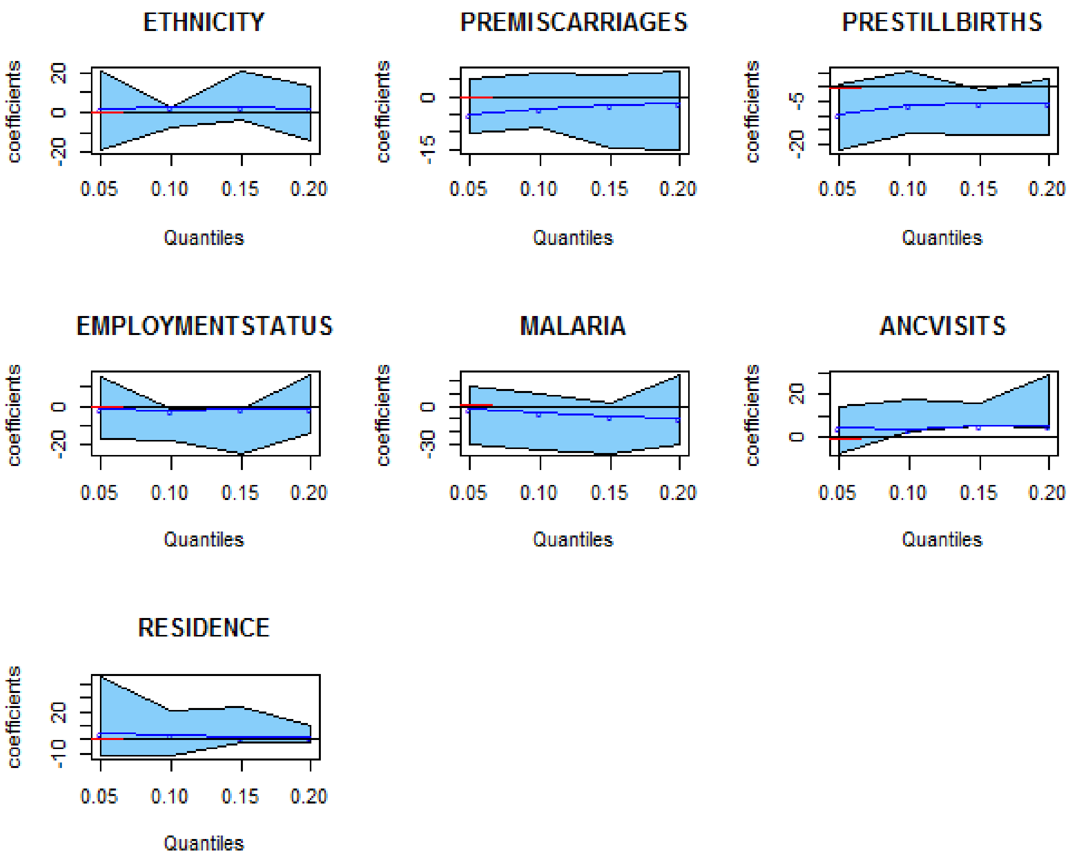

Fig. 2. CQR point estimates (blue curve), 95% CI (lighter blue region) and PHM estimates (red solid line) comparison on risk factor effects

A noteworthy finding pertaining to these results is the lack of any variable that had a statistically significant influence across all quantiles. Based on the data shown in Table 4, it is apparent that multiple risk factors have a significant impact on the length of survival of miscarriages. However, these risk factors exert their influence in different ways at different quantile points. The diverse variables suggest that, at a certain moment in the monitoring procedure, they will have a significant impact on the length of survival for miscarriages. The Fig. 1 depicts the estimated treatment effects (blue curve) observed across all quantiles, with a specific focus on the outcomes presented in Tables 4. The provided Figure additionally displays the 95% (lighter blue shaded region) bootstrap confidence intervals pertaining to the coefficient effects of the CQR model and the null effect (solid horizontal black line). The presented data indicates that there is no statistically significant variation in ethnicity effect across the entirety of the survival distribution (blue line and black line almost overlap). The influence of ethnicity exhibits a noteworthy adverse impact during the initial stages of pregnancy, yet this effect diminishes in significance as the pregnancy advances. The number of prior miscarriages has been found to have a statistically significant negative impact on women's likelihood of survival throughout the initial stages of pregnancy. The observed effect persists in the negative range across all quantiles, although it loses statistical significance in the higher quantiles. A noteworthy adverse impact on survival time was detected across all quantiles in relation to the number of previous stillbirths. Furthermore, the study reveals that occupation consistently exerts a negative and non-statistically significant impact on the duration of survival following miscarriages in women, regardless of their position within the quantiles. A considerable adverse impact during the latter stages of pregnancy was observed in relation to malaria, although no such effect was found during the early stages. For the number of ANC visits variable, it exhibited a statistically significant favorable effect across all quantile levels while place of residence plot shows non-significant effect.

In Fig. 2, a thorough assessment of the coefficient effects is depicted, comparing the results obtained from the Cox PH model (represented by the solid red line) and the censored quantile regression model (illustrated by the dotted blue line). The image presented in this study also showcases the 95% bootstrapped confidence intervals for the coefficient effects of the CQR model. Based on the results depicted in Fig. 2, it is evident that variables such as prior miscarriages, prior stillbirths, place of residence, and number of ANC visits have a statistically significant adverse effect on the duration of survival of miscarriages in the initial phases of pregnancy. The Cox PH model indicates that the observed factors do not exhibit statistically significant impacts. The figure shown illustrates that both the ethnicity and malaria status had a significant effect on survival length within the quantiles ranging from 0.1 to 0.15. Last but not the least occupation status is shown to have no effect across all the quantilaes. The Cox's proportional hazards model is frequently found to be consistent with the quantile regression estimates for specific covariate effects. Nevertheless, there are cases in which the estimates for some covariate effects exhibit significant disparities. An important feature of the quantile regression model is its capacity to incorporate cross-over effects, whereby the influence of a covariate can manifest as either positive or negative. In contrast, Cox estimates do not allow for the potential scenario when a covariate initially increases the hazard for a specific time period and subsequently decreases it.

The results obtained from the censored quantile regression model suggest that the occurrence of previous miscarriages is significantly linked to a negative effect on the length of time a pregnancy survives throughout its early stages. The findings from the CQR model indicated that the influence of prior stillbirths was significant during the initial phases of pregnancy, but did not exhibit statistical significance at any other point throughout the duration of the study. In addition, the variables of place of residence and malaria status demonstrated a significant influence on the length of survival throughout the early stages of pregnancy. The results of this study highlight the potential benefits of utilizing CQR models in the examination of survival data, in order to discover risk factors linked to survival outcomes. Upon conducting a comparison between the Cox model and the CQR model, it was noted that several factors, namely ethnicity, previous number of miscarriages, previous number of stillbirths, occupation status, and malaria infection, exhibited a statistically significant adverse effect on the duration of survival during the initial phases of pregnancy in the CQR model. Nevertheless, according to the Cox proportional hazards model, the variables' effects do not exhibit a statistically significant deviation from zero. The structure of the Cox's proportional hazards model is typically in line with the quantile regression estimates for specific covariate effects. Nevertheless, there exist significant discrepancies in the estimations of various covariate effects. The application of the censored quantile regression model, as opposed to the proportional hazards model, demonstrated notable discrepancies in the impacts of the coefficients. The Cox proportional hazards model demonstrated that the influence of risk factors remains remarkably stable throughout the entire spectrum of survival time. The quantile regression model, following the application of censoring, has indicated that a risk factor may demonstrate statistical significance at specific quantile points within the survival time distribution for women, while not exhibiting significance at other locations. The results of this study are consistent with the conclusions presented by Koenker and Billias in their examination of the Pennsylvania reemployment bonus studies (Koenker & Bilias, 2001). In addition, Koenker et al. suggest in their study that censored quantile regression could be a more flexible alternative to the proportional hazards model, since it offers additional insights into the impact of risk factors on survival length (Koenker et al., 2008). More so, Xue et al (2016) conducted a study that suggests the Cox proportional hazards model, commonly used for assessing time to event data, possesses certain drawbacks. The constraint of the assumption of a constant hazard ratio throughout time, commonly referred to as the proportionality assumption, plays a significant role in this context. Furthermore, the emphasis of this model lies in the modeling of the hazard rate as opposed to the direct estimation of the survival time. The conclusion has great importance as it highlights a limitation in the general proportional hazards model, which, despite being highly esteemed among experts, fails to sufficiently consider the dynamic influence of risk factors across the whole survival period.

4 Conclusion

The core of this thesis was to showcase the important insight that can be gained by using censored quantile regression model. The results of the study suggest that the application of quantile regression models in examining survival data, particularly in the context of the duration of miscarriage survival among pregnant women, demonstrates that the influence of different risk factors on miscarriage survival times varies across different time intervals. The CQR models place increased emphasis on the identification of significant risk variables during each designated time frame. The CQR model also underscored the dynamic nature by which risk factors can influence the duration of survival. The study revealed that the impact of a risk factor is considerable in the early stages of pregnancy, but loses relevance as the pregnancy progresses beyond a certain point. A significant advantage of this finding is that it empowers decision-makers to strategically formulate treatments by taking into account the efficacy of different risk factors. This technique has the potential to provide effective support for those receiving treatment and make a significant contribution to the prevention of miscarriages. The censored quantile regression model is specifically developed for analyzing the quantiles of the time to event. This particular paradigm presents a heightened level of flexibility and interpretability. The present study has provided a comprehensive understanding of the advantages of utilizing censored quantile regression models in order to uncover prognostic risk factors related to miscarriage.

Disclaimer (Artificial intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts

Ethical approval

The research acquired ethical clearance from the Institutional Ethics and Research Committee (MUCHS-MTRH IERC) of Moi University College of Health Sciences and Moi Teaching and Referral Hospital. The office of the chief executive for health County hospitals granted permission or approval to utilize medical data of pregnant women from specific facilities. Prior to conducting the review, authorization or permission was acquired from the appropriate authorities at the hospitals and Antenatal clinic. The utilization of encrypted hard disks was employed for the purpose of gathering datasets from various health facilities. The names of mothers were obtained from databases consisting of computer-based or paper-based records. These names were de-identified and no data was shared, ensuring the confidentiality and privacy of the information were upheld.

Acknowledgement

I am grateful to the health records section of Kakamega County Teaching and Referral Hospital (KCGTRH) for their cooperation during data collecting. I thank everyone who has helped my initiative in whatever way.

Competing Interests

Author has declared that no competing interests exist.

References

Aalen, O. (1978b). Nonparametric inference for a family of counting processes. *Annals of Statistics, 6*, 701–726.

Altman, D. (1991). *Practical Statistics for Medical Research.* London: Chapman & Hall/CRC.

Altshuler, B. (1970). Theory for the measurement of competing risks in animal experiments. *Mathematical Biosciences, 6*, 1-11.

American College of Obstetricians and Gynecologists' Committee on Practice Bulletins—Gynecology. ACOG Practice Bulletin No. 200. (2018, November). Early Pregnancy Loss. *Obstet Gynecol, 132*(5), 197-207.

Andersen AN, Wohlfahrt J, Christens P, Olsen J, Melbye M, Nybo A. (2000). General practice linkage study. *Bmj. 2000;320*, 320.

Andersen et al. (2002). Fever in pregnancy and risk of fetal death: a cohort study. *Lancet, 360*(9345), 1552-6.

Andersen, N.A. M., Vastrup, P., Wollfahrt, J., Andersen, P. K., Olsen, J. and Melbye, M. (2002). Fever in pregnancy and risk of fetal death: a cohort study. *Lancet*, *360*(9345), 1552-6.

Arck PC, Rücke M, Rose M, et al. (2008). Early risk factors for miscarriage: a prospective cohort study in pregnant women. *Reprod Biomed Online*, *17*(1), 101-113.

Armitage, P., Berry, G. and Matthews, J.N.S. (2002). *Statistical Methods in Medical Research* (4th ed. ed.). Oxford: Blackwells Science Ltd.

Austin, P.C., tu, J.V., Daily, P.A., Alter, D.A. (2005). The use of quantile regression in health care research a case study examining gender differences in the timeliness of thrombolytic therapy. *statistics in Medicine, 24*, 791-816.

Benson LS, Holt SK, Gore JL, Callegari LS, Chipman AK, Kessler L, Dalton VK. (2023, March). Early Pregnancy Loss Management in the Emergency Department vs Outpatient Setting. *JAMA Netw Open*, *6*(3).

Breslow, N. (1974). Covariance analysis of censored survival data. *Biometrics*, 89-99.

Breslow, N.E. and Crowley, J. (1974). A large Sample Study of the Life Life Table and Product Limit Estimates under random Censorship. *Annals os Statistics, 2*, 437-453.

Cai, Y. and Feng, WQ. (2005). Famine, social disruption, and involuntary fetal loss: Evidence from Chinese survey data. *Demography, 42*, 301-322.

Carey, V.J., Yong, F., Frenkel, L.M. and McKinney, R. (2004). Growth velocity assessment in paediatric aids smoothing penalized quantile regression and the definition of growth failure. *Statistics in Medicine, 23*, 509-526.

Carlson, E. and Mourgova, M. (2003). Demographic Consequences of Social Inequality in Pregnancy Outcomes.". *Genus LIX*, *2*, 11-28.

Chatenoud, L., Parrazin, F., Di cintio, E., Zanconato, G., Benzi, G., Bortolus, R. and La Vecchia, C. (1998). Paternal and maternal smoking habits before conception and during the first trimester: relation to spontaneous abortion. *Am Epidemiol, 8*(8), 520-6.

Clement EG, Horvath S, McAllister A, Koelper NC, Sammel MD, Schreiber CA. (2019, January). The Language of First-Trimester Nonviable Pregnancy: Patient-Reported Preferences and Clarity. *Obstet Gynecol.*, *133*(1), 149-154.

Cnattingius et al. (2000). Caffeine intake and the risk of first trimester spontaneous abortion. *N Engl J Med, 343*(25), 1839-45.

Cnattingius S., Signorrello L. B., Anneren G., Clausson B., Ekbom A., Ljunger E., Blot W. J., Maclaughlin J. K., Petersson G., Rane A. and Granath F. (2000). Caffeine intake and the risk of first trimester spontaneous abortion. *N Engl J Med, 343*(25), 1839-45.

Collet, D. (2003). *Modelling Binary data* (2nd ed. ed.). Boca Raton, Florida: Chapman & Hall/CRC.

Coste et al. (1991). Risk factors for spontaneous abortion: a case control study in France. *Hum Reprod, 6*(9), 1332-7.

Coste J., Job-Spira N and Fernandez H. (1991). Risk factors for spontaneous abortion: a case control study in France. *Hum Reprod, 6*(9), 1332-7.

Cournane, S., Conway, R., Creagh, D., Byrne, D., Sheehy, N., et al. (2016). Radiology imaging delays as independent predictors of length of hospital stay for emergency medical admissions. *Clinical Radiology, 71*, 912-918.

Cox, D. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, 34*, 187-220.

Cox, D.R. and Oakes, D. (1984). *Analysis of Survival Data.* London: Chapman & Hall/CRC.

Cox, D.R.and Snell, E.J. (1968). A general definition of residuals (with discussion). , A, 30, 248-275. *Journal of the Royal Society, A*(30), 248-275.

de Feijter M, van Gelder MMHJ,Vissers LCM, Kant AC, Woestenberg PJ. (2024). The risk of miscarriage after COVID-19 vaccination before and during pregnancy. *Pharmacoepidemiol Drug Saf., 33*(1).

De la Rochebrochard, E. and Thonneau, P. (2002). Paternal age and maternal age are risk factors for miscarriage; results of multicenter European study. *Hum Reprod , 17*(6), 1649-56.

Dellicour S, Desai M, Mason L, et al. (2007). Epidemiology and burden o fmalaria in pregnancy. Lancet Infect Dis 2007;7:93–104. *Lancet Infect Dis 2007;7:93–104, 7*, 93-104.

Dellicour S, Desai M, Mason L, et al. (2013). Exploring risk perception and attitudes to miscarriage and congenital anomaly in rural Western Kenya. *PLoS ONE, 8*.

Dellicour, S. A. (2016). Weekly miscarriage rates a community-based prospective cohort study in rural western Kenya.. . *BMJ 6:e011088*.

Ding, R., McCarthy, M.L., Desmond, J.S., Lee, J.S., Aronsky, D., et al. (2010). Characterizing waiting room time treatment time and boarding time in the emergency department using quantile regression. *Academic Emergency Medicine, 17*, 813-823.

Ellett, K., Buxton, E.J., and Luesley, D.M. (1992). The Effect of Ethnic Origin on the Rate of Spontaneous Late Mid-Trimester Abortion." 2(1). *Ethnicity and Disease, 2*(1), 84-86.

Faradmal, J., Roshanaei, G., Mafi, M., Sadighi-Pashaki, A. and Karami, M. (2016). Application of censored quantile regression to determine overall survival related factors in breast cance. *Journal of Research in Health Sciences, 16*, 36-40.

Feodor, N.S, Andersen, P.K., Strandberg-Larsen, K., and Nybo, A. A-M. (2014). Risk factors for miscarriage from a prevention perspective: a nationwide follow-up study. *BJOG, 121*, 1375-1385.

Fitzenberger B., Koenker R. and Machado J.A.F. (2002). *‘Economic applications of quantile regression (studies in empirical economics)’.* Springer-Verlag.

Fitzenberger, B., and Wilke, R. A. (2005). Using quantile regresion for duration analysis. *Springer, 2*(3), 103-118.

Flemming J. A, Nanji S., Wei X, Webber C., Groome P. et al. (2017). Association between the time to surgery and survival among patients with colon cancer A population based study. *European Journal of Surgical Oncology*.

Frazier T, Hogue CJR, Bonney EA, Yount KM, Pearce BD. (2018, january). Weathering the storm; a review of pre-pregnancy stress and risk of spontaneous abortion. *Psychoneuroendocrinology, 92*, 142-154.

Ganatra et al. (2015). Expanding access to medical abortion challenges and opportunities. 22(44 Suppl 1):. *Reprod Health Matters, 22*(44 Suppl1), 1-3.

Garcia-Enguidanos et al. (2002). Risk Factors in Miscarriage: a Review.102(2):111-119. *European Journal of Obstetrics, Gynecology and Reproductive Biology, 102*(2), 111-119.

Garcia-Enguidanos, A., M.E. Calle, J. Valero, S. Luna, and V. Dominguez-Rojas. (2002). Risk Factors in Miscarriage: a Review.102(2):111-119. *European Journal of Obstetrics, Gynecology and Reproductive Biology*, *102*(2), 111-119.

Grambsch, P.M. and Therneau, T.M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81*, 515-526.

Guendelmn et al. (1990). Generational differences in perinatal health among the Mexican population: fingings from HHANES 1982-84. *AM J Public Health, 80*(suppl), 61-65.

Hendriks E., MacNaughton H., MacKenzie M.C. (2019, February). First Trimester Bleeding: Evaluation and Management. *Am Fam Physician, 99*(3), 166-174.

Jackson T., Watkins E. (2021, March). Early pregnancy loss. *JAAPA, 34*(3), 222-27.

Joseph K.S., Kramer M.S. (2018, April). The fetuses-at-risk approach: survival analysis from a fetal perspective. *Acta Obstet Gynecol Scand., 97*(4), 454-465.

Kagereki, E.K., Wanjoya, A. and Mageto, T. (2019). Modelling Cases of Spontaneous Abortion Using Logistic Regression. *International Journal of Data Science and Analysis, 5*(6), 143-147.

Kalbfleisch, J.D. and Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data* (2nd ed. ed.). New York: Wiley.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*, 457-481.

Kenya Bureau of Statistics. (2019). *"2019 Kenya populayion and Housing census volume IV: Distribution of popualtion by Socio-economics charateristics".* Retrieved may 2, 2021, from https://www//housingfinance africa.org/appuploads/VOLUME-IV-KPHC-2019.PDF

Kicia, M., Skurzak, A., Wiktor, K., Iwanowicz-Palus, G., & Wiktor, H. (2015). Anxiety and stress in miscarriage. *Polish Journal of Public Health, 125*(3), 162-165.

Kline and Stein. (1984). Spontaneous abortion. *Perinatal Epidemiology. M.B. Bracken*, 23-51.

Koenker, R. (2008). Censored quantile regression redux. *Journal of statistical Soft-ware, 27*(6), 1-25.

Koenker, R., and Bassett Jr, G. (1978). Regression quantiles. *Econometrica, 46*, 33-50.

Koenker, R., and Bilias, Y. (2001). Quantile regression for duration data: a reappraisal of the pennsylvania reemployment bonus experiments. *In Economic applications of quantile regression*, 199-220.

Koenker, R., and Geling, O. (2001). Reappraising Medfly Longevity: A Quantile Regression Survival Analysis. *Journal of the American Statistical Association, 96*(1), 458-468.

Koenker, R., and Hallock, K. (2001). Quantile regression an introduction. *Journal of Economic Perspectives, 15*, 43-56.

Kong GWS, Chung TKH, Lai BPY, Lok IH. (2010). Gender comparison of psychological reaction after miscarriage - A 1-year longitudinal study. *BJOG An Int J Obstet Gynaecol, 117*, 1211-1219.

Lee, E.T. and Wang, J.W. (2013). *Statistical Methods for Survival Data Analysis* (4th ed. ed.). New York: Wiley.

Leong, J.K., Ghim, H., Rahul, M., John, C.A., Suan, T.B., Thiam Chye, T. and Truls Østbye. (2013). A prospective study of risk factors for first trimester miscarriage in Asian women with threatened miscarriage. *Singapore Med J, 54*(8), 425-431.

Lindbohm et al. (2002). effects ofeExposure to environmental tobacco smoke on reproductive health. *Scand J Work Eviron Health, 28*(2), 84-96.

Lindbohm M. L., Sallmén M., and Taskinen H. (2002). effects ofeExposure to environmental tobacco smoke on reproductive health. *Scand J Work Eviron Health, 28*(2), 84-96.

Magnus, M.C.,Wilcox, A.J., Nils-Halvdan, M.,Weinberg, C.R and Håberg, S.E. (2019). Role of maternal age and pregnancy history in risk of miscarriage: prospective register based study. *BMJ, 364*, 1869.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports, 50*, 163-170.

Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22*, 719–748.

Maraka S, Ospina NM, O'Keeffe DT, Espinosa De Ycaza AE, Gionfriddo MR, Erwin PJ, Coddington CC, Stan MN, Murad MH, Montori VM. (2016, April). Subclinical Hypothyroidism in Pregnancy: A Systematic Review and Meta-Analysis. *Thyroid, 26*(4), 580-590.

Meier, P. (1975). *Estimation of a distribution function from incomplete observations. In: Perspectives in Probability and Statistics.* (J. Gani, Ed.) london: Academic Press.

Metwally M, O. K. (2008, September). Does high body mass index increase the risk of miscarriage after spontaneous and assisted conception? A meta-analysis of the evidence. *Fertil Steril., 90*(3), 714-726.

Mills, T. A., Ricklesford, C., Cooke, A., Heazell, A. E. P., Whitworth, M., & Lavender, T. (2014). Parents' experiences and expectations of care in pregnancy after stillbirth or neonatal death: a metasynthesis. *BJOG: An International Journal of Obstetrics & Gynaecology, 121*(8), 943-950.

Mukherjee et al. (2013). Risk of miscarriage among Black and White women in a US: Prospective cohort study. *AM J Epidemiol, 177*(11), 1271-1278.

Neill, S. (2023, April). Management of Early Pregnancy Loss. *JAMA, 329*(16), 1399-1400.

Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics,, 14*, 945–965.

Nilsson et al. (2014). Risk factors of miscarriage from prevention perspective: a nation wide follow up study. *B JOG*, *121*(11), 1375-84.

Ogasawara et al. (2000). Embryonic karyotype of abortuses in relation to the number of previous miscarriages. *Fertil Steril, 73*(2), 300-4.

Ogasawara M., aoki K., Okada s., and suzumori K. (2000). Embryonic karyotype of abortuses in relation to the number of previous miscarriages. *Fertil Steril, 73*(2), 300-4.

Ortiz et al. (2021). Metabolomics in endometriosis: challenges and perspective for future studies. *Reproduction Fertil, 2*(2), R35-R50.

Parazzin et al. (1997). Determinants of risk factors of spontaneous abortions in the first trimester of pregnancy. *Epidemiology, 8*(6), 681-3.

Pasquali, R. (2006). Obesity, fat distribution and infertility. Maturitas. *Maturitas, 54*(4), 363-371.

Peck et al. (2010). A review of the epidemiologi evidence concerning the reproductive health effects of caffeine consumption: a 200-2009 update. *Food Chem Toxicol, 48*, 2449-2576.

Peck J.D., Leviton A., and Cowan L.D. (2010). A review of the epidemiologi evidence concerning the reproductive health effects of caffeine consumption: a 200-2009 update. *Food Chem Toxicol, 48*, 2449-2576.

Peng, L., and Huang, Y. (2008). Survival Analysis With Quantile Regression Models. *Journal of the American Statistical Association, 103*(1, 2, 3, 10, 13), 637-649.

Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant procedures. *Journal of the Royal Statistical Society,A, 135*, 185–207.

Pollard, A.H., Yusuf, F. and Pollard, G.N. (1990). *Demographic Techniques* (3rd ed. ed.). Sydney: Pergamon Press.

Portnoy, S. (2003). Censored Quantile Regression. *Journal of American Statistical Association, 98*, 1001-1012.

Powell, J. (1984). Least Absolute Deviations Estimation for the Censored Regression Model. *Journal of Econometrics, 25*, 303-325.

Powell, J. (1986). Censored Regression Quantiles. *Journal of Econometrics, 32*, 143-155.

Practice Committee of the American Society for Reproductive Medicine. (2012, November). Evaluation and treatment of recurrent pregnancy loss: a committee opinion. *Fertil Steril., 98*(5), 1103-1111.

Rasch, V. (2003). Cigarette, alcohol and caffeine consumption: risk factors for spontaneous abortion. *Acta Obstet Gynecol Scand, 82*(2), 182-8.

Reagan, L. (1991). Recurrent Miscarriage. *Br Med J, 302*, 543-544.

Regan et al. (1989). Influence of past reproductive performance on risk of spontaneous abortion. *Bmj, 299*(6698), 541-5.

Regan l., Braude P. R., and Trembath P. K. (1989). Influence of past reproductive performance on risk of spontaneous abortion. *Bmj, 299*(6698), 541-5.

Sanchez Ortiz, S.; Huerta, C.; Llorente-García, A.; Ortega, P.; Astasio, P.; Cea-Soriano, L. (2021). A Validation Study on the Frequency and Natural History of Miscarriages Using the Spanish Primary Care Database BIFAP. *Healthcare, 9*(5), 596.

Shapiro, S. and Bross, D. (1980). Risk Factors for Fetal Death in Studies of Vital Statistics Data: Inference and Limitations. (I. P. Hook, Ed.) *Human Embryonic and Fetal Death*, 89-106.

Sharp HT, Johnson JV, Lemieux LA, Currigan SM. (2017, April). Executive Summary of the reVITALize Initiative: Standardizing Gynecologic Data Definitions. *Obstet Gynecol, 129*(4), 603-607.

Simpson and Carson. (1993). Biological Causes of Fetal Loss. (R. Gray, Ed.) *Biomedical and Demographic Determinants of Reproduction*,, 287-315.

Simpson J. L. and Carson S. (1993). Biological Causes of Fetal Loss. (R. Gray, Ed.) *Biomedical and Demographic Determinants of Reproduction*,, 287-315.

Simpson, J.L.and Mills, J.L. (1986). Methodologic Problems in Determining Fetal Loss Rates. In G. S. B. Brambati (Ed.), *Chorionic Villus Sampling: Fetal Diagnosis of Genetic Diseases in the First Trimester* (p. 227). New York: M. Dekker.

Slama, R., Bouyer.J.,Windham, G., Fenster, L., Werwatz., A. and Swan, S.H. (2005). Influence of Paternal Age on the Risk of Spontaneous Abortion. *Am J Epidemiol, 161*(9), 816-823.

Stanton et al. (2006). Still Birth rates: delivering estimates in 190 countries. *Lancet*, *367*, 1487-1494.

Stringer, E.M., Kendall, M.A., Lockman, S., Campbell,T.B., Nielsen-Saines, K., Sawe, F., et al. (2018). Pregnancy outcomes among HIV-infected women who conceived on antiretroviral therapy. *PLoS ONE, 13*(7).

Sufian, M.A., Hamzi, W.,Hamzi, B.,Sagar, A.S.M.S., RahmanM.,Varadarajan, J., Hanumanthu, M.,Azad, M.A.K. (2024). Innovative Machine Learning Strategies for Early Detection and Prevention of Pregnancy Loss: The Vitamin D Connection and Gestational Health. *Diagnostics 2024, 14, 920, 14*, 920.

The R Foundation for Statistical computing. (2011). R version 2.13.1. (The R Foundation for Statistical computing) Retrieved from The R Foundation for Statistical Computing website: http://www.r-project.org/foundation

Therneau, T.M. and Grambsch, P.M. (2000). *Modelling Survival Data: Extending the Cox Model.* New York: Springer.

Wang, X., Chen, C., Wang, L., Chen, D., Guang, W. and French, J. (2003). Conception, Early Pregnancy Loss, and Time to Clinical Pregnancy: a Population-Based Prospective Study. *Fertility and Sterility*, *79*(3), 577-584.

Wen et al. (2001). The association of maternal caffeine consumption and nausea with spontaneous abortion. *Epidemiology*, *12*(1), 38-47.

WHO. (1977). Recommended Definitions, Terminology and Format for Statistical Tables Related to the Perinatal Period and Use of a New Certificate for Cause of Perinatal Deaths. Modifications Recommended by FIGO as Amended October 14, 1976. *Acta Obstet Gynecol Scand*, *56*(3), 247-253.

Wong EY, Ray R, Gao DL, et al. (2010). Physical activity, physical exertion, and miscarriage risk in Women textile Workers in Shanghai, china. *Am J Ind Med., 53*, 497-505.

Wood, J. (1994). *Dynamics of human Reproduction: Biology, Biometry, Demography.* New York: Aldine de Gruyter.

Woodward, M. (2014). *Epidemiology: Study Design and Data Analysis* (3rd ed. ed.). Boca Raton, Florida: Chapman & Hall/CRC.

World Health Organization. (1977). WHO recommended definitions, terminology and format for statistical tables related to perinatal period. *Acta Obstet Gynecal Scand, 56*, 247-53.

Wu HL, Marwah S, Wang P, Wang QM, Chen XW. (2017, May). Misoprostol for medical treatment of missed abortion: a systematic review and network meta-analysis. *Sci Rep., 7*(1), 1664.

Wu Y, Yu X, Li M, Zhu J, Yue J, Wang Y, Man YZhou C, Tong R and Wu X. (2024). Risk prediction model based on machine learning for predicting miscarriage among pregnant patients with immune abnormalities. *Front. Pharmacol. 15:1366529, 15*.

Xiong YQ, Tan J, Liu YM, He Q, Li L, Zou K, Sun X. (2019, May). The risk of maternal parvovirus B19 infection during pregnancy on fetal loss and fetal hydrops: A systematic review and meta-analysis. *J Clin Virol, 114*, 12-20.

Xue, X., Xie, X., and Strickler, H. D. (2016). A censored quantile regression approach for the analysis of time to event data. *Statistical Methods in Medical Research, 21*(6), 21-35.

Yazdani, A., Yaseri, M., Haghighat, S. et al. (2021). The comparison of censored quantile regression methods in prognosis factors of breast cancer survival. *Sci Rep, 11*.

Zinaman et al. (1996). "Estimates of Human Fertility and Pregnancy Loss. *Fertility and Sterility, 65*(3), 503-509.

Zinaman, M.J., Clegg E.D, Brown C.C, O'Connor J., and Selevan S.G. (1996). "Estimates of Human Fertility and Pregnancy Loss. *Fertility and Sterility, 65*(3), 503-509.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/126704