Enhancing Survival Analysis of Contraceptive Use with Time-Varying Covariates: Evidence from Longitudinal Data

Baffoe Samuel[1], Arori O. Wilfred[2], Anyango Cynthia Linet[3]

[1]Department of Applied Mathematics, Koforidua Technical University, Faculty of Applied Sciences and Technology, Koforidua, Ghana
[1,2,3]Department of Statistics and Actuarial Science, Maseno University, School of Mathematics, Statistics and Actuarial Science

ABSTRACT
Statistical methodologies for medical and health research have changed significantly, bringing out the dynamics pertaining to disease progression and treatment outcomes. Methods for analyzing survival data help understand the changes in subjects over time, including assessing the time to event. However, the standard survival models assume only time-invariant relationships, ignoring the clinical variable's time-varying nature, which is observed mostly in chronic diseases. This study addresses this limitation by creating a time-varying covariate model for longitudinal data. The model incorporates a shared random-effects structure for longitudinal and survival components, facilitating correlation between the two processes. We develop a Cox proportional hazard model that incorporates a time-varying covariate, validate its applicability to data, and compare the predictive accuracy of the proposed model to the standard model. Existing data from Performance Monitoring for Action (PMA) was used to compare the performance of the standard Cox model with the time-varying covariate model. Key findings indicate that age, education, intention of using contraception in the future, and method switching significantly influence the risk of discontinuation in contraceptive use. The time-varying model shows the best prediction values based on AIC, BIC, and the concordance index, demonstrating the advantage of employing such models in reproductive health studies. Informed by the policy implications, there should be some strategies targeting younger, educated women, as well as those who will use contraceptives in the future. Other than improving the methodologies in the field of survival analysis, this study introduces a more accurate and adaptable framework for clinical forecasting, ultimately leading to improved treatment outcomes.

Keywords: Time-varying covariate, Cox proportional hazards model, Contraceptive Discontinuation, Family Planning

INTRODUCTION

In health sciences, survival modelling is an important statistical method employed to analyze the time-to-event data that is useful in providing predictions on event outcomes which may include death, relapse from disease, or dropping out of the treatment program, among others (Chen, G.H., 2024). However, survival models like the Cox proportional hazards model often assume constant covariate effects that neglect the time-varying nature of many clinical variables (Rizopoulous, et al.,2017). Building upon prior findings on the risk of contraceptive discontinuance among Kenyan women by Baffoe et al. (2024), who employed survival analysis to investigate the determinants of stopping contraception, underscoring demographic factors such as education level and age that impact contraceptive use patterns, the purpose of this study is to improve the current method of survival analysis of contraceptive use by the use of time-varying covariates, analysing the longitudinal data of contraceptive use to get a better understanding of the factors that contribute to contraceptive continuation and discontinuation.

The Cox proportional hazard model is a framework on which most survival analysis models are built since it identifies the effect of covariates on the hazard function, which is the risk of occurrence of an event at a given time. Let T be the random variable representing the time until the event of interest occurs, and let X = (X_1, X_2, ..., X_p.) be a vector of p covariates or explanatory variables. The Cox model assumes that the hazard function $h_{t|x}$ as the product of a baseline hazard function $h_0(t)$ which is independent of covariates, and an exponential term dependent on covariates:

$$h(t|X) = h_0(t)\, exp\big(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p\big) = h_0(t)\, exp(\beta' \boldsymbol{X}) \qquad (1)$$

Where: $\beta' = (\beta_1, \beta_2, \ldots, \beta_p)$ represents the vector of regression coefficients associated with the covariates $X = (x_1, x_2, \ldots, x_p)$ represents the values of the vector of explanatory variables for a particular individual. The model assumes a constant hazard ratio (HR) over time given as:

$$HR = \frac{\hat{h}_0(t) exp\,(\sum_{i=1}^{k} \beta_i x_i^*)}{\hat{h}_0(t) exp(\sum_{i=1}^{k} \beta_i x_i)} = exp\left[\sum_{i=1}^{k} \beta_i (X_i^* - X_i)\right] \qquad (2)$$

Where HR > 1, indicates that members of the first group (X^*) are at a high risk of experiencing the event, and HR < 1 suggests that members of the second group (X) are at a high risk of experiencing the event and HR = 1, indicates equal risk for the event of interest. The parameters β_i are estimated by a partial likelihood function:

$$L(\beta) = \prod_{i=1}^{n} \left(\frac{exp\,(\boldsymbol{X}_i \beta)}{\sum_{j \epsilon R(t_i)} exp\,(\boldsymbol{X}_j \beta)} \right) \qquad (3)$$

Where $R(t_i)$ denotes the risk set at a time t_i, the log of the partial likelihood gives the sum over the risk set, we obtain:

$$l(\beta) = In\big(L(\beta)\big) = \sum_{j=1}^{r} \left\{ \boldsymbol{X}_i \beta - ln\left(\sum_{k \epsilon R(t_{(j)})} exp\,(\boldsymbol{X}_k \beta) \right) \right\} \qquad (4)$$

If tied events exist, then either Breslow's (1974) or Efron's (1977) approximation is employed. The Cox model is partly parametric since it does not impose a particular form on the baseline hazard function. It is still very flexible, can be easily interpreted, and is also capable of accommodating censored observations (Baffoe et al., 2024).

Time-Varying Covariate Models

A more descriptive analysis of the survival data is provided by the extension of the Cox model that includes time-varying covariates. This is achieved by introducing time-varying covariates so that instead of the hazard ratios being constant, it is adjusted for the covariates during the study period. For example, the flexible parametric survival model introduced by Royston and Parmar (2019) incorporates time-dependent effects:

$$ln\big(H(t;x)\big) = s(\ln(t)\,;\gamma_0) + \boldsymbol{X}\beta + \sum_{k=1}^{J} s(\ln(t)\,;\gamma_k)\boldsymbol{X}_k, \qquad (5)$$

Given H has the cumulative hazard, at the time t, $s(\ln(t)\,;\gamma_0)$ is a restricted cubic spline function of log time, K is the number of time-varying covariate effects an$s(\ln(t)\,;\gamma_k)$ is the spline function for the k^{th} time-dependent effect. Furthermore, Landmark analysis (Andersen et al., 2013) divides the follow-up period into windows to assess covariate impacts. These approaches make the application of the Cox model more useful by estimating time evolutions of the hazard ratios and are capable of quantifying time-varying effects of the covariates and survival endpoints (Baffoe et al., 2024).

Literature Review

In a study on contraceptive use and discontinuation conducted on rural women in North-West Tanzania, Safari et al. (2019) applied both adjusted and unadjusted multivariable logistic regression models in exploring the determinants of contraceptive use and discontinuation. The Cox proportional hazard model and life table analysis were used to show individual determinants of discontinuation rates among women using contraceptive methods in Tanzania, this approach sets the stage for recognizing the dynamic nature of contraceptive use and the necessity of considering time as a critical variable in survival analysis. Likewise, Austin et al. (2020) discussed issues related to superior censoring and time-varying covariates, and deficiencies of the standard Cox model, along with introducing a Fine-Gray sub-distribution hazard regression model for a better estimation of cumulative incidences over time. This approach shows how different factors affect both contraceptive use and duration of use but are not always mutually exclusive. In a similar line of reasoning, Prol et al. (2024) also examined contraceptive statistics in the US and the salient role of social factors and explored the differences in contraceptive methods utilization by age, race, education, marital status, and insurance coverage. They also argued that health equity and reproductive autonomy are issues that need to be addressed. Together, these studies call for a more comprehensive approach to the analysis of contraceptive behavior that employs the time-varying covariates and demographic variables and explains how individual and contextual factors work together to shape reproductive health decision-making. Baffoe et al (2024) showed that chances of contraceptive dropout reduce with increased age whereas, increased levels of education exhibited a higher risk of dropping out. These findings underscore the demographic influences on contraceptive dynamics and inform the methodological foundation for this study

Several methods have been proposed for handling flexible trend specifications in the Cox models as a way of dealing with non-linear trends, time-varying covariates, and real-time data. Hofner et al. (2011) proposed using penalized splines to capture non-linear trends, offering adaptability to diverse data patterns with a model:

$$h(t|i) = h_0(t)\exp\big(\beta_1 x_{i1}(t) + \beta_2 x_{i2}(t) + \ldots + f(time)\big) \qquad (6)$$

Assuming $f(time)$ as the non-linear function. Ibrahim et al. (2005) introduced a Bayesian framework for time-varying covariates, estimating parameters and their uncertainties through prior distributions but facing computational challenges with large datasets:

$$h(t|i) = h_0(t)\exp(\beta_1 x_{i1}(t) + \beta_2 x_{i2}(t) + \ldots) \qquad (7)$$

Henderson et al. (2000) extended the Cox model by including some of the longitudinal data as covariates: This approach allows for more meaningful interpretations of survival outcomes. However, the complexity of implementation and interpretation is now a stumbling block because the model's sensitivity to assumptions about the relationship between longitudinal data and survival components poses a challenge, a robust sensitivity analysis is recommended and different alternatives can be considered. the hazard rate within the sub-population defined by the set of covariates is given by

$$h(t|i) = h_0(t)\exp\big(\beta_1 x_{i1}(t) + \beta_2 x_{i2}(t) + \ldots + g(longitudinal\ data)\big) \qquad (8)$$

Tanner et al. (2023) also introduced dynamic updating with real-time data for better prediction of the scale parameter. However, the model faces challenges or becomes limited in practical settings when there is a delay or irregular updates.

$$h(t|i) = h_0(t)\exp\left(\sum_{j=1}^{p} \beta_j\, x_{ij}(t) + \delta(real - time\ data)\right) \qquad (9)$$

Zhang et al. (2018) proposed a "Group-Based Modeling of Time-Varying Covariates in Cox Models" to decrease subjectivity in group levels. This model introduces K groups with different β_{jk} structures for heterogeneous effects. However, the use of the proposed approach to define groups poses a potential bias within the subgroup identification.

$$h(t|i) = h_0(t)\exp\left(\sum_{k=1}^{k} \beta_{jk} x_{ik}(t) + \cdots\right) \qquad (10)$$

Altogether, these models illustrate the strength of advancing Cox modeling to adapt to the various survival data structures and the significance of the continual evolution of new methods to confront more computational, interpretational, and practical obstacles. Despite the developments that have been made concerning non-linear trends, time-varying covariates, and heterogeneity in the survival data, there is an important trade-off in terms of computational efficiency, sensitivity to assumptions, and definition of subgroups. However, these findings provided the methodological background for this study.

Methodology

Data Source and Variables

Baffoe et al. (2024) used data from Performance Monitoring for Action (PMA) surveys to analyze contraceptive discontinuation among Kenyan women. The Cox proportional hazard model was used to investigate the effects of demographic characteristics (age, education level, marital status, etc) on contraceptive use characteristics indicator (ever-used contraceptives). In this study, the time-varying nature of components including changes in contraceptive use, perceived intention of using contraceptives, and other behavioral characteristics are considered. This introduces a dynamic approach to modeling changes in predictor variables, hence a more time-sensitive method of capturing contraceptive user behaviors and how they influence the risk of discontinuation.

A time-varying Cox proportional hazard model is introduced. The model is envisaged to improve the ability to explain factors relating to discontinuation and help identify specific ways in which changing contraceptive methods and intended future use of contraceptives are likely to impact discontinuation. This methodological improvement corresponds to suggestions regarding longitudinal studies to capture the complex nature of contraceptive use and its discontinuation.

Time-varying Variables: The following changing factors concerning contraceptive behavioral patterns: change of method, perceived intention to use contraception, and other factors associated with contraceptive changes.

Outcome Variable: Stopping contraception and duration of contraception.

Modeling Approaches

Traditional Cox PH Model: This model looks at the relationship between demographic and contraceptive use factors and the risk of dropping out of contraceptive use while ignoring the fact that some of the covariates may be time variants.

$$h(t|X) = h_0(t)\exp(\beta_1 Age + \beta_2 Level\ of\ Education + \beta_3 Marital\ Status + \beta_4 Ever\ Use\ Contraceptive + \beta_5 Currently\ Using\ Contraceptive) \quad (11)$$

$h_0(t)$ is the baseline hazard function, β represents the coefficient for each variable.

Proposed Cox PH Model with Time-Varying Covariates: This model uses time-varying covariates like switching between users and non-users of contraceptives together with changes in the probability of using contraceptives at different times hence fitting the hazard rates more dynamically.

$$h(t|X(t)) = h_0(t)\exp\big(\beta_1 Age + \beta_2 Level\ of\ Education + \beta_3 Marital\ Status + \beta_4 Ever\ Use\ Contraceptive + \beta_5 Currently\ Using\ Contraceptive + \beta_6 Intention\ to\ Use\ in\ Future(t) + \beta_7 Changes\ in\ Contraceptive\ Methodes(Switching)(t) + \beta_8 Contraceptive\ Use\ Patterns(t)\big) \quad (12)$$

Model fit is tested with measures such as the Akaike Information Criterion (AIC), Bayesian Information criterion (BIC), and the concordance index (C-index). White robust standard errors control the errors at the participant ID level so that the estimated result is highly accurate.

The modified mathematical model for the time-varying covariates

The traditional Cox PH model does not consider a time-variant covariate. This modified model for time-variant covariate aims to fill this gap. A comprehensive integrated framework for jointly analyzing both survival and longitudinal data is presented by this model. This model offers new possibilities for standard statistical problems of this type, such as unmeasured confounding and missing data. It also shows how to deal with generalizable claims and shared random effects from covariates efficiently. When it comes to predictions on lifetime survival outcomes, the model reflects the dynamic nature of clinical covariates.

Components of the proposed Model: This model is an extension of the Cox proportional hazards model, which is an important method in survival analysis. The Cox model estimates hazard ratios and also allows the integration of time-varying covariates. Time-varying covariates are explicitly modeled to show the changing clinical parameters over a time of study. It is a model that will accept both continuous and categorical covariates—what might be called a full rendering of developing patient profiles. This model introduces the concept of dynamic treatment effects, meaning that treatments at different times may have different effects. This approach allows a more nuanced look at efficacy in the treatment of specific diseases - particularly with longitudinal studies reporting its effects over time and how people respond to those treatments in turn. This approach captures the time-varying covariates and clinical parameter changes over time with the look at longitudinal data. The joint modeling accounts for the interdependence between survival outcomes and repeated measurements. Robust methods have been developed to deal with censored observations to reduce bias. Methods such as inverse probability weighting are considered in addressing potential survivorship bias. There is rigorous internal and external dataset verification of the model; this sensitivity analysis will measure the effect of model assumptions, ensuring reliable and robust results.

Rationale for the Proposed Model: The model is built to be versatile and accommodate different disease domain characteristics. This model
is modified to handle uneven observations and irregularities in longitudinal studies to ensure an accurate representation of the time-varying characteristics of covariates. For explicitly modeling time-varying covariates, the model aims to enhance the predictive accuracy of survival outcomes. Precise prognoses for patients are essential for healthcare professionals to make an informed decision. The proposed model makes a methodological contribution to survival analysis and statistics. The model attempts to fill some gaps in the literature, brings a fresh look at joint models and time-varying regression coefficients, and introduces a higher-level statistical methodology beyond the current limits.

Survival Model with Time-Varying Covariates for Longitudinal Data

A time-varying covariate model for the survival analysis of longitudinal data is a model for the observations and time-to-event outcome analysis, where both aspects are accommodated in a complete model. The proposed model incorporates a shared random-effects structure in the longitudinal and survival settings, which allows for correlation between the two processes.

The Longitudinal Component:

$$Y_{ij}(t) = \mu(t) + b_{i0} + b_{i1}x_{ij}(t) + \epsilon_{ij}(t) \quad (13)$$

The Survival Component:

$$h(t|i) = h_0(t)exp(\beta_1 x_{i1}(t) + \beta_2 x_{i2}(t) + \ldots + \gamma(t) + \emptyset(t)) \quad (14)$$

Shared Random Effects:

$$b_i \sim \mathcal{N}(0, \Sigma_b) \quad (15)$$

Correlation Structure:

$$cor\left(\epsilon_{ij}(t), b_i\right) = \rho \quad (16)$$

Components of the Model

Longitudinal Component $(Y_{ij}(t))$: Describes the evolution of the longitudinal measurement over time. It incorporates a fixed effect $\left(\mu(t)\right)$, time-varying covariates $(x_{ij}(t))$, individual-specific random effects (b_{i0}), and error term $(\epsilon_{ij}(t))$.

Survival Component $\left(h(t|i)\right)$: Models the hazard function for the time-to-event outcome.

Includes time-varying covariates $(x_{i1}(t), x_{i2}(t), \ldots)$, shared random effects (b_i), and additional terms $\left(\gamma(t) \text{ and } \emptyset(t)\right)$.

Let's denote the vector β_γ representing the coefficients associated with the time-varying covariates $z(t)$.

Similarly, β_ϕ represents the coefficients associated with the time-varying covariates $w(t)$.

The vector representation for $\gamma(t)$:

$$\gamma(t) = \boldsymbol{\beta}_{\boldsymbol{\gamma}}^{\boldsymbol{T}} z(t) \quad (17)$$

$$\boldsymbol{\beta}_{\boldsymbol{\gamma}} = \begin{bmatrix} \beta_{\gamma 1} \\ \beta_{\gamma 2} \\ \vdots \\ \beta_{\gamma p} \end{bmatrix}$$

$\beta_{\gamma 1}, \beta_{\gamma 2}, \ldots, \beta_{\gamma p}$ are the coefficients associated with each time-varying covariate $z(t)$.

p is the number of time-varying covariates in $z(t)$.

The vector representation for $\phi(t)$:

$$\phi(t) = \boldsymbol{\beta}_{\phi}^{T} w(t) \tag{18}$$

$$\boldsymbol{\beta}_{\phi} = \begin{bmatrix} \beta_{\phi 1} \\ \beta_{\phi 2} \\ \vdots \\ \beta_{\phi q} \end{bmatrix}$$

$\beta_{\phi 1}, \beta_{\phi 2}, \ldots, \beta_{\phi q}$ are the coefficients associated with each time-varying covariate $w(t)$.
q is the number of time-varying covariates in $w(t)$.
These vectors represent the coefficients associated with the respective time-varying covariates $z(t)$ and $w(t)$, allowing for the calculation of $\gamma(t)$ and $\phi(t)$ within the survival model.

Shared Random Effects (b_i):
Captures individual-specific variability in both the longitudinal and survival processes.
Modeled as a multivariate normal distribution with a covariance matrix (Σ_b). The covariance matrix Σ_b represents the covariance structure of the shared random effects b_i. Since b_i follows a multivariate normal distribution with mean 0 and covariance matrix Σ_b, Σ_b will be a square matrix with a dimension equal to the number of shared random effects.
Let's denote the covariance matrix Σ_b as:

$$\boldsymbol{\Sigma_b} = \begin{bmatrix} \sigma_{b1}^2 & \sigma_{b1,b2} & \cdots & \sigma_{b1,bm} \\ \sigma_{b2b1} & \sigma_{b2}^2 & \cdots & \sigma_{b2,bm} \\ \vdots & \vdots & & \vdots \\ \sigma_{bm,b1} & \sigma_{bm,b2} & \cdots & \sigma_{bm}^2 \end{bmatrix}$$

Where:
σ_{bi}^2 represents the variance of the $i-th$ shared random effect $b_{i.}$
σ_{bij} represents the covariance between the $i-th$ and $j-th$ shared random effect b_i and b_j.
The matrix is symmetric since the covariance between b_i and b_j is the same as the covariance between b_j and b_i.
The diagonal elements (σ_{bi}^2) represent the variances of individual random effects, while off-diagonal elements (σ_{bij}) represent the covariances between pairs of random effects. These covariances capture the dependency structure among the shared random effects.

Correlation Structure:
The correlation structure introduces a correlation (ρ) between the random effects in the longitudinal and survival components.
The strong association between the individual-specific patterns of measurement and survival outcome would only occasionally be followed by longitudinal data no longer understating like gathering place.
This joint model includes time-varying covariates and shared random effects, this correlation pattern can be extended to later time points which incorporate survival outcomes for longitudinal data. The correlation structure takes into account possible dependencies between longitudinal and survival processes.

Cox Proportional-Hazards Model with Time-Varying Covariates
The hazard function $(h(t))$ in the Cox Proportional-Hazards model is given by:

$$h(t) = h_0(t) \cdot \exp(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k) \tag{19}$$

Where:
$h_0(t)$ is the baseline hazard function.

$X_1, X_2, ..., X_k$ *are the covariates.*

$\beta_1, \beta_2, ..., \beta_k$ are the corresponding coefficients.

Incorporating time-varying covariates $(Z_1(t), Z_2(t), ..., Z_m(t))$ into the model, the hazard function becomes:

$$h(t) = h_0(t) \cdot \exp\left(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \gamma_1 Z_1(t) + \gamma_2 Z_2(t) + \ldots + \gamma_m Z_m(t)\right) \quad (20)$$

Where:

$Z_1(t), Z_2(t), ..., Z_m(t)$ are time-varying covariates.

$\gamma_1, + \gamma_2, ..., \gamma_m$ are the corresponding coefficients for time-varying covariates.

Dynamic Treatment Effects:

Specify time-varying covariates based on the objective of the study:

1. Ever use the treatment (Contraceptives) (Yes/No):
   $z_{1i(t)} = 1$ if an individual *i* has ever used treatment over time *t*, else $z_{1i(t)} = 0$.
2. Age at First Use:
   $z_{2i(t)} = t - Age_i$ at which individual *i* first used treatment (Contraceptives) if
   $Age_i \leq t$, else $z_{2i(t)} = 0$
3. Duration of Use:
   $Z_{3i(t)} = t - \text{Start time}_t$, where $Start\ time_t$ is the time when the individual *i* started using treatment (Contraceptives) .
4. Treatment (Contraceptives) Method:
   Since this is a categorical variable, we can represent it as binary indicators for each method, or we could use a time-varying covariate representing the intensity of usage of a particular method over time.
5. Changes in Treatment Method (Contraceptive Methods):
   $z_{5i(t)} = 1$ if an individual *i* changes treatment methods between *t* and *t*+*dt*, else $z_{5i(t)} = 0$. (dt is a small time interval)
6. Intention to Use Treatment in the Future:
   $z_{6i(t)} = 1$ if individual *i* has the intention to use treatment in the future at time *t*, else
   $z_{6i(t)} = 0$
7. Time since Last treatment Use:
   $Z_{7i(t)} = t - \text{Last use time}_i$, where *Last use time i* is the time of the last reported use of a treatment method by individual *i*.
8. Cumulative Duration of Treatment Use:
   $Z_{8i(t)} = \int [\text{Start time}_i, t]\ z_{3i(dt)}$, that is the cumulative time spent using treatment from start time *i* to time *t*.
9. Treatment (Contraceptives) Method Switching:
   $z_{9i(t)} = 1$ if individual *i* switch treatment methods between consecutive time points, else $z_{9i(t)} = 0$
10. Treatment Use Patterns:
    This categorical variable can be represented using binary indicators for different patterns of use.
11. Time-varying Intention to Use Treatment:
    $z_{11i(t)} = 1$ if individual i change intention to use treatment between t and $t + dt$, else $z_{11i(t)} = 0$. (dt is a small time interval)

Interpretation of the Model

The extended Cox Proportional Hazards model with time-varying covariates and an additional time-varying component $\phi(t)$ can be interpreted by considering the impact of each variable on the hazard function. Let's break down the interpretation of the model:

$h(t \mid i)$: The hazard function for individual i at time t represents the risk of an event occurring at that time for the individual.

$h_0(t)$: Represents the baseline hazard function at a time t. The hazard when all covariates are zero. This is the reference for comparing the impact of other covariates.

$\beta_1, \beta_2, \dots, \beta_{(p\times q)}$: Coefficients associated with fixed covariates X_{ji} represent the log hazard ratio. For example, if β_1 is positive, it indicates an increase in the hazard for a one-unit increase in X_{1i}.

$X_{1i}, X_{2i}, \dots, X_{pi}$: Fixed covariates for individual i are time-invariant characteristics. The coefficients $(\beta_1, \beta_2, \dots, \beta_{p\times q})$ quantify the impact of these covariates on the hazard,

$\gamma_{1(t)}, \gamma_{2(t)}, \dots, \gamma_{q(t)}$: Coefficients associated with time-varying covariates $Z_{(ji(t))}$ represent the log hazard ratio for the effect of $Z_{(ji(t))}$ at time t. Positive values indicate an increase in hazards associated with changes in $Z_{(ji(t))}$.

$Z_{(1i(t))}, Z_{(2i(t))}, \dots, Z_{(qi(t))}$: Time-varying covariates for individual i capture changes over time. The coefficients $\gamma_{1(t)}, \gamma_{2(t)}, \dots, \gamma_{q(t)}$: quantify the impact of these covariates on the hazard at a given time.

$\emptyset(t)$: The additional time-varying covariate represents unobserved time-dependent effects not explicitly accounted for by measured covariates. The coefficient $\emptyset(t)$: quantifies the impact of this unobserved factor on the hazard at a time t.

Example:

If $\gamma_{(1(t))}$ is positive, it suggests that an increase in $Z_{(1i(t))}$ at time t is associated with an increased hazard of the event for individual i at that specific time

It's imperative to note that interpretation of the coefficients should be done in the context of the field of study with respect to the variables considered for the study. It is also important to ensure that the proportional hazards assumption is validated.

Partial Likelihood Function for the Model

Both longitudinal and survival components are incorporated into the joint model. The partial likelihood function is considered in terms of shared random effects structure, which is conditional on longitudinal data.

When an observed event with times T_i and a censoring indicator δ_i, then the partial likelihood function for the survival component is given as:

$$L(\beta, \gamma, \phi, \Sigma\_b \mid data)$$

$$= \prod_{i=1}^{n} \left[\frac{exp\left(\sum_{t_i}(\delta_i(\beta^T x_i(t_i) + \gamma(t_i) + \phi(t_i)\right) - log\left(\Sigma_j \exp\left(\beta^T x_i(t_j) + \gamma(t_j) + \phi(t_j)\right)\right))}{\sum_{k=1}^{n} exp\left(\sum_{t_k}(\delta_k(\beta^T x_k(t_k) + \gamma(t_k) + \phi(t_k)\right) - log(\Sigma_l \exp\left(\beta^T x_k(t_l) + \gamma(t_l) + \phi(t_l)\right)))} \right]^{\delta_i} \quad (21)$$

Here:

i indexes the individuals.

j indexes all persons at risk at time t_j

k indexes the event of interest
l indexes time points in a subset for comparisons
Possible range is $1 \leq j, k, l \ \leq n$
t_i indexes the observed event times for individual i.
$x_i(t)$ represents the time-varying covariate vector for individual i at time t.
β are the coefficients for the time-varying covariates.
$\gamma(t)$ and $\phi(t)$ are adaptational components in the survival model.
δ_i is the censoring indicator for individual i, where $\delta_i = 1$ if the event is observed and $\delta_i = 0$ if the event is censored.
Σ_b is the covariance matrix for the shared random effects.
The sum is over all observed event time t_i for each individual.
This partial likelihood function, gives the contribution to the overall likelihood of each individual's survival data based on the observed longitudinal data and the shared random effect structure. The provided partial likelihood focuses on the survival aspect and assumes that the longitudinal data is observed for each individual.

Assumptions
The extended Cox Proportional Hazards Model with time-varying covariates relies on several key assumptions to ensure reliable outcomes. First, observations have to be independent which implies that the hazard of one individual has to be independent from the hazard of another. Second, non-informative censoring also supposes that the risk of censoring does not influence the survival probability and does not introduce bias to the estimated hazard ratios. Third, no interaction is allowed between the covariates and time to preserve the dynamic that changes independently of the hazard rate. To avoid bias, it is important to specify the model correctly; more so, the covariates should not be perfectly collinear, whether static or time-varying. Data limitations include suitability for the proportional hazard framework, adequate follow-up time to document the events, and lack of competing risks that would necessitate another model. Subscribing to these assumptions improves the precision and reliability of survival consequences if time-varying covariates are applied.

Results
The findings of this study are based on the Cox proportional hazards (PH) model and are given the form of hazard ratios (HR), which is a critical measure in survival statistics. The HR measures the relationship between a covariate and the hazard rate of an event happening which in this case is stopping contraceptive use.
Standard Cox PH Model: Preliminary findings based on the analysis of the traditional model (Table 1) indicate several crucial findings pointing to the factors affecting the hazard of discontinuation of contraceptive use. Log_Age is shown to have a very strong negative relationship with the hazard (HR = 0.7649, p < 2e-16), implying that as age increases the probability of stopping contraceptives decreases. In the case of the level of education, there is a small but statistically significant impact on the likelihood of discontinuation (HR = 1.015, p = 0.0151), contrary to what might be expected in most subjects, but it increases with the level of education of the participant. Any prior contraceptive use increases the risk (HR = 1.353, p < 2e-16), that is, every use of contraceptives causes discontinuation to have a positive effect. However, there was no significant effect on the hazard if the woman is currently using contraceptives as

shown in the result (HR = 1.274e−09, p = 0.8743). The model has good accuracy based on the concordance index, which is 0.848 The model fit statistics are AIC = 238047.5 and BIC = 238.085.2.

Table 1: Cox Proportional-hazards Model Results (Traditional Model)

Variable	Coefficient	exp(coef)	Pr(>|z|)	exp(-coef)	lower .95	upper .95	Significanc
Age	-0.268	0.765	<2e-16	1.307	0.718	0.814	***
Level of education	0.015	1.015	0.0151	0.985	1.003	1.028	*
marital Status	-0.017	0.983	0.0925	1.018	0.963	1.003	
Ever use Contraceptive	0.303	1.353	<2e-16	0.739	1.308	1.4	***
Currently Using Contraceptive	-20.480	1.27e-09	0.8743	7.85e+08	8.06e-120	2.01e+101	

significance levels are denoted as follows: *p<0.05, **p<0.01, ***p<0.001.

Table 2: Statistical Measures and Model Performance Metrics for the Traditional Model

Measure	Value	Standard Error	df	p-value
Concordance	0.848	0.002		
Likelihood ratio test	24041		5	<2e - 16
Wald test	446.4		5	<2e - 16
Score (log-rank) test	17961		5	<2e - 16

\

Proposed Cox PH Model with Time-Varying Covariates:

As shown in Table 3 below, the findings of the proposed model enhance understanding of the factors that contribute to the hazard of cessation of contraceptive use. The result also shows that Log_Age has a negative influence on the probability of drop-out (HR = 0.7996, p-value < 2e-16) supporting the claim that older users are less likely to drop out of contraceptives. The level of education shows a strong positive relation (HR = 1.015, p = 1.47 × 10 −6), indicating that as the educational level rises the likelihood of course dropout also increases slightly. Marital status thus has a highly significant inverse effect (HR = 0.9735, p = 2.97e−06), confirming that marital status is a protective factor against the risk of discontinuation. Importantly, any ever use of contraceptives did not cause a significant increase (HR = 1.187, p = 0.584). On the other hand, currently using contraceptives presents the strongest significant negative impact with an HR of 6.182e−10 at <2e−16 percent level of significance. Finally, the probability of future contraception usage also has a negative impact on the continuation rate (HR = 0.7797, p < 2e−16). Switching has a significantly positive effect (HR = 5.779, p < 2e−16), which means that switching methods increase the chances of discontinuation significantly. However, there is no influence observed in the contraceptive use pattern (HR = 0.8892, p = 0.707). It verifies very good predictive accuracy, with concordance =0.858 and fit statistics AIC = 237513.6 and BIC = 237573.9.

Table 3: Cox Proportional-Hazards Model Results (Proposed Model)

Variable	Coefficient	exp(coef)	Robust SE	Pr(>|Z|)	exp(-coef)	lower .95	Upper. 95	Significance
log_Age	-0.2236	0.8	0.0166	<2e-16	1.251	0.774	0.826	***
Level of Education	0.0147	1.015	0.0031	1.47e-06	0.985	1.009	1.021	***
marital Status	-0.0269	0.974	0.0058	2.97e-06	1.027	0.963	0.984	***
Ever Use Contraceptive	0.1718	1.187	0.3134	0.584	0.842	0.642	2.195	
Currently Using Contraceptive	-21.2	6.18e-10	0.3119	<2e-16	1.62e+09	3.35e-10	1.14e-09	***
Intention to Use Contraceptive in Future	-0.2489	0.78	0.0162	<2e-16	1.283	0.755	0.805	***
Changes in Contraceptive Methods (Switching)	1.754	5.779	0.0138	<2e-16	0.173	5.625	5.937	***
Contraceptive Use Patterns	-0.1175	0.889	0.313	0.707	1.125	0.481	1.642	

Significance codes: 0'*'0.001'**'0.01'*'0.05'.' 0.1' '1**

Table 4: Statistical Measures and Model Performance Metrics for the Proposed Cox Model

Measure	Value	Standard Error	df	p-value
Concordance	0.858	0.002	8	p<2e-16
Likelihood ratio test	24581		8	p<2e-16
Wald test	782204		8	p<2e-16
Score (log-eank) Test	18267		8	p<2e-16
Robust Score Test	7777		8	p<2e-16

Model Comparison

The comparison between the standard Cox PH model and the proposed Cox PH models reveals several critical improvements with the time-varying covariates model:

Proportional Hazards Assumption Check: Some of the covariate's assumptions of the standard model were shown to have violated the proportional hazards assumption, which includes log_Age, Marital Status, and Ever Use Contraceptive, the results of which highlighted the need for the use of time-varying covariates to account for better contraceptive use.

Model Fit: The assessment of the Model fit, which includes the Concordance Index (C-index), the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC) also emphasizes the effectiveness of the proposed new methodology of employing Cox proportional hazards model with time-varying covariates against the basic technique. The C-index indicates the extent of the model's ability to predict an outcome and has a scale of 0.5, (implies random chance) to 1 (indicates a perfect prediction), which was higher for the proposed model (0.858) as compared to 0. 848 of the standard models – which points to better survival outcome discrimination for the proposed model. Similarly, the proposed model demonstrated a better-fit performance for reaching more decreased AIC and BIC values, 237,513.6 and 237,573.9 compared to a model of 238,047.5 for AIC and 238,085.2 for BIC. Burnham and Anderson on the other hand propose that values differing by 10 or more points on the AIC or BIC are good enough to support the argument of the better model over the other (Burnham & Anderson, 2002). The differences in these measurements exceeding 500 points in both criteria are significant to support the superiority of the proposed model which is more effective in achieving an optimum balance between goodness of fit and model parsimony. These findings suggest that the modeling of time-varying covariates improves the capability of the model to explain structural changes in contraceptive use behaviors which helps make the predictions more accurate and useful for targeting public health programs.

Discussion

As the findings of this study showed, consideration of time-varying covariates is useful in capturing the dynamics of contraceptive use; arguably providing a better understanding of the factors that explain contraceptive dropout. Comparing the standard Cox PH model results with the Cox Model with time-varying covariates, by identifying key predictors and the prediction indexes between the two models.

Key Findings and Implications

Impact of Age and Education: Similar to findings in Baffoe et al. (2020), the two models also showed that the probability of discontinuing contraceptive use decreases with age meaning that older persons are likely to continue to use contraceptives consistently. Likewise, and consistently with the previous analysis, education has a positive and a relatively small impact on discontinuation in both models. By extension, these findings suggest that young women and those with higher education may require special persuasion to continue using contraception. Such interventions could be subject-specific information campaigns with messages developed for a given age and educational level that resonate with findings by Prol et al. (2024).
Intention to Use Contraception: The time-varying model showed a strong relationship between future intention to use contraception and significantly reduced discontinuation rates, which underlines the need for continued counseling and encouragement. Using the theory of planned behavior this finding implies that enhancing an individual's behavioral intention to use contraceptives can greatly enhance consistent use of contraceptives, which can be seen as a crucial approach to long-term family planning for the general public which is in consistent with Safari et al, (2019) and Austin et al, (2020).
Switching Contraceptive Methods: The significant positive impact of the method switching on the probability of discontinuation indicated in the proposed model suggests that those who switch methods may need some support, this support the findings from Baffoe et al, (2024) Another way that public health programs could help to integrate these two methods is where such programs provide advice concerning possible side effects as well as the difficulties that women are likely to encounter while using new methods of contraceptives. Ensuring this support could help reduce the levels of dropout and increase overall contentment with the acceptability of contraceptive methods.

Policy and Programmatic Implications

The study's findings have implications for reproductive health policy and public health programs. Targeted efforts aimed at younger women, women with higher education, and those who declared an intention to continue using contraceptives may be appropriate for increasing long-term contraceptive use. Further, assisting patients who want to switch methods might guarantee contraceptive chain connectivity which is quite crucial in offsetting gap instability to give steady family planning results. These findings underscore the need for public health interventions to include point-of-care counseling and other services for women for whom contraceptive method changes may be desirable at different points in their lives.

Theoretical and methodological contributions.

This study provides the theoretical background on reproductive health research by explaining the importance of time-varying covariates in the analysis of contraceptive use. Methodologically, this research aims to improve the method and will fill gaps. Limitations pointed out in the previous literature, including computational issues, uncontrolled heterogeneity, and sensitivity to assumptions by illustrating how the inclusion of time-varying covariates improves model fit. The adjustments of the clustering by the participant ID through the robust standard error accounts for within-subject dependencies, reducing bias in parameter estimates, which enhances computational efficiency and reliability allowing the model to handle large datasets for complex relationships with greater accuracy. This is in agreement with the recommendation made by Henderson et al. (2000) and Tanner et al. (2023) concerning the improvements of the methods for developing a sound approach for altering survival data structures. These adjustments are not unique to contraceptive research but can be equally applied to other fields where longitudinal data is used These changes enhance the fit of the model to the data as evidenced in the reduced AIC/BIC values and enhanced concordance indices. This contribution builds on the theories by Rizopoulos et al. (2017) and Royston and Parmar (2019) and offers a more accurate interpretation of contraceptive discontinuation and dynamic covariate effects. The methodology used in this research makes the survival outcomes more realistic – a major advance over standard models.

Conclusion

This work offers a theoretical framework for analyzing contraceptive use processes by examining prospectively measured covariates in the Cox proportional hazard structure. This way, we obtain prospective predictors of contraceptive discontinuation and pre-identify important variables, including age, education level, intentions concerning the further use of contraceptives, and switching behavior, which must be considered in the further study of contraceptive utilization. The use of time-varying covariates improves model fit and prediction as it delivers a better estimation of the dynamic covariates that characterize contraceptive use.

Altogether, the proposed model provides insights for developing public health policy and programmatic approaches to decreasing rates of contraceptive disruption among users. As such, this model provides substantial evidence for the need for interventions to maintain consistent contraceptive use among women; especially the young and educated ones. Moreover, facilitating the process of the method change might represent a major source of continuity since method change has been associated with higher rates of discontinuation.

This work, therefore, expands the methodological literature of survival analysis by showing how time-varying covariate aids in identifying temporal phenomena in reproductive health data. The method employed may form a useful starting point for other similar research in a bid to enhance model fit, as well as extend the understanding of temporal characteristics in longitudinal data. Subsequent work can include cross-validation and external validation studies in other useful settings to establish the generalization of the findings and improvement of the model for use in other reproductive health and health-related sectors.

References:

1. Andersen, R., Fabi, V., Toftum, J., Corgnati, S. P., & Olesen, B. W. (2013). Window opening behaviour modeled from measurements in Danish dwellings. Building and Environment, 69, 101-113.
2. Baffoe, S., Arori, O. W., & Anyango, C. L. (2024). Risk of Discontinuance in Contraceptive Uptake Among Kenyan Women: A Survival Analysis Using the Cox Proportional Hazard Model. International Journal for Multidisciplinary Research, 6(6), 1-10.
3. Breslow, N. (1974). Covariance analysis of censored survival data. Biometrics, 89-99.
4. Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.). Springer-Verlag. https://doi.org/10.1007/b97636
5. C. Austin, P., Latouche, A., & P. Fine, J. (2020). A review of the use of time-varying covariates in the Fine-Gray subdistribution hazard competing risk regression model. ncbi.nlm.nih.gov
6. Chen, G. H. (2024). An Introduction to Deep Survival Analysis Models for Predicting Time-to-Event Outcomes. arXiv preprint arXiv:2410.01086.
7. Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. Journal of the American Statistical Association, 72(359), 557-565.
8. Henderson, R., Diggle, P., & Dobson, A. (2000). Joint modeling of longitudinal measurements and event time data. Biostatistics, 1(4), 465–480.
9. Hofner, B., Kneib, T., Hartl, W., & K¨ uchenhoff, H. (2011). Building Cox-type structured hazard regression models with time-varying effects. Statistical Modelling, 11(1), 3-24.
10. Ibrahim, J. G., Chen, M.-H., & Sinha, D. (2005). Bayesian survival analysis. Springer Science & Business Media.
11. Prol, B., Klein, S., Rennie, C., & Andelija, S. (2024). Respondent Demographics and Contraceptive Use Patterns in the United States: A National Survey of Family Growth Analysis. ncbi.nlm.nih.gov
12. Rizopoulos, D., Molenberghs, G., & Lesaffre, E. M. (2017). Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. Biometrical Journal, 59(6), 1261-1276.
13. Royston, P., Choodari-Oskooei, B., Parmar, M. K., & Rogers, J. K. (2019). Combined test versus log-rank/Cox test in 50 randomized trials. Trials, 20, 1-10.
14. Safari, W., Urassa, M., Mtenga, B., Changalucha, J., Beard, J., Church, K., Zaba, B., & Todd, J. (2019). Contraceptive use and discontinuation among women in rural North-West Tanzania. ncbi.nlm.nih.gov
15. Tanner, K. T., Keogh, R. H., Coupland, C. A., Hippisley-Cox, J., & Diaz-Ordaz, K. (2023). Dynamic updating of clinical survival prediction models in a changing environment. Diagnostic and Prognostic Research, 7(1), 24.
16. Zhang, Z., Reinikainen, J., Adeleke, K. A., Pieterse, M. E., & Groothuis-Oudshoorn, C. G. (2018). Time-varying covariates and coefficients in Cox regression models. Annals of translational medicine, 6(7)