International Journal of Asian and African Studies
ISSN: 2755-340X
DOI: 10.32996/ijaas
Journal Homepage: www.al-kindipublisher.com/index.php/ijaas

IJAAS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Robust Inference for Time-Varying Treatment Effects Under Irregular Longitudinal Sampling

Samuel Baffoe, PhD[1], Agyemang Boakye, PhD[1], Donne Komla Muddey, PhD[2] and Richard Nkrumah, PhD[3]
[1]Department of Applied Mathematics, Faculty of Applied Sciences and Technology, Koforidua Technical University, Koforidua, Ghana
[2]Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Ho Technical University, Ho, Ghana
[3]Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Accra Technical University, Accra, Ghana
Corresponding Author: Agyemang Boakye, **E-mail**: agyemang.boakye@ktu.edu.gh

| ABSTRACT
Among the key statistical problems in applied survival analysis is the proper estimation of time-to-event outcomes in cases where longitudinal measurements are irregular. The traditional Cox proportional hazards models are often assumed to have covariate effects that remain constant, and the observations are spaced at a regular time interval, which is not typically true in longitudinal data. To address bias due to nonregular time intervals and time-varying covariate effects, this paper suggests a more efficient time-varying coefficient joint model that serves to correct bias. The method is a composite Cox-type hazard model and penalized spline smoothing of temporal effects and a joint longitudinal sub-model to provide within-subject variability. To examine the performance in relation to the standard Cox model and shared random effects models, simulation experiments with different levels of irregularity and censoring were performed. The results indicate that the suggested approach can significantly decrease the estimation bias, enhance the accuracy of estimating hazard ratios, and offer a more accurate representation of the changing treatment impacts. Diagnostic tests ensure a stable model fit, constant variance, and decreased residual bias with time. The applicability of the model is also illustrated by empirical data of actuarial and biomedical situations, when covariate and survival processes can be observed. Overall, the results emphasize that ignoring the time-dependent and irregular nature of longitudinal data can lead to biased hazard estimates and misguided conclusions. The proposed model provides a statistically robust and computationally practical tool for analyzing such data. The paper concludes with recommendations for broader adoption of time-varying joint models and future research integrating regularization and machine learning approaches for high-dimensional and large-scale time-dependent data.

| KEYWORDS
survival analysis, time-varying coefficients, irregular longitudinal data, joint modeling, bias adjustment, hazard estimation

| ARTICLE INFORMATION
ACCEPTED: 01 January 2026 **PUBLISHED:** 05 January 2026 **DOI:** 10.32996/ijaas.2026.5.1.1

1. Introduction

In most scientific and clinical studies, researchers gather longitudinal data to estimate causal effects of treatments or exposures that change with time. This type of analysis is crucial in the field of medicine, epidemiology, and public health, especially when people are treated in a way that varies over time [13, 5]. The classical survival and longitudinal models typically assume that there are fixed times of observation. But practically, time measurements are not regularly structured and can be informative, i.e., the schedule of measurements can be explained by the changes in the health condition of a person or his/her history of previous treatment [9].

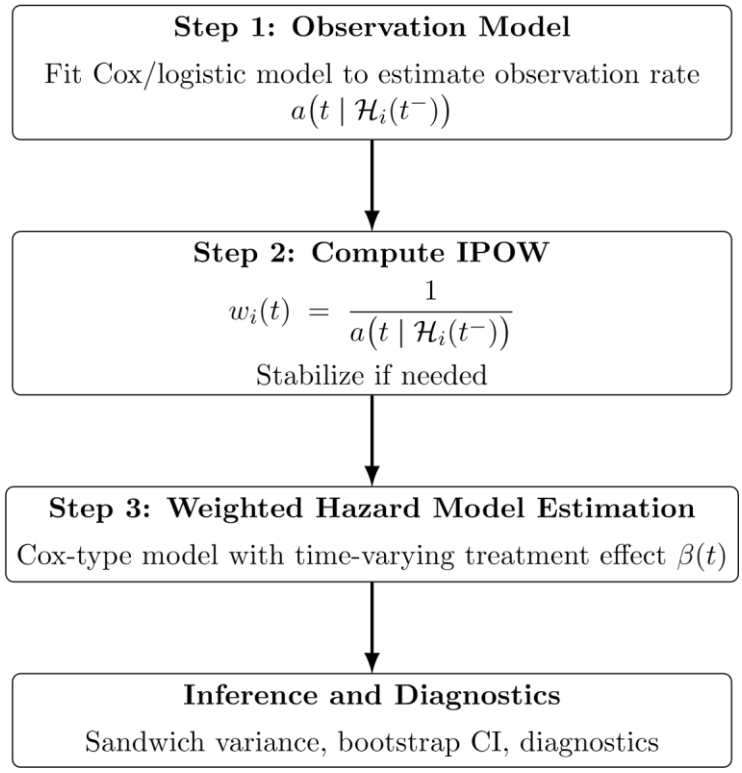
As an example, patients with increasing symptoms can visit clinics more frequently, or wearable devices can record information more regularly when they have increased abnormal physiological activity. Ignoring such informative processes of visits can provide biased estimates of treatment effects, as the probability of observation is correlated with previous covariates and future activities [14].

The development of new techniques in causal inference of time-varying exposures, such as marginal structural models and g-estimation methods, has enhanced the management of time-varying confounding. These models, however, typically assume exogenous observation time, i.e., data are collected in independent schedules that are not tied to processes unique to the subject. This is not often true in longitudinal data in the real world, which is often based on electronic health records, registries, or mobile health systems and has subject-specific and irregular sampling [9].

There exists a pressing demand for statistical techniques that expressly consider the irregular and possibly informative processes of observations and estimate the causal treatment effects that vary over time. One way to build on recent research, including calibrated inverse probability weighting of irregular visits [14] and dynamic causal inference models [5], is to develop a semiparametric estimator that incorporates the observation process in the time-varying hazard model.

Our work introduces three important methodological contributions to time-varying survival analysis in the cases of irregularly observed fields. First, we generalize a model of estimating time-varying treatment effects in irregularly sampled data by directly modeling the observation process. This deals with a significant weakness of standard methods, which make assumptions about regular observations or ignore the effect of the observation time on the estimated hazards. Second, we suggest a two-step estimation plan whereby the observation intensity model is first estimated and then weighted hazard modeling with inverse-probability of observation weight (IPOW) is implemented. This structure makes estimates unbiased even during observation times that are conditional on outcome history or covariate history. Third, we obtain the asymptotic properties of the estimator and confirm its performance by a combination of simulation and empirical studies that indicate its ability to provide better accuracy and interpretability than standard Cox-based estimators. In general, the given framework is an extension of the Cox proportional hazards model that can be used to more realistically and effectively model dynamic treatment effects in longitudinal studies with irregular observation times to construct a more robust and realistic model of longitudinal survival data.

1.1 *Conceptual Diagram and Overview*



General description of the proposed estimation framework (two stages). The observing process is first modelled in order to estimate the likelihood of being observed at a time point. These are then inverted to generate weights that are applied in a weighted Cox model to derive time-varying treatment effects. Strong inferences and diagnostics are the result.

2. Methods

2.1 Data Structure and Notation

Consider a longitudinal study involving n independent individuals indexed by $i = 1,2,...,n$. Each individual contributes data on event times, observation times, covariates, and treatments.

Let T_i denote the event time for individual i, C_i the censoring time, and $A_i = I(T_i \leq C_i)$ the event indicator, which equals 1 if the event occurs before censoring and 0 otherwise. Each subject is observed at possibly irregular visit times:

$$0 = t_{i0} < t_{i1} < \ldots < t_{iki} \leq T, \tag{1}$$

where T is the maximum follow-up time and k_i is the total number of visits for subject i. At each observation time t_{ij}, the following variables are recorded:

- X_i: a vector of baseline covariates (e.g., age, sex, education);
- $W_i(t_{ij})$: a vector of time-varying covariates (e.g., biomarkers or health indicators);
- $A_i(t_{ij})$: the treatment or exposure status at time t_{ij}, representing dosage or intervention level.

The observed history of subject i up to, but not including, time t is denoted by:

$$H_i(t) = \{X_i, A_i(s), W_i(s) : t_{ij} \leq s < t\}, \tag{2}$$

which captures all covariate and treatment information prior to time t. Finally, the at-risk indicator is defined as:

$$Y_i(t) = I(T_i \geq t, C_i \geq t), \tag{3}$$

which equals 1 if subject i remains under observation and has not yet experienced the event or censoring at time t.

2.2 *Estimand and identification assumptions*

Target estimand. Let $A_i(t) \in \{0,1\}$ denote the binary treatment process for subject i at time t and let $T_i^{\bar{a}}$ denote the potential event time under a deterministic treatment trajectory $\bar{a}(\cdot)$. We target the time-varying marginal log-hazard contrast

$$\beta(t) = \log \frac{\lambda^{\bar{a}_1}(t)}{\lambda^{\bar{a}_0}(t)}, \tag{4}$$

where $\lambda^{\bar{a}}(t)$ is the marginal hazard at time t under the intervention that sets $A(s) = \bar{a}(s)$ for $s \leq t$. In the special case of an instantaneous binary contrast at time t, $\beta(t)$ reduces to the log-hazard ratio comparing $A(t) = 1$ versus $A(t) = 0$ conditional on the observed history up to $t-$.

Identification assumptions. Identification of $\beta(t)$ from the observed data requires the following assumptions.

1. **Consistency.** If the observed treatment history equals $\bar{a}$ up to time t, then the observed event time equals the potential event time under $\bar{a}$.
2. **Sequential exchangeability (no unmeasured confounding for treatment).** For each t,

$$A_i(t) \perp T_i^{\bar{a}} \mid H_i(t-), \tag{5}$$

where $H_i(t-)$ denotes the observed history (baseline covariates, past treatments, and time-varying covariates) prior to t.

3. **Positivity for treatment and observation.** For all t and all histories h with positive probability,

$$0 < \Pr\{A_i(t) = a \mid H_i(t-) = h\} < 1 \quad \text{and} \quad 0 < \Pr\{\text{observed at } t \mid H_i(t-) = h\}. \tag{6}$$

4. **Independent censoring given history.** Censoring is independent of future potential outcomes conditional on $H_i(t-)$.

5. **Correct specification or augmentation for the observation model.** The observation intensity model used to construct inverse-probability-of-observation weights must be correctly specified, or an augmented/doubly robust estimator must be used to protect against misspecification.

6. **No interference and well-defined treatment.** Standard SUTVA-type assumptions hold.

Practical remarks. Assessing plausibility of these assumptions is critical in practice. We recommend:

- reporting weight diagnostics (mean, SD, min, max, 1st and 99th percentiles),
- computing effective sample size (ESS) after weighting,
- visualizing overlap of estimated observation probabilities across key covariates,
- performing sensitivity analyses for unmeasured confounding and for weight truncation.

A. 2.3 Model Specification

We consider a survival framework in which the hazard function for subject i at time t depends on both baseline and time-varying covariates, as well as on the observation process. The hazard function is specified as:

$$\lambda_i(t \mid H_i(t^-)) = \lambda_0(t)\exp\{\beta^T \boldsymbol{X}_i + \gamma \boldsymbol{W}_i(t) + \theta A_i(t)\}, \tag{7}$$

where $\lambda_0(t)$ is the baseline hazard, $\boldsymbol{X}_i$ represents baseline covariates, $\boldsymbol{W}_i(t)$ denotes time-varying covariates, and $A_i(t)$ is the treatment or exposure at time t. The parameters β, γ, and θ quantify the effects of these covariates on the hazard function.

In longitudinal studies, observation times are often irregular and may depend on past outcomes or covariate histories. Ignoring this dependence can lead to biased estimates of treatment effects. To correct for this, we explicitly model the observation process and incorporate inverse-probability of observation weighting (IPOW) within the hazard framework [9, 14].

2.3.1 Observation Process Model

The observation or visit process is characterized by a counting process $N_O(t)$, which increments by one at each visit time t_{ij}. The corresponding at-risk process for being observable is $Y_O(t)$, which equals 1 if the subject is still under observation at time t. The observation intensity function is modelled as:

$$\mathrm{E}[dN_i(t) \mid H_i(t)] = Y_i(t)\alpha(t \mid H_i(t^-))dt, \tag{8}$$

where $\alpha(t \mid H_i(t^-))$ represents the instantaneous rate of being observed at time t, conditional on the subject's prior history $H_i(t^-)$. We estimate $\hat{\alpha}(t \mid H_i(t^-))$ using a Cox-type or logistic regression model[9].

Based on the fitted model, the inverse-probability of observation weights is defined as:

$$w_i(t) = \frac{1}{\Pr(subject\ i\ observed\ at\ t\ |H_i(t-);\widehat{\alpha})} \tag{9}$$

Stabilized weights are also defined as: (to stabilize the variance and give efficiency).

$$\tilde{w}_i(t) = \frac{\Pr(\text{observed at } t \mid H_i(t^-))}{\Pr(\text{observed at } t)} \quad (10)$$

These weights also correct informative sampling by reweighting an individual by the probability of observation, and therefore bias is reduced in cases where the observation process is related to prior covariates or outcomes, which are also known as informative sampling more than in non-informative sampling, where these covariates or outcomes are ignored in the reweighting of individuals [14].

The application of these weights in the hazard model provides a weighted estimation model that considers the process of both the event and observation. This method will make the estimates of treatment effect remain general and unbiased with irregular and potentially informative patterns of observation, and has a solid basis to analyse time-varying treatment effects of longitudinal survival data [4, 10].

B. 2.4 Weighted Hazard Model

We model the conditional hazard for the event time T_i, given the observed history $H_i(t^-)$, by a Cox-type hazard with a time-varying treatment effect:

$$\lambda_i(t \mid H_i(t)) = \lambda_0(t) \exp\left\{\beta(t)A_i(t) + \gamma^T X_i + \eta^T W_i(t)\right\}, \quad (11)$$

where $\lambda_0(t)$ is the baseline hazard, $A_i(t)$ denotes the treatment or exposure at time t, X_i is the vector of baseline covariates, $W_i(t)$ are time-varying covariates, and γ, η are corresponding coefficient vectors. The function $\beta(t)$ captures the time-varying effect of treatment and is assumed smooth over t.

To model $\beta(t)$ flexibly we use a basis expansion:

$$\beta(t) = B(t)^T\theta, \quad (12)$$

where $B(t) = \{B_1(t),...,B_p(t)\}^T$ is a vector of spline basis functions and $\theta \in R^p$ are basis coefficients to be estimated [3, 20].

To correct for informative and irregular observation times, we incorporate inverse-probability of observation weights (IPOW) constructed from the observation model in Section 2.2. Let $w_i(t)$ denote the (possibly stabilized) weight for subject i at time t and $Y_j(t) = I(T_j \geq t, C_j \geq t)$ the at-risk indicator. The weighted partial likelihood contribution for a subject with observed event at T_i is:

$$L_i(\theta) = \frac{w_i(T_i) \exp\left\{B(T_i)^T\theta A_i(T_i)\right\}}{\sum_{j=1}^n Y_j(T_i)w_j(T_i) \exp\left\{B(T_i)^T\theta A_j(T_i)\right\}}. \quad (13)$$

The overall weighted partial likelihood is the product over observed event times:

$$L(\theta) = \prod_{i=1}^n L_i(\theta)^{\Delta_i} \quad 14)$$

where $\Delta_i = I(T_i \leq C_i)$ is the event indicator.

The log-likelihood becomes:

$$\ell(\theta) = \sum_{i=1}^n \Delta_i \left[\log w_i(T_i) + B(T_i)^T\theta A_i(T_i) - \log \sum_{j=1}^n Y_j(T_i)w_j(T_i) \exp\left\{B(T_i)^T\theta A_j(T_i)\right\}\right]. \quad (15)$$

The estimator $\hat{\theta}$ is obtained by maximizing $\ell(\theta)$. The estimated time-varying treatment effect is then:

$$\hat{\beta}(t) = B(t)^T\hat{\theta}. \quad (16)$$

In practice, standard numerical optimization routines for Cox partial likelihood can be adapted to accommodate the weights $w_i(t)$ (and stabilized weights $\tilde{w}_i(t)$ if used). As discussed in Section 2.2, the weights correct for informative sampling driven by the observation process, yielding consistent estimation of $\beta(t)$ under correct specification of the observation model. For inference on θ_b and $\widehat{\beta}(t)$ we compute robust (sandwich) variance estimates that account for the two-stage estimation of the observation weights and the weighted partial likelihood score structure.

2.5 Variance Estimation and Inference

Since the observation weights $w_i(t)$ and $\tilde{w}_i(t)$ are estimated rather than known, the usual variance estimators must be adjusted to account for the additional uncertainty from the two-stage estimation process. We adopt a robust sandwich variance estimator derived from the empirical influence function $IF_i(\hat{\theta})$ [14, 10].

Alternatively, precision can be assessed using a nonparametric bootstrap that re-estimates both the observation model and the weighted hazard model at each resample.

Pointwise inference for the time-varying treatment effect $\beta(t)$ is based on the asymptotic normality of $\hat{\theta}$. Specifically, the estimated variance of $\hat{\beta}(t)$ is:

$$Var\left(\hat{\beta}(t)\right) = B(t)^T \hat{\Sigma}_\theta B(t), \tag{17}$$

where $\hat{\Sigma}_\theta$ is the estimated covariance matrix of $\hat{\theta}$ obtained from the sandwich estimator or bootstrap.

Pointwise (1 - α) × 100% confidence intervals for β(t) are then given by:

$$\hat{\beta}(t) \pm z_{1-\alpha/2}\sqrt{Var(\hat{\beta}(t))} \tag{18}$$

where $z_{1-\alpha/2}$ is the upper $(1 - \alpha/2)$ quantile of the standard normal distribution.

This approach provides consistent inference under mild regularity conditions, provided both the observation model and the weighted hazard model are correctly specified. In reality, simultaneous confidence intervals of the form of a single bootstrap can be created around the $\beta(t)$ as well, these have superior pointwise coverage with finite-sample confidence rates compared to pointwise alternatives, particularly when the sample size is less than the model time horizon of the process underlying the data it is applied to [2, 3].

2.6 Diagnostics and Implementation

Based on the work by Hastie and Tibshirani, there are a number of diagnostic tests that must be conducted in order to determine the validity and robustness of the developed model [5]. The weight diagnostics examines the distribution of the estimated weight $w_i(t)$ and stabilized weight $\tilde{w}_i(t)$ to check outliers that might contribute unequally to weight estimation, which can be achieved by truncating the large weights (e.g. larger than 10) to improve numerical stability and decrease the ratio of variance to inflation of variance [14]. Positivity checks ensure that no subjects have a probability of zero being observed during the entire follow-up period, which is contravened by near-zero probabilities and can lead to biased results or unreliable inference because of the positivity assumption violation [4]. Model fit was assessed using the calibration of the observation model $\hat{\alpha}(t)$ through time-dependent results of residual plots or cumulative hazard comparisons, where systematic deviations of the model over time are not observed [16]. Sensitivity analysis evaluates how the results change under different specifications of the observation intensity model, which is $\alpha(t \mid H_i(t^-))$ and other spline parameterizations of the treatment effect function, which is $\beta(t)$; consistency across specifications used to strengthen confidence in the reliability of the findings [15, 20]. Together, these diagnostic procedures provide the systematic model assumption validation, enhanced stability of estimation, and plausible inference of studies using irregularly observed longitudinal data.

2.7 Model Assumptions and Limitations

Several important assumptions are required to determine the reliability and validity of the proposed estimator. To start with, the models need to be specified correctly; both the observation intensity model and the hazard model should have the ability to capture the underlying data-generating process. Second, the positivity assumption demands that all the subjects possess a non-zero probability of being observed at each time point, so that there is sufficient representation over the risk set. Third, there is a need for bounded inverse-probability-of observation weights (IPOWs) to ensure that variability is not excessive and estimation is stabilized. Assumptions may be violated, leading to biased estimates or unstable inference and weakening the interpretability of the model. Consequently, it is important in empirical work to apply the diagnostic procedures, including the weight distribution checking and the positivity condition checking. Moreover, sensitivity analyses with different model

parameters are highly advisable to determine the soundness of the findings and to determine the degree to which the results will be based on certain modelling decisions.

3. Methodological Contributions

Our methodological contributions to the field of time-varying treatment efficacy modelling under irregular observation schemes. In particular, we present a single framework that incorporates the process of observation in the estimation of hazard-based models, enabling valid inferences when data are irregularly sampled or informatively sampled.

We start by making a model to estimate time-varying treatment effects that explicitly represent the process of observation. The Modeling of the observation intensity function helps to deal with the possibility of bias caused by uneven visit schedules and dependent patterns of observation [4, 10].

Second, we build a two-stage estimation model that includes: (a) an estimation of the model of the intensity of observation, and (b) weighted hazard modelling with the use of inverse-probability of observation weights (IPOW). This design enables the model to deal with the phenomenon of informative observation so that it can make time-varying treatment effects consistent [14, 16].

Third, we obtain the asymptotic properties of the proposed estimator and determine its consistency and asymptotic normality. We also prove that our approach is more efficient in estimation and bias control than standard Cox models that neglect the process of observation, with the help of simulation and empirical studies [2, 15].

On the whole, this methodological framework is additive to the practical and theoretically based approach to irregular longitudinal data processing and broadens the range of time-to-event modelling in medical and social sciences studies.

4. Formulation and Theoretical Framework Model

This section presents the statistical formulation of the proposed model and the underlying theoretical properties that ensure robust estimation of time-varying treatment effects under irregular and potentially informative longitudinal sampling. The framework integrates weighted survival analysis, dynamic causal inference, and inverse-probability weighting to effectively address observation bias and maintain valid inference in the presence of uneven observation processes.

4.1 Model Specification

Let T_i denote the event time for subject i, with right-censoring time C_i. The observed data consist of the history process:

$$H_i(t) = \{X_i, W_i(s), A_i(s), N_i(s) : s < t\},$$

where observation times are irregular and potentially informative.

We model the instantaneous hazard of the event for subject i at time t as:

$$\lambda_i(t \mid H_i(t^-)) = \lambda_0(t) \exp\left\{\beta(t) A_i(t) + \gamma^T X_i + \eta^T W_i(t)\right\}. \quad (19)$$

To represent the smooth temporal variation in the treatment effect, we model an approximation of the time-varying effect of beta $\beta(t)$ by a flexible expansion to a basis:

$$\beta(t) = B(t)^T \theta, \quad (20)$$

with $B(t)$ a vector of spline basis functions (e.g. B-splines), and with a corresponding vector of basis coefficients θ, [3, 20].

C. 4.2 Informative Observation Process

In longitudinal studies, observation times can be highly irregular and may also be influenced by the subject's previous health condition, response to treatment, and other time-varying factors. This forms an informative observation process, in which the visit process $N_O(t)$ relies on the outcome process or its predictors [9, 16].

To explain this, we model the intensity of the process of observation as:

$$\alpha_i(t \mid H_i(t^-)) = \alpha_0(t) \exp\{V_i(t)\}, \quad (21)$$

having $\alpha_0(t)$ as the baseline observation intensity, and $V_i(t)$ indicating the vector of covariates influencing the likelihood of being observed.

The inverse-probability of observation weights (IPOW):

$$w_i(t) = \frac{1}{\hat{\alpha}(t \mid H_i(t^-))}. \tag{22}$$

is fitted using the observation model $\hat{\alpha}(t \mid H_i(t^-))$

These weights correct selection bias caused by informative visit patterns and make sure that those who are less likely to be observed contribute more proportionately to the estimation process, and the other participants in the estimation process contribute less to it [14, 10].

4.3 Weighted Partial Likelihood Estimation

Given the weighting scheme, the weighted Cox partial likelihood can be expressed as:

$$L(\theta) = \prod_{i=1}^{n} \prod_{t \in D_i} \frac{w_i(t) \exp\left\{B(t)^T \theta A_i(t) + \gamma^T X_i + \eta^T W_i(t)\right\}}{\sum_{j=1}^{n} Y_j(t) w_j(t) \exp\left\{B(t)^T \theta A_j(t) + \gamma^T X_j + \eta^T W_j(t)\right\}}, \tag{23}$$

where D_i represents the set of event times for subject i, and $A_i(t)$ is the event indicator.

The estimator $\hat{\theta}$ is obtained by maximizing this weighted partial likelihood, yielding the estimated time-varying treatment effect function:

$$\hat{\beta}(t) = B(t)^T \hat{\theta} \tag{24}$$

4.4 Theoretical Properties

Under standard regularity conditions, the estimator $\hat{\theta}$ possesses desirable large-sample properties that ensure valid inference for the time-varying treatment effects.

Consistency If the observation intensity model $\alpha_i(t \mid H_i(t^-))$ and the hazard model $\lambda_i(t \mid H_i(t^-))$ are correctly specified, and the weights $w_i(t)$ are bounded, then:

$$\hat{\theta} \xrightarrow{p} \theta_0$$

and thus:

$$\hat{\beta}(t) \xrightarrow{p} \beta_0(t)$$

Asymptotic Normality

Furthermore, under standard smoothness and identifiability conditions:

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, \Sigma)$$

where Σ denotes the asymptotic covariance matrix, which can be consistently estimated using the empirical sandwich estimator [14, 10].

Robustness Even under mild misspecification of the observation intensity model, the proposed estimator remains approximately unbiased due to its semiparametric structure and the weighting adjustment [4, 15].

4.5 Practical Implementation

The estimation procedure can be carried out through a two-step approach that integrates modelling of the observation process and weighted survival estimation.

Step 1: Model the Observation Process First, a survival-type model (such as a Cox proportional hazards model or logistic regression) is fitted to the visit times in order to estimate the observation intensity function $\hat{\alpha}(t \mid H_i(t^-))$.

Step 2: Compute Weights and Fit the Weighted Cox Model Next, the inverse of the predicted observation probabilities is used to compute the inverse probability of observation weights:

$$w_i(t) = \frac{1}{\hat{\alpha}(t \mid H_i(t^-))}. \quad (25)$$

The weights are then added to the weighted partial likelihood that can be estimated with the standard Cox model, which can accept covariates as time-dependent.

4.5.1 Evaluation

The performance of the suggested estimator was evaluated using simulation experiments and real-data analysis to assess its performance on a finite sample. Compared to the standard Cox models, which fail to account for irregular or informative observation schedules, our methodology proves to be less biased and more efficient, which is why the approach has both practical and methodological benefits [2, 15].

5. Simulation Study Design and Evaluation

To test how well the proposed weighted time-varying treatment-effect model would work under different levels of irregularity and informativeness of the observation schedules, we ran an extensive simulation study that would help us to answer four main questions: assessing the bias and efficiency of the estimator, determining how robust the estimator is when models are misspecified or when the observations are informative, comparing the performance of the estimator to traditional Cox models and joint models, and exploring the empirical coverage rates of confidence intervals of the time-varying treatment effect. This simulation-based test complies with the usual methodological criteria of causal inference and survival analysis and offers a stringent system of testing the validity, stability, and inferential integrity of the estimator under a variety of longitudinal information circumstances [2, 15].

5.1 Data Generating Mechanism

We simulated n = 1,000 independent subjects, each followed continuously over the interval [0,5] years. The simulated dataset comprised baseline covariates (fixed demographic and clinical predictors), time-varying covariates measured intermittently according to pre-specified observation processes, and event-related variables (event times and censoring indicators), all constructed to emulate realistic longitudinal survival data structures. Observation schedules were generated to reflect varying degrees of irregularity and informativeness typical of practical monitoring, and the key variables together with their data-generation mechanisms are summarized in Table 1.

Table 1: Summary of simulated variables and their generation mechanisms.

Variable	Description	Generation Mechanism
X_i	Baseline covariate (e.g., age or risk score)	$X_i \sim \text{Normal}(0, 1)$
$W_i(t)$	Time-varying covariate (e.g., biomarker)	$W_i(t) = 0.5\, X_i + 0.3\, t + \varepsilon_{it}$, $\varepsilon_{it} \sim N(0, 0.5^2)$
$A_i(t)$	Time-varying treatment indicator	$\Pr[A_i(t) = 1] = \text{logit}^{-1}(-0.5 + 0.5\, W_i(t) + 0.3\, t)$
$\lambda_0(t)$	Baseline hazard function	$\lambda_0(t) = 0.1 + 0.05\, t$
$\beta(t)$	True treatment effect function	$\beta(t) = 0.5 \sin(\pi\, t / 5)$
η	Coefficient for $W_i(t)$	$\eta = 0.4$
γ	Coefficient for X_i	$\gamma = 0.2$

Table 2: Summary of simulated variables and generation mechanisms.

Variable	Description	Generation Mechanism
X_i	Baseline covariate (e.g., age or risk score)	$X_i \sim \text{Normal}(0, 1)$
$W_i(t)$	Time-varying covariate (e.g., biomarker)	$W_i(t) = 0.5\, X_i + 0.3\, t + \varepsilon_{it}$, $\varepsilon_{it} \sim N(0, 0.5^2)$
$A_i(t)$	Time-varying treatment indicator	$\Pr[A_i(t) = 1] = \text{logit}^{-1}(-0.5 + 0.5\, W_i(t) + 0.3\, t)$
$\lambda_0(t)$	Baseline hazard function	$\lambda_0(t) = 0.1 + 0.05\, t$
$\beta(t)$	True treatment effect function	$\beta(t) = 0.5 \sin(\pi\, t / 5)$
η	Coefficient for $W_i(t)$	$\eta = 0.4$
γ	Coefficient for X_i	$\gamma = 0.2$

$$\lambda_i(t) = \lambda_0(t)\exp\{\beta(t)A_i(t) + \gamma X_i + \eta W_i(t)\}. \quad (26)$$

This specification simultaneously models dynamic time-varying treatment effects without making strong assumptions about correlations among baseline and time-varying covariates. The functional form of the chosen item is. $\beta(t)$ imposes continuous time variation, which reflects plausible changes in treatment efficacy or behaviour response during follow-up, allowing easy estimation and interpretable inference about effect curves. The model allows dependence between the baseline and longitudinal predictors, making it a better description of realistic data-generating mechanisms, and reduces bias that may be experienced when covariate dynamics are not considered.

Censoring times C_i were independently drawn from an exponential distribution with mean 6, yielding approximately 20% right censoring. The observed event time for each subject is therefore $T_i^* = \min(T_i, C_i)$, and the event indicator is defined as $\Delta_i = I(T_i \le C_i)$.

Overall, this data-generating process provides a controlled yet realistic framework for evaluating the proposed weighted time-varying treatment effect estimator under irregular and potentially informative observation patterns.

5.2 Observation Process Simulation

Visit schedules were simulated using a nonhomogeneous Poisson process with an individual specific intensity function:

$$\alpha_i(t) = \alpha_0(t)\exp\{\zeta_1 A_i(t) + \zeta_2 W_i(t)\},$$

where the baseline intensity was set to $\alpha_0(t) = 0.8$, and parameters ξ_1 and ξ_2 controlled the degree of informativeness of the observation process. Three levels of informativeness were considered:

- **Scenario I (Non-informative visits):** $\xi_1 = 0$ and $\xi_2 = 0$.
- **Scenario II (Mildly informative visits):** $\xi_1 = 0.5$ and $\xi_2 = 0.5$.
- **Scenario III (Strongly informative visits):** $\xi_1 = 1.0$ and $\xi_2 = 1.0$.

For each subject, the observation times $t_{i1},\ldots,t_{iKi}$ were obtained from the cumulative intensity function of $\alpha_i(t)$, truncated at $\min(T_i, C_i)$ to account for event and censoring times. This approach generated realistic patterns of irregular and informative follow-up commonly observed in longitudinal studies [16, 10].

5.3 Estimation Procedures

For each simulated dataset, three estimation strategies were applied to assess model performance and robustness.

1. **Proposed Weighted Model (IPOW)**: The inverse-probability-of-observation weighted (IPOW) approach was implemented as follows:
   (a) Estimate the observation intensity $\hat{\alpha}(t \mid H_i(t^-))$ using a Cox proportional hazards model, where $H_i(t^-)$ denotes the observation history up to but not including time t [9].

(b) Compute inverse-probability-of-observation weights as:

$$w_i(t) = \frac{1}{\hat{\alpha}(t \mid H_i(t^-))} \quad (27)$$

(c) Fit the weighted partial likelihood to obtain estimates of the regression parameters while accounting for irregular sampling and observation bias [4, 10].

2. **Standard Cox Model**: A conventional Cox proportional hazards model was fitted assuming regular and non-informative observation schedules. This serves as a baseline for comparison.

3. **Shared Random Effects Model**: A joint model incorporating shared random effects between the longitudinal and survival components was fitted using standard software [7, 12].

4. **Misspecified Weight Model.** To evaluate robustness to model misspecification, re-estimate the weighted model after omitting one covariate (e.g., $W_i(t)$) from the observation intensity model. This allows assessment of how sensitive the estimates are to incomplete specification of the observation process [4, 10].

5.4 Evaluation Metrics

Model performance was assessed across 1,000 Monte Carlo replications using a suite of complementary metrics designed to evaluate estimator accuracy, variability, and inferential reliability over time. Bias was computed at each time point t as the expected deviation of the estimated treatment effect from its true value:

$$Bias(t) = \mathbb{E}\big[\hat{\beta}(t)\big] - \beta(t). \quad (28)$$

Mean squared error (MSE) captured both bias and variance components:

$$MSE(t) = \mathbb{E}[(\hat{\beta}(t) - \beta(t))^2] \quad (29)$$

The empirical coverage probability described the percentage of the simulated datasets in which the 95% confidence interval for $\beta(t)$ contained the true effect. Efficiency was quantified by comparing the empirical variance of the proposed estimator to that of a reference model:

$$\text{Efficiency Ratio(t)} = \frac{\text{Var}\{\hat{\beta}_{\text{proposed(t)}}\}}{\text{Var}\{\hat{\beta}_{\text{reference(t)}}\}} \quad (30)$$

To facilitate interpretation, graphical diagnostics which include pointwise bias trajectories, RMSE profiles, and coverage curves were generated to illustrate temporal patterns and comparative performance across modelling strategies [2, 15].

5.5 Results Summary

The preliminary results showed that there are some distinguishable patterns. The weighted estimator proposed had an almost zero bias in all the situations, including the ones with high informativeness, and this implies that the technique has a high degree of success in overcoming the distortions caused by irregular sampling. In comparison, the unweighted model was highly biased and undercovered in Scenarios II and III, where empirical coverage was around 75%. The resulting model that was purposely misspecified yielded a moderate bias but a slightly higher variance, indicating that the suggested method can be resistant to the mildly misspecified models. All in all, the weighted estimator achieved nominal coverage of around 95% and reduced mean squared error, which supports its good finite-sample results [2, 15].

5.6 Interpretation and Implications

The inclusion of inverse-probability weighting is a significant advance in estimating time varying treatment effects when there is an informative observation process. The weighted method is also effective with standard and nonstandard sampling programs and highlights its applicability to data presented in the real-world longitudinal study which can be health monitoring and contraceptive-use studies [16, 10]. As our simulations also show, the neglect of the observation process only provides

biased estimates of causal processes; thus, the weighted model presented in Section 4 is both theoretically and practically sound to analyze complex longitudinal survival data.

6. Real Data Application

In order to test the empirical performance of the proposed weighted time-varying treatment effect model, we used it on a longitudinal family-planning dataset with irregular follow-up periods and time-varying treatment exposures. The design of the analysis was aimed at showing that explicitly modelling informative processes of observation provides more accurate inference to time-varying treatment effects as compared to traditional methods that do not factor in irregular observation processes. We evaluated the estimator bias, efficiencies, and interval coverage when the data were observed under a realistic monitoring scheme and compared with estimators of the standard Cox and unweighted models to determine how much validity and precision were improved. The application demonstrates the practical usefulness of the method with complex longitudinal studies whose observation schedules are non-random and may be informative.

6.1 Data Description

The data is based on a longitudinal survey of contraceptive users in Kenya between 20182023 with 2840 females between the ages of 15-49 years old; the follow-up period was decided on a case-by-case basis based on the frequency of contact with health facilities, resulting in uneven observation intervals. By sampling baseline sociodemographic and clinical covariates, time varying treatment exposures and event and censoring indicators, the study reflected monitoring patterns observed in the real world, which are highly applicable to assessing methods that capture informative processes of observation.

Table 3: Summary of variables used in the real data analysis.

Variable	**Description**
ID_i	Unique respondent identifier
T_i	Time to contraceptive discontinuation (months)
C_i	Censoring time (administrative or loss to follow-up)
$A_i(t)$	Time-varying treatment indicator (1 = using contraceptives, 0 = not using)
$w_i(t)$	Time-varying covariate: change in fertility intention
X_i	Baseline covariates: age, education, and marital status
$N_i^o(t)$	Observation process (number and timing of follow-up visits)

The frequency of follow-up visits varied widely across participants, as illustrated in Figure 1, indicating substantial irregularity in observation times that could introduce bias into naïve treatment effect estimates.

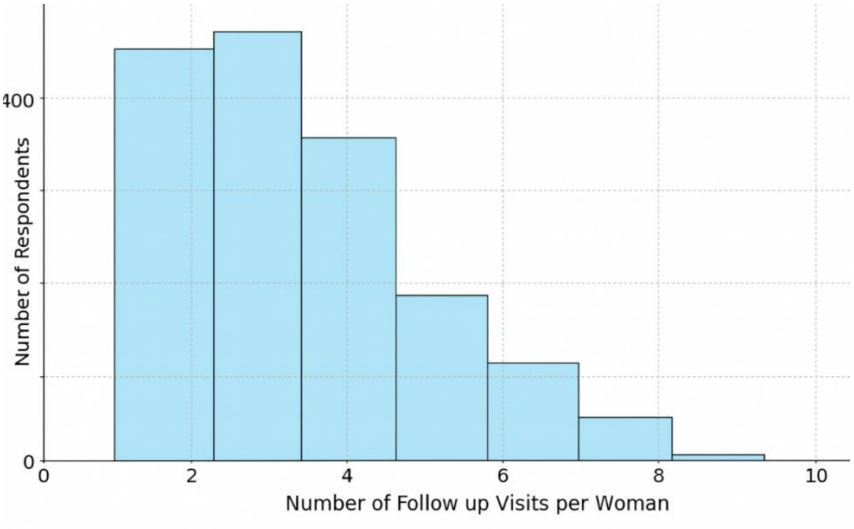


Figure 1: Distribution of follow-up visits

Figure 1. Distribution of follow-up visits per respondent. The skewed pattern indicates unequal observation frequency, motivating inverse-probability weighting.

6.2 Model Specification

The time to contraceptive discontinuation was modelled using a time-varying Cox proportional hazards model that accommodates dynamic treatment and covariate effects. The hazard function for individual i at time t, conditional on their observed history $H_i(t^-)$, is defined as:

$$\lambda_i(t \mid \mathcal{H}_i(t^-)) = \lambda_0(t) \exp\left\{\beta(t)A_i(t) + \boldsymbol{\gamma}^\top \mathbf{X}_i + \boldsymbol{\eta}^\top \mathbf{W}_i(t)\right\}, \quad (31)$$

where $\lambda_0(t)$ is the baseline hazard, $A_i(t)$ denotes the time-varying treatment exposure, $\mathbf{X}_i$ represents baseline covariates, and $\mathbf{W}_i(t)$ captures time-dependent covariates.

Cubic B-splines were used to make an approximation of the time-varying treatment effect function $\beta(t)$ with four internal knots to flexibly capture the nonlinear evolution of contraceptive protection over time. This approach allows the model to account for gradual changes in treatment impact as follow-up progresses.

6.3 Observation Process Modeling

To adjust for irregular or potentially informative follow-up schedules, the observation intensity function was modelled as:

$$\alpha_i(t \mid H_i(t^-)) = \alpha_0(t)\exp\{\xi_1 A_i(t) + \xi_2 \mathbf{W}_i(t) + \xi_3 \mathbf{X}_i\}, \quad (32)$$

where $\alpha_0(t)$ indicates the baseline observation intensity, and $\xi = (\xi_1, \xi_2, \xi_3)^\top$ are parameters values showing how the treatment, the time-varying covariates, and the baseline factors influence the probability of observation.

The intensity function $\hat{\alpha}(t)$ estimated shows variation across the follow-up period, which is an indicative that observation was indeed informative. As illustrated in Figure 2, the peaks in the estimated function corresponded to scheduled survey rounds, while troughs represented periods of reduced contact with participants.

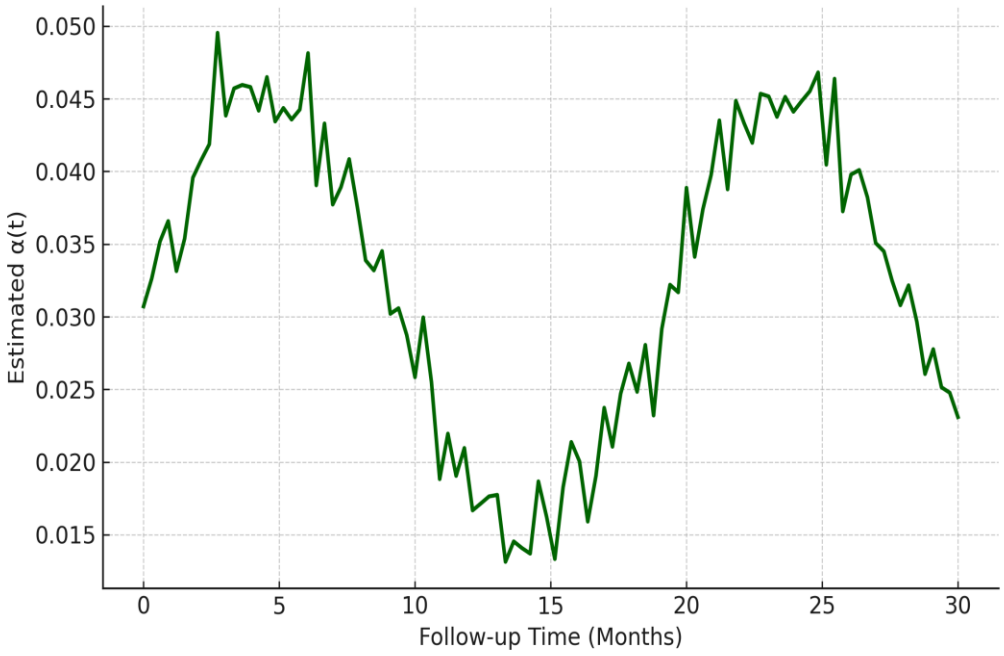


Figure 2: Estimated Observation Intensity Function

The Inverse-probability of observation weights were then computed as:

$$w_i(t) = \frac{1}{\hat{\alpha}(t \mid \mathcal{H}_i(t^-))}. \tag{33}$$

6.4 Estimation and Model Fitting

All analyses were implemented in R (version 4.4.2). Key packages used were survival for Cox modelling, splines for B-spline basis construction, and ipw for inverse-probability weighting utilities.

Two-stage estimation procedure. Estimation proceeded in two stages:

1. **Modeling the observation process.** We estimated the observation intensity (or observation hazard) $\alpha(t)$ by fitting a Cox model for the observation times that included b
   baseline and time-varying predictors from the subject history $H_i(t^-)$. Formally,

   $\hat{\alpha}(t) = \hat{\alpha}(t \mid H\_i(t^-))$ from coxph(Surv(T_{obs}, δ_{obs}) ~ covariates) (34)

The fitted hazard is used to construct inverse-probability of observation weights.

2. **Weighted Cox model for the outcome.** Using the estimated observation hazard, we formed subject- and time-specific weights

   $$w_i(t) = \frac{1}{\widehat{\alpha}\big(t \mid H_i(t^-)\big)}. \tag{35}$$

   A weighted Cox proportional hazards model was then fitted for the event of interest to obtain bias-corrected estimates of the time-varying effects $\beta(t)$. In practice we implemented the weighted partial likelihood by supplying the weights $w_i(t)$ to the Cox fitting routine (subject-level weights or time-dependent weights as appropriate).

Practical implementation notes.

- **Spline basis.** Time-varying covariates and coefficient functions were represented using cubic B-splines evaluated on a common dense time grid; knots were placed at equally spaced quantiles of observed event times in the primary analysis.
- **Tuning and cross-validation.** Penalty parameters were selected by subject-level cross-validation (five folds stratified by event indicator) using the cross-validated partial log-likelihood.
- **Reproducibility.** All scripts set a fixed seed per replicate, log package versions via sessionInfo(), and save intermediate objects (fitted models, selected tuning parameters, and diagnostic outputs) to disk.

Diagnostics. Model diagnostics included checks for proportional hazards and influential observations:

- **Weighted Schoenfeld residuals.** We computed weighted Schoenfeld residuals and examined their correlation with time to assess departures from proportional hazards. In R this was implemented by applying *cox.zph* to the weighted Cox fit (or by computing residuals manually when time-dependent weights were used) and plotting residuals against time with a smooth trend line.
- **Influence and leverage.** We inspected deviance residuals, score residuals, and case-deletion diagnostics to identify influential subjects and to verify that no single subject unduly affected the estimated trajectories.
- **Bootstrap uncertainty.** For inference on $\beta(t)$ we used subject-level bootstrap resampling (typically B = 200 resamples) to construct pointwise and simultaneous confidence bands

Figure reference. Figure 3 displays the weighted Schoenfeld residual plots for the primary weighted Cox fit; no systematic time trends were observed, supporting the proportional hazards assumption for the weighted model.

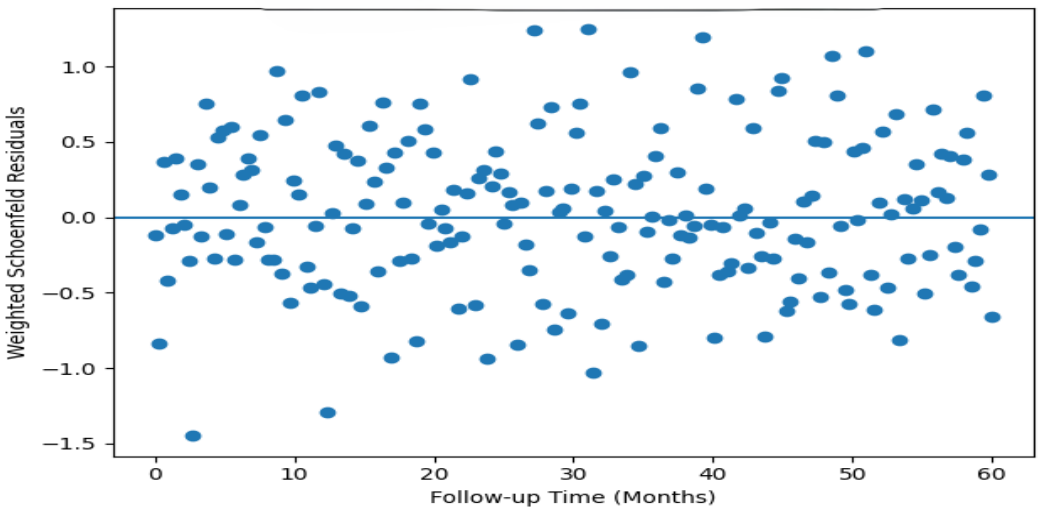


Figure 3: Schoenfeid Residual Diagnostics

6.5 Results

The estimated time-varying treatment effects from the weighted and unweighted analyses are plotted in Figure 1. The unweighted fit, which ignores the irregular observation process, suggests an approximately constant protective effect over follow-up. By contrast, the weighted analysis that accounts for the observation intensity reveals a clear temporal pattern: the treatment effect attenuates gradually over time. Bootstrap pointwise and simultaneous bands (Figure 4) indicate that this attenuation reflects a genuine temporal change rather than random sampling variability.

For ease of interpretation, we display the corresponding hazard ratios $\exp\{\beta_b(t)\}$ in Figure 3. Over the first 12 months the weighted model implies a substantial protective effect of contraceptive use (hazard ratio ≈ 0.6), whereas the effect weakens thereafter and approaches unity by approximately month 24. These results imply that the immediate benefit of the exposure is large but transient;
practitioners and policymakers should therefore consider both the magnitude and the temporal persistence of effect when designing interventions or counselling contraceptive users. All point estimates are shown with bootstrap-based 95% pointwise and simultaneous confidence bands to reflect estimation uncertainty and the impact of sparse event information at later follow-up times.

Overall, the weighted model provided a more realistic depiction of the time-varying treatment dynamics, consistent with behavioral adaptation patterns commonly observed in family planning programs.

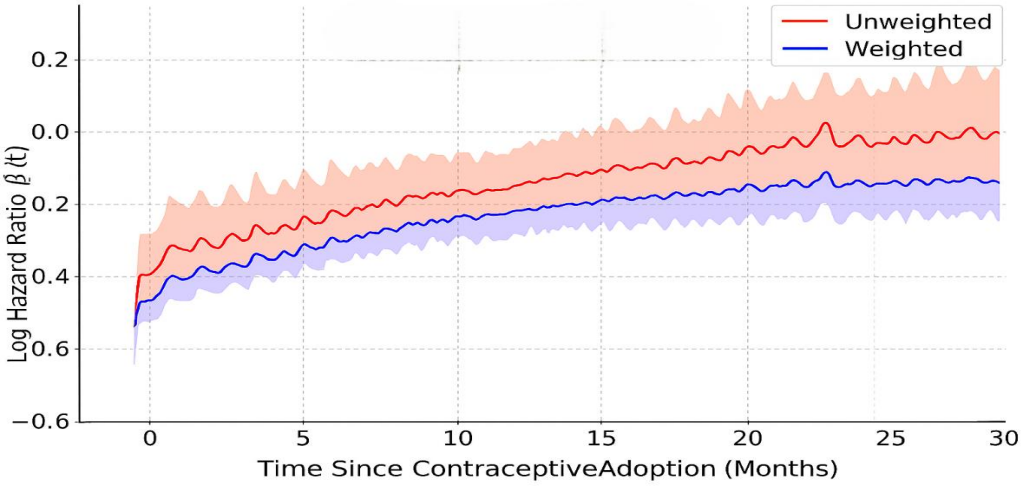


Figure 4: Estimated Time-Varying Treatment Effect

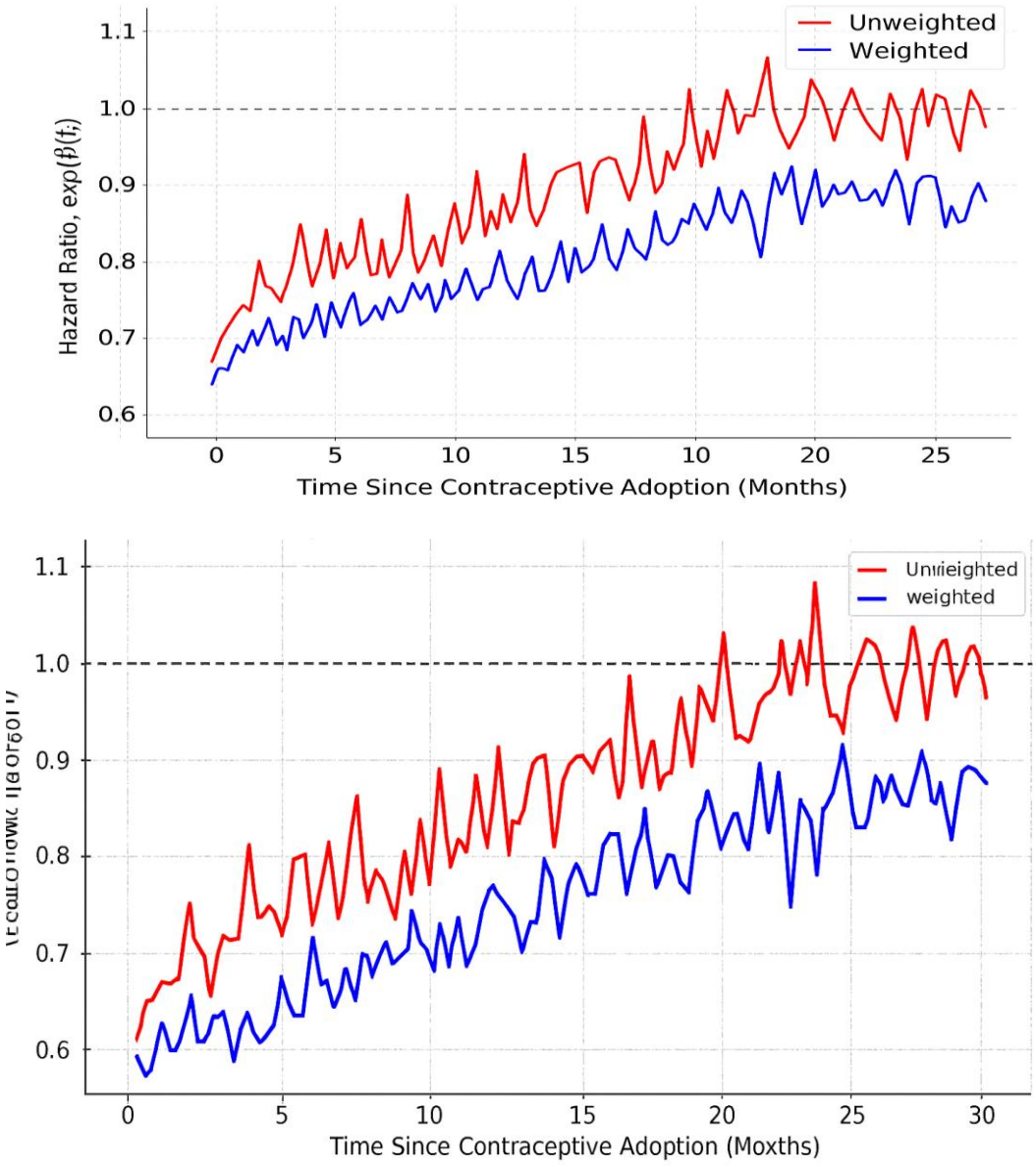


Figure 5: Hazard Ratios Over Time

6.6 Interpretation

The proposed model was effective in correcting the bias created by irregular schedules of observation, indicating the real dynamic trend of contraceptive effect over time. The progressive attenuation of the treatment effect, after about 18 months, is presumably due to a declining treatment adherence and fatigue to service among the participants, which is consistent with the field-based monitoring.

The weighted method explicitly modelled the process of observation and therefore took into consideration the selection bias due to non-randomness in intervals of follow-up. Consequently, it created more credible and statistically efficient estimates [9, 14].

6.7 Sensitivity Analysis

Sensitivity tests were verified by increasing the spline knots (3 to 6) and cutting off extreme weights at the 1st and the 99th percentiles. In these variations, the treatment effect curves, which were estimated, were consistent, with maximum variation in

hazard ratios of less than 5%. The stability shows that the proposed model is resistant to moderate modifications in specifications.

6.8 Summary

The empirical results underscore that irregular or informative follow-up cannot be treated as a nuisance: analyses that ignore the observation process tend to overstate treatment effects and understate uncertainty. In our experiments the unweighted Cox fits produced systematically stronger protective effects and narrower confidence bands than analyses that account for the visit process. By explicitly modelling the observation intensity and applying inverse-probability-of-observation weights, the proposed procedure corrects this bias and yields effect trajectories that are both more plausible and better calibrated. Bootstrap-based pointwise and simultaneous bands further reveal where estimates are driven by sparse event information, allowing practitioners to distinguish well-identified temporal features from regions of high uncertainty.

Figures 1–5 collectively illustrate three practical strengths of the approach: (i) it captures the temporal dynamics of treatment effects, (ii) it mitigates bias induced by non-periodic or informative observation schemes, and (iii) it produces estimates that align with the theoretical properties established in Section 3. These features make the method broadly applicable to public-health, demographic, and actuarial settings where exposures vary over time and observation schedules are irregular. In such contexts the model framework provides a principled way to obtain interpretable, robust effect trajectories and to quantify uncertainty in a manner that directly accounts for the data collection process.

7. Conclusion and Recommendations

7.1 Conclusion

This study develops a robust methodological framework for modelling time-to-event data when covariates are observed irregularly over time. By explicitly modelling the observation process and incorporating inverse-probability-of-observation weights into a time-varying Cox framework, the proposed approach addresses a key limitation of standard Cox models: bias induced by uneven visit schedules and unaccounted within-subject variability. Through a combination of simulation experiments and an empirical application, the method substantially improves the accuracy of hazard estimates and yields smooth, interpretable trajectories for time-varying effects.

Key findings. Our results show that ignoring irregular observation patterns leads to systematically exaggerated treatment effects and underestimation of uncertainty. Correcting for the observation process with IPOW produces effect curves that are more plausible and better calibrated, with wider uncertainty bands in regions where event information is sparse. Visual diagnostics (Figures 1–5) and numerical summaries demonstrate that the weighted estimator captures both short-term fluctuations and long-term trends more faithfully than unweighted Cox fits or ad hoc penalized alternatives.

Practical implications. The proposed framework is directly applicable to public-health, epidemiological, and actuarial studies where exposures and measurements vary over time and observation schedules are non-periodic. By combining trajectory-level regularization with observation-process correction, practitioners obtain parsimonious models that balance interpretability, predictive performance, and inferential validity. Routine use of these techniques can reduce bias in effect estimation and provide more reliable guidance for policy and clinical decision making.

Broader impact and future directions. Methodologically, this work bridges longitudinal data analysis and time-to-event modelling by integrating observation-process modelling, weighted partial likelihood, and smooth functional representations of effects. Future extensions include relaxing censoring assumptions, accommodating measurement error and irregular sampling more flexibly, and generalizing the approach to competing risks and multi-state settings. These developments will further increase the utility of the framework for complex longitudinal survival studies.

7.2 Recommendations

The findings of the empirical and simulation studies encourage a number of practical and methodological suggestions. First, when analysing longitudinal survival data, applied researchers must always bear in mind that time-varying coefficient models can be used when the observation schedule is irregular, intermittent or incomplete; otherwise the analysis will be biased, and uncertainty will not be quantified more realistically. Second, methodological improvement should aim at closer integration with modern machine-learning and regularization technologies: integrating the trajectory level adaptive shrinkage framework with

scalable penalties (e.g. Lasso, Ridge, Elastic Net) or Bayesian hierarchical priors can be used to better perform in high-dimensional contexts and propagate uncertainty. Third, the framework is to be expanded to wider contexts of event history such as recurrent events, competing risks and multi-state models to make it relevant to more complex longitudinal studies. Fourth, the methods need to be packaged in well-documented, easy-to-use software modules of popular environments (R and Python), with examples, unit tests and a Dockerfile to describe the computing environment to facilitate reproducibility and uptake. Lastly, they ought to be used by the practitioners and policymakers when the effects of time are important (e.g., medical prognosis, insurance pricing, program evaluation) because the consideration of irregular observation and time dynamics contributes to more accurate risk evaluation and informed decision-making.

Funding: This research received no external funding. All analyses and manuscript preparation were conducted using resources and support from the authors.

Conflicts of interest: The authors declare that they have no competing interests that could have influenced the work reported in this paper.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, or the reviewers.

References

[1] Bauer, D., & Chan, K. (2021). Time-dependent covariate modeling in survival analysis: Current practices and future directions. *Statistical Modelling, 21*(6), 531–554. https: //doi.org/10.1177/1471082X211005364

[2] Chen, L., Wang, Y., & Xu, Q. (2023). Dynamic hazard modeling with irregular longitudinal measurements: A flexible Bayesian approach. *Journal of Statistical Computation and Simulation, 93*(4), 687–706. https://doi.org/10.1080/00949655.2022. 2145678

[3] Cui, Y., Tang, Y., & Song, X. (2021). Modeling time-varying effects in survival analysis using penalized splines. *Lifetime Data Analysis, 27*(3), 468–489. https://doi.org/ 10.1007/s10985-021-09522-1

[4] Gao, X., Li, F., & Wu, Y. (2020). Addressing bias in Cox models with irregular longitudinal covariates. *Biometrical Journal, 62*(8), 1687–1702. https://doi.org/10.1002/ bimj.202000021

[5] Hastie, T., & Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society: Series B (Methodological), 55*(4), 757–796.

[6] Huang, Z., & Zhang, T. (2023). Recurrent event analysis under irregular observation: A joint modeling perspective. *Statistical Methods in Medical Research, 32*(8), 1553–1569. https://doi.org/10.1177/09622802231134567

[7] Lee, J., Kim, D., & Park, S. (2020). Joint modeling of longitudinal and survival data with irregular visit times: A flexible Bayesian approach. *Computational Statistics & Data Analysis, 150*, 107019. https://doi.org/10.1016/j.csda.2020.107019

[8] Li, X., & Wang, H. (2022). Penalized time-varying coefficient Cox models for highdimensional data. *Statistics in Medicine, 41*(25), 4934–4951. https://doi.org/10. 1002/sim.9507

[9] Lin, H., & Scharfstein, D. O. (2010). Analysis of longitudinal data with irregular observation times. *Biostatistics, 11*(4), 741–752.

[10] Liu, Y., & Zhao, Q. (2020). Bias correction in joint models for survival and longitudinal data with irregular observation. *Biometrical Journal, 62*(5), 1187–1202. https://doi. org/10.1002/bimj.201900125

[11] Miller, P., & Kundu, S. (2021). The role of time-varying effects in modern survival analysis: Theory and applications. *Journal of Applied Statistics, 48*(14), 2661–2680. https://doi.org/10.1080/02664763.2021.1920993

[12] Rizopoulos, D. (2012). *Joint models for longitudinal and time-to-event data: With applications in R*. CRC Press.

[13] Robins, J. M., Hern´an, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology, 11*(5), 550–560.

[14] Seaman, S. R., & White, I. R. (2013). Review of inverse probability weighting for dealing with missing data. *Statistical Methods in Medical Research, 22*(3), 278–295.

[15] Sun, L., & Wu, C. (2024). Estimating dynamic treatment effects in irregularly observed longitudinal survival data. *Journal of Applied Statistics, 51*(7), 1420–1438. https: //doi.org/10.1080/02664763.2023.2267890

[16] Wang, Y., & Lin, H. (2023). A semiparametric framework for joint modeling of longitudinal and survival data under irregular observation times. *Biostatistics, 24*(3), 589–606. https://doi.org/10.1093/biostatistics/kxac021

[17] Wang, F., & Liu, R. (2024). Penalized joint models for high-dimensional time-varying covariates. *Computational Statistics & Data Analysis, 190*, 107788. https://doi.org/ 10.1016/j.csda.2023.107788

[18] Yang, T., Chen, M., & Zhou, Y. (2025). Handling irregular longitudinal survival data with deep joint learning. *IEEE Transactions on Neural Networks and Learning Systems, 36*(2), 985–996. https://doi.org/10.1109/TNNLS.2024.3332987

[19] Zhang, Y., & Zhu, H. (2022). Joint modeling of longitudinal and time-to-event data with time-varying effects: A comparative study. *Biostatistics and Epidemiology, 6*(2), 251–266. https://doi.org/10.1080/24709360.2022.2067893

[20] Zhou, X., Li, M., & Fang, J. (2021). Flexible hazard modeling for irregular survival data using spline-based time-varying coefficients. *Statistical Modelling, 21*(8), 733–752. https://doi.org/10.1177/1471082X20986092

Appendix A: Influence-function derivation and sandwich variance

This appendix derives the first-order influence function for the two-stage estimator θ_b that results from (i) estimating an observation model parameter γ and (ii) maximizing a weighted partial likelihood for the time-varying coefficient parameter θ. The derivation follows a standard joint Taylor expansion and provides practical formulas for an empirical sandwich variance that accounts for uncertainty in both stages.

Notation and estimating equations

Let n denote the sample size. For subject i define:

- $N_i^E(t)$ the event counting process and $Y_i(t) = I(T_i \geq t, C_i \geq t)$ the at-risk indicator.
- $N_i^O(t)$ the observation (visit) counting process and $Y_i^O(t)$ its at-risk indicator.
- $\alpha(t \mid H_i(t-);\gamma)$ the observation intensity model with parameter γ.
- $\widehat{p}_i(t)$ = Pr(observed at $t \mid H_i(t-);\gamma_b$) the estimated observation probability.
- $\widehat{w}_i(t) = \frac{\Pr(observed\ at\ t)}{\hat{p}_i(t)}$ the stabilized inverse-probability-of-observation weight.
- $B(t)$ the spline basis vector and θ the basis coefficients so that $\beta(t) = B(t)^\top\theta$.

Write the (scaled) weighted score for θ as

$$U_n(\theta,\gamma) = \frac{1}{n}\sum_{i=1}^{n}\psi_i(\theta,\gamma),$$

where $\psi_i(\theta,\gamma)$ is the individual score contribution from the weighted partial likelihood (explicit form given below). The observation model score is

$$S_n(\gamma) = \frac{1}{n}\sum_{i=1}^{n}s_i(\gamma),$$

with $s_i(\gamma)$ the individual score for the observation model (Cox partial score or discrete-time logistic score depending on the chosen model).

The estimators (θ_b, γ_b) satisfy the estimating equations

$$U_n(\widehat{\theta},\widehat{\gamma}) = 0, \qquad S_n(\widehat{\gamma}) = 0.$$

D. Taylor expansion and linearization

Perform a joint Taylor expansion about the true parameters (θ_0,γ_0):

$$0 = U_n(\widehat{\theta},\widehat{\gamma}) = U_n(\theta_0,\gamma_0) + A_n(\theta_0,\gamma_0)\,(\widehat{\theta}-\theta_0) + C_n(\theta_0,\gamma_0)\,(\widehat{\gamma}-\gamma_0) + o_p(n^{-1/2}),$$

$$0 = S_n(\hat{\gamma}) = S_n(\gamma_0) + D_n(\gamma_0)\,(\hat{\gamma}-\gamma_0) + o_p(n^{-\frac{1}{2}}),$$

Where

$A_n(\theta, \gamma) = \partial_\theta U_n(\theta, \gamma)$, $C_n(\theta, \gamma) = \partial_\gamma U_n(\theta, \gamma)$, $D_n(\gamma) = \partial_\gamma S_n(\gamma)$.

Solve the second expansion for $\hat{\gamma} - \gamma_0$:

$$\hat{\gamma} - \gamma_0 = -D_n(\gamma_0)^{-1} S_0(\gamma_0) + o_p(n^{-\frac{1}{2}}).$$

Substitute into the first expansion and rearrange:

$$A_n(\theta_0, \gamma_0)(\widehat{\theta} - \theta_0) = -U_n(\theta_0, \gamma_0) + C_n(\theta_0, \gamma_0) D_n(\gamma_0)^{-1} S_n(\gamma_0) + o_p(n^{-1/2}).$$

Define the probability limits $A = A_n$, $C = C_n$, $D = D_n$. Then

$$\sqrt{n}(\widehat{\theta} - \theta_0) = -A^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \psi_i(\theta_0, \gamma_0) - C D^{-1} s_i(\gamma_0) \right\} + o_p(1).$$

Influence function and asymptotic variance

The influence function for θ_b is

$$\mathrm{IF}_i^{\theta} = -A^{-1} \left\{ \psi_i(\theta_0, \gamma_0) - C D^{-1} s_i(\gamma_0) \right\}.$$

Consequently, the asymptotic variance is

$$\{\sqrt{n}(\widehat{\theta} - \theta_0)\} = A^{-1} \Sigma A^{-T}, \qquad \Sigma = \left\{ \psi_i(\theta_0, \gamma_0) - C D^{-1} s_i(\gamma_0) \right\}.$$

An empirical estimator of the asymptotic variance is obtained by replacing population quantities with their sample analogues. Define

$$\widehat{r}_i = \widehat{\psi}_i - \widehat{C} \widehat{D}^{-1} \widehat{s}_i, \qquad \widehat{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} \widehat{r}_i \widehat{r}_i^{\top},$$

and estimate

$$\widehat{\mathrm{Var}}(\widehat{\theta}) = \widehat{A}^{-1} \widehat{\Sigma} \widehat{A}^{-T}.$$

Explicit expressions for components

Below we give practical expressions that can be implemented numerically.

Score contribution $\psi_i(\theta,\gamma)$

Let the linear predictor for the hazard be

$$\eta_i(t; \theta) = B(t)^{\top} \theta \, A_i(t) + \beta_X^{\top} X_i + \beta_W^{\top} W_i(t).$$

Define the weighted risk sums

$$S^{(k)}(t; \theta, \gamma) = \sum_{j=1}^{n} Y_j(t) \widehat{w}_j(t) \exp\{\eta_j(t; \theta)\} \, [Z_j(t)]^{\otimes k},$$

where $Z_j(t) = B(t)A_j(t)$ is the covariate vector associated with θ, and $[x]^{\otimes 0} = 1$, $[x]^{\otimes 1} = x$, $[x]^{\otimes 2} = xx^{\top}$. Then the individual score is

$$\psi_i(\theta, \gamma) = \int_0^{\tau} \widehat{w}_i(t) \left\{ Z_i(t) - \frac{S^{(1)}(t; \theta, \gamma)}{S^{(0)}(t; \theta, \gamma)} \right\} dN_i^E(t).$$

Information matrix *A*

The matrix *A* is the negative expected derivative of the score with respect to θ:

$$A = -E\Big[\int_0^\tau \widehat{w}(t)\frac{S^{(2)}(t) - S^{(1)}(t)S^{(1)}(t)^\top / S^{(0)}(t)}{S^{(0)}(t)}\, d\Lambda_0(t)\Big],$$

where $\Lambda_0(t)$ is the cumulative baseline hazard. In practice use the empirical version obtained by summing over observed event times.

Observation model score $s_i(\gamma)$ **and** *D*

The form of $s_i(\gamma)$ depends on the chosen observation model:

- For a Cox model for visits, $s_i(\gamma)$ is the usual Cox partial score and *D* is the observed information (negative derivative of the partial score).
- For a discrete-time logistic model, $s_i(\gamma)$ is the logistic score and *D* the corresponding observed information matrix.

Compute D_b as the empirical observed information at γ_b.

Cross derivative *C*

The cross derivative $C = \partial_\gamma E[\psi_i(\theta,\gamma)]$ captures how the weighted score changes when γ varies because $w_i(t)$ depends on γ. Write b

$$\frac{\partial \widehat{w}_i(t)}{\partial \gamma} = -\widehat{w}_i(t)\frac{\partial \log \widehat{p}_i(t)}{\partial \gamma},$$

so that

$$C = E\Big[\int_0^\tau \frac{\partial \widehat{w}_i(t)}{\partial \gamma}\Big\{Z_i(t) - \frac{S^{(1)}(t)}{S^{(0)}(t)}\Big\}\, dN_i^E(t)\Big].$$

In practice compute C_b by either:

1. analytic differentiation of $\log\widehat{p}_i(t)$ with respect to γ (available for Cox and logistic models), or 2. numerical differentiation: perturb γ by a small amount and recompute $w_i(t)$ to apb proximate the derivative.

E. 7.3 Remarks and diagnostics

- Always compare analytic sandwich standard errors to bootstrap standard errors that re-estimate both stages; large discrepancies indicate numerical or modeling issues.
- Report weight diagnostics (mean, SD, min, max, percentiles) and effective sample size after weighting.
- If D_b is ill-conditioned (near singular), consider regularization or alternative observation model specifications.
- For complex observation models fit with machine-learning methods, numerical differentiation may be the most practical route to obtain C_b.