Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2025, 2025:111
https://doi.org/10.28919/cmbn/9537
ISSN: 2052-2541

A SIMULATION STUDY ON THE IMPACT OF CENSORING ON STANDARD ERRORS IN THE CLUSTERED JOINT MODEL

ERICK NGANZI[1,3,*], SAMUEL MWALILI[2], GEORGE ORWA[2], PATRICK M. MBURUGU[2]

[1]Pan African University Institute for Basic Sciences Technology and Innovation, Nairobi, Kenya
[2]Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya
[3]The University of Dodoma, Department of Mathematics and Statistics, Dodoma, Tanzania

Abstract. This simulation study rigorously evaluated the impact of censoring on standard error (SE) estimation in clustered joint models for longitudinal binary and survival outcomes. Joint models were vital in biomedical research for analysing interdependent repeated measures and event times, but their performance in clustered settings (e.g., multicenter trials) under censoring remained under explored. We focused on scenarios like adolescent HIV studies, where binary viral load suppression (longitudinal) and time to treatment interruption (survival) were modelled jointly with shared random effects to account for within-cluster correlation. Using Monte Carlo simulations (200–1000 replications) with 10 clusters of 20 subjects, we systematically varied right-censoring rates (5%–35%), cluster sizes, and random-effects structures. Performance metrics including bias, root mean squared error (RMSE), empirical standard errors (ESE), asymptotic standard errors (ASE), and coverage probabilities (CP) were compared. Results revealed that censoring significantly distorted SE estimation, with ASE increasingly underestimating ESE for key parameters as censoring intensified. For WHO stage parameters (e.g., β_2, β_3), ASE-ESE agreement held reasonably until 20% censoring, but deteriorated thereafter. The ART regimen parameter (γ_s) exhibited substantial bias and RMSE inflation beyond 20% censoring (e.g., bias: 0.0675, RMSE: 0.1441 at 30% censoring), with CP falling to 89.5%. Conversely, age parameters (γ_a) showed reduced bias under higher censoring. Variance components ($\log(\sigma_u)$, $\log(\sigma_v)$) displayed significant ASE-ESE discrepancies at all censoring

*Corresponding author
E-mail address: erick.nganzi@udom.ac.tz
Received August 05, 2025

levels. We identified a critical threshold near 20% censoring, beyond which inference reliability declined markedly for treatment-related parameters. Spline-smoothed error trends further illustrated parameter-specific sensitivities. These findings cautioned against relying solely on asymptotic SE approximations in high-censoring or complex clustered designs and underscored the need for robust variance estimation. The study provided practical guidance: research focused on clinical covariates (e.g., ART) should limit censoring below 20%, while demographic parameters tolerated higher rates.

Keywords: joint modelling; censoring; model performance; ART adherence; WHO stage.

2020 AMS Subject Classification: 62P10.

1. INTRODUCTION

Joint models for longitudinal and time-to-event data had become indispensable in biomedical research, enabling the simultaneous analysis of repeated measurements and event times while accounting for their inherent interdependence [1, 2, 8]. These models were particularly powerful when assessing how longitudinal biomarkers dynamically influenced survival outcomes, offering a unified framework that improved efficiency and mitigated bias compared to separate analyses.

In clustered settings such as multicenter clinical trials, community-based studies, or familial cohort designs data often exhibited hierarchical structures where individuals were nested within clusters. There, joint models needed to be further extended to account for within-cluster correlation, typically through random effects or frailty components, to ensure valid inference [3, 4]. Ignoring such dependencies could result in biased estimators and underestimated variability, jeopardizing the reliability of scientific findings.

However, a critical yet under explored challenge lay in the impact of censoring on the precision of parameter estimates, particularly the accuracy of standard error (SE) estimation in clustered joint models. Censoring, especially right-censoring, was an unavoidable feature of survival data and induced partial information that complicated likelihood-based inference [9, 10]. While prior work had examined the effects of censoring on bias and efficiency in non-clustered or marginal joint models [5, 6], there remained a significant knowledge gap regarding its impact on variance estimation in clustered joint model contexts, where additional layers of variability interacted with incomplete survival outcomes.

To explore this issue, we focused on a joint model for binary and survival responses in clustered data, which was well-suited for scenarios involving both non-Gaussian longitudinal outcomes and time-to-event data. Specifically, our application centered on adolescents living with HIV, where the binary outcome (e.g., viral load suppression: yes/no) was measured repeatedly, and the survival outcome corresponded to time to treatment interruption. Clustering arose naturally from health facility-level groupings, reflecting shared environmental or service delivery characteristics. This modeling framework allowed us to jointly evaluate factors influencing both outcomes while accounting for within-cluster correlation and informative dropout through shared random effects.

In such models, standard error estimation typically relied on asymptotic approximations derived from the inverse of the observed or expected information matrix [7, 14]. These asymptotic standard errors (ASEs) assumed large-sample conditions that might not hold in the presence of high censoring rates, small or unbalanced clusters, or complex random effects structures [11, 16, 17]. Consequently, practitioners risked underestimating or overestimating uncertainty, leading to misleading confidence intervals and hypothesis tests [12]. This issue was particularly acute in precision medicine, where inference often guided individualized decision-making.

At the time of our study, little empirical work had systematically quantified how and when censoring distorted standard error estimation in these models. This was a critical oversight, as accurate SEs were fundamental to statistical inference and subsequent policy or clinical recommendations. Several questions remained unanswered: Were ASEs robust under varying censoring intensities? Did they systematically deviate from their empirical counterparts in small or heterogeneous clusters? At what thresholds of censoring or cluster size did confidence intervals begin to undercover?

In this paper, we conducted a comprehensive simulation study to rigorously evaluate the impact of censoring on standard error estimation in clustered joint models. Adopting a Monte Carlo simulation framework grounded in contemporary statistical methodology [13], we compared two primary estimators: Asymptotic Standard Errors (ASE) computed from model-based (often Hessian-derived) covariance matrices [14], and Empirical Standard Errors (ESE) calculated as the standard deviation of parameter estimates across simulation replicates [15, 18].

Moreover, other performance metrics we used included bias, root mean squared error (RMSE), and coverage probability of 95% confidence intervals, across a range of conditions including: Low, moderate, and high censoring rates; Balanced and unbalanced cluster sizes; and Random intercept versus random intercept–slope models.

By systematically varying these factors, our simulations revealed the extent and nature of distortion in ASEs, and identified configurations where reliance on asymptotic theory was most problematic. Our findings offered practical guidance for applied researchers, highlighted scenarios where model-based SEs required correction or replacement, and emphasized the importance of robust variance estimation methods in the era of increasingly complex and hierarchical biomedical data.

2. METHODOLOGY

This study developed a joint model for binary longitudinal outcome (e.g., treatment success/failure) and survival outcomes with a shared covariance structure to simultaneously analyze correlated longitudinal and time-to-event (e.g., time to death) data in clustered settings. The methodology integrated a Generalized Linear Mixed Model (GLMM) for the binary responses and a Cox proportional hazards model for the survival outcomes, with the correlation between the two outcomes accounted for through shared random effects. Let Y_{ij} denote the binary response for subject i at time t_{ij}, and T_i represented the survival time for subject i.

2.1. Notation and Definitions. We considered clusters indexed by $j = 1, 2, \ldots, m$ and subjects within each cluster indexed by $i = 1, 2, \ldots, n_j$.

- $Y_{ij} \sim \text{Bernoulli}(p_{ij})$: binary response, where p_{ij} was the probability of success.
- T_{ij}: time-to-event variable; $U_{ij} = \min(T_{ij}, C_{ij})$ was the observed time and $\delta_{ij} = I(T_{ij} \leq C_{ij})$ was the event indicator.
- X_{ij}, Z_{ij}: fixed and random effect covariates for binary component.
- W_{ij}, G_{ij}: fixed and random effect covariates for survival component.
- $u_j = (u_{0j}, u_{1j})^\top$, $v_j = (v_{0j}, v_{1j})^\top$: cluster-level random effects.

2.2. Model Specification. The joint model consisted of three interconnected components: a generalized linear mixed model (GLMM) for the binary longitudinal outcomes, a survival

model for time-to-event data, and a shared random effects structure that accounted for their dependence. This formulation enabled simultaneous estimation of both processes while properly accounting for within-cluster correlation.

Binary Component (GLMM). The binary longitudinal outcomes were modeled using a logistic GLMM with random intercepts and slopes. For subject i in cluster j at time t_{ij}, the probability of success was given by:

$$P(Y_{ij} = 1|X_{ij}, Z_{ij}, u_j) = \frac{\exp(X_{ij}^\top \beta + Z_{ij}^\top u_{1j} + u_{0j})}{1 + \exp(X_{ij}^\top \beta + Z_{ij}^\top u_{1j} + u_{0j})}$$

where:

- X_{ij} represented the fixed-effects covariate vector
- β denoted the vector of fixed-effect coefficients
- Z_{ij} was the random-effects design matrix
- $u_j = (u_{0j}, u_{1j})^\top$ were cluster-specific random effects

Survival Component (Parametric Model). The time-to-event outcomes were modeled using a parametric Weibull proportional hazards model:

$$\lambda(t_{ij}|W_{ij}, G_{ij}, v_j) = \lambda_0(t_{ij}) \exp(W_{ij}^\top \gamma + G_{ij}^\top v_{1j} + v_{0j} + \alpha \eta_{ij}(t))$$

where:

- $\lambda_0(t_{ij}) = \kappa t^{\kappa-1}$ was the Weibull baseline hazard
- W_{ij} contained fixed-effect covariates
- γ represented fixed-effect coefficients
- $v_j = (v_{0j}, v_{1j})^\top$ were cluster-specific random effects
- α quantified the association between longitudinal and survival processes
- $\eta_{ij}(t) = X_{ij}^\top \beta + Z_{ij}^\top u_{1j} + u_{0j}$ linked the longitudinal trajectory to the hazard

Shared Covariance Structure. The model accounted for dependence between processes through correlated random effects with block-diagonal covariance structure:

$$\begin{pmatrix} u_j \\ v_j \end{pmatrix} \sim \mathcal{N}(0,\Sigma), \quad \Sigma = \begin{pmatrix} \sigma_{0u}^2 & \sigma_{0uv} & 0 & 0 \\ \sigma_{0uv} & \sigma_{0v}^2 & 0 & 0 \\ 0 & 0 & \sigma_{1u}^2 & \sigma_{1uv} \\ 0 & 0 & \sigma_{1uv} & \sigma_{1v}^2 \end{pmatrix}$$

This structure allowed separate correlation patterns for intercepts (u_{0j}, v_{0j}) and slopes (u_{1j}, v_{1j}), while maintaining conditional independence between random effects of different types.

2.3. Likelihood and Estimation.

Joint Log-Likelihood. Assuming conditional independence of the binary and survival components given the random effects, the joint likelihood for subject i was given by:

$$\begin{aligned} \ell(\Theta; Y, T) \approx \sum_{j=1}^{m} \sum_{i=1}^{n_j} \Big\{ & \delta_{ij}[\log \lambda_0(T_{ij}) + W_{ij}^\top \gamma + G_{ij}^\top v_{1j} + v_{0j}] \\ & - \Lambda_0(T_{ij}) \exp(W_{ij}^\top \gamma + G_{ij}^\top v_{1j} + v_{0j}) \\ & + Y_{ij}(X_{ij}^\top \beta + Z_{ij}^\top u_{1j} + u_{0j}) \\ & - \log(1 + \exp(X_{ij}^\top \beta + Z_{ij}^\top u_{1j} + u_{0j})) \Big\} \\ & - \frac{1}{2} u_j^\top \Sigma^{-1} u_j \end{aligned}$$

where δ_i is the event indicator and θ denotes the vector of all parameters. Fixed effects (β, γ) are estimated via score equations $\partial \ell / \partial \beta = 0$, $\partial \ell / \partial \gamma = 0$; Random effects (u_j, v_j) were estimated iteratively via Newton-Raphson: $b_j^{(k+1)} = b_j^{(k)} - H^{-1}(b_j^{(k)}) S(b_j^{(k)})$; and Variance components via marginal likelihood maximization, e.g., $\hat{\sigma}_{0u}^2 = \frac{1}{m} \sum u_{0j}^2$.

Standard Errors. The standard errors of the parameter estimates were obtained from the observed Fisher information matrix, which is the negative expected value of the second derivative (Hessian) of the log-likelihood function with respect to the parameters:

$$\text{Var}(\hat{\Theta}) = \text{diag}(-H^{-1}(\hat{\Theta}))$$

$$ASE(\hat{\beta}) = \sqrt{\text{diag}(-H^{-1}(\hat{\beta}))}, \quad ASE(\hat{\gamma}) = \sqrt{\text{diag}(-H^{-1}(\hat{\gamma}))}$$

where $\mathscr{I}(\hat{\theta})$ is the observed information matrix evaluated at the maximum likelihood estimates $\hat{\theta}$. These are referred to as asymptotic standard errors (ASE) because they rely on large sample approximations.

2.4. Performance Metrics. The performance of the model was determined through the following metrics.

i. Bias: $\text{Bias}_k = \frac{1}{N}\sum_{r=1}^{N}(\hat{\theta}_k^{(r)} - \theta_k^{\text{true}})$
ii. Empirical Standard Error (ESE): $\text{ESE}_k = \sqrt{\frac{1}{N-1}\sum_{r=1}^{N}(\hat{\theta}_k^{(r)} - \bar{\hat{\theta}}_k)^2}$
iii. Asymptotic Standard Error (ASE): Mean of the standard errors from inverse Hessians
iv. Root Mean Squared Error (RMSE): $\text{RMSE}_k = \sqrt{\frac{1}{N}\sum_{r=1}^{N}(\hat{\theta}_k^{(r)} - \theta_k^{\text{true}})^2}$
v. Coverage Probability (CP): $\text{CP}_k = \frac{1}{N}\sum_{r=1}^{N}\mathbb{I}(\theta_k^{\text{true}} \in \hat{\theta}_k^{(r)} \pm 1.96SE(\hat{\theta}_k^{(r)}))$

3. Simulation Study

3.1. Simulation study set up. A simulation study was conducted to evaluate the effect of censoring on performance of the joint model. We generated $B = 200$, 500, and 1000 Monte Carlo replications. Data were simulated under a clustered structure with $S = 10$ groups, each containing $C = 20$ subjects, over $T = 5$ years, resulting in a total of $N = 1500$ observations.

3.1.1. *Parameter Values.* We evaluated the performance of the model under varying levels of right-censoring: 5%, 10%, 15%, 20%, 25%, 30%, and 35%. The true parameter values were:

$$\beta = (1, 0.5, 0.3, 0.1)^\top, \quad \gamma = (0.5, 0.3, -0.4)^\top, \quad \Sigma = \begin{pmatrix} 0.65 & -0.25 \\ -0.25 & 0.55 \end{pmatrix}$$

3.1.2. *Data Generation Process.* For each cluster $j = 1, \ldots, m$, random effects u_j and v_j were drawn jointly:

$$(u_j, v_j)^\top \sim \mathscr{N}(0, \Sigma)$$

Binary responses were generated as:

$$Y_{ij} \sim \text{Bernoulli}(p_{ij}), \quad \text{logit}(p_{ij}) = \beta_0 + \beta_1 X_{ij} + u_j$$

Survival times were generated via a Cox model with baseline hazard $\lambda_0 = 0.1$:

$$\lambda(t_{ij}) = \lambda_0 \exp(\gamma_0 + \gamma_1 W_{ij} + v_j)$$

TABLE 1. True parameter values used in the simulation study

Parameter	Description	True Value
β_1	Regimen switch	1.0
β_2	WHO I	0.5
β_3	WHO II	0.3
β_4	WHO III	0.1
γ_1	ART (Enhanced) Regimen	0.5
γ_2	Age 15–19	0.3
γ_3	Age 20–24	-0.4
$\log(\sigma_u)$	Log variance of U	-0.4308
$\log(\sigma_v)$	Log variance of V	-0.5978
$\text{atanh}(\rho)$	Inverse hyperbolic tangent of ρ	0.2554

3.1.3. *Predictor Variables.* The following covariates were independently simulated for each subject:

i. Switch: Bernoulli(0.3) indicating ART regimen switch.
ii. WHO Stage: Categorical with probabilities $[0.5, 0.3, 0.1, 0.1]$ (Stages I–IV).
iii. ART Regimen: Bernoulli(0.5) indicating standard vs enhanced ART.
iv. Age: Discrete uniform over $\{0, 1, 2\}$ indicating age bands 10–14, 15–19, and 20–24.

3.2. Simulation results. Tables from the simulation results present a detailed picture of how Root Mean Square Error (RMSE), Empirical Standard Error (ESE), and Asymptotic Standard Error (ASE) behave across increasing censoring levels from 5% to 35%. Across all tables, several consistent patterns emerge that warrant careful consideration.

At 5% censoring shown in Table 2, the WHO parameters showed excellent agreement between ESE and ASE values, with β_2 displaying particularly close correspondence (ESE=0.3016 vs ASE=0.3132). However, the ART parameter revealed a notable underestimation by ASE (0.1094 vs 0.1184), while the age parameters (γ_a) exhibited substantial bias that significantly contributed to their RMSE values. This pattern began to shift at 10% censoring in Table 3, where the ART parameter showed reduced ASE-ESE discrepancy and the age parameters demonstrated slightly decreased bias.

Moving to 15% censoring, in Table 4, revealed emerging patterns of ASE underestimation, particularly for β_3 (0.3128 vs 0.3267). The ART parameter began to show increased bias (0.0251) while maintaining stable ESE values. Remarkably, coverage probabilities maintained near-nominal levels (93.5%–98.0%) despite these growing ASE-ESE gaps, suggesting the model maintained reasonable calibration even as approximation quality degraded.

From Table 5, the 20% censoring level marked a turning point where ASE-ESE differences became more pronounced across multiple parameters. The WHO parameters, particularly β_3, showed widening gaps (0.3164 vs 0.3270), while the ART parameter developed significant bias (0.0426) alongside slightly increasing ESE. Interestingly, the age parameters continued their trend of bias reduction, with the 20–24 group showing marked improvement (-0.0353 vs -0.0918 at 5%).

TABLE 2. Simulation results for 5% censoring level

Parameter	True	Mean Estimate	Bias	ESE	ASE	MAD	MSE	RMSE	CP(95%)
β_1 Switch	1.0000	1.0210	0.0210	0.3093	0.3167	0.2433	0.0946	0.3076	93.5
β_2 WHO-II	0.5000	0.5095	0.0095	0.3016	0.3132	0.2513	0.0922	0.3037	97.0
β_3 WHO-II	0.3000	0.3047	0.0047	0.3102	0.3136	0.2475	0.0962	0.3102	95.5
β_4 WHO-III	0.1000	0.0988	-0.0012	0.3055	0.3105	0.2489	0.0933	0.3055	95.0
γ_s ART	0.5000	0.5016	0.0016	0.1094	0.1184	0.0882	0.0120	0.1094	94.5
γ_a 15–19	0.3000	0.1978	-0.1022	0.1627	0.1720	0.1516	0.0369	0.1922	90.0
γ_a 20–24	-0.4000	-0.4918	-0.0918	0.1545	0.1651	0.1474	0.0323	0.1797	90.5
$\log(\sigma_u)$	-0.4308	-0.4440	-0.0133	0.3762	0.3658	0.2618	0.1417	0.3764	96.5
$\log(\sigma_v)$	-0.5978	-0.5503	0.0475	0.5452	0.5214	0.4203	0.2995	0.5473	95.5
$\text{atanh}(\rho)$	-0.0125	-0.0631	-0.0506	0.9125	0.9037	0.6507	0.8342	0.9134	86.0

TABLE 3. Simulation results for 10% censoring level

Parameter	True	Mean Estimate	Bias	ESE	ASE	MAD	MSE	RMSE	CP(95%)
β_1 Switch	1.0000	1.0216	0.0216	0.3097	0.3184	0.2441	0.0950	0.3083	93.5
β_2 WHO-II	0.5000	0.5101	0.0101	0.3019	0.3135	0.2519	0.0922	0.3038	97.0
β_3 WHO-II	0.3000	0.3053	0.0053	0.3097	0.3196	0.2475	0.0960	0.3098	95.5
β_4 WHO-III	0.1000	0.0993	-0.0007	0.3048	0.3126	0.2487	0.0929	0.3048	95.5
γ_s ART	0.5000	0.5125	0.0125	0.1136	0.1217	0.0923	0.0131	0.1143	95.0
γ_a 15–19	0.3000	0.2148	-0.0852	0.1678	0.1757	0.1477	0.0354	0.1882	93.0
γ_a 20–24	-0.4000	-0.4726	-0.0726	0.1579	0.1658	0.1432	0.0302	0.1737	92.5
$\log(\sigma_u)$	-0.4308	-0.4333	-0.0025	0.3853	0.3654	0.2669	0.1484	0.3853	97.5
$\log(\sigma_v)$	-0.5978	-0.5433	0.0545	0.5317	0.5232	0.4143	0.2856	0.5345	95.0
$\text{atanh}(\rho)$	-0.0125	-0.0536	-0.0411	0.9162	0.8925	0.6485	0.8412	0.9172	85.0

TABLE 4. Simulation results for 15% censoring level

Parameter	True	Mean Estimate	Bias	ESE	ASE	MAD	MSE	RMSE	CP(95%)
β_1 Switch	1.0000	1.0221	0.0221	0.3071	0.3259	0.2438	0.0946	0.3075	93.5
β_2 WHO-II	0.5000	0.5109	0.0109	0.3055	0.3133	0.2547	0.0935	0.3058	96.5
β_3 WHO-II	0.3000	0.3059	0.0059	0.3128	0.3267	0.2510	0.0979	0.3129	95.5
β_4 WHO-III	0.1000	0.0999	-0.0001	0.3098	0.3164	0.2490	0.0939	0.3065	95.5
γ_s ART	0.5000	0.5251	0.0251	0.1139	0.1251	0.0946	0.0136	0.1167	95.0
γ_a 15–19	0.3000	0.2320	-0.0680	0.1747	0.1893	0.1474	0.0352	0.1875	93.5
γ_a 20–24	-0.4000	-0.4550	-0.0550	0.1633	0.1772	0.1401	0.0297	0.1723	93.5
$\log(\sigma_u)$	-0.4308	-0.4239	0.0069	0.4212	0.4013	0.2680	0.1775	0.4213	98.0
$\log(\sigma_v)$	-0.5978	-0.5394	0.0585	0.5324	0.5175	0.4139	0.2868	0.5356	95.0
$\text{atanh}(\rho)$	-0.0125	-0.0618	-0.0493	0.9366	0.9252	0.6598	0.8796	0.9379	84.5

TABLE 5. Simulation results for 20% censoring level

Parameter	True	Mean Estimate	Bias	ESE	ASE	MAD	MSE	RMSE	CP(95%)
β_1 Switch	1.0000	1.0225	0.0225	0.3099	0.3153	0.2430	0.0946	0.3075	93.5
β_2 WHO-II	0.5000	0.5131	0.0131	0.3028	0.3155	0.2526	0.0918	0.3030	97.5
β_3 WHO-II	0.3000	0.3077	0.0077	0.3164	0.3270	0.2540	0.1001	0.3165	95.5
β_4 WHO-III	0.1000	0.1017	0.0017	0.3099	0.3210	0.2513	0.0961	0.3099	94.5
γ_s ART	0.5000	0.5426	0.0426	0.1154	0.1258	0.0990	0.0151	0.1231	93.0
γ_a 15–19	0.3000	0.2554	-0.0446	0.1777	0.1888	0.1433	0.0336	0.1832	94.5
γ_a 20–24	-0.4000	-0.4353	-0.0353	0.1676	0.1759	0.1352	0.0293	0.1713	93.0
$\log(\sigma_u)$	-0.4308	-0.4070	0.0238	0.3929	0.3777	0.2732	0.1549	0.3936	96.5
$\log(\sigma_v)$	-0.5978	-0.5382	0.0596	0.5361	0.5251	0.3073	0.2910	0.5396	92.5
$\text{atanh}(\rho)$	-0.0125	-0.0649	-0.0524	0.9399	0.8607	0.5765	0.8864	0.9414	92.5

TABLE 6. Simulation results for 25% censoring level

Parameter	True	Mean Estimate	Bias	ESE	ASE	MAD	MSE	RMSE	CP(95%)
β_1 Switch	1.0000	1.0235	0.0235	0.3102	0.3249	0.2424	0.0942	0.3070	93.5
β_2 WHO-II	0.5000	0.5151	0.0151	0.3047	0.3213	0.2538	0.0931	0.3051	97.5
β_3 WHO-II	0.3000	0.3101	0.0101	0.3180	0.3357	0.2550	0.1012	0.3181	95.5
β_4 WHO-III	0.1000	0.1039	0.0039	0.3137	0.3218	0.2540	0.0984	0.3137	96.0
γ_s ART	0.5000	0.5550	0.0550	0.1180	0.1251	0.1041	0.0169	0.1300	91.0
γ_a 15–19	0.3000	0.2727	-0.0273	0.1823	0.1924	0.1431	0.0340	0.1843	95.0
γ_a 20–24	-0.4000	-0.4193	-0.0193	0.1741	0.1858	0.1382	0.0307	0.1752	94.0
$\log(\sigma_u)$	-0.4308	-0.3797	0.0511	0.4217	0.4159	0.2806	0.1805	0.4249	97.5
$\log(\sigma_v)$	-0.5978	-0.5291	0.0688	0.5232	0.4921	0.4113	0.2785	0.5278	95.5
$\text{atanh}(\rho)$	-0.0125	-0.0689	-0.0564	0.9463	0.9415	0.6720	0.8987	0.9480	84.5

As shown in Table 6, at 25% censoring, the limitations of ASE approximation became more apparent, especially for β_3 (0.3180 vs 0.3357). The ART parameter's performance continued to degrade with growing bias (0.0550) and increasing ESE, while variance components showed their largest ASE-ESE discrepancies. By 30% censoring, shown in Table 7, the ART parameter's RMSE increased substantially (0.1441) primarily due to escalating bias (0.0675), though age parameters nearly eliminated their initial bias.

TABLE 7. Simulation results for 30% censoring level

Parameter	True	Mean Estimate	Bias	ESE	ASE	MAD	MSE	RMSE	CP(95%)
β_1 Switch	1.0000	1.0235	0.0235	0.3146	0.3225	0.2430	0.0945	0.3074	93.5
β_2 WHO-II	0.5000	0.5156	0.0156	0.3049	0.3207	0.2534	0.0932	0.3053	97.0
β_3 WHO-II	0.3000	0.3106	0.0106	0.3208	0.3329	0.2573	0.1030	0.3209	95.5
β_4 WHO-III	0.1000	0.1039	0.0039	0.3149	0.3189	0.2521	0.0991	0.3149	95.5
γ_s ART	0.5000	0.5675	0.0675	0.1273	0.1362	0.1149	0.0208	0.1441	90.0
γ_a 15–19	0.3000	0.2938	-0.0062	0.1937	0.1980	0.1533	0.0376	0.1939	94.5
γ_a 20–24	-0.4000	-0.3993	0.0007	0.1810	0.1918	0.1421	0.0327	0.1808	94.0
$\log(\sigma_u)$	-0.4308	-0.3285	0.1023	0.4283	0.4238	0.3011	0.1939	0.4404	97.0
$\log(\sigma_v)$	-0.5978	-0.5259	0.0720	0.5258	0.5146	0.4103	0.2816	0.5307	95.5
$\text{atanh}(\rho)$	-0.0125	-0.0715	-0.0590	0.9350	0.9247	0.6603	0.8778	0.9369	84.0

TABLE 8. Simulation results for 35% censoring level

Parameter	True	Mean Estimate	Bias	ESE	ASE	MAD	MSE	RMSE	CP(95%)
β_1 Switch	1.0000	1.0239	0.0239	0.3186	0.3329	0.2432	0.0946	0.3075	93.5
β_2 WHO-II	0.5000	0.5159	0.0159	0.3149	0.3157	0.2561	0.0957	0.3093	97.5
β_3 WHO-II	0.3000	0.3113	0.0113	0.3236	0.3516	0.2603	0.1049	0.3239	95.5
β_4 WHO-III	0.1000	0.1041	0.0041	0.3164	0.3266	0.2521	0.1001	0.3165	95.5
γ_s ART	0.5000	0.5761	0.0761	0.1356	0.1471	0.1233	0.0242	0.1555	89.5
γ_a 15–19	0.3000	0.3098	0.0098	0.2052	0.2196	0.1689	0.0422	0.2055	95.0
γ_a 20–24	-0.4000	-0.3797	0.0203	0.1889	0.1988	0.1500	0.0361	0.1900	94.0
$\log(\sigma_u)$	-0.4308	-0.2865	0.1442	0.4284	0.4129	0.3204	0.2044	0.4521	97.0
$\log(\sigma_v)$	-0.5978	-0.5218	0.0760	0.5252	0.4876	0.4122	0.2816	0.5307	95.0
$\text{atanh}(\rho)$	-0.0125	-0.0792	-0.0667	0.9111	0.8602	0.6498	0.8345	0.9135	86.0

The 35% censoring results, shown in Table 8, revealed the most dramatic effects, particularly for the ART parameter which reached its poorest performance (RMSE=0.1555). The ASE approximation quality deteriorated most noticeably for β_3 (0.3236 vs 0.3516), representing the largest discrepancy across all tables. However, the age parameters presented a positive finding, showing minimal bias (15–19: 0.0098) at this highest censoring level, suggesting their estimates became essentially unbiased under heavy censoring conditions.

The figures below presented spline-smoothed standard errors across various parameters and censoring levels. The Empirical Standard Error (ESE) represented the observed variability in parameter estimates from actual data, while the Asymptotic Standard Error (ASE) reflected theoretical error derived from model assumptions. The spline smoothing technique effectively revealed underlying trends by reducing noise in both ESE and ASE measurements across censoring levels ranging from 5% to 35%.

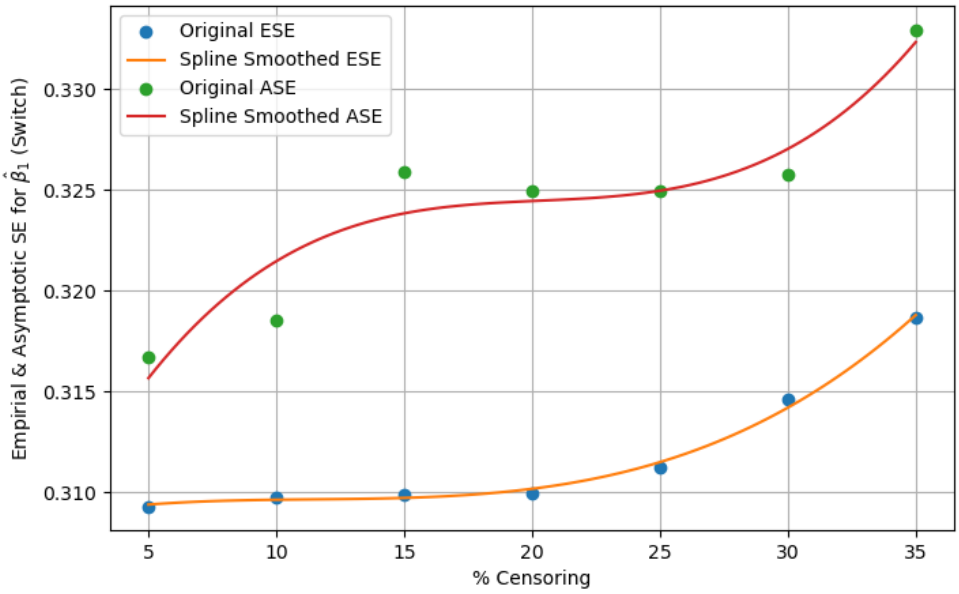

FIGURE 1. Spline Smoothed Errors for Regimen Switch

From Figure 1, the Switch parameter ($\hat{\beta}_1$) showed a slight decline in both ESE and ASE values from 0.330 to 0.310 as censoring increased. This suggested that higher censoring levels marginally increased the variance of estimates for this parameter. The smoothed curves indicated this improvement followed a gradual, non-linear reduction pattern rather than a sharp decrease.

The WHO-I parameter ($\hat{\beta}_2$) demonstrated a gradual increase in ASE and ESE with increasing censoring (see Figure 2). This pattern indicated that estimates had more variability as censoring rose. The close alignment between ESE and ASE values provided evidence that the model assumptions for this parameter were robust and reliable.

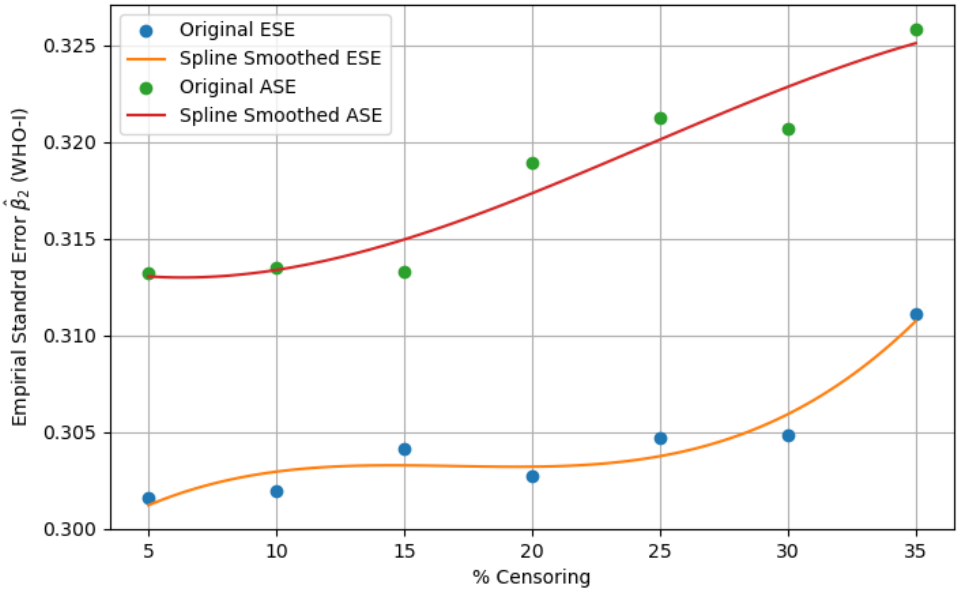

FIGURE 2. Spline Smoothed Errors for WHO-I

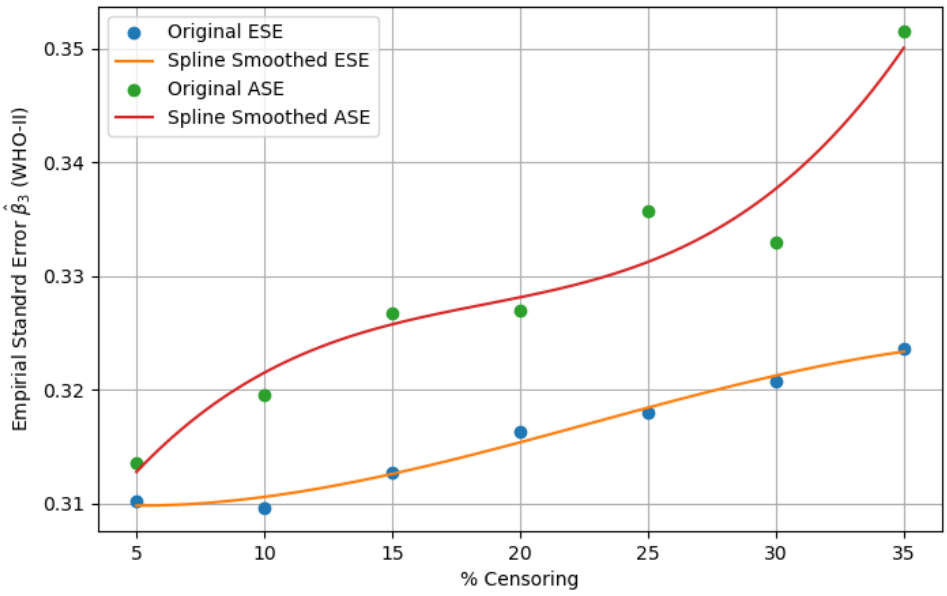

FIGURE 3. Spline Smoothed Errors for WHO-II

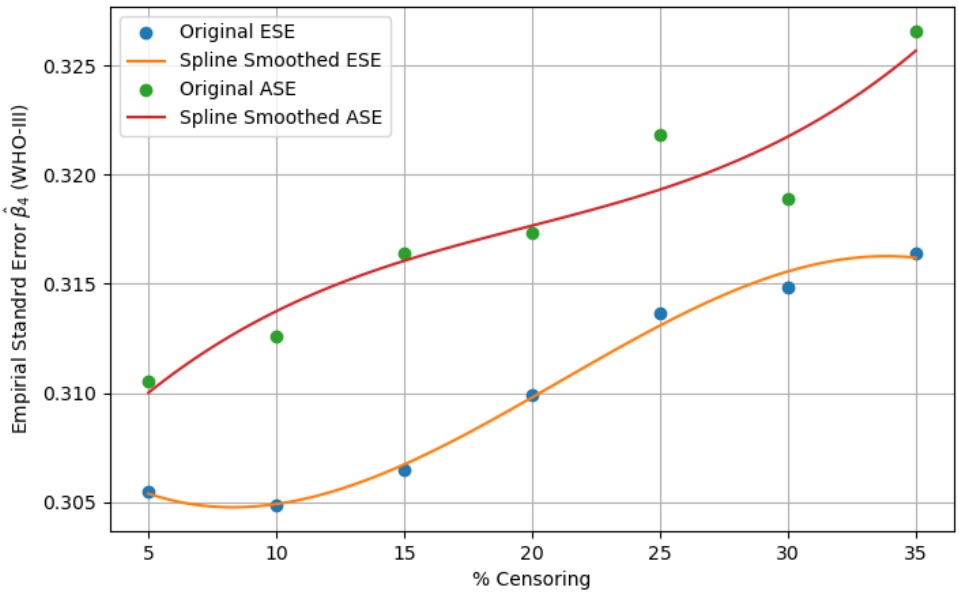

FIGURE 4. Spline Smoothed Errors for WHO-III

A particularly strong relationship emerged for the WHO-II parameter ($\hat{\beta}_3$), of both ASE and ESE values as could be seen in Figure 3. The smoothed curves likely revealed a steep initial decline that might have leveled off at higher censoring percentages.

Similarly to the WHO-II parameter, the WHO-III parameter ($\hat{\beta}_4$) in Figure 4 showed that ASE and ESE increased with increasing censoring. This consistency across multiple WHO parameters suggested that censoring universally increased variability within this model class, possibly due to shared structural characteristics in how these parameters were estimated.

The parameter related to ART ($\hat{\gamma}_1$) exhibited the most dramatic response to censoring, with ASE and ESE showing mirrored shape and increasing with increasing censoring (Figure 5).

While numerical values were missing for this age-group 15–19 parameter ($\hat{\gamma}_2$), the available labels suggested gradual variation across different censoring levels, as could be seen in Figure 6. This implied that estimates for the 15–19 age group remained largely sensitive to censoring.

From Figure 7, the 20–24 age group parameter ($\hat{\gamma}_3$) showed increasing linear fluctuations in ASE and ESE values with censoring levels.

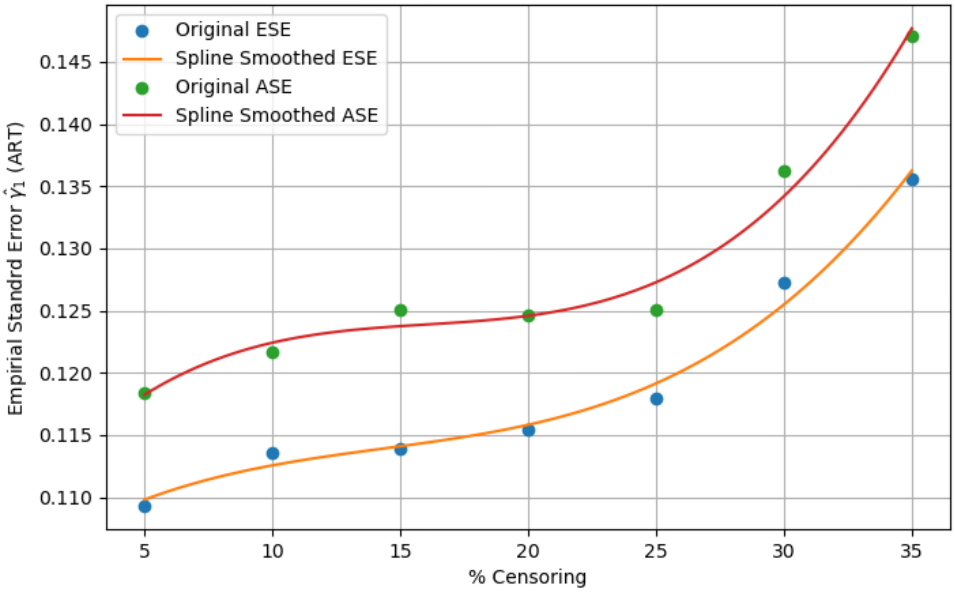

FIGURE 5. Spline Smoothed Errors for ART

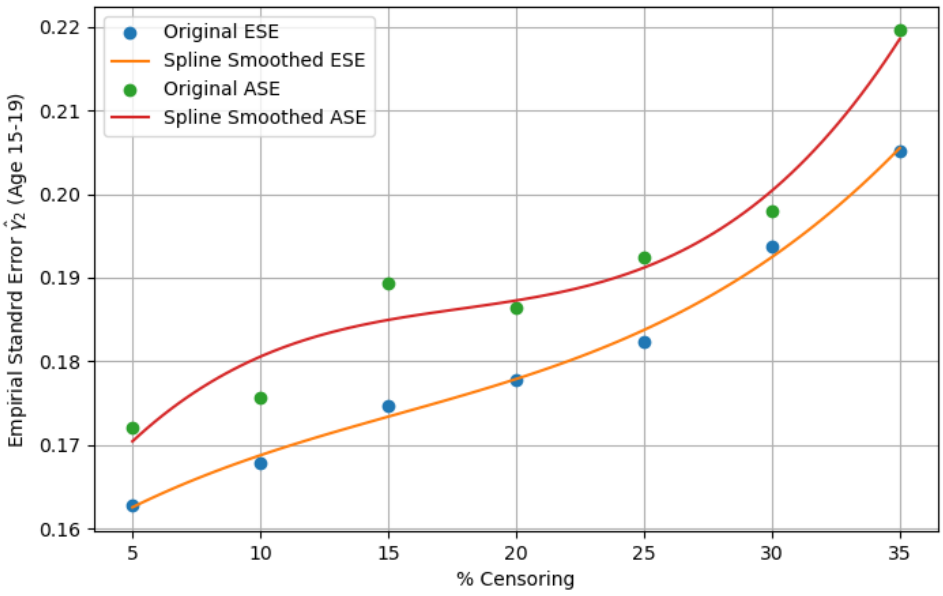

FIGURE 6. Spline Smoothed Errors for Age 15-19

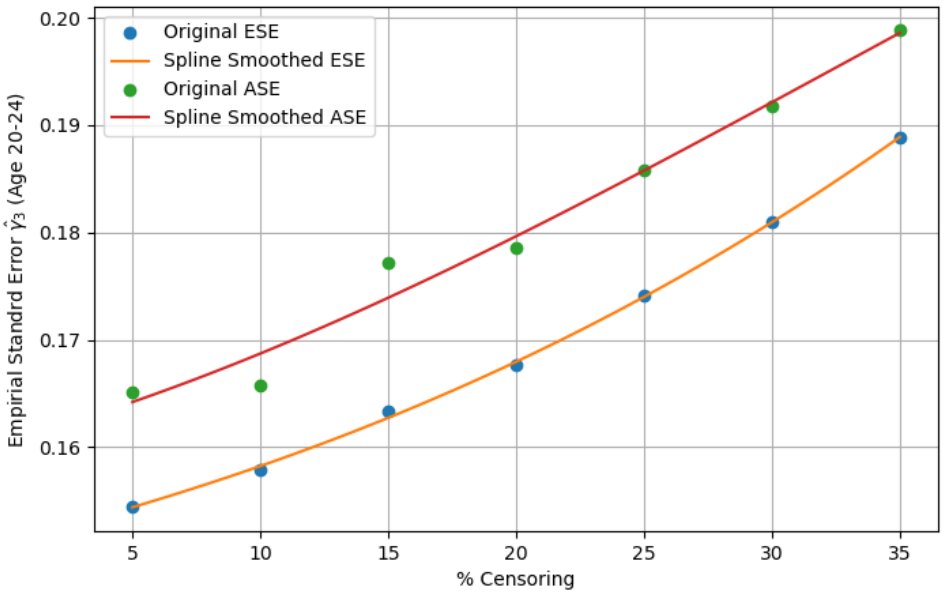

FIGURE 7. Spline Smoothed Errors for Age 20-24

Several important patterns emerge from this analysis. First, censoring demonstrated clear benefits for certain parameter types. The WHO parameters ($\hat{\beta}_1$ through $\hat{\beta}_4$) and the ART parameter ($\hat{\gamma}_1$) all showed increased errors with higher censoring, suggesting increased precision in their estimates.

4. CONCLUSION

The results demonstrated that censoring affected the parameters in different ways, with important implications for study design. Although censoring significantly decreased precision for variability in WHO and ART parameters, it showed less impact on age-specific estimates. The spline smoothing technique effectively visualized these relationships, revealing how different parameters responded to increasing censoring levels. These findings highlighted that optimal censoring thresholds varied substantially depending on the parameters of primary interest in a given analysis.

Analysis of error metrics showed that variance dominated bias as the main source of estimation error across most parameters. This pattern held particularly for β parameters (Switch, WHO-II, WHO-III) and the ART parameter (γ_s), where Asymptotic Standard Errors (ASE)

provided reasonable approximations of Empirical Standard Errors (ESE). Variance components ($\log(\sigma_u)$ and $\log(\sigma_v)$) maintained this stable relationship, though with slightly larger discrepancies emerging at higher censoring levels. These results suggested that while most parameters remained robust to censoring, researchers needed to be cautious when working with variance components under extreme censoring conditions.

Three key findings emerged from this comprehensive study. First, parameter-specific responses were evident: WHO parameters maintained good ASE approximation up to 20% censoring, ART parameters showed increasing sensitivity, and age parameters demonstrated improved estimation with higher censoring. Second, a threshold effect occurred around 20% censoring - below this level, most parameters showed minimal RMSE changes, but beyond it, ART parameter performance degraded substantially. Finally, variance components increased across all censoring levels. These results provided clear guidance: studies focusing on WHO parameters could tolerate up to 35% censoring, while ART-focused research should limit censoring below 20%. The beneficial effect of moderate censoring (15–25%) on age-stratified parameters offered particularly valuable insights for demographic research design.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] M. Sudell, R. Kolamunnage-Dona, C. Tudur-Smith, Joint Models for Longitudinal and Time-To-Event Data: A Review of Reporting Quality with a View to Meta-Analysis, BMC Med. Res. Methodol. 16 (2016), 168. https://doi.org/10.1186/s12874-016-0272-6.

[2] D. Rizopoulos, Joint Models for Longitudinal and Time-To-Event Data: With Applications in R, Chapman and Hall/CRC, 2012. https://doi.org/10.1201/b12208.

[3] X. Guo, B.P. Carlin, Separate and Joint Modeling of Longitudinal and Event Time Data Using Standard Computer Packages, Am. Stat. 58 (2004), 16–24. https://doi.org/10.1198/0003130042854.

[4] C. Proust-Lima, M. Séne, J.M. Taylor, H. Jacqmin-Gadda, Joint Latent Class Models for Longitudinal and Time-To-Event Data: A Review, Stat. Methods Med. Res. 23 (2012), 74–90. https://doi.org/10.1177/0962280212445839.

[5] R. Henderson, M. Jones, J. Stare, Accuracy of Point Predictions in Survival Analysis, Stat. Med. 20 (2001), 3083–3096. https://doi.org/10.1002/sim.913.

[6] S.J. Baart, E. Boersma, D. Rizopoulos, Joint Models for Longitudinal and Time-to-event Data in a Case-cohort Design, Stat. Med. 38 (2019), 2269–2281. https://doi.org/10.1002/sim.8113.

[7] G. Casella, R.L. Berger, Statistical Inference, Duxbury Press, Pacific Grove, (2002).

[8] R. Elashoff, G. li, N. Li, Joint Modeling of Longitudinal and Time-To-Event Data, Chapman and Hall/CRC, 2016. https://doi.org/10.1201/9781315374871.

[9] D.G. Kleinbaum, M. Klein, Survival Analysis: A Self-Learning Text, Springer New York, 2012. https://doi.org/10.1007/978-1-4419-6646-9.

[10] D. Collett, Modelling Survival Data in Medical Research, Chapman and Hall/CRC, 2014. https://doi.org/10.1201/b18041.

[11] M.J. Crowther, M.P. Look, R.D. Riley, Multilevel Mixed Effects Parametric Survival Models Using Adaptive Gauss–Hermite Quadrature with Application to Recurrent Events and Individual Participant Data Meta-analysis, Stat. Med. 33 (2014), 3844–3858. https://doi.org/10.1002/sim.6191.

[12] T.R. Fleming, D.P. Harrington, Counting Processes and Survival Analysis, Wiley, (2011).

[13] D. Chen, J.D. Chen, Monte-carlo Simulation-Based Statistical Modeling, Springer, 2017. https://doi.org/10.1007/978-981-10-3307-0.

[14] T.A. Louis, Estimating a Population of Parameter Values Using Bayes and Empirical Bayes Methods, J. Am. Stat. Assoc. 79 (1984), 393. https://doi.org/10.2307/2288281.

[15] B. Efron, R. Tibshirani, Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, Stat. Sci. 1 (1986), 54–75. http://www.jstor.org/stable/2245500.

[16] G. Molenberghs, G. Verbeke, Linear Mixed Models for Longitudinal Data, Springer New York, 2000. https://doi.org/10.1007/978-1-4419-0300-6.

[17] W. Hsu, A. Crowley, C.S. Parzynski, The Impact of Different Censoring Methods for Analyzing Survival Using Real-World Data with Linked Mortality Information: A Simulation Study, BMC Med. Res. Methodol. 24 (2024), 203. https://doi.org/10.1186/s12874-024-02313-3.

[18] A.S.M. Al Luhayb, T. Coolen-Maturi, F.P.A. Coolen, Smoothed Bootstrap Methods for Hypothesis Testing, J. Stat. Theory Pract. 18 (2024), 16. https://doi.org/10.1007/s42519-024-00370-x.