Demystifying Clone-Censor-Weight Method in Target Trial Emulation: A Real-World Study of HPV Vaccination Strategies

Tuo Lin[1*], Yongqiu Li[2], Zhuochao Huang[3], Toni T. Gui[1], Wei (Vivian) Wang[4], Yi Guo[2*]

[1]Department of Biostatistics,

University of Florida, Gainesville, FL

[2]Department of Health Outcomes and Biomedical Informatics,

University of Florida, Gainesville, FL

[3]Department of Statistics,

University of Florida, Gainesville, FL

[4]Merck & Co., Inc., Rahway, NJ

[*]Corresponding author: Tuo Lin, tuolin@ufl.edu & Yi Guo, yiguo@ufl.edu

April 21, 2026

Abstract

Target trial emulation (TTE) offers a principled way to estimate treatment effects using real-world observational data, but analyses of time-varying treatment strategies remain vulnerable to immortal time bias. The clone-censor-weight (CCW) approach is increasingly used to address this problem, yet key aspects of its causal interpretation and implementation remain unclear. In this work, we emulate a target trial using electronic health records (EHRs) to compare completion of a 3-dose 9-valent human papillomavirus vaccination (HPV) series within 12 months versus remaining partially vaccinated among vaccine initiators. We link CCW to the classic potential outcome framework in causal inference, evaluate the role of different weighting mechanisms, and account for within-subject correlation induced by cloning using cluster-robust variance estimation. Our study provides practical guidance for applying CCW in real-world comparative effectiveness studies to address immortal time bias and supports more rigorous and interpretable treatment effect estimation in TTE.

1 Introduction

The rapid digitization of healthcare systems and advances in data storage technologies have led to the generation of massive volumes of longitudinal real-world clinical data from electronic health records (EHRs), medical insurance claims, and other large-scale databases [56, 21, 66]. These data provide unprecedented opportunities to evaluate treatment effectiveness in routine clinical practice and to generate evidence that may not be feasible or fully captured in randomized control trials (RCTs) [1, 32]. However, drawing credible causal conclusions from observational healthcare data remains methodologically challenging, particularly in the presence of time-varying treatments, confounding, and predefined grace periods [30, 51].

Target trial emulation (TTE) provides a principled framework for emulating the design and causal interpretation of an RCT using EHRs when conducting an actual trial is not feasible [28, 13, 18]. Since these data are observational, robust causal inference methods are required to estimate treatment effects and properly quantify the uncertainty [12, 27]. For survival outcomes, conventional approaches such as propensity score weighted analysis can lead to immortal time bias when comparing treatment strategies that initiate at different times during follow-up [60, 68, 40]. Simply put, patients must stay event-free long enough to receive later-stage treatments, which preferentially selects for longer event-free and can artificially inflate estimated treatment effects.

To address this problem, Hernán proposed the clone-censor-weight (CCW) framework [24]. In this framework, each individual is cloned and assigned to all treatment strategies at baseline; clones are then artificially censored when their observed treatment history deviates from the assigned strategy; and inverse probability of censoring weights (IPCW) are applied to account for this artificial censoring. By combining cloning, censoring, and weighting, CCW provides an implementable solution to immortal time bias and supports unbiased estimation of per-protocol treatment effects under time-varying treatment strategies in TTE, as demonstrated across a wide range of published applications [69, 57, 69, 4].

Despite its growing use in pharmacoepidemiology and comparative effectiveness research, several methodological questions of the CCW approach remain unresolved. While Hernán and colleagues have provided narrative guidance on implementation [24, 20, 19, 43], many investigators, including our clinical and statistical collaborators, remain puzzled by the underlying statistical principles,

and, in some cases, have applied the method incorrectly. In particular, the connection between CCW and the classic potential outcome framework in causal inference has not been systematically explored [53, 59, 34], leading to confusion about whether to incorporate inverse probability of treatment weights (IPTW) in addition to IPCW [25, 38, 50], and how to account for correlations among cloned copies from the same individual [71, 41, 39].

In this study, we leverage a real-world example to clarify and strengthen the methodological understanding of the CCW approach within the TTE framework. Specifically, persistent high-risk HPV infection can progress to precancerous cervical lesions and ultimately invasive cervical cancer [45, 5, 54, 37, 9]. We investigate the comparative effectiveness of 9-valent human papillomavirus (9vHPV) vaccine dosage strategies using EHR data [61, 35, 33]. Our primary objective was to estimate the causal effect of completing a 3-dose 9vHPV within 12 months versus remaining partially vaccinated (receiving 1–2 doses or completing 3 doses after 12 months) on the risk of incident composite HPV-related cervical disease (persistent hrHPV infection, CIN2+, or cervical cancer) among 9vHPV initiators. The proposed analytic framework explicitly emulates a target trial comparing longitudinal vaccination strategies, mitigates immortal time bias arising from time-varying vaccine initiation, and links the CCW procedures to the classical potential outcomes framework to demystify its causal interpretation.

Empirically, we evaluate the performance of CCW in this applied setting by examining covariate balance, weight stability, and sensitivity to model specification. We compare CCW-based estimates with those obtained from naive analyses to quantify the magnitude of bias and the potential for misleading conclusion when immortal time is not properly addressed. Additionally, we investigate key implementation considerations, including whether additional IPTW is necessary and the use of appropriate cluster-robust sandwich variance estimator to account for within-individual correlation induced by cloning.

By grounding the theoretical foundation in a clinically relevant HPV vaccination study, this work provides a transparent and reproducible roadmap for implementing CCW in real-world data analyses [65, 3]. Our study clarifies common misconceptions in CCW implementation, formalizes the statistical underpinnings of the approach, and offers practical guidance for investigators conducting TTE using longitudinal EHRs. Together, these contributions advance the rigor and interpretability of causal effect estimation in large-scale digital health studies, supporting more reliable real-world

evidence generation.

2 Results

2.1 Cohort Selection and Baseline Characteristics

Our study cohort was constructed using EHR data from the UF Health Integrated Data Repository (IDR), with records spanning January 1, 2012 to February 28, 2025. After applying prespecified eligibility criteria, 3,905 individuals were included in the final analytic cohort (Figure 1). Among these, 924 (23.7%) completed the 3-dose vaccination within 12 months, and 2981 (76.3%) remained partially vaccinated.

Baseline characteristics of the study population are summarized in Table 1. Several characteristics differed between groups: individuals who completed the vaccination series were older at baseline, whereas the partially vaccinated group had a higher proportion of Non-Hispanic Black participants and more recent first vaccination dates (all $p < 0.001$). Visit frequency did not differ significantly between groups. Overall, these findings indicate notable baseline imbalances between vaccination strategies prior to adjustment.

2.2 CCW Mitigates Immortal Time Bias

We evaluated CCW performance in the analytic cohort. The weighted Kaplan–Meier (KM) curves constructed from the cloned and weighted data showed a faster decline in survival probability among partially vaccinated individuals compared with those completed the 3-dose series (Figure 2). In contrast, without applying the CCW approach, the raw KM curves suggested the opposite conclusion: survival probability declined faster among individual who completed the 3-dose series than among those who were partially vaccinated. This finding was consistent with the naive Cox model fitted to the observed (uncloned, unweighted) data, which yielded a hazard ratio (HR) of 1.85 (95% CI: 0.95–3.57) (Table 2), with the partially vaccinated group as reference, thereby misleadingly suggesting a higher hazard among fully vaccinated individuals. The apparent discrepancy between the naive Cox estimate and the weighted KM curves reflects immortal time bias in the unadjusted analysis.

Investigations of unstabilized IPCW showed right-skewed distributions (Figure 3), with most

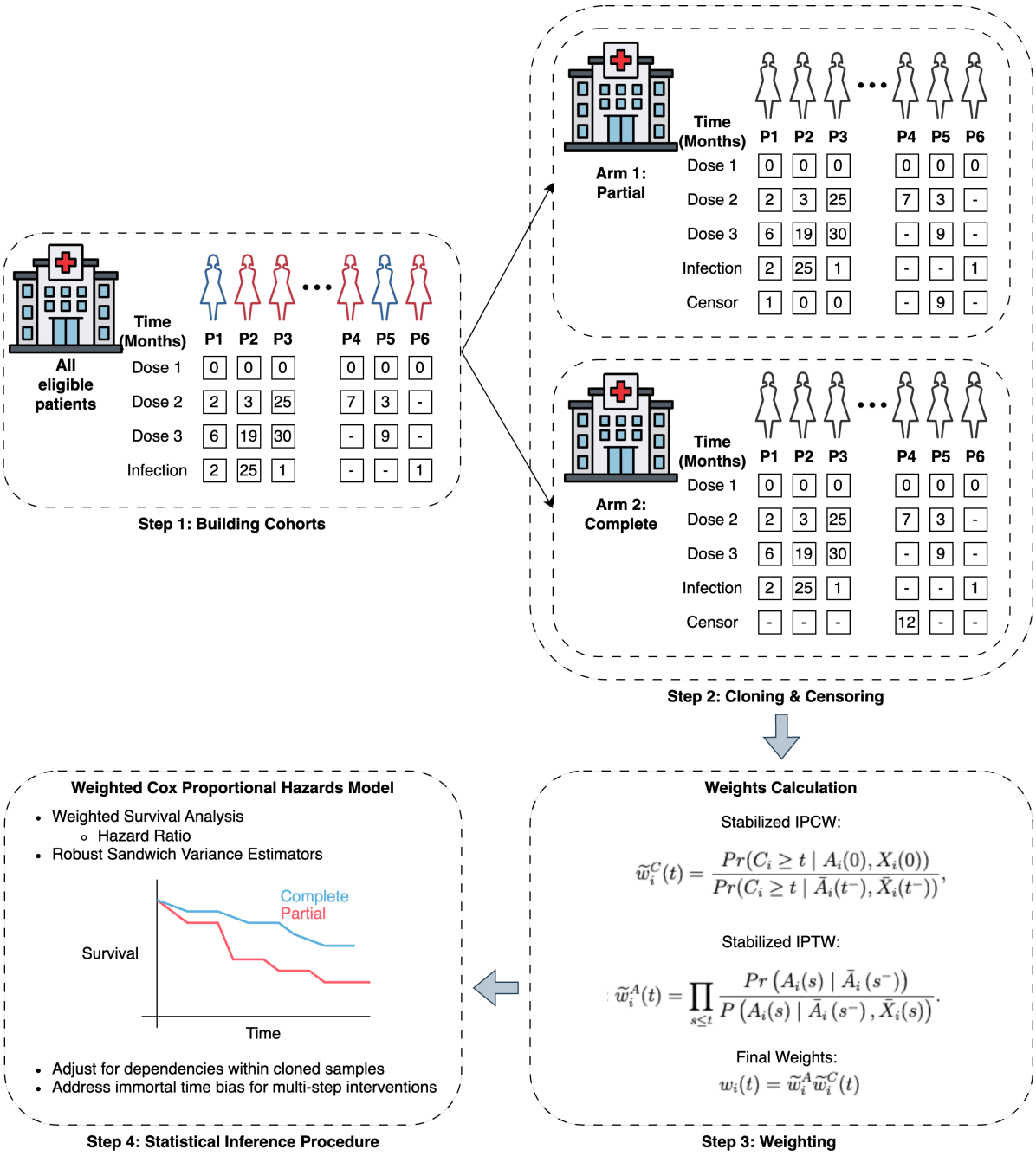

Figure 1: Overview of CCW framework in TTE.

Table 1: Baseline characteristics of the study population

Characteristic	Complete (N=923)	Partial (N=2979)	Overall (N=3902)	P-value
Age [Mean (SD)]	21.0 (3.3)	19.3 (3.3)	19.7 (3.3)	< 0.001
Race/ethnicity				< 0.001
Hispanic	122 (13.2%)	420 (14.1%)	542 (13.9%)	
NHB	103 (11.2%)	705 (23.7%)	808 (20.7%)	
NHW	415 (45.0%)	1401 (47.0%)	1816 (46.5%)	
Other	274 (29.7%)	417 (14.0%)	691 (17.7%)	
Unknown/Missing	9 (1.0%)	36 (1.2%)	45 (1.2%)	
Visit frequency				0.580
1–4	351 (38.0%)	1183 (39.7%)	1534 (39.3%)	
5–8	178 (19.3%)	601 (20.2%)	779 (20.0%)	
9–11	105 (11.4%)	313 (10.5%)	418 (10.7%)	
≥ 12	289 (31.3%)	882 (29.6%)	1171 (30.0%)	
Calendar year of first vaccine				< 0.001
2015–2017	268 (29.0%)	726 (24.4%)	994 (25.5%)	
2018–2020	438 (47.5%)	1345 (45.1%)	1783 (45.7%)	
2021–2023	217 (23.5%)	908 (30.5%)	1125 (28.8%)	

Notes: NHB = Non-Hispanic Black; NHW = Non-Hispanic White. Percentages are column percentages. P-values comparing Complete vs Partial are obtained from t-test for age and chi-squared test for categorical variables.

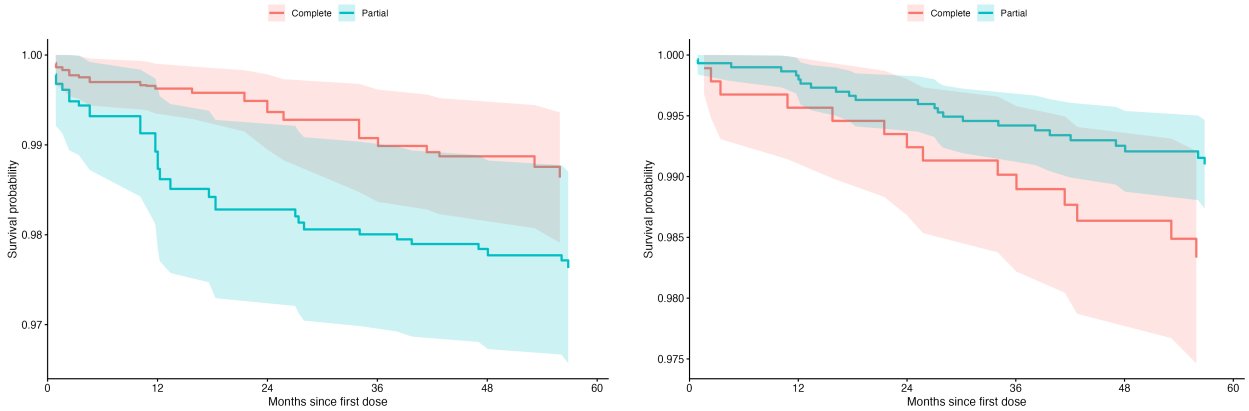

Figure 2: Kaplan–Meier curves for time to incident HPV-related cervical disease by vaccination strategy. The left figure includes the weighted KM curves and the right figure includes the naive KM curves.

weights concentrated near small values and a small proportion of individuals having very large weights. After stabilization, the IPCW still included some values greater than 15. These extreme weights correspond to underrepresented observations that would otherwise contribute little to analysis. Failing to account for them can yield biased point and confidence intervals. After applying CCW with stabilized IPCW, the weighted Cox model yielded an HR of 0.44 (95% CI: 0.18–1.04), aligning with the weighted KM patterns and the expected protective effect of complete vaccination. Collectively, these results suggest that CCW can effectively mitigate the immortal time bias arising from time-dependent dose completion and produce clinically plausible effect estimates.

Table 2: Estimated per-protocol effect of complete vs. partial vaccination (reference)

Scenario (All weights were stabilized)	**HR (95% CI)**	**P-value**
Naive analysis	1.85 (0.95–3.57)	0.068
CCW + IPCW	0.44 (0.18–1.04)	0.063
CCW + IPCW + IPTW	0.42 (0.17–1.03)	0.058
CCW + IPCW + sandwich	0.44 (0.33–0.57)	< 0.001
CCW + IPCW + IPTW + sandwich	0.42 (0.32–0.55)	< 0.001
CCW + IPCW + bootstrap	0.44 (0.34–0.62)	< 0.001
CCW + IPCW + IPTW + bootstrap	0.42 (0.31–0.57)	< 0.001

2.3 IPTW Improves Baseline Confounding Control Beyond IPCW

To account for measured baseline imbalances between vaccination strategies, we applied stabilized IPTW in addition to stabilized IPCW within the CCW framework. Unless otherwise specified, all main analyses reported below and Table 2 used stabilized weights, except for the sensitivity analyses. The estimated effect was similar to IPCW alone: CCW with IPCW + IPTW together yielded an HR of 0.42 (95% CI: 0.17-1.03) (Table 2). We also examined the distribution of the unstabilized and stabilized IPTW and most stabilized weights were below 2.5 (Figure 3), consistent with the minimal change in the hazard ratio after adding IPTW. Although the estimate changed little, IPTW was retained in the primary specification to ensure balance of measured baseline covariates and to support a per-protocol interpretation of the CCW method. More pronounced differences would be expected in settings where the IPTW distribution is more dispersed.

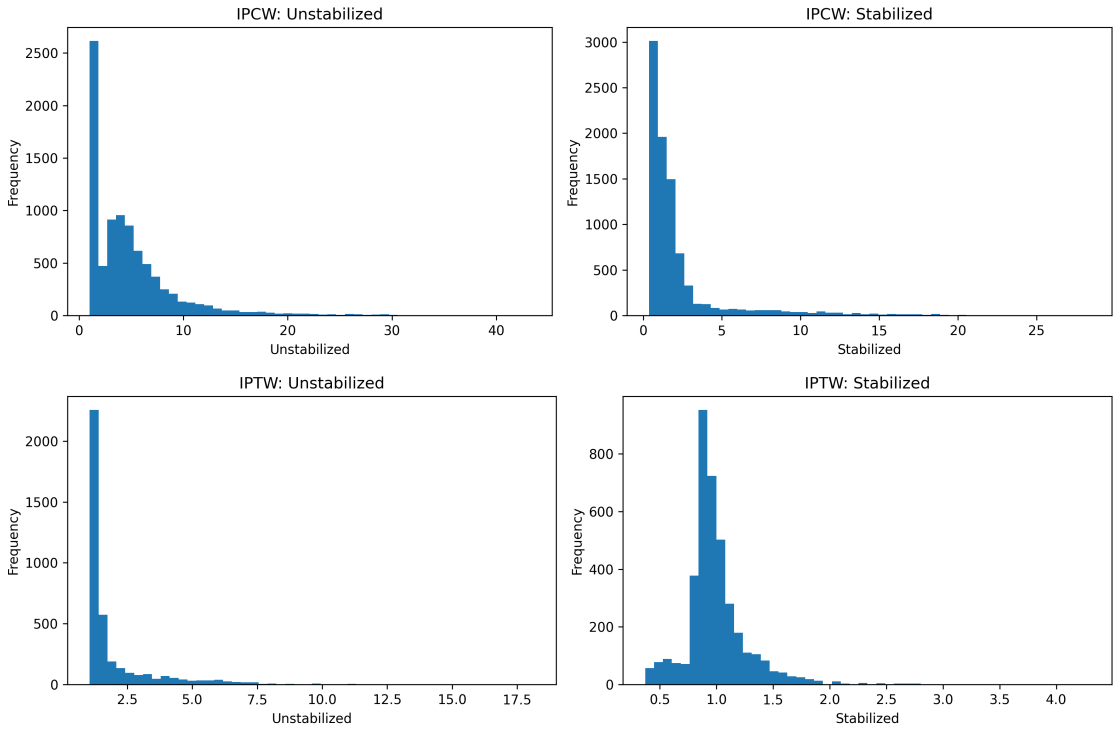

Figure 3: Histogram of unstabilized and stabilized IPCW and IPTW.

2.4 Robust Variance Estimation Under CCW Cloning

When each individual was cloned and assigned to different treatment groups, within-subject correlation was introduced in the expanded dataset. We accounted for such dependence using robust variance estimation via sandwich variance estimators. Point estimates were unchanged, but inference differed materially: CCW with IPCW and a sandwich variance estimator yielded an HR of 0.44 (95% CI: 0.33–0.57), and CCW with IPCW + IPTW and a sandwich variance estimator yielded an HR of 0.42 (95% CI: 0.32–0.55) (Table 2). In comparison, analyses using non-robust variance estimation produced non-significant results despite same hazard ratios, potentially leading to different scientific conclusions regarding whether complete 3-dose vaccination is needed. This contrast highlights the importance of robust inference to account for cloning-induced dependence in CCW.

We assessed robustness of the CCW findings using an alternative variance estimation approach based on the nonparametric bootstrap. Bootstrap resampling was conducted at the individual level, carrying forward all associated clones. Results were consistent with the primary sandwich-based analyses: the HR for CCW with IPCW was 0.44 (95% CI: 0.34–0.62), and the HR for CCW with

IPCW+IPTW was 0.42 (95% CI: 0.31–0.57) (Table 2). Given that the bootstrap is commonly used for complex data structures where asymptotic variance derivations are challenging, the close agreement between sandwich- and bootstrap-based inference supports the validity of our robust variance estimation. However, bootstrap inference is computationally intensive (e.g., requiring approximately 45 minutes in our study), whereas the sandwich variance estimator provides a substantially more efficient approach (e.g., only a few seconds in our study) for routine implementation.

2.5 Sensitivity Analysis

Since the weighting component is fundamental to the CCW approach and our primary interest was in estimating the HR, we focused our sensitivity analyses on alternative weighting specifications within weighted Cox models. Using unstabilized weights with a weighted Cox model and robust sandwich variance yielded an HR of 0.56 (95% CI: 0.40–0.79). This results was consistent in direction with the primary analysis, although the estimated effect was somewhat attenuated. Additionally, using stabilized weights estimated from a more flexible random forest (RF) method produced an HR of 0.42 (95% CI: 0.26–0.68), with a minimum leaf size of 200. We further examined several other specifications for minimum leaf sizes of 20, 50, and 100, which produced estimated HRs of 0.11 (95% CI: 0.07–0.17), 0.18 (95% CI: 0.11–0.30), and 0.30 (95% CI: 0.19–0.48), respectively. These results indicate that the estimated effect was consistent in direction across RF tuning parameter settings, but its magnitude varied. In particular, larger minimum leaf sizes yielded estimates closer to the primary weighted Cox result.

3 Discussion

In target trial emulation (TTE) with time-to-event outcomes and time-varying treatment initiations, immortal time bias remains one of the most pervasive threats to valid causal inference [60, 28]. In this work, we systematically reviewed and formalized the clone–censor–weight (CCW) approach as a principled solution to this problem within the TTE framework. By framing CCW as part of the TTE design, we highlight its role as a design-stage solution to avoid immortal time bias, rather than a post hoc modeling adjustment. Cloning plays a role analogous to creating an unobserved potential outcome trajectory; strategy-specific censoring enforces protocol adherence; and inverse

probability weighting reconstructs the counterfactual risk sets that would have been observed under sustained adherence [29, 13]. This causal framework distinguishes CCW from naive per-protocol comparisons that inadvertently use post-baseline information to define treatment groups.

Although CCW has been increasingly adopted in real-world biomedical research [69, 57, 69, 4], several methodological ambiguities persist. First, confusion often arises regarding whether inverse probability of treatment weighting (IPTW) is required in addition to inverse probability of censoring weighting (IPCW) [24, 19]. The IPTW is used to address baseline confounding in treatment assignment, whereas IPCW plays the primary role in correcting for protocol deviation–induced censoring in CCW analyses. Second, failing to account for the within-individual correlation induced by cloning can lead to biased interval estimation. Proper inference requires robust sandwich variance estimators clustered at the individual level or bootstrap resampling procedures that are conducted prior to cloning [41, 39, 16]. Clarifying these implementation details is essential for promoting rigorous CCW practice and supporting interpretable conclusions.

The real-world example presented in this study highlights both the necessity and the practicality of CCW in biomedical research settings where RCTs are not feasible. From a translational perspective, completion of a 3-dose series of HPV vaccination is influenced by many factors including clinical accessibility, workflows, and structural barriers [14, 36]. These factors also associate with healthcare patterns, such as screening intensity, healthcare utilization, and baseline vulnerability, making bias control essential if causal estimates are to inform clinical practice [58]. CCW clarifies how to attribute early risk time fairly across strategies and how to handle deviations in a way that preserves interpretability for clinicians and guideline stakeholders.

From a clinical and public health perspective, this framework enables more credible comparisons of HPV vaccine dosing strategies using routinely collected EHRs by addressing immortal time bias that can occur when vaccination doses are initiated during follow-up. By leveraging surrogate endpoints along the HPV–cervical disease pathway (e.g., hrHPV persistence and CIN2+) when invasive cancer outcomes are not yet observable, these analyses can expedite evidence generation to inform timely vaccination policy and implementation, helping programs optimize dose timeliness through targeted outreach and resource allocation to high-risk populations [37, 9]. Importantly, completing the 3-dose 9vHPV series is expected to optimize protection against the oncogenic HPV types, consistent with current dosing recommendations [7, 8]. Quantifying the incremental benefit of

full-series completion versus partial vaccination can inform community outreach and programmatic prioritization aimed at reduce HPV-related disparities [47, 2].

Beyond traditional regression models, CCW can accommodate modern machine learning (ML) approaches. Model agnostic ML approaches can be incorporated into the estimation of causal effects and nuisance functions such as the probabilities of censoring and baseline treatment assignment (e.g., the RF in our sensitivity analysis), particularly in high-dimensional EHR data [48, 10]. However, ML application does not obviate the need for careful design configuration and the underlying assumption validity of the CCW framework. Inference in such settings should rely on resampling techniques, such as the nonparametric bootstrap performed at the individual level prior to cloning, to properly account for sampling variability [16].

Our study is not without limitations. First, our work focused on CCW for comparing pre-specified dosing strategies with a fixed decision rule at time zero (e.g., complete 3-dose series within 12-month window vs not), rather than adaptive strategies that update treatment decisions in response to evolving patient outcomes and treatment history [44]. Extending CCW to such dynamic treatment regimen setting remains an important direction for future research, potentially linking CCW with structural nested models or reinforcement learning frameworks [64, 72]. Second, as in all observational data analyses, unmeasured confounding may persist and can cause bias if not adjusted properly [34, 63]. Thus, developing methods to adjust for unmeasured confounders represents an important topic for future research.

As digital health and large-scale EHRs become increasingly central to clinical evidence generation, methodological rigor is a prerequisite for analytic scalability. By formalizing and clarifying the CCW framework, we aim to provide a practical foundation for TTE workflows and strengthen the bridge between causal inference theory and real-world digital medicine applications.

4 Methods

4.1 Rationale for CCW and Immortal Time Bias

4.1.1 Immortal Time Bias

The CCW has been proposed and widely used in real-world studies to address immortal time bias. "Immortal time" refers to a period of follow-up during which the outcome cannot occur by design [60, 68]. In our study comparing complete versus partial vaccination strategies, the time between initiation and completion can constitute immortal time for the complete group: by definition, individuals must remain event-free to receive subsequent doses and be classified as complete. Immortal time bias arises when this pre-completion, event-free time is incorrectly attributed to the complete regimen.

4.1.2 Intent-to-treat and Per-protocol Analysis

In RCTs, the immortal time bias is generally not an issue under the intent-to-treat (ITT) principle, since treatment assignment is predetermined by randomization at baseline [22]. However, immortal time bias can arise under per-protocol (PP) principle when protocol adherence is defined using post-baseline information [29]. Table 3 provides an explanation using our study data. In a hypothetical RCT, patients would be assigned at baseline to one of two vaccination strategies. Suppose Patient 5 is randomized to the partial group. Under an ITT analysis, Patient 5 would remain in the partial group regardless of whether they complete the third dose within 12 months. By contrast, a PP analysis that classifies individuals based on eventual completion within 12 months would reassign Patient 5 to the complete group because dose 3 is received at 6 months, even though the infection occurs at 2 months. This classification relies on future dosing information to define group membership and attributes pre-completion follow-up time to the complete strategy, even though completion has not yet occurred at the time of infection.

4.1.3 Why CCW for Target Trial Emulation

In observational EHR data, treatment strategies are typically realized over follow-up rather than assigned at baseline; hence, an ITT analysis is often not directly applicable, and PP analyses are

Patient	Dose 1	Dose 2	Dose 3	Infection	Received Intervention
1	0 mo	7 mo	–	–	Partial
2	0 mo	3 mo	19 mo	25 mo	Partial
3	0 mo	3 mo	9 mo	–	Complete
4	0 mo	–	–	1 mo	Partial
5	0 mo	2 mo	6 mo	2 mo	Complete
...	...	...	...	...	...

Table 3: Selected examples of the raw data from our study. "mo" denotes months.

commonly used [26, 13]. To restore a trial-like structure and clearly define the causal estimand, target trial emulation (TTE) addresses this challenge by explicitly specifying eligibility criteria, time zero, treatment strategies, follow-up, and censoring rules to mimic a randomized trial using real-world data [6, 28]. When a baseline assignment rule exists (e.g., guideline- or policy-driven eligibility for a dosing strategy), an ITT analysis can be defined using baseline information only. In our setting, however, dosing strategies (e.g., completion of a 3-dose series vaccination within 12-month grace period) are determined by post-baseline dosing information and follow-up history rather than baseline assignment. Consequently, naive PP analyses that group individuals by eventual completion status are susceptible to immortal time bias. CCW was developed to address this problem.

4.2 Target Trial Emulation Configuration

Given the prolonged latency (typically 15-20 years) between HPV infection and the development of invasive cervical cancer [67], direct observation of cancer outcomes would require extended medical records beyond the scope of our data. We therefore defined clinically meaningful surrogate endpoints that capture early disease progression along the HPV–cervical neoplasia pathway [49, 31]. The primary composite outcome was the first occurrence of any of the following: (1) persistent high-risk HPV (hrHPV) infection, defined as two positive hrHPV DNA tests separated by ≥ 6 months; (2) cervical intraepithelial neoplasia grade 2 or higher (CIN2+); or (3) cervical cancer diagnosis. The event date was defined as the earliest date at which any component of the composite outcome was observed.

We emulated a hypothetical randomized trial using our study data by prespecifying the target

trial protocol elements a priori. These elements included eligibility criteria, treatment strategies, the assignment framework, time zero, follow-up and grace periods, censoring rules, and outcome definitions (Table 4). Individuals meeting the inclusion and exclusion criteria entered the emulated trial at time zero, yielding an analytic cohort of $N = 3902$ (Figure 4). By aligning the observational analysis with a prespecified target trial protocol, we aimed to optimize the causal interpretability of the effects of different 9vHPV vaccine dosing strategies on the risk of HPV-related cervical disease.

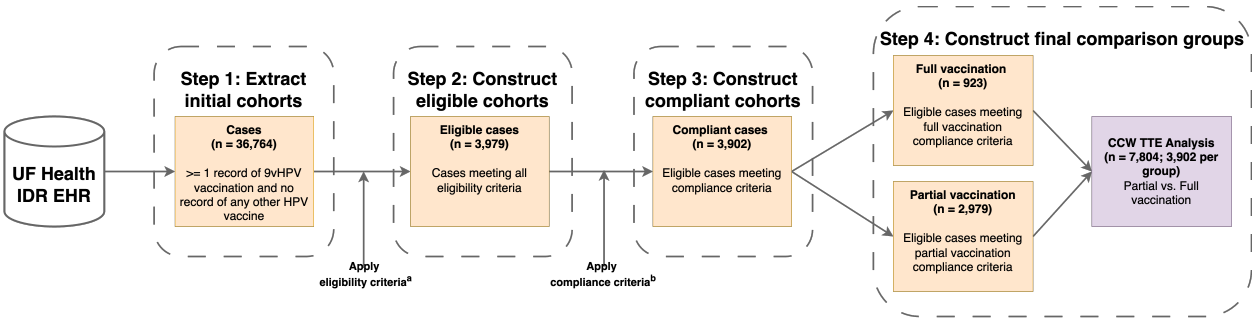

Figure 4: Data workflow for the study. After cloning, the final analyzable sample included 7,804 observations from 3,902 independent subjects.

4.3 Cloning

In CCW, the first step is to create a cloned copy of each eligible individual and assign each "clone" to different hypothetical (or potential) treatment strategies, thus mimicking the baseline assignment of an emulated target trial rather than the treatments eventually received.

In our real example, before receiving the third dose (i.e., completing the dose regimen), an individual's eventual classification as complete versus partial is not yet determined. Classifying individuals based on eventual completion (e.g., labeling those who later complete as "complete" from time zero) would use post-baseline information and would incorrectly attribute early follow-up time. Such procedure can induce immortal time bias, particularly when infections occur before third dose. Under the target trial definition, individuals are assigned at baseline to hypothetical strategies, and prior to the completion of third dose they remain compatible with both strategies because completion has not yet been fully ascertained.

To mitigate this bias, we consider to create a cloned copy of each individual under the alternative treatment strategy to the one observed, while retaining the individual's observed dosing history and covariate profile (Table 5). This approach closely aligns with the classic potential outcome

Table 4: Target trial emulation protocol for 9vHPV vaccination comparison study using clone-censor-weight method

Component	**Specification**
Aim	Compare partially vaccinated (1–2 doses) vs. completing full vaccination within 12 months (3 doses) on the risk of HPV-related cervical disease outcomes among 9vHPV initiators.
Eligibility Criteria	**Included if they:** • were aged 15 to 26 years at the index date • were recorded as female in the EHR • had at least one healthcare encounter during the baseline period (i.e., the year prior to the index date) • had at least one healthcare encounter after the grace period (i.e., the year following the index date) to ensure ongoing clinical activity for outcome assessment • had a grace period that ended at least one year before the end of the study period to ensure sufficient time for outcome assessment **Excluded if they had any baseline record of:** • HIV infection • hysterectomy • immunosuppressive conditions • a positive hrHPV test • abnormal cervical cytology or histology (including CIN2+ and AIS) • cervical cancer • pregnancy • any sexually transmitted infection (STI), including herpesvirus, gonorrhea, chlamydia, trichomoniasis, or other STIs **Compliance criteria:** • receive three or fewer 9vHPV doses • meet ACIP-recommended dose intervals: at least 4 weeks between doses 1 and 2, at least 12 weeks between doses 2 and 3, and at least 5 months between doses 1 and 3
Time Zero (t_0)	Date of first recorded 9vHPV dose in the EHR.
Follow-up and Grace Periods	• Follow-up from initiation until the earliest of outcome, artificial censoring due to deviation from assigned strategy, death, $t_0 + 5$ years, or study end date (02/28/2025) • Grace period is defined as the 12-month window following t_0
Treatment Strategies	**Strategy 1: Partially vaccinated individuals** • receive 1 or 2 doses of 9vHPV • no receipt of a 3rd dose at any time during follow-up • deviation and artificial censoring: if a 3rd dose is received, artificially censored at the 3rd dose date **Strategy 2: Fully vaccinated individuals** • receive exactly 3 doses of 9vHPV • completion requirement: 3rd dose occurs by $t_0 + 365$ days • deviation and artificial censoring: if the 3rd dose is not received by t_0+365 days, artificially censored at t_0+365 days; any late completion after the grace period is treated as protocol deviation
Assignment Framework	Every observation will be copied and assigned to both strategies per CCW procedures.
Outcomes	First occurrence of any of the following after t_0: • persistent hrHPV infection (two positive hrHPV DNA tests separated by $\geq$ 6 months) • CIN2+ • cervical cancer
Causal Contrast	Per-protocol (PP) effect indicated by hazard ratio (HR).
Censoring Rules	Censoring present in the original data will be retained; artificial censoring will be applied at protocol deviation per CCW procedures.

framework in causal inference, where each individual is conceptualized as having counterfactual outcomes under multiple possible treatment strategies over time. At time zero, all individuals have well-defined potential outcomes that are infection-free and uncensored under each treatment strategy. As time progresses, infection occurrence and treatment initiation cause divergence across these potential outcome trajectories, highlighting the need for censoring and weighting.

Patient	**Dose 1**	**Dose 2**	**Dose 3**	**Infection/ Censoring**	**Assigned Strategy**
1	0 mo	7 mo	–	Censored at 12 mo	Complete
1	0 mo	7 mo	–	–	Partial
2	0 mo	3 mo	19 mo	Censored at 12 mo	Complete
2	0 mo	3 mo	19 mo	Infection at 25 mo	Partial
3	0 mo	3 mo	9 mo	–	Complete
3	0 mo	3 mo	9 mo	Censored at 9 mo	Partial
4	0 mo	–	–	Infection at 1 mo	Complete
4	0 mo	–	–	Infection at 1 mo	Partial
5	0 mo	2 mo	6 mo	Infection at 2 mo	Complete
5	0 mo	2 mo	6 mo	Infection at 2 mo	Partial
...	...	...	...	...	...

Table 5: Illustration of cloning and censoring under CCW. Each individual is replicated and assigned at baseline to each treatment strategy. Clones are censored when observed treatment strategy becomes incompatible with the assigned protocol; otherwise, infection events are recorded when they occur. "mo" denotes months.

4.4 Censoring

After cloning, the second step is to censor each cloned copy at the time it deviates from its assigned protocol. This step ensures that disease occurrence and treatment initiation are properly represented over time.

For example, in our data, before month 12 (the prespecified grace period for completing the series), patient 1 is still eligible for both treatment strategies, so follow-up is retained under both. By month 12, however, the patient has not received the third dose, making the observed history incompatible with the complete strategy. Consequently, the clone assigned to the complete strategy is artificially censored at month 12 due to deviation from the strategy definition, and outcomes beyond that time under the complete strategy are not identifiable from the observed data. Patient 2 experiences infection at month 25. Because this patient did not complete three doses within

the 12-month grace period, the clone assigned to the complete strategy is censored at month 12. The infection at month 25 is therefore only observed for the clone in partial strategy. The same compatibility-based censoring rules are applied to all individuals, as showcased in Table 5.

4.5 Weighting

4.5.1 Inverse Probability of Censoring Weighting

The third step is to weight the uncensored individuals to recover the target population under each treatment strategy. This is accomplished using the inverse probability of censoring weighting (IPCW) approach, which accounts for the probability of uncensored observations conditional on treatment and covariate history up to each time point [52, 11].

Causal estimand is in general hard to define using the hazard ratio [23, 15]. Under the target trial framework each individual can be represented by counterfactual counting and at-risk processes under each treatment strategy [62, 55]. We therefore interpret the Cox hazard ratio as a model-based summary of the per-protocol effect comparing the two strategies, rather than as a universally well-behaved causal estimand. However, censoring introduces missingness in these processes. Thus, the IPCW is introduced to up-/down-weight the contribution of observed data in the presence of right censoring. Let $\bar{A}_i(t^-)$ denote the treatment history before t (not including t), $\bar{A}_i(0)$ the baseline treatment assignment, $\bar{X}_i(t^-)$ the covariates history before t, $\bar{X}_i(0)$ the baseline covariates, and C_i the censoring time. The IPCW at time t is defined as

$$w_i^C(t) = \frac{1}{Pr(C_i \geq t \mid \bar{A}_i(t^-), \bar{X}_i(t^-))},$$

which corresponds to the inverse of the probability that the observation remains uncensored up to time t. A common used stabilized weight [11, 27], which reduces the influence of extreme values, is defined as

$$\widetilde{w}_i^C(t) = \frac{Pr(C_i \geq t \mid A_i(0), X_i(0))}{Pr(C_i \geq t \mid \bar{A}_i(t^-), \bar{X}_i(t^-))},$$

where the numerator is changed to the conditional probability conditioning on baseline covariates and treatment.

After adjusting the weights, we construct a weighted counting process and its associated weighted martingale [17]. The weighted martingale represents the counterfactual process under a specific treatment strategy. Thus, applying standard Cox partial likelihood methods with time-varying covariates (Table 3) to the weighted data yield consistent estimates of the hazard ratio [17].

Patient	**Time start**	**Time end**	**Event**	**Strategy**
1	0 mo	12 mo	Censored at 12 mo	Complete
1	0 mo	Max time	Censored at study end	Partial
2	0 mo	12 mo	Censored at 12 mo	Complete
2	0 mo	25 mo	Infection at 25 mo	Partial
3	0 mo	9 mo	Complete at 9 mo	Complete
3	9 mo	Max time	Censored at study end	Complete
3	0 mo	9 mo	Censored at 9 mo	Partial
4	0 mo	1 mo	Infection at 1 mo	Complete
4	0 mo	1 mo	Infection at 1 mo	Partial
5	0 mo	2 mo	Infection at 2 mo	Complete
5	0 mo	2 mo	Infection at 2 mo	Partial

Table 6: Start/end-time format for Cox regression under CCW with Complete vs Partial strategies. Intervals are split at dose receipt when modeling time-varying dose status, and at strategy-specific censoring or infection.

4.5.2 Inverse Probability of Treatment Weighting

Although CCW clones each individual across treatment strategies and makes a false impression of perfect balance at baseline, each clone is actually censored when it deviates from its assigned strategy. Hence, the resulting comparison between treatment strategies will be biased if treatment assignment at baseline is subject to selection bias. As in other observational designs, CCW needs inverse probability of treatment weighting (IPTW) to address this baseline selection bias in treatment assignment [51, 27], in addition to the IPCW used to account for artificial censoring as discussed in the previous subsection. Let $\bar{X}_i(t)$ be the covariates history up to t (including t). The IPTW and the stabilized IPTW are defined as

$$\text{IPTW} : w_i^A(t) = \prod_{s \leq t} \frac{1}{P\left(A_i(s) \mid \bar{A}_i\left(s^-\right), \bar{X}_i(s)\right)},$$

$$\text{Stabilized IPTW} : \widetilde{w}_i^A(t) = \prod_{s \leq t} \frac{Pr\left(A_i(s) \mid \bar{A}_i\left(s^-\right)\right)}{P\left(A_i(s) \mid \bar{A}_i\left(s^-\right), \bar{X}_i(s)\right)}.$$

When the treatment is not time-varying (i.e., strategy is defined at baseline and does not depend on intermediate decisions), there is no longitudinal treatment process $\{A_i(s) : s \leq t\}$ to model. In this setting, the stabilized IPTW simplifies to

$$\widetilde{w}_i^A = \frac{P(A_i)}{P(A_i \mid X_i(0))},$$

where A_i denote the treatment received. Note here we adjust for baseline covariates $X_i(0)$ rather than time-varying covariates measured at follow-ups. This is because conditioning on post-treatment confounding variables could block the causal pathway and thus introduce bias in causal estimation.

The IPTW reflects the probability of receiving the treatment actually observed, conditional on the subjects' covariates. When a subject is cloned, both clones should use the same treatment weights because such weights is determined by the treatment the subject truly received, not by the counterfactual assignment of the clone.

The final weights are obtained as the product of the treatment and censoring weights:

$$w_i(t) = \widetilde{w}_i^A \widetilde{w}_i^C(t),$$

where $\widetilde{w}_i^A$ is the stabilized IPTW shared across cloned samples in different treatment groups, while $\widetilde{w}_i^C(t)$ is the clone-specific, time-dependent IPCW accounting for censoring.

4.6 Statistical inference procedure

After processing the data through cloning, censoring, and weighting, a weighted Cox model can be applied to estimate the hazard ratio comparing the treatment strategies. Other methods such as weighted Kaplan-Meier estimators or weighted accelerated failure time models can also be applied to estimate widely used causal estimands, including restricted mean survival time and cumulative incidence differences [46, 70]. However, a common pitfall is failing to account for the correlation induced by cloning when quantifying the uncertainty of the estimate. Robust sandwich variance estimators or bootstrap methods should be applied to adjust for this correlation [71, 41, 39]. Without such adjustments, the point estimate remains consistent, but standard errors may be biased, leading to invalid confidence intervals and p-values [42]. For weighted Cox model, robust sand-

wich variance estimators provide the fastest and simplest way to implement. Other nonparametric models rely on computationally intense bootstrap to obtain variance estimates [16].

5 Data Description

Our study was conducted using EHR data from the UF Health IDR extracted from the period between January 1, 2012 and February 28, 2025. The UF Health IDR is a clinical data warehouse that consolidates information from UF Health's clinical and administrative systems, including the Epic EHR system. It contains over one billion observational records from more than two million patients, ranging from structured data such as demographics, diagnoses, procedures, laboratory tests, and medications, to unstructured clinical narratives such as discharge summaries, provider notes, and pathology reports.

Patient-level EHR data used in this study are maintained in the UF Health IDR and are not publicly available. Access to the minimum dataset is available to qualified investigators through the UF Health IDR, subject to institutional review and data governance approvals.

References

[1] Victoria L Bartlett, Sanket S Dhruva, Nilay D Shah, Patrick Ryan, and Joseph S Ross. Feasibility of using real-world data to replicate clinical trial evidence. *JAMA network open*, 2(10):e1912869, 2019.

[2] Deborah Bateson, Yin Ling Woo, and Jayashri Kulkarni. Elimination of cervical cancer: ensuring equity. *The Lancet Public Health*, 8(4):e248–e249, 2023.

[3] Eric I Benchimol, Liam Smeeth, Astrid Guttmann, Katie Harron, David Moher, Irene Petersen, Henrik T Sørensen, Erik von Elm, Sinéad M Langan, and RECORD Working Committee. The reporting of studies conducted using observational routinely-collected health data (record) statement. *PLoS medicine*, 12(10):e1001885, 2015.

[4] Abhishek Bhatia, Alexander J Preiss, Xuya Xiao, M Daniel Brannock, G Caleb Alexander, Robert F Chew, Hannah Davis, Megan Fitzgerald, Elaine Hill, Elizabeth P Kelly, et al. Effect of nirmatrelvir/ritonavir (paxlovid) on hospitalization among adults with covid-19: an electronic health record-based target trial emulation from n3c. *PLoS medicine*, 22(1):e1004493, 2025.

[5] Francesc X Bosch, Attila Lorincz, Nubia Muñoz, CJLM Meijer, and Keerti V Shah. The causal relation between human papillomavirus and cervical cancer. *Journal of clinical pathology*, 55(4):244–265, 2002.

[6] Aidan G Cashin, Harrison J Hansford, Miguel A Hernán, Sonja A Swanson, Hopin Lee, Matthew D Jones, Issa J Dahabreh, Barbra A Dickerman, Matthias Egger, Xabier Garcia-Albeniz, et al. Transparent reporting of observational studies emulating a target trial—the target statement. *Jama*, 334(12):1084–1093, 2025.

[7] Centers for Disease Control and Prevention. Human papillomavirus (hpv) vaccination: Information for health care providers, November 2021. Accessed 2026-03-02.

[8] Centers for Disease Control and Prevention. Hpv vaccine recommendations, July 2024. Accessed 2026-03-02.

[9] Hui-Chi Chen, Mark Schiffman, Ching-Yu Lin, Mei-Hung Pan, San-Lin You, Li-Chung Chuang,

Chang-Yao Hsieh, Kai-Li Liaw, Ann W Hsing, Chien-Jen Chen, et al. Persistence of type-specific human papillomavirus infection and increased long-term risk of cervical cancer. *Journal of the National Cancer Institute*, 103(18):1387–1396, 2011.

[10] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 2018.

[11] Stephen R Cole and Miguel A Hernán. Constructing inverse probability weights for marginal structural models. *American journal of epidemiology*, 168(6):656–664, 2008.

[12] Issa J Dahabreh and Kirsten Bibbins-Domingo. Causal inference about the effects of interventions from observational studies in medical journals. *Jama*, 331(21):1845–1853, 2024.

[13] Goodarz Danaei, Luis Alberto García Rodríguez, Oscar Fernández Cantero, Roger W Logan, and Miguel A Hernán. Electronic medical records can be used to emulate target trials of sustained treatment strategies. *Journal of clinical epidemiology*, 96:12–22, 2018.

[14] Cam Escoffery, Kara Riehman, Lesley Watson, A Sandy Priess, Marcie Fisher Borne, Sean Nathaniel Halpin, Carlie Rhiness, Emily Wiggins, and Michelle C Kegler. Facilitators and barriers to the implementation of the hpv vacs (vaccinate adolescents against cancers) program: a consolidated framework for implementation research analysis. *Preventing Chronic Disease*, 16:E85, 2019.

[15] Michael P Fay and Fan Li. Causal interpretation of the hazard ratio in randomized clinical trials. *Clinical Trials*, 21(5):623–635, 2024.

[16] Christopher A Field and Alan H Welsh. Bootstrapping clustered data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(3):369–390, 2007.

[17] Thomas R Fleming and David P Harrington. *Counting processes and survival analysis*. John Wiley & Sons, 2013.

[18] Edouard L Fu. Target trial emulation to improve causal inference from observational data: what, why, and how? *Journal of the American Society of Nephrology*, 34(8):1305–1314, 2023.

[19] Charles E Gaber, Armen A Ghazarian, Paula D Strassle, Tatiane B Ribeiro, Maribel Salas, Camille Maringe, Xabier Garcia-Albeniz, Richard Wyss, Wei Du, and Jennifer L Lund. Demystifying the clone-censor-weight method for causal research using observational data: A primer for cancer researchers. *Cancer Medicine*, 13(23):e70461, 2024.

[20] Charles E Gaber, Kent A Hanson, Sodam Kim, Jennifer L Lund, Todd A Lee, and Eleanor J Murray. The clone-censor-weight method in pharmacoepidemiologic research: foundations and methodological implementation. *Current Epidemiology Reports*, 11(3):164–174, 2024.

[21] Cynthia J Girman, Mary E Ritchey, and Vincent Lo Re III. Real-world data: assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products. *Pharmacoepidemiology and drug safety*, 31(7):717, 2022.

[22] Sandeep K Gupta. Intention-to-treat concept: a review. *Perspectives in clinical research*, 2(3):109–112, 2011.

[23] Miguel A Hernán. The hazards of hazard ratios. *Epidemiology*, 21(1):13–15, 2010.

[24] Miguel A Hernán. How to estimate the effect of treatment duration on survival outcomes using observational data. *Bmj*, 360, 2018.

[25] Miguel A Hernán, Babette Brumback, and James M Robins. Marginal structural models to estimate the causal effect of zidovudine on the survival of hiv-positive men. *Epidemiology*, 11(5):561–570, 2000.

[26] Miguel A Hernán and Sonia Hernández-Díaz. Beyond the intention-to-treat in comparative effectiveness research. *Clinical trials*, 9(1):48–55, 2012.

[27] Miguel A Hernán and James M Robins. *Causal inference*. CRC Boca Raton, FL, 2010.

[28] Miguel A Hernán and James M Robins. Using big data to emulate a target trial when a randomized trial is not available. *American journal of epidemiology*, 183(8):758–764, 2016.

[29] Miguel A Hernán, James M Robins, et al. Per-protocol analyses of pragmatic trials. *N Engl J Med*, 377(14):1391–1398, 2017.

[30] Miguel A Hernán, Brian C Sauer, Sonia Hernández-Díaz, Robert Platt, and Ian Shrier. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *Journal of clinical epidemiology*, 79:70–75, 2016.

[31] Rolando Herrero, Sholom Wacholder, Ana C Rodríguez, Diane Solomon, Paula González, Aimee R Kreimer, Carolina Porras, John Schussler, Silvia Jiménez, Mark E Sherman, et al. Prevention of persistent human papillomavirus infection by an hpv16/18 vaccine: a community-based randomized clinical trial in guanacaste, costa rica. *Cancer discovery*, 1(5):408–419, 2011.

[32] William R Hersh, Mark G Weiner, Peter J Embi, Judith R Logan, Philip RO Payne, Elmer V Bernstam, Harold P Lehmann, George Hripcsak, Timothy H Hartzog, James J Cimino, et al. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Medical care*, 51:S30–S37, 2013.

[33] Warner K Huh, Elmar A Joura, Anna R Giuliano, Ole-Erik Iversen, Rosires Pereira de Andrade, Kevin A Ault, Deborah Bartholomew, Ramon M Cestero, Edison N Fedrizzi, Angelica L Hirschberg, et al. Final efficacy, immunogenicity, and safety analyses of a nine-valent human papillomavirus vaccine in women aged 16–26 years: a randomised, double-blind trial. *The Lancet*, 390(10108):2143–2159, 2017.

[34] Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge university press, 2015.

[35] Elmar A Joura, Anna R Giuliano, Ole-Erik Iversen, Celine Bouchard, Constance Mao, Jesper Mehlsen, Edson D Moreira Jr, Yuen Ngan, Lone Kjeld Petersen, Eduardo Lazcano-Ponce, et al. A 9-valent hpv vaccine against infection and intraepithelial neoplasia in women. *New England Journal of Medicine*, 372(8):711–723, 2015.

[36] Aruna Kamineni, Paula R Blasi, Gabrielle D Gundersen, Malia Oliver, John B Dunn, Denise A Galloway, and Margaret M Madeleine. Barriers to human papillomavirus vaccine series completion among insured individuals in an integrated healthcare setting. *Infectious Diseases: Research and Treatment*, 14:11786337211018712, 2021.

[37] Jill Koshiol, Lisa Lindsay, Jeanne M Pimenta, Charles Poole, David Jenkins, and Jennifer S

Smith. Persistent human papillomavirus infection and cervical neoplasia: a systematic review and meta-analysis. *American journal of epidemiology*, 168(2):123–137, 2008.

[38] Mark J Laan and James M Robins. *Unified methods for censored longitudinal data and causality*. Springer, 2003.

[39] Eric W Lee, Lee-Jen Wei, David A Amato, and Sue Leurgans. Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In *Survival analysis: state of the art*, pages 237–247. Springer, 1992.

[40] Linda E Lévesque, James A Hanley, Abbas Kezouh, and Samy Suissa. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *Bmj*, 340, 2010.

[41] Danyu Y Lin and Lee-Jen Wei. The robust inference for the cox proportional hazards model. *Journal of the American statistical Association*, 84(408):1074–1078, 1989.

[42] Tuo Lin, Ruohui Chen, Jinyuan Liu, Tsungchin Wu, Toni T Gui, Yangyi Li, Xinyi Huang, Kun Yang, Guanqing Chen, Tian Chen, et al. On frequency and probability weights: An in-depth look at duelling weights. *International Statistical Review*, 2024.

[43] Camille Maringe, Sara Benitez Majano, Aimilia Exarchakou, Matthew Smith, Bernard Rachet, Aurélien Belot, and Clémence Leyrat. Reflection on modern methods: trial emulation in the presence of immortal-time bias. assessing the benefit of major surgery for elderly lung cancer patients using observational data. *International journal of epidemiology*, 49(5):1719–1729, 2020.

[44] Susan A Murphy. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65(2):331–355, 2003.

[45] National Cancer Institute. Cervical cancer: Causes, risk factors, and prevention. National Cancer Institute (NCI) website, 2025. Accessed 2026-03-02.

[46] Ai Ni, Zihan Lin, and Bo Lu. Stratified restricted mean survival time model for marginal causal effect in observational survival data. *Annals of epidemiology*, 64:149–154, 2021.

[47] Linda M Niccolai and Caitlin E Hansen. Practice-and community-based interventions to increase human papillomavirus vaccine coverage: a systematic review. *JAMA pediatrics*, 169(7):686–692, 2015.

[48] Romain Pirracchio, Maya L Petersen, and Mark Van Der Laan. Improving propensity score estimators' robustness to model misspecification using super learner. *American journal of epidemiology*, 181(2):108–119, 2015.

[49] S Rachel Skinner, Cosette M Wheeler, Barbara Romanowski, Xavier Castellsagué, Eduardo Lazcano-Ponce, M Rowena Del Rosario-Raymundo, Carlos Vallejos, Galina Minkina, Daniel Pereira Da Silva, Shelly McNeil, et al. Progression of hpv infection to detectable cervical lesions or clearance in adult women: Analysis of the control arm of the viviane study. *International journal of cancer*, 138(10):2428–2438, 2016.

[50] James M Robins and Dianne M Finkelstein. Correcting for noncompliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted (ipcw) log-rank tests. *Biometrics*, 56(3):779–788, 2000.

[51] James M Robins, Miguel Angel Hernan, and Babette Brumback. Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560, 2000.

[52] James M Robins, Andrea Rotnitzky, and Lue Ping Zhao. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the american statistical association*, 90(429):106–121, 1995.

[53] Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

[54] Mark Schiffman, Philip E Castle, Jose Jeronimo, Ana C Rodriguez, and Sholom Wacholder. Human papillomavirus and cervical cancer. *The lancet*, 370(9590):890–907, 2007.

[55] Shaun Seaman, Oliver Dukes, Ruth Keogh, and Stijn Vansteelandt. Adjusting for time-varying confounders in survival analysis using structural nested cumulative survival time models. *Biometrics*, 76(2):472–483, 2020.

[56] Rachel E Sherman, Steven A Anderson, Gerald J Dal Pan, Gerry W Gray, Thomas Gross, Nina L Hunter, Lisa LaVange, Danica Marinac-Dabic, Peter W Marks, Melissa A Robb, et al. Real-world evidence—what is it and what can it tell us? *New England Journal of Medicine*, 375(23):2293–2297, 2016.

[57] Kayoko Shioda, Alexander Breskin, Pravara Harati, Allison T Chamberlain, Toshiaki Komura, Benjamin A Lopman, and Elizabeth T Rogawski McQuade. Comparative effectiveness of alternative intervals between first and second doses of the mrna covid-19 vaccines. *Nature communications*, 15(1):1214, 2024.

[58] Hannah R Simons, Zoe D Unger, Priscilla M Lopez, and Julia E Kohn. Predictors of human papillomavirus vaccine completion among female and male vaccine initiators in family planning centers. *American journal of public health*, 105(12):2541–2548, 2015.

[59] Jerzy Splawa-Neyman, Dorota M Dabrowska, and Terrence P Speed. On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, pages 465–472, 1990.

[60] Samy Suissa. Immortal time bias in pharmacoepidemiology. *American journal of epidemiology*, 167(4), 2008.

[61] U.S. Food and Drug Administration. Gardasil 9 (human papillomavirus 9-valent vaccine, recombinant): Prescribing information. Package insert (FDA), March 2025. Revised 03/2025.

[62] Mark J Van der Laan, Sherri Rose, et al. *Targeted learning: causal inference for observational and experimental data*, volume 4. Springer, 2011.

[63] Tyler J VanderWeele and Peng Ding. Sensitivity analysis in observational research: introducing the e-value. *Annals of internal medicine*, 167(4):268–274, 2017.

[64] Stijn Vansteelandt and Marshall Joffe. Structural nested models and g-estimation: the partially realized promise. *Statistical science*, pages 707–731, 2014.

[65] Erik Von Elm, Douglas G Altman, Matthias Egger, Stuart J Pocock, Peter C Gøtzsche, and Jan P Vandenbroucke. The strengthening the reporting of observational studies in epidemiology

(strobe) statement: guidelines for reporting observational studies. *The lancet*, 370(9596):1453–1457, 2007.

[66] Wenjuan Wang, Davide Ferrari, Gabriel Haddon-Hill, and Vasa Curcin. Electronic health records as source of research data. *Machine Learning for Brain Disorders*, pages 331–354, 2023.

[67] World Health Organization. Human papillomavirus (hpv) and cancer. WHO Fact sheet, December 2025. Accessed 2026-03-03.

[68] Kabir Yadav and Roger J Lewis. Immortal time bias in observational studies. *Jama*, 325(7):686–687, 2021.

[69] Kazuki Yoshida, Jun Liu, Rishi J Desai, Robert J Glynn, Daniel H Solomon, and Seoyoung C Kim. Comparative safety of gout treatment strategies on cardiovascular outcomes using observational data: clone-censor-weight target trial emulation approach. *Epidemiology*, 34(4):544–553, 2023.

[70] Jessica G Young, Mats J Stensrud, Eric J Tchetgen Tchetgen, and Miguel A Hernán. A causal framework for classical statistical estimands in failure-time settings with competing events. *Statistics in medicine*, 39(8):1199–1236, 2020.

[71] Scott L Zeger, Kung-Yee Liang, and Paul S Albert. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, pages 1049–1060, 1988.

[72] Ying-Qi Zhao, Donglin Zeng, Eric B Laber, and Michael R Kosorok. New statistical learning methods for estimating optimal dynamic treatment regimes. *Journal of the American Statistical Association*, 110(510):583–598, 2015.