MISCLASSIFICATION OF TIME-AT-RISK DUE TO FREE DRUG SAMPLE USE: A SIMULATION STUDY

Zilu Zhang, MS,[1] Efe Eworuke, PhD,[2] Margie R. Goulding, PhD,[2] Michael D. Nguyen, MD,[2] Bruce Fireman, MS,[3] Judith C. Maro, PhD[1]

1 Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA; 2 Food and Drug Administration, Silver Spring, MD, USA; 3 Kaiser Permanente Northern California, Oakland, CA, USA

OBJECTIVES

- We examined the potential bias associated with misclassified time-at-risk due to free sample use of study drugs using a simulation framework.

BACKGROUND

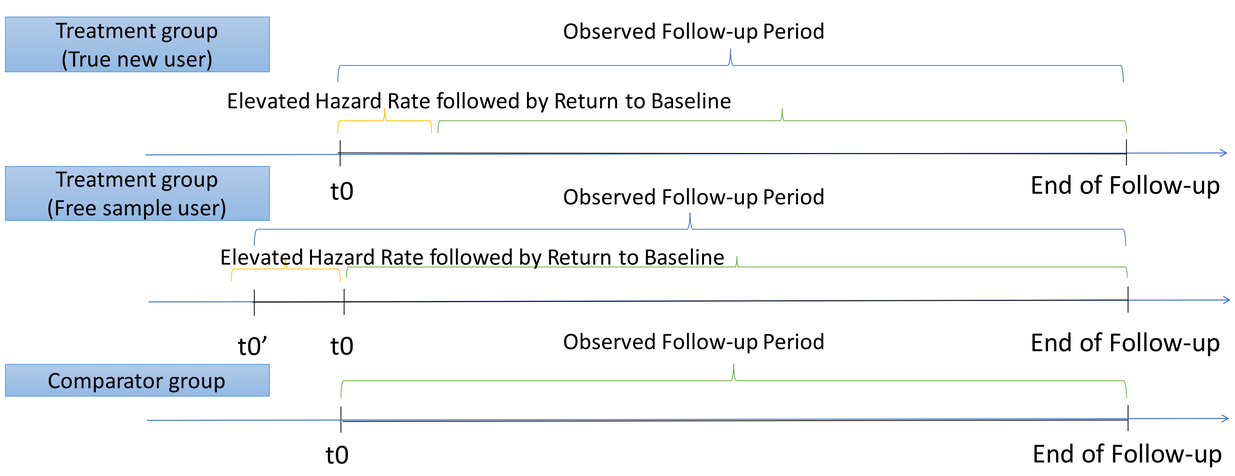
- In epidemiologic studies relying on claims data, the use of free samples of study drugs -- which are generally unobserved in claims data -- may result in both selection bias and misclassification of time-at-risk.
- New users that experience an event while using free samples may discontinue use and never be observed in a claims-based study population (selection bias).
- Apparent new users may be free sample users when they are first observed in claims data with misclassified time-at-risk.
- Misclassified time-at-risk is important when monitoring for acute onset adverse events with a time-varying hazard that is elevated immediately following new use.

METHODS

Simulation Assumptions:

- We assumed that an outcome event that occurs during free sample use causes the user to cease treatment and therefore, precludes entry to our observed cohort.
- We also assumed there is no free sample use in the comparator drug group, e.g. comparator drug is generic.

Figure 1. Follow-up timeline for the simulated cohort



METHODS CONT.

We simulated follow-up for angioedema following initiation of angiotensin-converting enzyme inhibitors (ACEI) as compared to beta-blockers based on aggregate data in the Sentinel Distributed Database. We investigated various scenarios with different settings of misclassification of time-at risk (Table 1). We adjust the data generating mechanism so that risk of an outcome during follow-up in the comparator group is 15%. The total simulated sample size is 1,000,000 persons in a 1:1 matching scenario with 500 replicates. The Hazard Ratios are estimated using Cox proportional hazard model.

Simulation Algorithm:

We simulated the survival cohort following the steps:

We simulated the exposure level indicator follows a binomial distribution with half target drug user and half comparator drug user in the true cohort:

$$A \sim \text{Binomial}(N, p = 0.5)$$

Where $N = n_{treatment} + n_{comparator}$

Among the treatment group, we simulated a binary indicator for true new drug user follows a binomial distribution:

$$A_{newuser} \sim \text{Binomial}(n_{treatment}, p = p_{newuser})$$

Where $n_{treatment} = n_{newuser} + n_{freesampleuser}$

We generated the linear combinations based on the intercept and the exposure level, and then created the event time which follows the exponential distribution:

$$T^E = -\frac{\log(u)}{\lambda * \exp(\text{linPred}_0 + A * b_1)}$$

Where λ is the scale parameter of the exponential distribution (i.e. baseline hazard) which is presumed to be 0.05.

We created the censoring time with similar settings with flat rate:

$$T^C \sim \exp(RateC)$$

Where RateC is set to be $\frac{1}{1000} = 0.001$.

Therefore the follow-up time will be $min(T^E, T^C)$. We censored the subject when $T^C < T^E$.

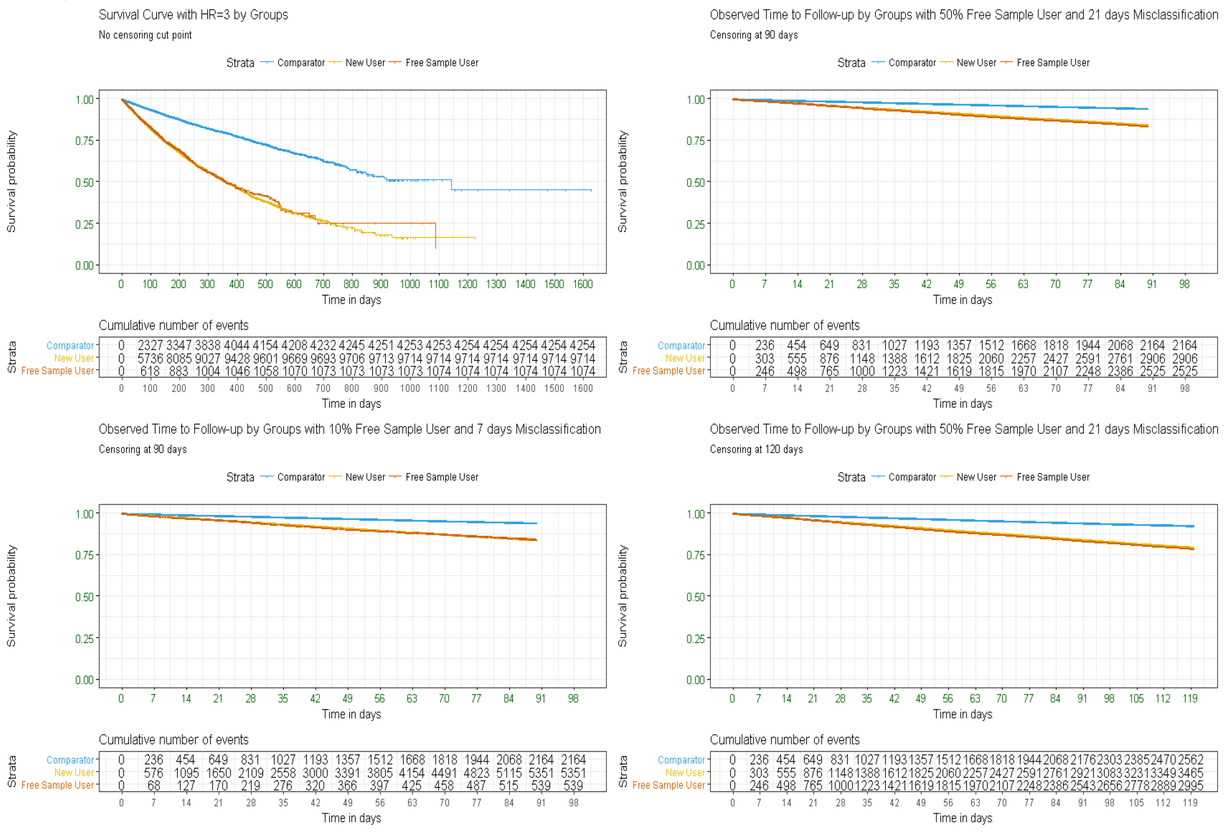
We estimated hazard ratios among users of the study drug who take free samples and have misclassified days-at-risk (Figure 1). We repeated this simulation with various specified proportions of free sample users (10%, 25% and 50%) and various degrees of misclassification in days-at-risk (7, 14, 21, 28) and compared all the HR estimates with the true hazard ratio (e.g. HR=3). We censored the observed cohort at 90 days which is the end of follow-up.

In some scenarios, we limited the days of the imposed elevated risk. That is, we assume a higher risk immediately following study drug exposure and through a pre-specified risk window and lowered the risk thereafter (e.g. HR=3 during first 14 days and HR=1 thereafter). To estimate the HR during the risk interval, we included in the Cox model a risk-window-by-treatment interaction term.

RESULTS

- The result from 500 Monte Carlo simulations shows that if the risk is constant in both groups throughout the follow-up, then the misclassified time-at-risk would not bias the hazard ratio (HR) estimates, but selection bias is still a concern. The observed hazard ratio is very close to the true hazard ratio (e.g. HR=3, and the bias is 0.0003 for 10% of free sample user and 21 misclassified days). Similarly, when we vary the proportion of free-sample use, days-at-risk and censoring point, the estimated HRs are always close to the true HR (Figure 2).
- With a constant hazard, selection bias would arise if free sample users discontinue use before appearing in the claims data and discontinuation is associated with risk.

Figure 2. Example survival curves for 4 simulation scenario with constant risk over time, sample size is 10,000



Misclassified time-at-risk will introduce bias when:

- *Baseline risk changes over time*
- *The hazard ratio changes over time*
- *Censoring is informative (related to risk)*

1. We examined various scenarios with elevated risk during the early follow-up period. The bias for the scenarios with a more modestly elevated hazard ratio (true HR=3 and 1 thereafter) is smaller than the scenarios with a large hazard ratio (true HR=5 and 1 thereafter) when controlling for other factors (Table 1).
2. There will be more bias when there is a greater number of misclassified days-at-risk that occur during the elevated risk period (Table 1).
3. The direction of the bias is always towards the null (Table 1).

Table 1: Results of estimated hazard ratio and bias of selected misclassification scenarios

Selected result with 500 replicates	Observed HR during the imposed risk window	True HR during the imposed risk window	Bias during the imposed risk window
25% free sample user, 21 misclassified days, HR=3	3.001	3.000	0.001
25% free sample user, 28 imposed risk window, 7 misclassified days, HR=3 and 1 thereafter	2.586	3.000	-0.414
25% free sample user, 28 imposed risk window, 7 misclassified days, HR=5 and 1 thereafter	3.677	5.000	-1.323
25% free sample user, 28 imposed risk window, 21 misclassified days, HR=3 and 1 thereafter	2.514	3.000	-0.486

- Moreover, the misclassification would bias our HR estimates either when there is an increasing baseline risk, or when there is informative censoring (censoring relies on the exposure level).

DISCUSSION

In this simulation setting, we assume there is no user in the comparator drug group who received free samples before their initial claim. The simulated population is based on a matched cohort and did not take into account problems related to residual confounding.

If the risk increases over time then the bias will result in estimates that deviate from the null.

CONCLUSION

- The presence of misclassified new users with an undocumented history of free sample use will bias estimates of the hazard ratio.
- A quantitative bias analysis framework allows investigators to vary the extent of the misclassification to account for the potential bias when there are baseline risk changes over time, the hazard ratio changes over time, or censoring is informative.

REFERENCES

1. Jonsson Funk M, Landi S. Misclassification in Administrative Claims Data: Quantifying the Impact on Treatment Effect Estimates. Curr Epidemiol Rep 2014; **1**: 175–185. doi:10.1007/s40471-014-0027-z.
2. Hampp C, Greene P, Pinheiro S. Use of Prescription Drug Samples in the USA: A Descriptive Study with Considerations for Pharmacoepidemiology. Drug Saf 2016; **39**: 261–270. doi:10.1007/s40264-015-0382-9.
3. Jacobus S, Schneeweiss S, Chan KA. Exposure Misclassification as a Result of Free Sample Drug Utilization in Automated Claims Databases and its Effect on a Pharmacoepidemiology Study of Selective COX-2 Inhibitors. *Pharmacoepidemiol Drug Saf* 2004; **13**: 695–702. doi:10.1002/pds.981.
4. Ball R, Robb M, Anderson S, Dal Pan G. The FDA's Sentinel Initiative—A Comprehensive Approach to Medical Product Surveillance. *Clin Pharmacol Ther* 2016; **99**: 265–268. doi:10.1002/cpt.320.
5. Gamble JM, McAlister F, Johnson J, Eurich, D. Quantifying the Impact of Drug Exposure Misclassification due to Restrictive Drug Coverage in Administrative Database: A simulation Cohort Study. Value in Health 2012; **15**: 191-197.

ACKNOWLEDGEMENTS/DISCLOSURES

- This work was supported by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I.
- The authors have no conflicts of interest to disclose.